

# **1** Trend 1: New forms of accountability for a new era of government

---

This chapter introduces the trend on the increasing adoption of Artificial Intelligence in the design and delivery of policies and services. It highlights the need and drive towards ensuring that the algorithms and underlying data avoid bias and discrimination, and that public servants understand data ethics. Finally, it provides practical examples and global case studies to help governments and their partners understand, test and embed new approaches to AI accountability in an effort to enhance transparency and reinforce trust with citizens.

---

Governments are increasingly adopting Artificial Intelligence in the design and delivery of policies and services. This is accompanied by efforts to ensure that the algorithms and underlying data avoid bias and discrimination and that public servants understand data ethics. Several forward-thinking governments and external ecosystems actors are promoting algorithmic accountability, emphasising transparency and explainability, with a view to building trust with citizens. Beyond algorithms, governments are promoting new concepts of transparency with the evolution of Rules as Code—open and transparent machine-consumable versions of government rules – and shedding light on the Internet of Things, which has embedded often-invisible sensors in public spaces. While promising, innovative policy efforts in these areas are often scattered and lack coherence, limiting the potential for collective learning and the scaling of good ideas. This underlines the need for further work on these topics, including fostering international alignment and comparability.

## Algorithmic accountability

Artificial Intelligence (AI) is reshaping economies, promising to generate productivity gains, improve efficiency and lower costs. As governments determine national strategic priorities, public investments and regulations, they hold a unique position in relation to AI. Many have acknowledged the economic importance and potential of AI, with AI strategies and policies now in place in more than [60 countries](#) worldwide.

The [OECD.AI](#) Policy Observatory has taken the lead in advancing OECD's AI-related efforts. An important milestone was the adoption of the [OECD AI Principles](#) in 2019. This pioneering set of intergovernmental standards on AI stresses the importance of ensuring that AI systems embody human-centred values, such as fairness, transparency, explainability and accountability, among others.

The majority of national AI strategies recognise the value of adopting AI in the public sector, alongside the need to mitigate its risks (OECD/CAF, 2022<sup>[1]</sup>; Berryhill et al., 2019<sup>[2]</sup>). In fact, governments are increasingly using AI for public sector innovation and transformation, redefining how they design and deliver policies and services. While the potential benefits of AI in the public sector are significant, attaining them is not an easy task. The field is complex and has a steep learning curve, and the purpose and context of government presents unique challenges. In addition, as in other sectors, public sector algorithms and the data that underpin them are vulnerable to bias, which may cause harm, and often lack transparency.

The OECD Open and Innovative Government Division (OIG) has undertaken extensive work on the use and implications of AI and Machine Learning (ML) algorithms in the public sector to help governments maximise the positive potential impacts of AI use and to minimise the negative or otherwise unintended consequences (see examples [here](#), [here](#) and [here](#)). Other organisations, including the European Commission, have also [reviewed](#) and [examined](#) the expanding landscape of AI in the public sector.

However, the rapid growth in government adoption of AI and algorithmic approaches underlines the need to ensure they are used in a responsible, ethical, trustworthy and human-centric manner. Perhaps more than any other sector, governments have a higher duty of care to ensure that no harm occurs as a result of AI adoption. Such potential consequences include the perpetuation of “Matthew effects”, whereby “privileged individuals gain more advantages, while those who are already disadvantaged suffer further” (Herzog, 2021<sup>[3]</sup>). For instance:

- The “Toeslagenaffaire” was a [child benefits scandal](#) in the Netherlands, where the use of an algorithm resulted in tens of thousands of often-vulnerable families being wrongfully accused of fraud and even hundreds of children being [separated from their families](#), resulting in the [collapse](#) of the government.

- Australia’s “[robodebt scheme](#)” leveraged a data-matching algorithm to calculate overpayments to welfare recipients, resulting in 470 000 incorrect debt notices totalling EUR 775 million being sent. This led to a national scandal and a [Royal Commission](#) after many welfare recipients were required to pay undue debts.
- In the United States, the use of facial recognition algorithms by police has resulted in [wrongful arrests](#), while [bias](#) has been uncovered in criminal risk assessment algorithms that help guide sentencing decisions, resulting in harsher sentences for Black defendants.
- Serbia’s 2021 Social Card law allows for the collection of data on social assistance beneficiaries using an algorithm to examine their socio-economic status. As a consequence, over 22 000 people have lost benefits without an explanation, resulting in legal petitions by a network of advocacy groups (Caruso, 2022<sup>[4]</sup>).

Government have sought to address this issue in a variety of ways, including outright bans on some types of algorithms. For instance, in Washington, DC, the proposed “[Stop Discrimination by Algorithms Act](#)” prohibits the use of certain types of data in algorithmic decision making, and at least [17 cities](#) in the United States and even entire countries, such as [Morocco](#), have implemented bans on government usage of facial recognition. However, a number of have since [backtracked](#), and OECD OPSI-MBRCGI’s prior report on [Public Provider versus Big Brother](#) shows that while authoritarian governments have employed algorithms as a means of social control (e.g. China’s [Social Credit System](#)), others have applied them in legitimate ways to deliver better outcomes for the public. Some even argue that algorithmic decision making can counteract unaccountable processes and offers “a viable solution to counter the rise of populist rhetoric in the governance arena” (Cavaliere and Romeo, 2022<sup>[5]</sup>).

While algorithms can indeed introduce bias and discrimination, so can humans. Indeed, algorithms can systematise the human bias observed in human decisions (Salvi del Pero, Wyckoff and Vourc’h, 2022<sup>[6]</sup>). The key to prevention is having the right safeguards and processes in place to ensure ethical and trustworthy development and use of AI technologies and to mitigate potential risks and biases, as emphasised by the 2023 [European Declaration on Digital Rights and Principles for the Digital Decade](#). One example of this approach is **algorithmic accountability**.

*Algorithmic accountability means “ensuring that those that build, procure and use algorithms are eventually answerable for their impacts.”*

*Source: The Ada Lovelace Institute, AI Now Institute and Open Government Partnership*

Broadly speaking, accountability in AI means that AI actors must ensure that their AI systems are trustworthy. To achieve this, accountable actors need to govern and manage risks throughout their AI systems (OECD, forthcoming-a, *Towards accountability in AI*). The concept of algorithmic accountability more specifically is rooted in “[transparency and explainability](#)” and broader “[accountability](#)”, values that are integral to the [OECD AI Principles](#). However, current legal and regulatory frameworks around the world lack clarity regarding these values, especially about the use of algorithms in public administrations. For instance, the European Union (EU)’s General Data Protection Regulation (GDPR) provides rules and remedies related to algorithmic decisions, but the question of whether explainability is also a requirement has given rise to much debate (Busuioc, 2021<sup>[7]</sup>). The EU’s Digital Services Act ([DSA](#)) (passed in July 2022), Canada’s proposed Artificial Intelligence and Data Act ([AIDA](#)), and the United States’ proposed Algorithmic Accountability Act ([AAA](#)) all include requirements for enhanced transparency for algorithms, but are generally aimed at companies, leaving the question of how public administrations should use algorithms open to interpretation. The proposed EU Artificial Intelligence Act ([AI Act](#)) and the related EU [AI Liability Directive](#), however, offer significant potential for algorithmic accountability in the public sector (Box 1.1).

### Box 1.1. Algorithmic accountability in the proposed AI Act and AI Liability Directive

Proposed in 2021, the **AI Act** is the first piece of regulation that specifically addresses the risks of AI. The Act tackles gaps in current European legal frameworks by adopting a risk-based approach. It sets four levels of risk for AI: unacceptable risk, high risk, limited risk, and minimal or no risk.

All forms of AI deemed to present an unacceptable risk will be banned as a threat to people's rights, safety and livelihoods. Those in the high risk category, such as biometric identification systems, will be subject to stricter obligations, which include appropriate human oversight measures, high-quality training datasets, and risk assessment and mitigation mechanisms. Limited risks AI systems will comply with lighter obligations that focus on transparency and ensure that users are aware that they are interacting with a machine. The use of minimal or no risk AI systems, which constitute the majority of those currently used in the EU, will be free or restriction.

To establish a shared framework to address the legal consequences of harms caused by AI systems, in September 2022 the Commission proposed the **AI Liability Directive**. With this policy the Commission wants to ensure that victims of harm caused by AI are not less protected than those of traditional technologies. The policy would decrease the burden of proof for victims, establish a "presumption of causality" against the developer, provider or user of the AI system, and make it simpler for victims to obtain information about high-risk systems – as defined by the AI Act – in court.

Source: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>,  
[https://ec.europa.eu/commission/presscorner/detail/en/ip\\_22\\_5807](https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5807).

As the international landscape continues to evolve and solidify, a number of forward-thinking governments worldwide are promoting algorithmic accountability, led largely by oversight and auditing entities, as well policy-making bodies often located at the centre of government. External ecosystems actors are also taking note and are working to ensure the use of algorithmic approaches in government meet the higher duty of care required of the public sector. However, despite these promising approaches, more needs to be done to build alignment among disparate definitions and practices around the world.

### ***From the inside-out: Innovative government efforts in algorithmic accountability***

Independent **oversight entities** have a critical role to play in auditing the use of algorithms in the public sector. Such algorithmic accountability can be seen in a variety of examples from around the world:

- In a first for the Latin American region, the independent [Chilean Transparency Council](#) is developing an open and participatory design for a binding "[General Instruction on Algorithmic Transparency](#)" for public entities. A public consultation is expected for 2023.
- The Netherlands Court of Auditors (NCA) has made significant advances in both front-end and back-end aspects of algorithmic accountability. In 2021, it developed an [audit framework](#) that assesses whether algorithms meet quality criteria. 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.
- In 2022, Spain created an independent Spanish Artificial Intelligence Supervision Agency (Box 1.2), and the Netherlands [launched](#) a similar entity in 2023. The draft AI Act (Box 1.1) also calls for a supervisory European Artificial Intelligence Board (EAIB).
- In an example of a successful [cross-border collaboration](#), in 2020 the Supreme Audit Institutions (SAIs) of Finland, Germany, the Netherlands, Norway and the United Kingdom (UK) collectively issued [Auditing machine learning algorithms: A white paper for public auditors](#).
- The United State Government Accountability Office (GAO) in 2021 issued [Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities](#).

### Box 1.2. Spanish Artificial Intelligence Supervision Agency (AESIA)

[AESIA](#) was enacted by law in mid-2022 and is the first dedicated national government agency of the EU to implement a direct mandate on supervising, monitoring and building rules on AI, both for the public sector and beyond. The agency was established in response to a proposed AI Act requirement on implementing national supervisory authorities to ensure the application and implementation of the rule of law concerning AI, and to help achieve Spain's [National AI Strategy](#). The new agency seeks to build a tailored Spanish vision and jurisprudence that could serve as a model for future European AI agencies.

The development of AESIA is a two-step process. First, the design of auditing and implementing guides is key to mainstreaming its vision, and to gathering evidence directly from users needed to build reliable and human-centred regulatory tools. Second, by creating new sandboxes (and incorporating [existing ones](#)) with a focus on AI and algorithmic accountability, the agency will be able to test its own rules with a view to achieving objectivity and ensuring that the balance between protecting human and digital rights, and economic interests, is maintained, and that the legitimate interests of all parties are met.

AESIA is expected to be fully operational by late 2023.

Source: [www.boe.es/buscar/act.php?id=BOE-A-2021-21653](http://www.boe.es/buscar/act.php?id=BOE-A-2021-21653), Interview with AESIA officials.

To obtain a complete picture of new forms of accountability, such as algorithmic accountability, it is necessary to look at other players in the public sector innovation and accountability ecosystems. Perhaps the most relevant of these are **policy-making offices** which set the rules that public sector organisations must follow. One recent example is the October 2022 US White House [Blueprint for an AI Bill of Rights](#), which includes five principles and associated practices to protect against harm – although the blueprint has received criticism for excluding law enforcement from its scope. Similarly, Spain's [Charter on Digital Rights](#) includes 28 sets of rights, many of which relate directly to ethical AI and algorithmic accountability, such as “conditions of transparency, auditability, explainability, traceability, human oversight and governance”.

Figure 1.1. AI Bill of Rights Principles in the United States



Source: [www.whitehouse.gov/ostp/ai-bill-of-rights](http://www.whitehouse.gov/ostp/ai-bill-of-rights).

Additional relevant examples include:

- In late 2021, the UK Cabinet Office's Central Digital and Data Office issued one of the world's first [national algorithmic transparency standards](#), which is being piloted with a handful of agencies (see full case study later in this publication). Relatedly, the United Kingdom, through [The Alan Turing Institute](#) (see p. 167 of OPSI's [AI primer](#) for a case study on its Public Policy Programme), has also created an excellent [AI Standards Hub](#) to advance trustworthy AI through standards such as the Algorithmic Transparency Standard.
- Canada's [Directive on Automated Decision Making](#), issued by the Treasury Board Secretariat, requires agencies using or considering any algorithm that may yield automated decisions to complete an Algorithmic Impact Assessment. This questionnaire calculates a risk score which in turn prescribes actions that must be taken (OPSI's [report](#) on AI in the public sector includes a full case study). Canada's Algorithmic Impact Assessment has inspired similar mechanisms in [Mexico](#) and [Uruguay](#).
- France's Etalab has issued [guidance](#) on Accountability for Public Algorithms, which [sets out](#) how public organisations should report on their use to promote transparency and accountability. The guidance proposes six principles for the accountability of algorithms in the public sector, among other elements.
- The Netherlands' Ministry of Interior and Kingdom Relations has created a Fundamental Rights and Algorithms Impact Assessment ([FRAIA](#)), which facilitates an interdisciplinary dialogue to help map the risks to human rights from the use of algorithms and determine measures to address these risks.
- At the local level, policy offices in the cities of [Helsinki, Finland](#) and [Amsterdam, the Netherlands](#) have developed AI registers to publicly catalogue AI systems and algorithms, while a policy office in Barcelona, Spain, has developed a [strategy](#) for ethical use of algorithms in the city. Based on Helsinki and Amsterdam's work, in 2023 nine cities have collaborated through the [Eurocities](#) network to create an algorithmic transparency [standard](#).

In addition to these internal government approaches, countries have adhered to non-binding international recommendations and principles for responsible and ethical AI that could guide this work. Such examples include the aforementioned [OECD AI Principles](#) and UNESCO's [Recommendation on the Ethics of AI](#). The development of such instruments continues, for example through the Council of Europe, which has a [committee](#) dedicated to AI (CAI) that is developing a [Legal Instrument on Artificial Intelligence, Human Rights, Democracy and the Rule of Law](#). In regard to accountability, the OECD.AI Policy Observatory is working to make principles more concrete through its [Working Group on Tools & Accountability](#) and collaboration around the prototype [OECD-NIST Catalogue of AI Tools & Metrics](#).

Alongside these initiatives scoped specifically around AI and algorithms, the application of broader open-by-default approaches can help make governments algorithms more accountable to their people. In this regard, the [OECD Good Practice Principles for Data Ethics in the Public Sector](#) underscore the need to make source code openly available for public scrutiny and audit and the need for more control over data sources informing AI systems (see Box 1.3). Other examples in this area include the Open Source Software [initiative](#) implemented by Canada in the context of its OGP Action Plan, as well as France's application of [open government](#) in the context of public algorithms.

### Box 1.3. OECD Good Practices Principles for Data Ethics in the Public Sector

- Manage data with integrity.
- Be aware of and observe relevant government-wide arrangements for trustworthy data access, sharing and use.
- Incorporate data ethical considerations into governmental, organisational and public sector decision-making processes.
- Monitor and retain control over data inputs, in particular those used to inform the development and training of AI systems, and adopt a risk-based approach to the automation of decisions.
- Be specific about the purpose of data use, especially in the case of personal data.
- Define boundaries for data access, sharing and use.
- Be clear, inclusive and open.
- Publish open data and source code.
- Broaden individuals' and collectives' control over their data.
- Be accountable and proactive in managing risks.

Source: <https://oe.cd/dataethics>.

### ***From the outside-in: Broader ecosystems strengthening accountability***

While innovative and moving in the right direction, government algorithmic accountability efforts are currently scattered and lack coherence, which limits the potential for collective learning and the scaling of good ideas and successful approaches. The first step in bringing the global discussion on public sector algorithmic accountability into alignment is understanding the different approaches and developing a baseline for action. Some excellent work has already been done in this area, with the joint report of the independent Ada Lovelace Institute, AI Now Institute and OGP representing “the first global study of the initial wave of algorithmic accountability policy for the public sector” (Ada Lovelace Institute, AI Now Institute and Open Government Partnership, 2021<sup>[8]</sup>). Their work surfaced over 40 specific initiatives, identified challenges and successes of policies from the perspectives of those who created them, and synthesised some findings on the subject.

### Box 1.4. Six determinants for the effective deployment of algorithmic accountability

1. Clear **institutional incentives and binding legal frameworks** can support consistent and effective enforcement of accountability mechanisms, supported by reputational pressure from media coverage and civil society activism.
2. Algorithmic accountability policies need to clearly define the **objects of governance** as well as establish **shared terminologies** across government departments.
3. Setting the appropriate **scope of policy application** supports their adoption. Existing approaches for determining scope such as risk-based tiering will need to evolve to prevent under- and over-inclusive application.
4. Policy mechanisms that focus on transparency must be **detailed and audience appropriate** to underpin accountability.

5. Public participation supports policies that meet the needs of affected communities. Policies should **prioritise public participation** as a core policy goal, supported by appropriate resources and formal public engagement strategies.
6. Policies benefit from **institutional co-ordination** across sectors and levels of governance to create consistency in application and leverage diverse expertise.

Source: <https://www.adalovelaceinstitute.org/report/algorithmic-accountability-public-sector>.

Additional relevant work has been conducted by the International Organization for Standardization (ISO) and the Institute of Electrical and Electronics Engineers (IEEE) to develop technical standards or quality specifications approved by a recognised standardisation body. These can be powerful tools to ensure that AI systems are safe and trustworthy, and include, for instance, [ISO/IEC TR 24028:2020](#) on trustworthiness in AI and IEEE's Ethics Certification Program for Autonomous and Intelligent Systems ([ECPAIS](#)). Furthermore, the Association for Computing Machinery ([ACM](#)), the world's largest scientific and educational computing society, through its global Technology Policy Council, has issued a set of [Principles for Responsible Algorithmic Systems](#), which focus on relevant issues such as legitimacy and competency, minimising harm, transparency, explainability, contestability and accountability. The principles are accompanied by guidance on how to apply them while considering governance and trade-offs.

External actors in the accountability ecosystem are also working to hold governments accountable, or to assist them in doing so. Accountability ecosystems encompass “the actors, processes and contextual factors, and the relationships between these elements, that constitute and influence government responsiveness and accountability, both positively and negatively” (Halloran, 2017<sup>[9]</sup>). This shift towards transparency and accountability combined with ever-growing [Civic Tech](#), [Public Interest Tech](#) and [GovTech](#) movements, have expanded accountability ecosystems to incorporate a complex fabric of civil society organisations, academic institutions, private companies and individual members of the public. When leveraged well through partnerships, external ecosystem actors can even help governments compensate for a lack of institutional capacity in this space, as seen in the OECD's work with cities (OECD, 2021<sup>[10]</sup>).

As governments continue to push for more transparency in source code and algorithms, the interactions within these broader accountability ecosystem actors are poised to grow. A cluster of interesting examples of this dynamic can be seen in the Netherlands, which is shaping up to be a leader in algorithmic accountability both inside and outside government. [Algorithm Audit](#) is a Dutch nonprofit organisation that strives for “ethics beyond compliance”. It builds and shares knowledge about ethical algorithms, and includes independent audit commissions that shed light on ethical issues that arise in concrete use cases of algorithmic tools and methods. In another example, the [Foundation for Public Code](#)'s “codebase stewards” help governments publish transparent code in alignment with its [Standard for Public Code](#), which aims to enhance trustworthy codebases.

Additional relevant examples include:

- **European Digital Rights (EDRi)**, the biggest European network defending rights and freedoms online, consisting of 47 non-governmental organisation members and dozens of observers.
- [AI Sur](#), a consortium of organisations that work in civil society and academia in Latin America, which seek to strengthen human rights in the digital environment of the region.
- [AlgorithmWatch](#), a non-profit research and advocacy organisation committed to watching, unpacking and analysing automated decision-making systems and their impact on society.



The emergence of a growing body of GovTech startups (see Box 1.5 for a definition) is also helping governments and other organisations achieve algorithmic accountability (Kaye, 2022<sup>[11]</sup>). [Forbes](#) has listed the rise of GovTech startups as one of the five biggest tech trends transforming government in 2022, and there are signs of these companies entering the algorithmic accountability space. For instance, [Arthur](#), [Fiddler](#), [Truera](#), [Parity](#) and others are actively working with organisations on explainable AI, model monitoring, bias identification and other relevant issues. While most activities so far appear to support private sector companies, the public sector potential is significant, as is evident in the [selection](#) of Arthur by the United States Department of Defense (DoD) to monitor AI accuracy, explainability and fairness in line with the DoD's [Ethical AI Principles](#).

### Box 1.5. Definition of GovTech

GovTech is the ecosystem in which governments co-operate with startups, SMEs and other actors that use data intelligence, digital technologies and innovative methodologies to provide products and services to solve public problems.

They propose new forms of public-private partnerships for absorbing digital innovations and data insights to increase effectiveness, efficiency and transparency in the delivery of public services.

Source: <http://scioteca.caf.com/handle/123456789/1580>.

The emergence of external accountability ecosystem actors is a positive development. One of the most positive outcomes of algorithmic accountability policies and processes, such as the [Open Government Data](#) efforts that preceded them, is to empower non-governmental actors to scrutinise and shed light on public sector activities. As governments continue to empower these players through the provision of open data and algorithms and develop accountability mechanisms for better responsiveness, OPSI and the MBRCGI expect to see continued growth of these types of initiatives in the near term.

### ***Additional action needed going forward***

Governments and other ecosystems actors have made tremendous progress in this area in just a few years. A spectrum of approaches is unfolding with efforts exhibiting differing levels of maturity. For instance, most standards and principles around the world represent high-level, non-binding recommendations, but concrete laws like the EU's AI Act and US Algorithmic Accountability Act are coming into focus and have the potential to catalyse and align progress in this area.

In addition, most algorithmic accountability initiatives now focus on aspects of transparency, with many also incorporating elements of risk-based mitigation approaches. Fewer, though, demonstrate the ability for hands-on auditing of algorithms, which would close the loop on front-end accountability efforts to help ensure trustworthy use of AI in real-world use cases. Recent research from the Stanford Institute for Human-Centred AI ([HAI](#)) identifies nine useful considerations for algorithm auditing that can help inform these efforts (Metaxa and Hancock, 2022<sup>[12]</sup>) (Box 1.6).

### Box 1.6. Nine considerations for algorithm auditing

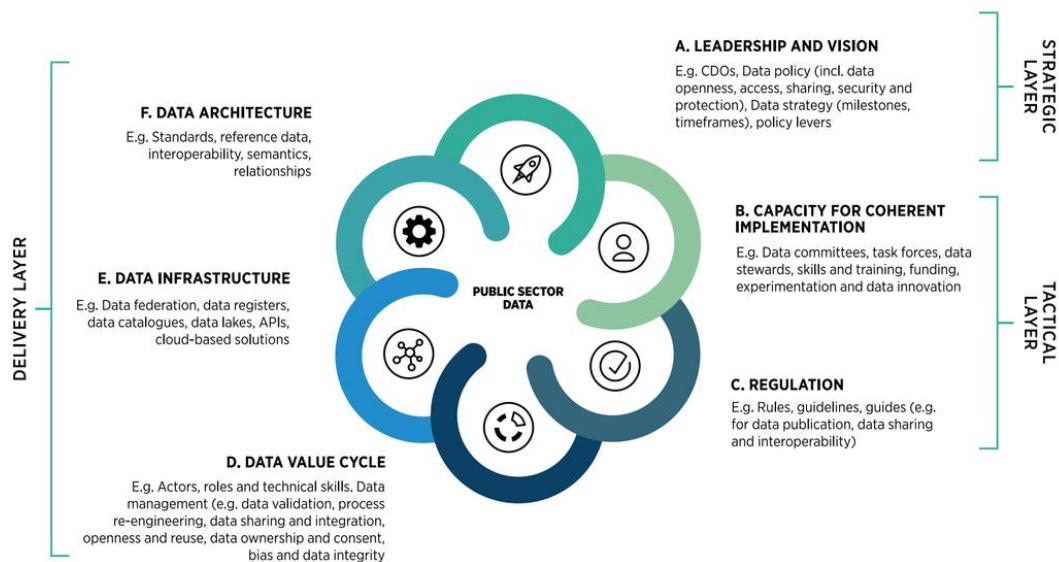
1. **Legal and ethical considerations** include relevant laws, the terms of service of different platforms, users involved with or implicated by audits, and personal and institutional ethical views and processes.
2. **Selecting a research topic** can include weighing discrimination and bias issues and political considerations (e.g. political polarisation, a technology's political impacts).
3. **Choosing an algorithm** to audit includes factoring in international considerations (e.g. which algorithms are popular where) and comparative factors (e.g. auditing one versus multiple algorithms and then comparing them).
4. **Temporal considerations** include how often the algorithm is updated and how the data might change before, during and after an audit is conducted.
5. **Collecting data** requires consideration of the possible available data sources and how analysing the data might scale.
6. **Measuring personalisation** involves considering how personalisation might change algorithms from person to person and how that might impact audits
7. **Interface attributes** require examination of the relationship between interfaces and metadata (e.g. how searches are displayed on a webpage)
8. **Analysing data** involves filtering the data, merging it with external data and choosing points of comparison.
9. **Communicating findings** requires considering the wider public discourse concerning the algorithms.

Source: <https://hai.stanford.edu/policy-brief-using-algorithm-audits-understand-ai>.

In addition to deepening and iterating their efforts, going forward, governments should work to ensure that public servants involved in building, buying or implementing algorithmic systems are informed about the AI and data ethics principles discussed in this section, and how they can play their part as stewards in ensuring such systems are accountable and serve the public good, alongside other actions in the accountability ecosystem. Such essential efforts range from basic definitional areas up to more sophisticated concepts and approaches. The challenges here have been cited in several studies which found that “in notable cases government employees did not identify regulated algorithmic surveillance technologies as reliant on algorithmic or machine learning systems, highlighting definitional gaps that could hinder future efforts toward algorithmic regulation” (Young, Katell and Krafft, 2019<sup>[13]</sup>). Furthermore, “definitional ambiguity hampers the possibility of conversation about this urgent topic of public concern” (Krafft et al., 2020<sup>[14]</sup>). AI Now’s [Algorithmic Accountability Policy Toolkit](#) can assist in this effort. It provides “legal and policy advocates with a basic understanding of government use of algorithms including, a breakdown of key concepts and questions that may come up when engaging with this issue, an overview of existing research, and summaries of algorithmic systems currently used in government”.

Finally, while this section has focused generally on algorithms and AI systems, governments must also pay close attention to issues related to the underlying data that are used to train modern AI systems. These are touched on in the OECD Good Practices Principles for Data Ethics in the Public Sector (Box 1.3) and can be seen in the penumbras of examples in this section. To achieve this in a holistic manner, governments must develop and implement robust data governance frameworks and processes across different layers (Figure 1.2).

Figure 1.2. OECD Framework for Data Governance in the Public Sector



Source: <https://oe.cd/ddps>.

As can be seen in Figure 1.2, data governance in the public sector comprises a broad, cross-cutting set of factors that serve as the foundation of a data-driven public sector, the use of data to increase public value and the role of data in building public trust (OECD, 2019<sup>[15]</sup>). Data governance intersects directly with and supports algorithmic accountability by helping to ensure the integrity and appropriateness of the underlying data itself along with algorithmic code and risk management processes. Good data governance is inextricably linked with algorithmic accountability but can also be supported by innovative governance techniques. For instance, data audits represent a powerful tool for auditors to assess the quality of data used by AI systems from different perspectives. For instance, auditors can assess if the data source in itself is trustworthy and whether the data are representative of the phenomena to which the AI algorithm is applied. Such data audits have been employed by governments, such as [Ireland's Valuation Office](#), to ensure accurate evaluations of commercial property. In fact, in 2022 Ireland's Office of Government Procurement developed an Open Data and Data Management [framework](#) that includes data auditing as its primary focus.

The efforts discussed in this trend are building a strong, cohesive foundation to take this innovative area of work to the next level, although much remains to be done. Research is pointing to challenges as governments and private sector organisations move from fragmented and cursory algorithmic accountability efforts to systems approaches that can provide for explainability and auditability, all supported by quality data governance. For instance, without stronger definitions and processes in this space, there is the risk of false assurances through “audit washing” where inadequately designed reviews fail to surface true problems (Goodman and Trehu, 2022<sup>[16]</sup>).

With the AI Act and other international and domestic rules looming, both governments and businesses will need to make rapid progress at data, code and process levels. Although governments have trailed behind the private sector for many activities related to AI, they also have the potential to be global leaders and practice shapers when it comes to algorithmic accountability. OPSI believes that leading governments are ready to come together to build a common understanding and vocabularies on algorithmic accountability in the public sector, as well as guiding principles for the design and implementation of governmental approaches which could result in tangible policy outcomes. OPSI intends to engage in additional work in this area in the belief that standardisation and alignment of algorithmic accountability initiatives is crucial to enable comparability, while still leaving room for contextual and cultural adaptation.

## Case Study: Algorithmic Transparency Recording Standard (United Kingdom)

Algorithmic tools are increasingly being used in the public sector to support high-impact decisions affecting individuals, for example in policing, social welfare, healthcare and recruitment. Research on public attitudes consistently highlights transparency as a key driver of public trust; therefore, building practical mechanisms for transparency is crucial to gaining and maintaining trust in governments' use of data and algorithms. In the United Kingdom (UK), for example, the OECD Trust Survey shows that only 52% of people trust their government to use their personal data for legitimate purposes (OECD, 2022<sup>[17]</sup>), while 78% of respondents to a UK survey on government data sharing wanted a detailed description of how their personal information is shared.

The United Kingdom's [Algorithmic Transparency Recording Standard](#) (ATRS) helps public sector bodies openly publish clear information about the algorithmic tools they use and why they are using them. The ATRS is one of the world's first policies to promote transparency in the use of algorithmic tools in government decision making, and it is positioned to serve as a key driver of responsible innovation and public trust in government.

### **Problem**

In the UK, as in many other countries, algorithms are used by public sector organisations to support decision making and can have a profound impact on the lives of citizens and residents. Recent experiences have shown that their implementation without adequate safeguards can result in discrimination or encroach on civil rights. A recent British example of problematic implementation in the public sector is the [failure of the A-level algorithm](#) in 2020.

The [Data Ethics Framework](#) was established in 2016 to address such risks, laying the foundations of responsible data use in public sector organisations, helping them to address ethical considerations within their projects and encouraging responsible innovation. In 2019, the government commissioned the UK Centre for Data Ethics and Innovation (CDEI) to conduct a [review into bias in algorithmic decision making](#), which confirmed that algorithms can lead to biased decisions resulting in significant negative impacts on people's lives. The CDEI further identified ways to address these risks through policy interventions, emphasising the importance of transparency.

The public has a democratic right to explanations and information about how the government operates and makes decisions, in order to understand the actions taken, appeal decisions and hold responsible decision makers to account. This is codified in the UK GDPR and emphasised in the OECD AI Principles of "[transparency and explainability](#)" and "[accountability](#)", adhered to by 46 countries. Nonetheless, there is still a lack of available information on how and why government bodies are using algorithmic tools, and in the absence of a standardised manner of presenting relevant data, citizens are unable to easily access this information. Lastly, public bodies that would like to be more transparent about how they are using algorithmic tools often struggle with how to communicate this complex information in an accessible manner. These are global challenges, and due to their persistence, many governments have adopted principles for ethical and trustworthy AI, but few have implemented them in meaningful ways.

### **An innovative solution**

The Algorithmic Transparency Recording Standard (ATRS), jointly developed by CDEI and the Central Digital and Data Office (CDDO), establishes a standardised way for public organisations to proactively and transparently publish information about how they are using algorithmic approaches in decision making. The ambition of this project is to increase public awareness and understanding of the use of algorithms in the public sector, while enhancing the capacities of the public sector to benefit from data and automation, thereby ensuring safer implementation of algorithms and easing the diffusion of best practices. Greater

algorithmic transparency is essential to enable public scrutiny and improved accountability of public sector decision-making processes involving algorithms.

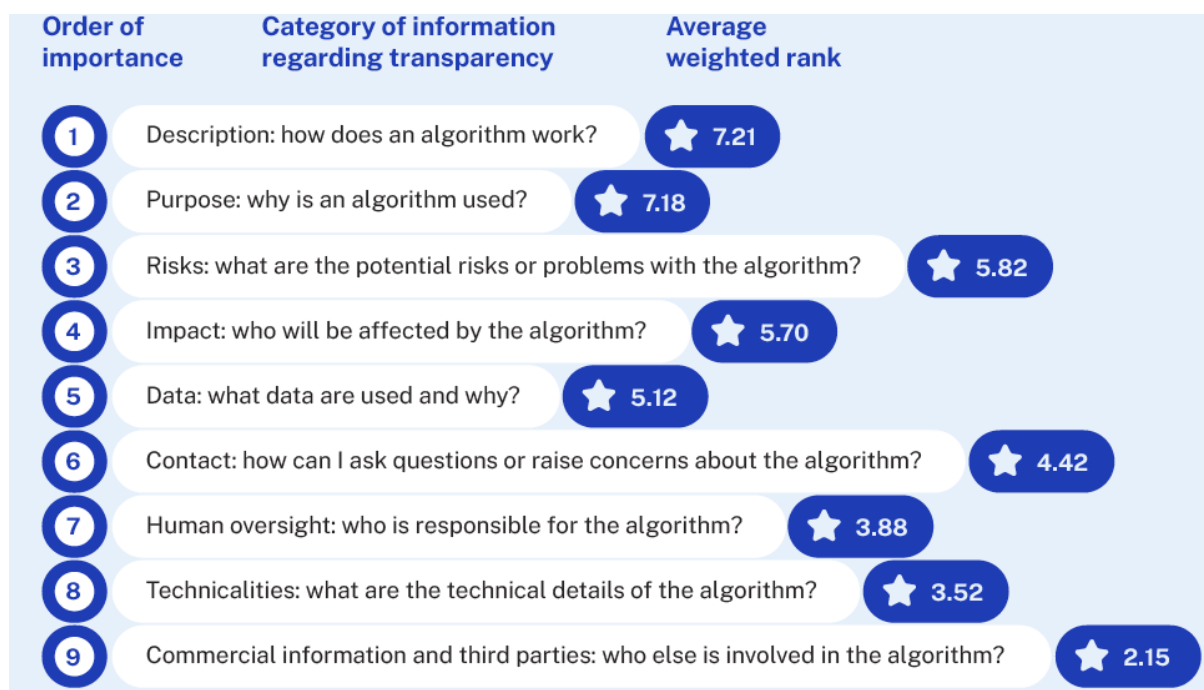
Work around the Standard comprises two elements. The first is the ATRS itself, which provides a structured schema that public sector organisations use to record and report information about the algorithms they use. The ATRS is divided into two reporting tiers. **Tier 1** is aimed at a general audience, and includes simple, concise details on how and why an algorithmic tool is being used, along with instructions on how to access more information. **Tier 2** is aimed at more technical or interested audiences, and is divided into five categories:

1. Information on who is responsible for the algorithm.
2. A description of the algorithm and the rationale for its use.
3. Details on the wider decision-making process and human oversight.
4. Technical specifications and data.
5. A breakdown of risks, mitigations and impact assessments conducted.

In addition to the ATRS, an important second element is the [implementation guidance](#). This helps organisations identify if the ATRS applies to their activities, as well as how to report information correctly.

The design and development of the ATRS has been underpinned by extensive collaboration with public sector, industry and academic stakeholders as well as citizen engagement. The CDEI and CDDO worked with BritainThinks to engage with a diverse range of members of the public over a three-week period, spending time to gradually build up participants' understanding and knowledge about algorithm use and discuss their expectations for transparency (see Figure 1.3 for the results of a survey on the importance of transparency categories in relation to algorithmic decision making in the public sector). This co-design process – which included working through prototypes to develop a practical approach to transparency that reflected expectations – led to the two-tier structure of the Standard and informed objectives for implementation.

**Figure 1.3. Respondents' rankings of importance for different aspects of transparency**



Source: [www.gov.uk/government/publications/cdei-publishes-commissioned-research-on-algorithmic-transparency-in-the-public-sector](http://www.gov.uk/government/publications/cdei-publishes-commissioned-research-on-algorithmic-transparency-in-the-public-sector).

The [first version](#) was published in November 2021 and piloted with ten public sector organisations through mid-2022, ranging from central government offices to local police departments. To date, six completed [transparency reports](#) have been published using the ATRS. For instance, it is now possible to retrieve accurate information on DARAT (Domestic Abuse Risk Assessment Tool), an algorithm that is being developed to help police officers in some areas predict the likelihood of future incidents of domestic abuse. The report provides information about many aspects of the algorithm such as the identity and responsibilities of members of the project team and technical details of the model. Based on feedback and lessons learned from the initial pilots, CDEI and CDDO launched an [updated version](#) in October 2022 on GitHub, which enabled anyone to open a two-way dialogue and propose changes for future iterations of the ATRS. This version was published officially on [gov.uk](#) in January 2023.

Going forward, in the short to medium term, the project team is investigating better ways of hosting and disseminating transparency reports, scaling from the pilot phase to full rollout by applying the ATRS to more and higher impact use cases (e.g. medical technology, criminal justice applications, benefits entitlements), and considering how the Standard could be embedded into public procurement practices to further reinforce transparency and accountability. In the long term, the project team believes the ATRS – with leading work from other OECD countries – could form the basis for a global standard on algorithmic reporting.

### ***Novelty***

The ATRS is one of the world's first initiatives of its kind and is leading the way internationally. Increasing algorithmic transparency has been at the forefront of AI ethics conversations globally, but much AI ethics work has been conceptual and theoretical, with only limited practical application, especially in the public sector. The Standard is a comprehensive policy and one of the very few undertaken by a national government to enhance transparency on the use of algorithmic tools in government decision making.

### ***Results and impact***

As noted above, ten pilots have been conducted, resulting in six published transparency reports so far. The pilots have demonstrated widespread support for algorithmic transparency from pilot partners, who highlighted the benefits of the ATRS both in terms of helping public servants gain confidence and knowledge about algorithmic approaches, and public accountability. Consultation with members of the public and suppliers of algorithmic tools revealed widespread support for the ATRS (97% of suppliers supported the initiative).

An additional positive impact of the ATRS has been the increased attention paid by senior leaders to understanding the importance of algorithmic transparency. Public transparency around the uses of algorithms has encouraged greater awareness within organisations, and helped combat the mindset that algorithms are solely a matter of importance for data scientists.

### ***Challenges and lessons learned***

This innovation faced two main challenges. First, it proved difficult to articulate the importance of transparency and to build momentum for using the ATRS. The team therefore engaged widely within government and made the benefits clear. With private suppliers, the team hosted roundtable discussions to gather views and incorporate them into the policy development process. The second challenge concerned the need to involve different types of stakeholders in the development and iteration of the ATRS. This was addressed by carefully designing the engagement process to ensure the representation of a broad range of perspectives among participants.

The project team learned many lessons. First, they found that many public sector teams would like to be more transparent and consider ethical questions, but might lack the guidance, capabilities or resources to

do so. To support teams in such efforts, the project team holds coaching calls with interested organisations and published guidance on common questions. They also found that initiatives like this can encourage a proactive culture in the public sector around embedding ethics into data and automation projects from the start.

Second, the team found that placing public engagement activities early on in the project lifecycle enabled them to act on the findings in meaningful ways, using the insights to develop the initial two-tiered design. Furthermore, these activities helped them to understand that the general public may not necessarily be interested in examining the content of each transparency report, but are reassured that this information is available openly and can be accessed by experts who can scrutinise it on their behalf – a finding that has informed the implementation approach taken by the team.

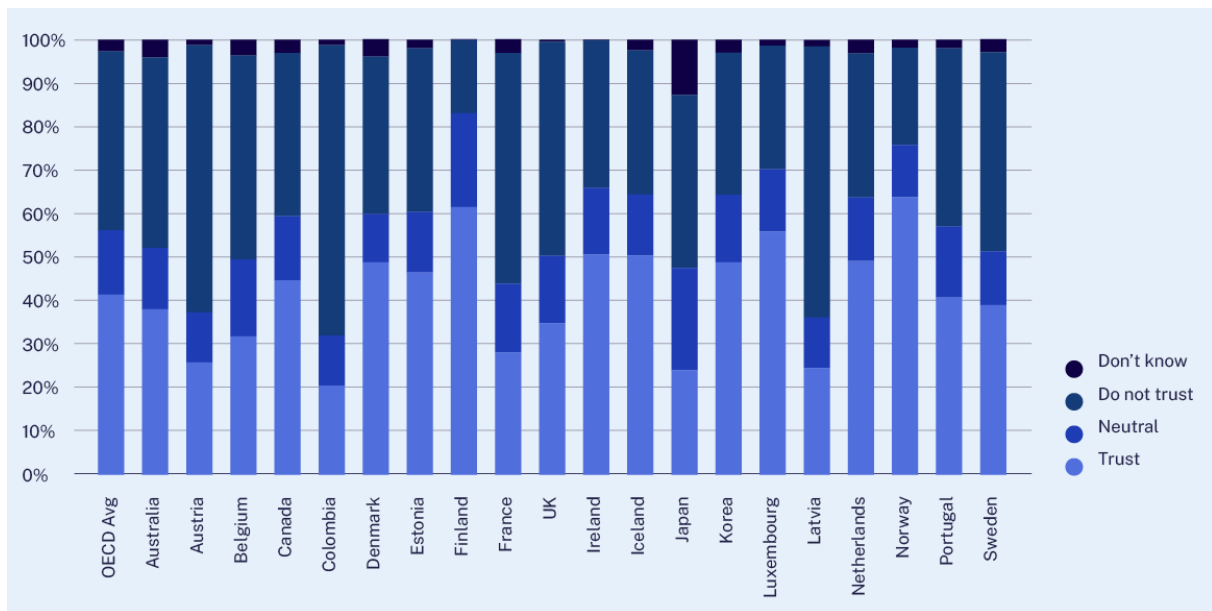
### ***Replicability***

There has been significant interest in replicating this innovation. The ATRS has featured in various international fora and working groups such as the Open Government Partnership's [Open Algorithms Network](#). The team has also been in contact with officials from different national governments to discuss aligning policies on algorithmic transparency, such as through a Tech Partnership between the UK and Estonia. Even some private companies, such as [Wolt](#), have leveraged the ATRS as inspiration in their own transparency policies. The problem posed by the opacity of automated decision-making systems is being recognised worldwide and, in this context, the ATRS appears to be a simple and effective innovation that is easily replicable. The aim is to see this innovation scaled internationally, becoming the standard for algorithmic transparency in the public sector, and perhaps beyond.

### **New aspects of transparency**

Intersecting with the theme of algorithmic accountability, governments are building new dimensions to their [open government](#) approaches, inching closer to visions of [radical transparency](#) and helping to build trust with citizens, which has been at a near record low in recent years (Figure 1.4) (OECD, 2022<sub>[17]</sub>). Public trust helps countries govern on a daily basis and respond to the major challenges of today and tomorrow, and is also an equally important outcome of governance, albeit not an automatic nor necessary one. Thus, governments need to *invest* in trust. Transparency is not the only way to achieve this (e.g. citizen engagement is also important, as discussed later in this report), but is a crucial factor (OECD, 2022<sub>[17]</sub>). It has assumed even greater importance in recent years, as aspects of transparency enable people to better understand and comply with government actions (e.g. COVID-19 responses).

Figure 1.4. Just over four in ten people trust their national government



Source: OECD Trust Survey (<http://oe.cd/trust>). Data available at <https://stat.link/jlkt6v>.

OPSI and the MBRCGI first explored [transparency](#) in the [2017 Global Trends](#) report. The OECD has covered many different angles of public sector transparency more broadly, such as efforts related to [Open Government](#), [Open State](#), Open Government Data (OGD), promoting [Civic Space](#), [anti-corruption and integrity](#), as well as specialised issues including transparency in the use of COVID-19 [recovery funds](#) among others. Indeed, one of the key focus areas in the recently issued [OECD Good Practice Principles for Public Service Design and Delivery in the Digital Age](#) is “be accountable and transparent in the design and delivery of public services to reinforce and strengthen public trust” (OECD, 2022<sub>[18]</sub>).

When looking at the latest public sector innovation efforts, two leading themes become apparent. The first is the advancement of the Rules as Code concept, which has gained significant traction in the last few years. The second is heightened transparency around the thousands of monitors and sensors embedded in daily life, the existence of which is unknown to most people.

### **Bringing about Rules as Code 2.0**

New technologies and approaches are leading to new aspects of transparency which empower the public while enhancing the accountability of governments. One area seeing growth in innovative applications is [Rules as Code](#) (Box 1.7), with some dubbing the new horizon [Rules as Code 2.0](#). While RaC offers a number of potential benefits, including better policy outcomes, improved consistency, and enhanced interoperability and efficiency (Mohun and Roberts, 2020<sub>[19]</sub>), [advocates](#) have [also](#) highlighted the importance of transparency, as RaC has made the rule-creation process more transparent in some cases, and enabled the creation of applications, tools and services that help people understand government obligations and entitlements. This can help bolster important elements of the OECD Recommendation on Regulatory Policy and Governance ([2012](#)), which serves as the OECD’s guiding framework on good regulatory and rulemaking practices.



### Box 1.7. Rules as Code (RaC)

Rules as Code (RaC) is a new take on one of the core functions of government: rulemaking. Fundamentally, RaC proposes to create an official, machine-consumable version of some types of government rules, to exist alongside the existing natural language counterpart. More than simply a technocratic solution, RaC represents a shift in how governments create some types of rules, and how third parties consume them.

Currently, governments typically produce human-readable rules that are individually consumed and interpreted by people and businesses. Each regulated entity, for example, must translate laws into machine-consumable formats for use in business rule systems. RaC could instead see official, machine-consumable versions of these rules produced by governments, concurrently with the natural language versions. This could allow businesses to consume machine consumable versions directly from government, while reducing the need for individual interpretation and translation for some types of rules.

Source: <https://oecd-opsi.org/publications/cracking-the-code>.

Since OPSI and MBRCGI's initial coverage of Rules as Code in the [2019](#) Global Trends report and OPSI's subsequent in-depth [primer](#) on the topic, the concept has reached new levels of adoption by innovative approaches within government, as it begins to embed a “new linguistic layer” (Azhar, 2022<sup>[20]</sup>) that transparently expresses rules in ways that both humans and machines can understand.

The Australian Government Department of Finance has sponsored a [project](#) that looked at how RaC could be delivered as a shared utility to deliver simpler, personalised digital user journeys for citizens. “My COVID Vaccination Status” served as the initial use case, drawing from publicly available COVID rules. The effort focused on the questions “Am I up to date with my COVID vaccinations?” and “Do I have to be vaccinated for my job?”, using a built simulator website to provide a simple, citizen-centric user journey to provide answers. This project represents a global first in use of RaC as a central, shared, open source service hosted on a common platform, allowing government offices and third parties enhanced access to information and the ability to build additional innovations on top. The project has helped demonstrate a path for scalable RaC architecture that can take this approach to new heights.

Nearby, New Zealand is rolling out an [ambitious project](#) to help people in need better understand their legal eligibility for assistance – a process that can be incredibly difficult, as the relevant rules are embedded in different complex laws. Grassroots community organisations are implementing a “Know Your Benefits” tool to address social injustice by helping people better understand their rights. The tool leverages codified rules to help citizens and residents gain access to support to which they are entitled, and to invoke their right to an explanation about any decision affecting them.

Other, additional efforts have surfaced in this space:

- Belgium's [Aviation Portal](#) translates the vast set of aviation laws and agreements into a single online aircraft registration platform.
- The UK Department for Work and Pensions has initiated an [effort](#) to generate human and machine-consumable legislation in pursuit of a Universal Credit Navigator to clarify benefits eligibility.
- Many projects are underway in different levels of government in the United States in areas such as benefits eligibility and policy interpretation, as showcased in Georgetown University's Beek Center [Rules as Code Demo Day](#).

In general, these approaches involve processes in which a multi-disciplinary team works to co-create a machine consumable version of rules which will exist in parallel with the human readable form (e.g. a

narrative PDF). However, a new take on this concept provides a hint of potential future developments. The Portuguese government's *Imprensa Nacional-Casa da Moeda* (INCM – National Printing House and Mint) has created a [functional prototype](#) of laws related to retirement that applies AI to decoding laws to make them consumable by digital systems. AI can be used increasingly in this space, optimally alongside and as tools of the aforementioned multi-disciplinary teams, to accelerate the RaC movement. Like the efforts discussed earlier in this trend, such approaches should be done in a way that is consistent with the [OECD AI Principles](#) and other applicable frameworks.

However, RaC is not a cure-all when it comes to putting in place good rulemaking practices and ensuring positive regulatory outcomes. For instance, the effects of a single regulation or rule may be dependent on a range of external factors, and its scope is currently best applied to relatively straightforward legal provisions. Yet, OPSI believes that Rules as Code has the potential to be truly transformative. In addition to OPSI's RaC primer, innovators wanting to learn more can leverage the Australian Society for Computers and Law (AUSCL)'s excellent and free series of [Masterclass sessions](#). Those who want to start digging into the models and code can check out [OpenFisca](#), the free and open source (FOSS) software powering many RaC projects around the world, and [Blawx](#). Interesting personal perspectives can also be found on blogs by [Hamish Fraser](#) and [Regan Meloche](#).

### ***The Internet of (Transparent) Things***

Smart devices and the Internet of Things (IoT) have become pervasive, yet in some ways remain invisible. There are over 11 billion IoT connected devices around the world, with more than 29 billion expected by 2030 as 5G technology continues to roll out (Transforma Insight, 2022<sup>[21]</sup>). The potential public sector benefits are significant ([OECD, 2021a](#)), especially through the creation of smart cities – cities that leverage digitalisation and engage stakeholders to improve people's well-being and build more inclusive, sustainable and resilient societies (OECD, 2020<sup>[22]</sup>). In fact, four in five people believe that IoT can be used to “create smart cities, improve traffic management, digital signage, waste management, and more” (Telecoms.org Intelligence, 2019<sup>[23]</sup>). The research for this report surfaced several notable examples:

- Singapore's [Smart Nation Sensor Platform](#) deploys sensors, cameras and other sensing devices to provide real-time data on the functioning of urban systems (ITF, 2020<sup>[24]</sup>). Also in Singapore, [RATSENSE](#) uses infrared sensors and data analytics to capture real-time data on rodent movements, providing city officials with location-based infestation information.
- In Berlin, [CityLAB Berlin](#) is developing an ambitious smart city strategy, and the local government's COMo project is using sensors to [measure carbon dioxide](#) to improve air quality and mitigate the spread of COVID-19.
- Seoul, Korea is pursuing a “[Smart Station](#)” initiative as the future of the urban subway system. A control tower will leverage IoT sensors, AI image analysis and deep learning to manage subway operations for all metro lines.
- In Tokyo, Japan, the installation of sensors on water pipelines has saved more than 100 million of litres per year by reducing leaks (OECD, 2020<sup>[25]</sup>).

While research shows that the vast majority of people support the use of sensors in public areas for public benefit, and that citizens have a fairly high level of trust in government with regard to smart cities data collection (Mossberger, Cho and Cheong, 2022<sup>[26]</sup>), IoT sensors and smart cities have raised significant concerns about “invasion of privacy, power consumption, and poor data security” (Joshi, 2019<sup>[27]</sup>), protection and ownership over personal data (OECD, forthcoming-b, *Governance of Smart City Data for Sustainable, Inclusive and Resilient Cities*) as well as other ethical considerations (Ziosi et al., 2022<sup>[28]</sup>). For example, San Diego's smart streetlights are designed to gather traffic data, but have also been used by police hundreds of times (Holder, 2020<sup>[29]</sup>), including to investigate protestors following the murder of George Floyd (Marx, 2020<sup>[30]</sup>), triggering surveillance fears. Less than half of the 250 cities surveyed in a 2022 [Global Review of Smart Cities Governance Practices](#) by UN-Habitat, CAF – the Development Bank

of Latin America and academic partners report legislative tools for ethics in smart city initiatives, with those that do exist being more prevalent in higher income countries.

In many cities, sensors are ubiquitous in public spaces, with opacity surrounding their purpose, the data they collect and the reason why. Individuals may even [be sensors themselves](#), depending on their activities and the terms accepted on their mobile device. These are important issues to think about, as “democracy requires safe spaces, or commons, for people to organically and spontaneously convene” (Williams, 2021<sup>[31]</sup>). The San Diego case mentioned above and others like it may serve as cautionary trends, for “if our public spaces become places where one fears punishment, how will that affect collective action and political movements?” (Williams, 2021<sup>[31]</sup>).

IoT and [Smart Cities](#) have been well documented by the OECD, with their concepts achieving a level of integration in many cities and countries that have arguably transferred them out of the innovation space and into steady-state. However, the new levels of personal agency, privacy protection and transparency being introduced to help ensure ethical application of smart city initiatives represent an emerging innovative element.

With regard to privacy protection, New York City's [IoT Strategy](#) offers a framework for classifying IoT data into “tiers” based on the level of risk:

- **Tier 1** data are not linked to individuals, and thus present minimal privacy risks (e.g. temperature, air quality, chemical detection).
- **Tier 2** data are context dependent and need to be evaluated based on their implementation (e.g. traffic counts, utility usage, infrastructure utilisation).
- **Tier 3** data almost always consist of sensitive or restricted information (e.g. location data, license plates, biometrics, health data).

While useful for conceptualisation, the tiers have not been adopted as a formal classification structure in government. Nonetheless, ensuring digital security is a fundamental part of digital and data strategies at both the city and country level, with some states according digital security a top priority in their digital government agenda. For example, Korea and the United Kingdom have both developed specific digital security strategies (OECD, forthcoming-b).

One of the more dynamic approaches is found in Aizuwakamatsu, Japan, which has adopted an “opt-in” stance to its city smart city initiatives, allowing residents to choose if they want to provide personal information in exchange for digital services (OECD, forthcoming-b). This represents “a markedly different approach to the mandatory initiatives in other smart cities that have been held back by data privacy” (Smart Cities Council, 2021<sup>[32]</sup>). Though this option applies only to public initiatives, it is difficult to envision how this approach would work with smart city elements that are more passive and not necessarily tied directly to specific residents.

Some of the most digitally advanced and innovative governments have taken new steps to make their IoT and smart city efforts open and transparent in order to foster accountability and public trust in government. One leading effort is the City of Amsterdam's mandatory **Sensor Register** and its associated Sensor Register Map, as discussed in the full case study following this section.

Other areas, such as [Innisfil](#), Canada; [Angers-Loire](#), France; and [Boston](#) and [Washington, DC](#) in the United States are leveraging Digital Trust for Places and Routines ([DTPR](#)) (Box 1.8), which has the potential to serve as a re-usable standard for other governments.

### Box 1.8. Digital Trust for Places and Routines (DTPR)

The open-source communication standard Digital Trust for Places & Routines (DTPR) aims to improve the transparency, legibility and accountability of information about digital technology.

In 2019, experts in cities around the world took part in co-design sessions to collaborate and prototype an initial open communication standard for digital transparency in the public sphere. In 2020, this standard underwent numerous cycles of online expert charrettes and small meetings, iterative prototype development, and long-term inclusive usability and concept testing.

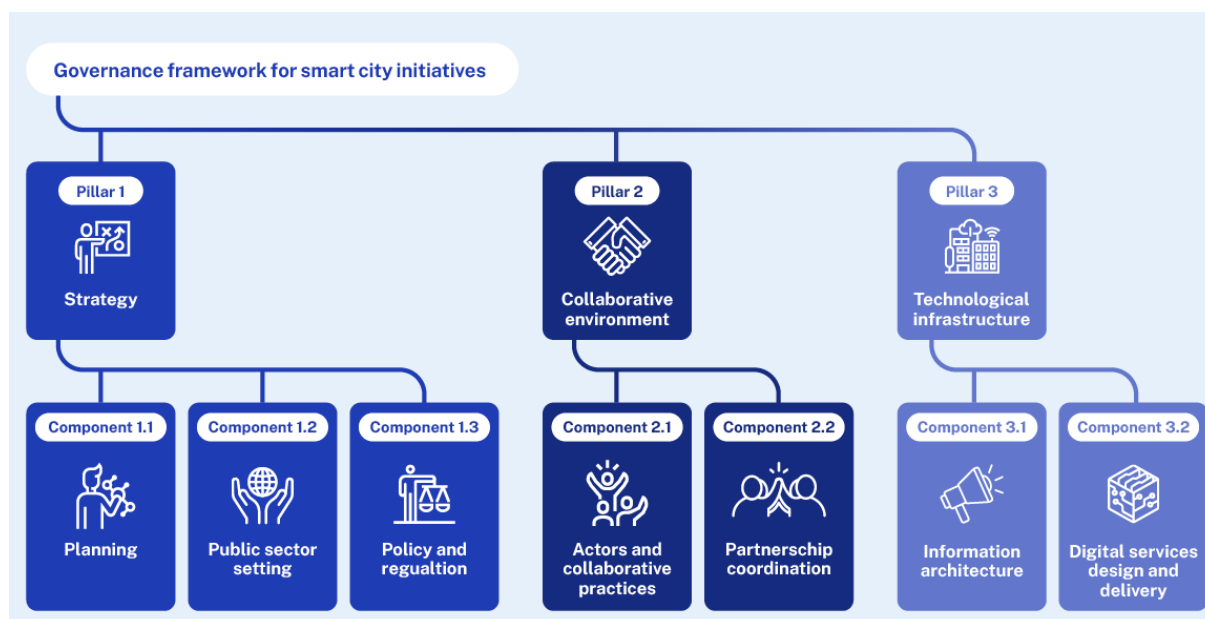
The final product, the DTPR, is a taxonomy of concepts related to digital technology and data practices, accompanied by a collection of symbols to communicate those concepts swiftly and effectively, including through physical signs or digital communication channels. Use of the DTPR provides a public, legible explanation of city technologies and their data footprints, enabling public input on city technologies and allowing the effectiveness of city technologies to be measured and evaluated.

Source: <https://dtptr.helpfulplaces.com>, [www.weforum.org/agenda/2022/06/smart-cities-public-spaces-data](http://www.weforum.org/agenda/2022/06/smart-cities-public-spaces-data).

Recent research has found that “cities today lack the basic building blocks to safeguard their interests and ensure the longevity of their smart city” (WEF, 2021<sup>[33]</sup>). As governments continue to deploy IoT sensors and pursue smart city strategies, they should follow the lead of the cities cited above, as “public trust in smart technology is crucial for successfully designing, managing and maintaining public assets, infrastructure and spaces” (WEF, 2022<sup>[34]</sup>). This is easier said than done, however, as governments face difficulties in establishing solid governance over such efforts – an important but often overlooked enabler of digital maturity that can help them move towards a more open and transparent approach.

The aforementioned [Global Review of Smart City Governance Practices](#) provides a number of recommendations and a valuable governance framework (Figure 1.5), Pillars 2 and 3 of which include elements relevant to transparency. In addition, researcher Rebecca Williams offers 10 calls to action for cities to consider in her report [Whose Streets? Our Streets! \(Tech Edition\)](#). These include “mandating transparency and legibility for public technology & data” and “challenge data narratives” to ensuring “that community members can test and vet government data collection and the narratives they reinforce”, and imagining “new democratic rights in the wake of new technologies.” Ethical use of the technologies discussed in this section go far beyond transparency alone, and the guidance in these resources can provide food for thought on moving towards a more comprehensive approach with transparency as a key pillar.

Figure 1.5. Governance framework for smart city initiatives



Source: <https://unhabitat.org/global-review-of-smart-city-governance-practices>.

### Case Study: Sensor Register (Amsterdam, Netherlands)

The Sensor Register is a tool of the City of Amsterdam used to obtain, combine and share publicly transparent information on all sensors placed for professional use in public spaces of the city. The Register is the result of an innovative [Regulation](#) which mandates the registration of all sensors of private, public, research and nonprofit organisations that collect data from public spaces. The registered sensors are visualised on an online map that allows any member of the public to see consult information on the sensor, including the kind of data it collects and processes and the responsible party. In addition, stickers are placed on sensors that collect sensitive information, providing details about their activity and showing an URL that directs to the online map, where citizens and residents are able to retrieve more information.

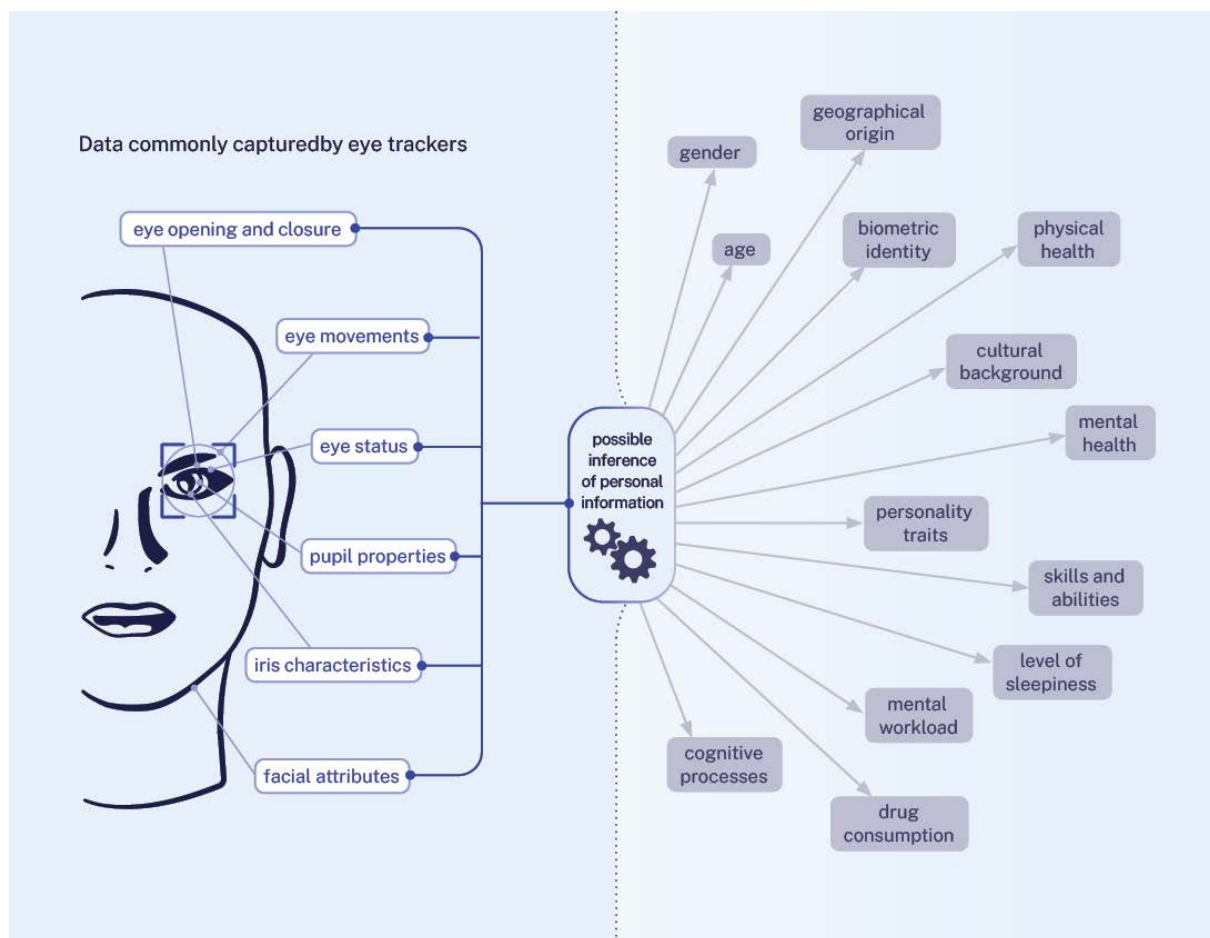
#### **Problem**

The widespread diffusion of new technologies capable of capturing and processing citizens' data in public spaces has given rise to heated debates about the threat of surveillance. Cities are becoming "smart cities" with a growing number of sensors collecting data and informing automated decision-making systems. An increasing number of billboards now have cameras installed that read spectators' glances, faces or body movements in reaction to the exhibited content. Such information when processed by advanced data analytics can reveal much more about a user than they may wish or expect to give away, as shown in Figure 1.6. These sensors were installed frequently without the city administration or passers-by being informed, as was the case in [many Dutch cities](#).

To address these emerging issues, it has become imperative to elaborate new policies to safeguard the dignity of citizens and residents and to avoid excessive and undesirable intrusion into people's lives. The GDPR has made important progress in this area, demanding transparency with respect to how personal data are collected and processed in public space. Expanding this idea, many civil society organisations and policy makers, including members of the City of Amsterdam, also asserted that citizens hold digital rights which extend beyond personal data to also include, for example, air quality and noise. This concept is based on the idea that citizens have the right to know what happens in public space, which belongs to

everyone. As [Beryl Dreijer](#), former Privacy Officer at the City of Amsterdam and project leader of the Sensor Register noted, “the municipality does not have the authority to prohibit the installation of sensors in public spaces”, but it can work to ensure their transparent and fair implementation, allowing citizens to be informed about what happens in public space, and thereby nurturing a fruitful debate on this issue. The City’s [Privacy Policy](#) has begun to codify residents’ digital rights, stating that people should be able to move about public spaces without surveillance, and seeks to put in place concrete mechanisms to achieve this end.

**Figure 1.6. Capture and analysis of gaze data enables the inference of personal information**



Source: (Kröger et al., 2020).

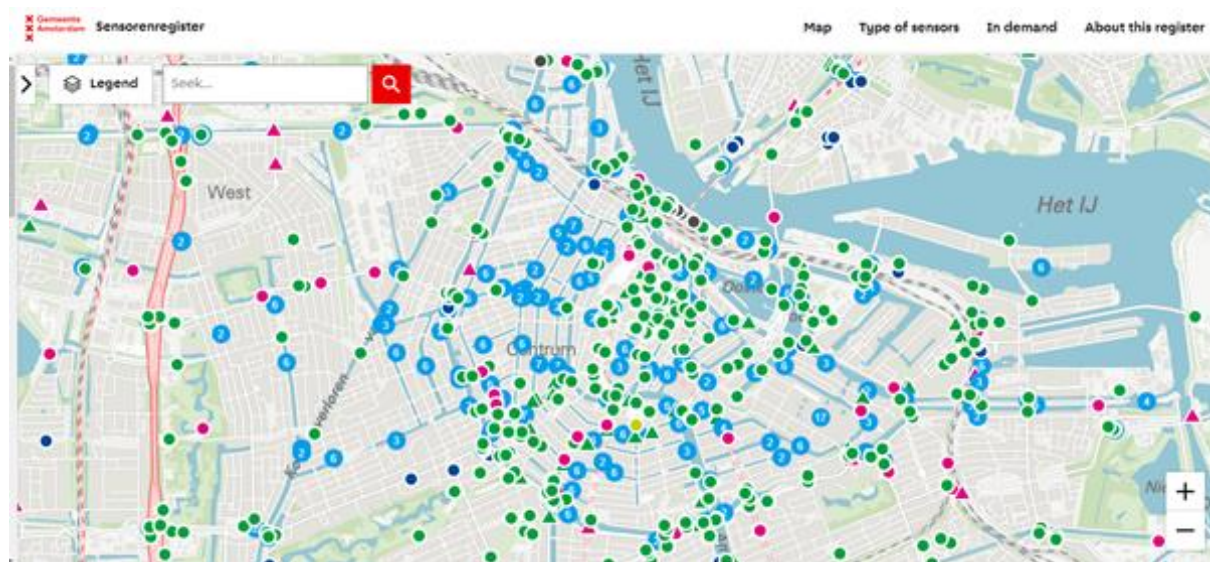
### **An innovative solution**

After building a public register of all government sensors, in 2021, the City of Amsterdam decided to pass an unprecedented [Regulation](#), requiring *all* parties that collect data in public space for a professional purpose to report their sensors and indicate which data are – or can be – collected by them. The Regulation imposes this requirement on public, private and research actors and non-profit organisations, and acts on all sensors placed in public space, excluding those for personal use such as smart doorbells.

Building on this adopted Regulation, and with the aim of ensuring transparency and privacy, different departments of the City of Amsterdam collaborated to develop the [Sensor Register Map](#), an online tool that allows anyone to view all sensors placed in public space. The Regulation defines sensors as follows: *an artificial sensor that is used or can be used to make observations and to process them digitally or to*

have them processed. Various types of sensors required to be registered are shown on the map including optical sensors (cameras), sound sensors, chemical sensors, pressure sensors and temperature sensors (an exhaustive list of the types of sensors covered by the Regulation and displayed on the Sensor Register Map is available on the [website](#)).

Figure 1.7. The Sensor Register Map

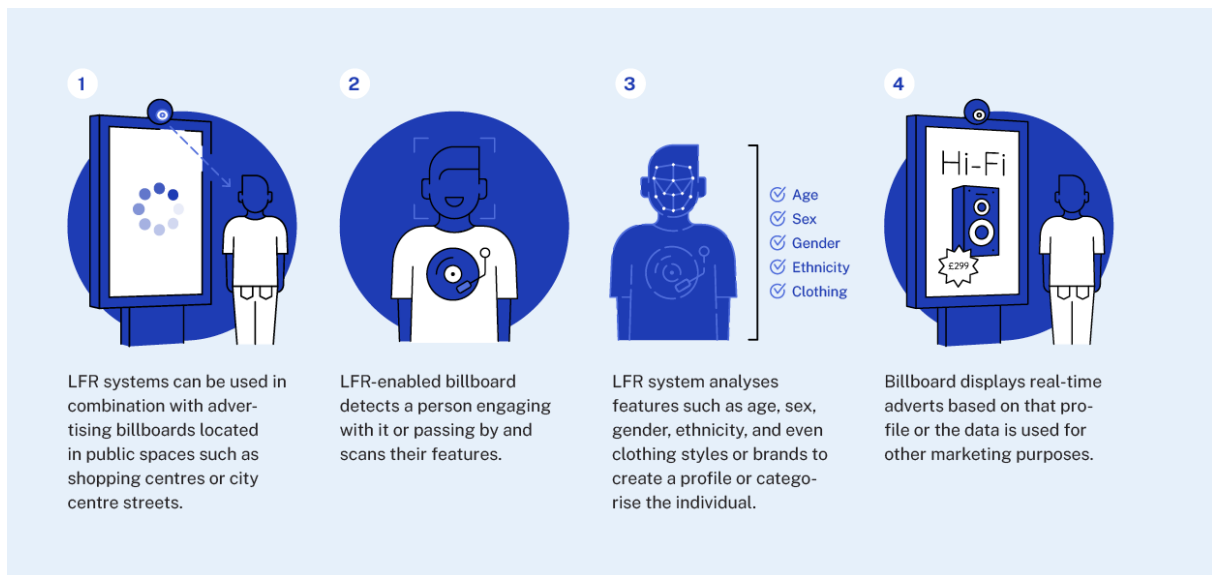


Source: Screenshot from <https://sensorenregister.amsterdam.nl> (retrieved 21 November 2022).

On each sensor that processes [sensitive data](#), a sticker is attached indicating why it is there and what it does, along with a URL to the Sensor Map where further information can be found. At the moment, only sensors working with sensitive data are required to have a sticker, but the plan is to extend this requirement to other sensors to inform all passers-by about the project. The decision to use a URL instead of a QR code was deliberate because the latter can easily be hacked to direct users to another page where they could be misled or subject to fraud.

Public spaces will become increasingly populated with sensors. In the United Kingdom, [London's King's Cross station uses facial recognition to track tens of thousands of people](#). These tools could be used to infer citizens' gender, sexual preference, emotional state and socioeconomic status, as [stated by the UK Information Commissioner's Office](#) (Figure 1.8). Under this scenario, citizens and residents become "unwitting objects of pervasive privacy infringements from which they do not have the chance to opt out", as a recent [Nesta report](#), funded by the City of Amsterdam, warned. In this context, the Regulation and the Map are intended to spark a debate about the role these technologies should play in communities, by increasing the awareness of citizens and residents, which is the first step in enabling them to critically address this issue. The City of Amsterdam is looking to take even stronger action in the future, declaring in its [coalition agreement 2022-2026](#), that "there will be a ban on biometric surveillance techniques, such as facial recognition".

Figure 1.8. How live facial recognition (LFR) can be used for marketing and advertising



Source: <https://ico.org.uk/media/for-organisations/documents/2619985/ico-opinion-the-use-of-lfr-in-public-places-20210618.pdf>.

### Novelty

This innovation is a novelty in the international context. Although attempts are being made to ensure that inevitable digitalisation is inspired by transparency and openness, this innovation is novel because it focuses on sensors in public space, underpinning the development of a new, broad understanding of digital rights. Furthermore, this project ensures that transparency is not restricted to imposing reporting requirements but also gives the public the possibility to easily access the information they want via the online map.

### Results and impact

Following publication of the first data on sensors, the project team received an influx of phone calls and emails from people saying that there were other cameras installed on the streets which were not represented on the map, which may result in field visits from the Amsterdam team. These immediate results demonstrated the interest of people in this issue. Indeed, contrary to the team's expectations, the project showed that many people care about digital rights and the potential dangers of new technologies in public spaces. In recent months, the innovation has garnered the attention of a researcher from Carnegie Mellon, who travelled to Amsterdam to understand more about the project to help inform their own sensor mapping [efforts](#). The University of Amsterdam Institute for Information Law is currently developing a report due for publication later this year which will evaluate the regulation and the Sensor Register.

The Sensor Register project caught the attention of citizens and residents, and the City of Amsterdam is now looking to expand its work on similar topics. For instance, it is collaborating with the Responsible Sensing Lab on [Shuttercam](#) to design cameras that gather only the required type or amount of data necessary to operate and, in this way, safeguard the right of citizens to walk around freely and unobserved.



## **Challenges and lessons learned**

The Sensor Register Map is the result of an initiative at the cutting edge of legal and digital rights recognition. As the first of its kind, despite the supportive and favourable political climate, unforeseen challenges have emerged. Beyond achieving registration compliance with businesses and nonprofits, the three main ones are:

- How to deal with moving cameras such as Google Street View and [debt collectors'](#) cars that move around in the public space of the city? Actors such as these are capable of capturing thousands of photos, which may present the same problems as the sensors, but are not required to report their data collection activity under the Regulation.
- How can mobile sensors be displayed on the map? The Regulation mandates reporting requirements for vehicles or vessels, but it is difficult to report information on such sensors on the Map.
- How should body cameras of enforcers and drones be regulated? The case of the latter is particularly complex because the Regulation does not directly cover sensors that are not connected to the ground. Such types of sensors are not reported to the city and cannot be displayed on the Map, though usage is fairly limited by rules due to proximity to the airport.

The Amsterdam team is working with researchers to explore some of these challenges. With respect to the success factors behind this innovation, the project team emphasises the fundamental role of the Regulation. Without this, the register would have been limited to public sector sensors. The Regulation widened the possibilities allowing the City to mandate transparency on all sensors, including those placed by private, research and nonprofit organisations.

## **Replicability**

This innovation is highly replicable. Although it is clear that Amsterdam has a political climate attentive to transparent and privacy-friendly digitalisation, the Regulation and Sensor Register Map are easily exportable to other contexts. Such a move is important given the pervasiveness of sensors in public space around the world and the relevance of an informed debate. As mentioned above, the Association of Dutch Municipalities is considering whether to scale the Register to the national level, as has already happened with Amsterdam's [register of algorithms used by public bodies](#).

## **Institutionalising innovative accountability**

The approaches discussed in the previous two themes are positive developments, illustrating how governments are connecting the concepts of innovation and accountability. Bringing together these worlds, however, has been a longstanding challenge in the public sector.

In talking with public servants anywhere in the world about innovation, perhaps the most commonly cited challenge is "risk aversion". Many feel that trying new things in the public sector is difficult because of the negative incentives built into the system, and this sentiment can permeate the culture of government. The main issue that comes up tends to be accountability mechanisms and entities such as oversight and auditing agencies. Innovation is fundamentally an iterative and risky process. Yet, audit processes can sometimes adopt a more rigid interpretation of what risks could have been foreseen and should have been planned for. Both accountability and oversight processes sometimes seem to be predicated on the idea that a right answer existed that could have been known beforehand.

To be clear, such functions are very important for governments. They help ensure confidence in the integrity of the public sector, identify where things could be done better and create guidance about how to avoid repeating errors in the future. Like innovation, at the end of the day, accountability is about achieving

better outcomes. The interplay between accountability and innovation is multifaceted but is not yet evolving rapidly enough to match the disruptive nature of new approaches and technologies in the public sector. Some governments have sought to better balance these two seeming counterweights. Back in 2009, the National Audit Offices in both the [United Kingdom](#) and [Australia](#) published guides on how to promote innovation. However, some governments are adopting a fresh perspective on accountability and putting in place processes where new ideas, methods and approaches can flourish while also reinforcing key principles of efficient, effective and trustworthy government.

One of the most systematic approaches identified for this report is the Government of Canada's Management Accountability Framework, in particular its "Innovation Area of Management" (see Box 1.9).

### Box 1.9. Management Accountability Framework – Innovation Area of Management (Canada)

The Management Accountability Framework (MAF) – a tool used by the Treasury Board of Canada Secretariat (TBS) to monitor the management performance of federal departments and agencies – has been expanded with the integration of a [new Area of Management \(AoM\) dedicated to innovation](#). Thanks to this tool, it is possible to assess a public organisation's ability to plan, generate and use rigorous evidence to inform decision making on high-impact innovations and, in this way, support excellence in innovation management.

The assessment is carried out with a [set of questions](#) that cover the various dimensions of innovation management. By answering each question, the organisations are assigned points. The questions investigate an organisation's ability to (1) commit resources to generate evidence to support innovation, (2) use rigorous methods of comparison to support innovation, (3) engage in innovation projects of high potential, and (4) use the evidence retrieved to inform executive-level decision making. The sum of the points scored by an organisation in each question represents a measure of the maturity level of an organisation in the field of innovation management. By collecting the results and establishing a dialogue with public organisations' teams, TBS highlights notable work, supports the diffusion of best practices, and provides expertise and guidance to departments willing to increase their innovation management maturity.

Extensive consultations and engagement sessions were held with federal departments and agencies – the primary users of the MAF – in order to develop the Innovation AoM. This co-development proved essential to consider various perspectives, validate terminology and ensure that the approach was in line with how departments operate. The new AoM has been implemented for the first time in the 2022-23 MAF reporting cycle and will be included in future cycles to measure progress over time.

Source: <https://oe.cd/maf-innovation>, Interview with Government of Canada officials on 24 November 2022.

Another interesting approach is the Accountability Incubator, based in Pakistan, which seeks to infuse government with accountable practices by tapping into the potential of young people (Box 1.10).

### Box 1.10. Accountability Incubator

The Accountability Incubator is a creative peer learning programme for young civic activists and change-makers who want to fight corruption and build accountability. It was developed to provide long-term support, networks and skills to people who are often overlooked by or left out of traditional civil society programmes. It is innovative in that it uses creative tools, a long-term approach and the very latest thinking to shape governance globally.

The Accountability Incubator provides tailored and hands-on support to activists and change-makers (termed “accountapreneurs”) from civil society around the world. The support they receive takes a hybrid format as useful resources are shared online and in person at Accountability Labs in eight countries – Belize, Kenya, Liberia, Nepal, Nigeria, Pakistan, Uganda and Zimbabwe. The year-long programme provides accountapreneurs with everything they need – from learning opportunities to communications support, stipends and pilot ideas – to super-charge accountability within their communities and find new ways to solve entrenched problems of governance.

The overarching objective of the innovation is to build a new generation of accountability change-makers that can create a more prosperous, inclusive and fair society. So far, the project has supported over 200 accountapreneurs who have developed more than 150 new ideas for accountability. Some of these ideas include setting up the first tool to crowdsource information on public services in Nepal, establishing a film school for women to fight corruption in Liberia and building a local media outlet for verified news on issues of democracy in Pakistan.

Source: <https://oecd-opsi.org/innovations/accountability-incubator>.

Other identified efforts have tended to focus on specific aspects bridging innovation and accountability, but with a more specific tech focus than the Canada and Pakistan cases. These include:

- The [Digital Transformation and Artificial Intelligence Competency Framework for Civil Servants](#) by UNESCO, the Broadband Commission for Sustainable Development and the International Telecommunication Union (ITU), which aims to promote the accountable use of innovative technology with a key focus on promoting trustworthy, inclusive and human rights-centric implementation of AI among civil servants.
- A national certification [programme](#) in Ireland to upskill civil servants on the ethical application of AI in government.

Although these efforts are positive and emit signals pointing to growing activities in the future, they remain few. More needs to be known about the relationship between relevant aspects of accountability (e.g. oversight, audit) and how they might be changed for the better. How can accountability structures (ultimately about trying to get better outcomes from government) be used to drive innovation inside public sector organisations, rather than hinder it? Is there room to respect important tenants of the audit process (e.g. independence) while also forging closer ties and partnerships between oversight functions and those that they look over? How can new forms of accountability that integrate users into the processes of monitoring and assessing public actions (e.g. participatory audits) drive innovation in government?

Little work has gone into answering these questions, signalling the need for deeper research and analysis at a global level. The OECD Open and Innovative Government Division (OIG) is exploring avenues to fill this gap and expand upon this field of study.

Although not discussed as a specific sub-theme in this report, it is important to note that leveraging innovative approaches to transform accountability, oversight and auditing functions themselves is also an area ripe for deeper analysis, as discussed by the [OECD Auditors Alliance](#). Some examples of this include Chile’s development of a data-driven [Office of the Comptroller General](#) (CGR), as well as the creation of accountability innovation hubs and labs in [Belgium](#), [Brazil](#), the [Netherlands](#), [Norway](#), the [United States](#) and the [European Court of Auditors](#) (Otia and Bracci, 2022<sup>[35]</sup>).

## References

- Ada Lovelace Institute, AI Now Institute and Open Government Partnership (2021), *Algorithmic accountability for the public sector*, <http://www.opengovpartnership.org/documents/>. [8]
- Azhar, A. (2022), *Sunday commentary: Policy as code*, [20]  
<https://www.exponentialview.co/p/sunday-commentary-policy-as-code>.
- Berryhill, J. et al. (2019), *Hello, World: Artificial Intelligence and its Use in the Public Sector*, [2]  
<http://oe.cd/helloworld>.
- Busuioc, M. (2021), "Accountable Artificial Intelligence: Holding Algorithms to Account", *Public Administration Review*, Vol. 81, pp. 825-836, <https://doi.org/10.1111/puar.13293>. [7]
- Caruso, F. (2022), *Serbia, algorithmic discrimination rehearsals*, [4]  
<https://www.balcanicaucaso.org/eng/Areas/Serbia/Serbia-algorithmic-discrimination-rehearsals-222242> (accessed on 1 February 2023).
- Cavaliere, P. and G. Romeo (2022), "From Poisons to Antidotes: Algorithms as Democracy Boosters", *European Journal of Risk Regulation*, Vol. 13/3, pp. 421-442, <https://doi.org/10.1017/err.2021.57>. [5]
- Goodman, E. and J. Trehu (2022), "AI Audit Washing and Accountability", *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.4227350>. [16]
- Halloran, B. (2017), *Strengthening Accountability Ecosystems: A Discussion Paper*, [9]  
<https://www.transparency-initiative.org/wp-content/uploads/2017/03/strengthening-accountability-ecosystems.pdf>.
- Holder, S. (2020), *In San Diego, 'Smart' Streetlights Spark Surveillance Reform*, [29]  
<https://www.bloomberg.com/news/articles/2020-08-06/a-surveillance-standoff-over-smart-streetlights>.
- ITF (2020), *Leveraging Digital Technology and Data for Human-centric Smart Cities*, [24]  
<https://www.itf-oecd.org/sites/default/files/docs/data-human-centric-cities-mobility-g20.pdf>.
- Joshi, N. (2019), *Exposing the dark side of smart cities*, <https://www.allerin.com/blog/exposing-the-dark-side-of-smart-cities>. [27]
- Kaye, K. (2022), *A new wave of AI auditing startups wants to prove responsibility can be profitable*, <https://www.protocol.com/enterprise/ai-audit-2022>. [11]
- Krafft, P. et al. (2020), "Defining AI in Policy versus Practice", *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, <https://doi.org/10.1145/3375627.3375835>. [14]
- Marx, J. (2020), *Police Used Smart Streetlight Footage to Investigate Protesters*, [30]  
<https://perma.cc/9Q5F-RTPN>.
- Metaxa, D. and J. Hancock (2022), *Using Algorithm Audits to Understand AI*, [12]  
<https://hai.stanford.edu/policy-brief-using-algorithm-audits-understand-ai>.
- Mohun, J. and A. Roberts (2020), "Cracking the code: Rulemaking for humans and machines", *OECD Working Papers on Public Governance*, No. 42, OECD Publishing, Paris, [19]  
<https://doi.org/10.1787/3afe6ba5-en>.

- Mossberger, K., S. Cho and P. Cheong (2022), "The Public Good and Public Attitudes Toward Data Sharing Through IoT", *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.4183676>. [26]
- OECD (2022), *Building Trust to Reinforce Democracy: Main Findings from the 2021 OECD Survey on Drivers of Trust in Public Institutions*, Building Trust in Public Institutions, OECD Publishing, Paris, <https://doi.org/10.1787/b407f99c-en>. [17]
- 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>. [18]
- OECD (2021), *Innovation and Data Use in Cities: A Road to Increased Well-being*, OECD Publishing, Paris, <https://doi.org/10.1787/9f53286f-en>. [10]
- OECD (2020), *Measuring Smart Cities' Performance*, <https://www.oecd.org/cfe/cities/Smart-cities-measurement-framework-scoping.pdf>. [22]
- OECD (2020), *Smart Cities and Inclusive Growth*, [https://www.oecd.org/cfe/cities/OECD\\_Policy\\_Paper\\_Smart\\_Cities\\_and\\_Inclusive\\_Growth.pdf](https://www.oecd.org/cfe/cities/OECD_Policy_Paper_Smart_Cities_and_Inclusive_Growth.pdf). [25]
- 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>. [15]
- OECD/CAF (2022), *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the*, <https://doi.org/10.1787/1f334543-en>. [1]
- Otia, J. and E. Bracci (2022), "Digital transformation and the public sector auditing: The SAI's perspective", *Financial Accountability & Management*, Vol. 38/2, pp. 252-280, <https://doi.org/10.1111/faam.12317>. [35]
- Salvi del Pero, A., P. Wyckoff and A. Vourc'h (2022), *Using Artificial Intelligence in the workplace*, OECD Publishing, <https://doi.org/10.1787/1815199X>. [6]
- Smart Cities Council (2021), *Smart city in Japan offers residents quake, privacy protection (Indian cities can take a cue)*, <https://www.smartcitiescouncil.com/article/smart-city-japan-offers-residents-quake-privacy-protection-indian-cities-can-take-cue>. [32]
- Telecoms.org Intelligence (2019), *Annual Industry Survey 2019 Report*, [https://itig-iraq.iq/wp-content/uploads/2019/12/Telecoms.com\\_Annual\\_Industry\\_Survey\\_FINAL.pdf](https://itig-iraq.iq/wp-content/uploads/2019/12/Telecoms.com_Annual_Industry_Survey_FINAL.pdf). [23]
- Transforma Insight (2022), *Global IoT connections to hit 29.4 billion in 2030*, <https://transformainsights.com/news/global-iot-connections-294>. [21]
- Véliz, C. (ed.) (2021), *Algorithmic Bias and Access to Opportunities*, Oxford Academic Press, <https://doi.org/10.1093/oxfordhb/9780198857815.013.21>. [3]
- WEF (2022), *3 ways cities can improve digital trust in public spaces*, <https://www.weforum.org/agenda/2022/06/smart-cities-public-spaces-data/>. [34]
- WEF (2021), *Governing Smart Cities: Policy Benchmarks for Ethical and Responsible Smart City Development*, [https://www3.weforum.org/docs/WEF\\_Governing\\_Smart\\_Cities\\_2021.pdf](https://www3.weforum.org/docs/WEF_Governing_Smart_Cities_2021.pdf). [33]
- Williams, R. (2021), *Whose Streets? Our Streets (Tech Edition)*, <https://www.belfercenter.org/sites/default/files/2021-08/WhoseStreets.pdf>. [31]

- Young, M., M. Katell and P. Krafft (2019), "Municipal surveillance regulation and algorithmic accountability", *Big Data & Society*, Vol. 6/2, p. 205395171986849, <https://doi.org/10.1177/2053951719868492>. [13]
- Ziosi, M. et al. (2022), "Smart cities: reviewing the debate about their ethical implications", *AI & SOCIETY*, <https://doi.org/10.1007/s00146-022-01558-0>. [28]



**From:**  
**Global Trends in Government Innovation 2023**

**Access the complete publication at:**

<https://doi.org/10.1787/0655b570-en>

**Please cite this chapter as:**

OECD (2023), "Trend 1: New forms of accountability for a new era of government", in *Global Trends in Government Innovation 2023*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/1b12de43-en>

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. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

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