ie UNIVERSITY
CENTER FOR THE GOVERNANCE OF CHANGE

BETTER POLICYMAKING IN THE AGE OF AI

RONY MEDAGLIA

OCTOBER 2025

TABLE OF CONTENTS

ABSTRACT

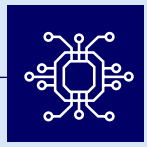
Artificial intelligence (AI) is transforming how governments design and deliver policy. AI tools such as predictive models, generative AI, and emerging agentic systems offer opportunities to broaden participation, improve efficiency, and strengthen evidence use. At the same time, they raise significant risks for fairness, accountability, and sustainability.

This paper examines how AI is shaping policymaking across the four main stages of the democratic policy cycle. Drawing on relevant research and empirical examples, we discuss how, in the agenda-setting phase, AI can support citizen deliberation and manage large-scale public input, but risks excluding underrepresented groups or being misused for public opinion manipulation. In policy formulation, AI can help simulate impacts and draft legislation, while also introducing dangers of bias, hallucinations, and loss of accountability. In policy implementation, AI can streamline administration and personalize services, yet may deepen divides and cloud institutional accountability. In policy evaluation, AI can accelerate analyses and stakeholder input, but automation bias and opacity threaten reliability. Beyond stage-specific risks, governments must also address broader concerns of workforce deskilling, dependence on dominant vendors, and the environmental costs of large-scale AI infrastructures.

TO LEVERAGE THESE OPPORTUNITIES AND MITIGATE THESE RISKS, THE PAPER OFFERS THREE PATHWAYS FOR ACTION:

01
Building AI literacy as a foundation for policymaking

02
Embedding accountability and rule-of-law principles within AI usage and deployments



03
Investing in open, sustainable, and sovereign AI models and infrastructure

Together, these measures can help ensure that AI enhances inclusiveness, transparency, and trust in public decision-making, positioning it as an engine of democratic prosperity.

1. INTRODUCTION

The swift diffusion of artificial intelligence (AI) systems opens new scenarios in the transformation how public institutions design and deliver policy. Both traditional AI applications (those providing algorithmic classification, recommendations and predictions based on training data), **generative AI** (GenAI) (creating textual and multimedia content based on patterns in training data) and novel **agentic AI** applications (that can plan, decide, and act autonomously over multiple steps rather than merely reacting to prompts) (Ilves et al., 2025) can be in fact applied to some of the core functions in public policymaking, and to support the pursuit of diverse policy goals (Valle-Cruz et al., 2020).

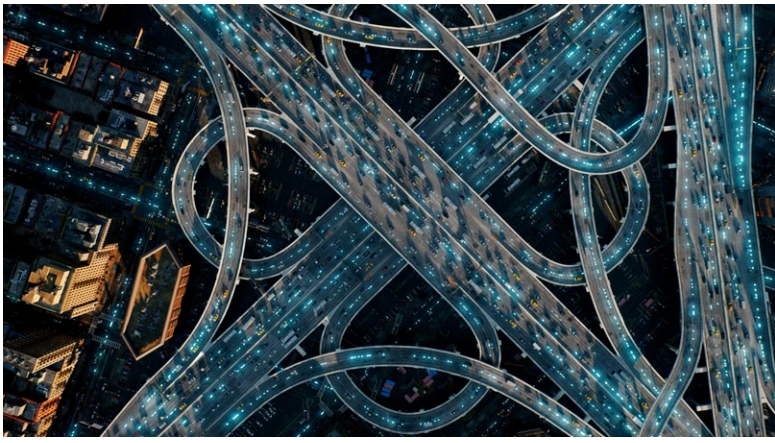
For instance, while the use of the latest applications of GenAI is still largely perceived by public servants as mostly a tool for **efficiency gains** (Tangi et al., 2025), there is growing interest in the use of AI to support also value-based goals such as improving the **inclusiveness**, the **transparency**, and the **fairness** of policymaking.

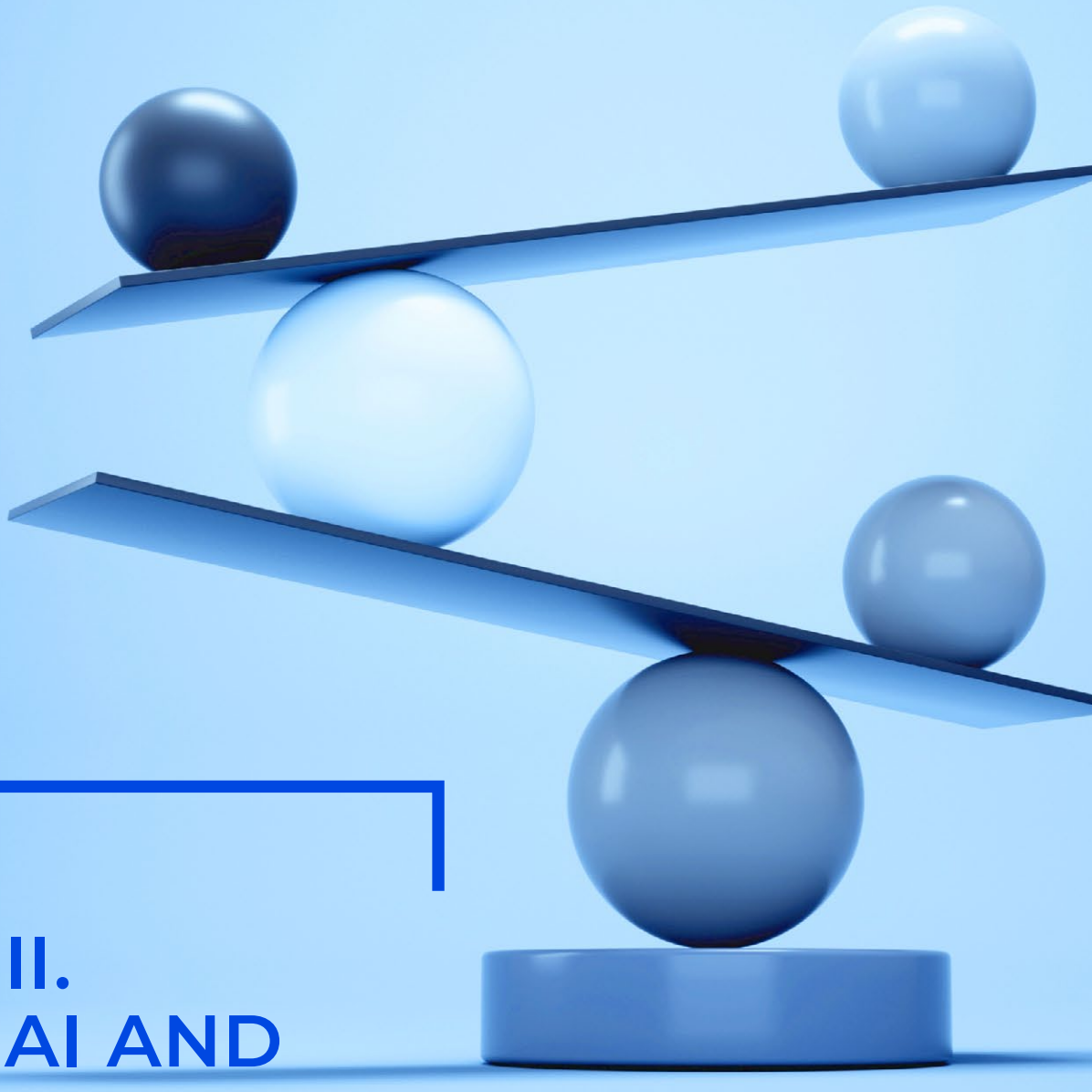
Different understandings of the potential of AI for policy making are often dependent on different **contextual factors**. These include the type of policy domain in which AI is applied, such as healthcare, transportation, or education (Fatima et al., 2020) and the differing strategic priorities pursued in national AI strategy, including efficiency, service delivery, citizen engagement, and the creation of public value (Viscusi et al., 2020; Wilson, 2022).

As a result, there are diverse interpretations of the role of AI and digital transformation across national contexts (van Noordt et al., 2025). For example, the **United States** portrays AI primarily as a driver of innovation, market competitiveness, and national security, whereas **China** highlights its potential for social stability, economic growth, and national rejuvenation, in addition to national security (Hine & Floridi, 2024). **Spain,** as a paradigmatic example of the European Union in this regard, frames digital transformation as an opportunity for fostering public value, trust, and government openness (The Ministry for Digital Transformation and the Civil Service, 2025).

Existing and envisioned uses of AI in policymaking feature many advantages, but also risks are to be considered. AI systems can, for example, exacerbate exclusion, reduce transparency, and have negative impacts on the natural environment, if not properly governed.

This paper aims at discussing uses of AI for policymaking, following the stages of the democratic policy lifecycle of agenda-setting, policy formulation, policy implementation, and policy evaluation. Across these four stages, we discuss opportunities and risks, drawing on recent exemplary cases and on the findings of relevant research. Based on this discussion, the paper also provides actionable recommendations that leverage the opportunities and mitigate risks in the use of AI for policymaking.





II.
AI AND POLICYMAKING: OPPORTUNITIES AND RISKS

Artificial intelligence has the potential to influence virtually every stage of the policymaking process, and major efforts have already started to be put into envisioning, designing, and experimenting with AI throughout the stages of the policy lifecycle.

The policy lifecycle is a classic framework to discuss policymaking activities (Jann & Wegrich, 2017). The framework includes an **agenda setting** stage, where policy problems are selectively identified; a **policy formulation** stage, where these problems are transformed into government programs; a **policy implementation** stage, where adopted policies are executed; and a **policy evaluation** stage, where policies are assessed against success criteria.

While each stage in this framework is characterized by a core focus, it is to be noted that in real-life different stages often overlap with each other, or can run in parallel; that stages do not necessarily follow each other in order, since a sequence of stages might be repeated multiple times; and that not all policies cover all the four stages in the cycle.

At each stage of the policy cycle, opportunities provided by AI co-exist with risks and limitations. While some of these opportunities, risks and limitations are unique to one or more of the stages, some are shared across all the lifecycle of a policy, and that can thus be considered as general features of the adoption of AI for policymaking.

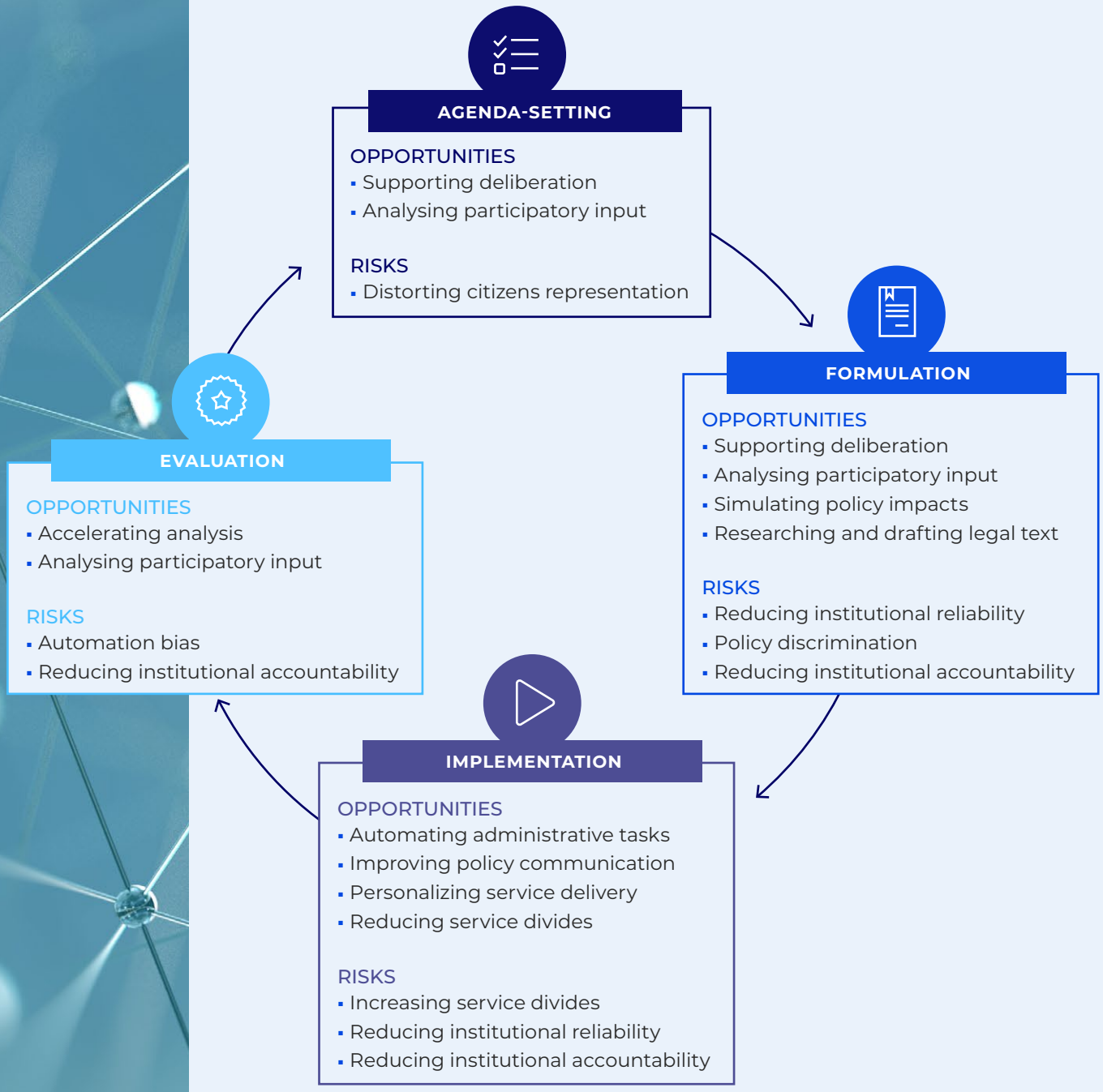
Table 1 and Figure 1 provide an overview of opportunities, risks and limitations in policymaking.

Table 1. Opportunities, risks and limitations of AI in policymaking

POLICY CYCLE STAGE	OPPORTUNITIES	RISKS AND LIMITATIONS	
Policy agenda-setting	▪ Supporting deliberation ▪ Analysing participatory input	▪ Distorting citizens representation	▪ Deskilling of public servants ▪ Vendor lock-in ▪ World-model deficiency ▪ Environmental degradation
Policy formulation	▪ Supporting deliberation ▪ Analysing participatory input ▪ Simulating policy impacts ▪ Researching and drafting legal text	▪ Reducing institutional reliability ▪ Policy discrimination ▪ Reducing institutional accountability	
Policy implementation	▪ Automating administrative tasks ▪ Improving policy communication ▪ Personalizing service delivery ▪ Reducing service divides	▪ Increasing service divides ▪ Reducing institutional reliability ▪ Reducing institutional accountability	
Policy evaluation	▪ Accelerating analysis ▪ Analysing participatory input	▪ Automation bias ▪ Reducing institutional accountability	

Source: Author's own elaboration.

Figure 1.
Opportunities and risks of AI in policymaking



The following sub-sections discuss opportunities and risks in using AI in connection to each of the stages of the policymaking cycle.

2.1 AGENDA-SETTING

Agenda-setting is the stage where recognized social problems are selectively elevated onto a public agenda. Different actors (e.g., political, societal, lobbyists and experts) frame issues, exploit media attention, and mobilize support to influence which issues gain priority while others are excluded (Jann & Wegrich, 2017).

> A key activity related to improving the agenda-setting phase where AI starts to be experimented with is **supporting public deliberation**.

Ideal public deliberation consist of an inclusive and reasoned exchange of arguments among citizens aimed at reaching mutual understanding (Habermas, 1996), and can be considered at the heart of a fair and sustainable agenda-setting process in policymaking.

For example, in Taiwan an AI tool known as Pol.is is used to collect and visualize the views of participants in offline and online discussion of which public problems to address. This becomes the basis for determining the extent of consensus about the nature of a policy issue, for example the regulation of ridesharing apps, telemedicine, and online alcohol sales (CrowdLaw for Congress, 2025).

In recent times, however, the sphere of public discourse has rapidly seen a degradation of the quality of the public debate, mainly in connection the fragmentation of worldviews due to the emergence of social media filter bubbles, the spread of digital disinformation and misinformation (also as a result of covert foreign influence operations), and the consequent increase in opinion polarization. To mitigate these phenomena, AI can be used as a support to public deliberation of a higher quality, as the input of a democratic and equitable agenda-setting process.

For example, a series of experiments with over 5000 participants has recently shown that an AI system based on large language models (LLMs) can help citizens debating controversial topics find a common ground (Tessler et al., 2024).

The AI-based "Habermas Machine" used in the experiment outperforms human mediators in generating shared group opinion statements, while still incorporating minority critiques. Drawing on the same affordance of LLMs, AI can thus also be envisioned to support politicians, high-level public officials, and other stakeholders (e.g., experts, business managers) involved in deliberations in the agenda-setting phase.

Artificial intelligence can also be used for **analysing participatory input** by actors involved in setting a policy agenda. In "Decide Madrid", a platform for citizens to propose policies to the city council, the application of Natural Language Processing (NLP) and machine learning (ML) enabled both the grouping of citizens, and the summarizing and clustering of topics proposed, mitigating the issue of information overload that these projects often suffer from (Arana-Catania et al., 2021). Agentic AI carries also great promises for this purpose. An AI agent, together with human oversight, could synthesize consultation inputs, categorize stakeholders, surface underrepresented perspectives, draft options and highlights risks in participation inputs.

With regards to the risks, using AI in this context can however result in **distorting citizen representation**. Even in the presence of well-intentioned initiatives, AI tools may in fact over-weight inputs from digitally savvy groups involved in deliberation, reinforcing the exclusion of underrepresented populations. Moreover, drawing on the capabilities of AI systems to efficiently analyze vast amounts of opinions expressed digitally by citizens, these systems can potentially be weaponized by governments—especially in polarized or authoritarian environments—to monitor dissent or suppress opposition under the guise of policy efficiency or security (Ünver, 2024).

Lastly, the use of AI in policymaking introduces a subtle but potent risk of manipulation in agenda setting.

Algorithmic systems optimized for engagement can steer collective attention and emotions toward specific issues, effectively constructing rather than reflecting public priorities (Zuiderwijk et al., 2021). Such influence can operate below the threshold of awareness, shaping salience and perception through affective cues and selective amplification (Glickman & Sharot, 2025). This raises critical concerns about transparency and democratic agency in AI-mediated governance, where the boundaries between persuasion and manipulation become increasingly blurred.

2.2 POLICY FORMULATION

The policy formulation stage transforms problems and demands into government programs. It involves defining objectives, considering alternative actions, negotiating among actors, integrating expert advice, and developing legislation (Jann & Wegrich, 2017).

Deliberative practices and the analysis of participatory input supported by AI also apply at this stage of the policy cycle, in a similar fashion to the agenda-setting phase.

> AI tools can enable feedback analysis from public consultations and support citizen deliberation for formulating policies (Bono Rossello et al., 2025).

In addition, AI can also be used for **simulating policy impacts**. Simulations can be used to anticipate the evolution of the implementation of policies from real metrics—a sort of "digital twin" for policies. For example, using GenAI, stakeholders in a project in the United Kingdom have been helped to visualise complex scientific evidence, generate visuals and "policy canvases" in workshops, to explore ideas in real time and refine policy options before decisions are made (UK Policy Lab, 2024).

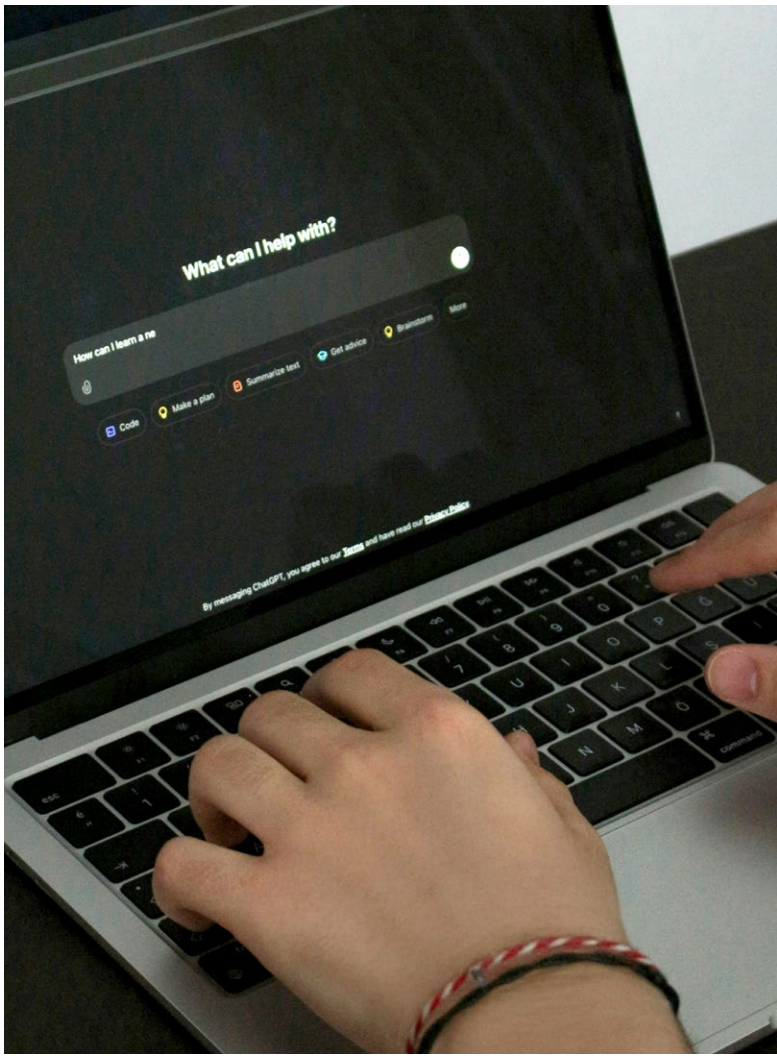
An area of policy formulation that carries a vast potential with regards to AI use is related to the legislation and regulation processes. For example, in a recent study a large language model has been used to generate and re-write scenarios simulating policy effects of new regulations—in this case, the European Union's Artificial Intelligence Act—on diverse areas, such as labour, well-being, social cohesion, and security (Barnett et al., 2024).

In addition, AI can support different aspects of **researching and drafting legal texts**. In the researching phase, AI applications can help parse, summarize, and analyse large corpuses of existing legal text to prepare the drafting of new laws. Subsequently, LLMs can help lawmakers ensure legal consistency and coherence in legal texts by detecting ambiguities, identifying connections with previous acts, and flagging obligations, rights, permissions, or penalties. Moreover, LLMs can detect drafting patterns, provide guidance on terminology, and highlight common errors. Lastly, LLMs can evaluate aspects of legal drafts, such as the digital-readiness of a legal text, their interoperability aspects, or compliance with policy mandates, such as gender-neutrality (Fitsilis & Mikros, 2024).

The use of LLMs in this context, however, also carries risks. Since LLMs draw on statistical and not on rule-based representations of the data they are trained on, they have an inherent tendency to hallucinate (Hannigan et al., 2024). For example, an AI model recently deployed by the European Parliament to facilitate citizen queries of its historical archives was found to give incorrect answers, including misidentifying the first President of the European Commission (Shrishak, 2025). Hallucinations in this context can end in **reducing institutional reliability** of legal authorities using AI for policy formulation.

Additionally, the use of AI tools in policy formulation can lead to **policy discrimination**. Predictive analytics and algorithms risk reinforcing bias and discrimination if training data is unrepresentative or historically skewed (Rooy, 2025). For example, The French National Family Allowance Fund (CNAF) applied a risk-scoring algorithm to detect overpayment or fraud in benefits. Civil society groups have challenged this system, arguing that it discriminates against people with disabilities, single mothers, and those in vulnerable low-income situations because the scoring criteria assign higher risk based on these attributes (Amnesty International, 2024).

Lastly, complex AI techniques, such as neural networks, make predictions through layers of statistical correlations that are difficult for humans to interpret, creating a "black box" effect that obscures how decisions are reached (Asatiani et al., 2021; Sun & Medaglia, 2019). This lack of transparency is problematic in policy formulation: opacity in algorithmic decision-making can thus result in reducing institutional accountability and undermines due process (Veale & Brass, 2019).



2.3 POLICY IMPLEMENTATION

The implementation stage consists of executing adopted policies, including specifying programs, allocating resources, and delivering services (Jann & Wegrich, 2017). A key element in this phase is improving the access of citizens to government.

> The most immediate use of AI to support the implementation of a policy is via **automating administrative tasks**.

Because AI systems can handle high-volume, low-complexity tasks such as claims processing, document verification, or eligibility screening, they allow reducing back office administrative process delays and human workloads. For example, AI can effectively support audits of unemployment insurance claims and improve both efficiency and fairness if used as a decision-support tool (Young et al., 2022). Generative AI, in particular, has the potential to free up valuable public sector time by supporting both back-office and frontline workers (Hashem et al., 2025).

The most relevant impacts of AI in policy implementation, however, can be expected in transforming the relationship between public authorities, as service providers, and citizens. When facing citizens, AI can support improving policy communication. For example, AI tools can help translate complex policy texts into more digestible formats. Under the Digital Europe Programme of the Commission, the WebText and Accessible Text language services simplify complex written content into short sentences and plain language (European Commission, 2025a).

With regards to **personalizing service delivery**, AI tools can tailor recommendations based on behavioural data and user profiles, reducing inaccuracy and inefficiencies that are typical of standardized service programs. For example, in a pilot in Finland, AI helped young people caring for family members to identify their hidden needs and automatically suggested tailored services, improving both the youths' awareness of support options and the system's ability to deliver timely assistance (Kopponen et al., 2024).

AI tools have also the potential of **reducing service divides**, by diminishing the barriers to access. For example, chatbots based on natural language and conversational input potentially allow more citizens to use public services, given their intuitive interfaces (Alishani et al., 2025). Paradoxically, AI systems deployed for service provision can at the same time lead to increasing service divides. AI-based services may in fact still leave behind those without access or skills, despite the potential of some AI applications for increased user-friendliness. AI systems still presume internet access, digital identity, and some literacy, which can further marginalize elderly, rural, low-income populations, or minorities (Carter et al., 2020).

A related risk of using AI for policy implementation activities is reducing institutional reliability, also linked to the hallucinations that can occur when service chatbots are not rule-based, but are powered by LLMs. Delegating service delivery choices to AI tools may hinder the ability to unambiguously identify who is responsible when service decisions cause harm, thus also reducing institutional accountability. When services fail due to flawed AI outputs, unclear chains of responsibility make redress difficult, especially where contractors or third-party systems are involved. AI-based services can undermine perceived legitimacy of public authorities, if citizens cannot understand how decisions are made, and public servants are unable to justify outcomes generated by black-boxed algorithms (Rachovitsa & Johann, 2022).

2.4 POLICY EVALUATION

The **evaluation** stage assesses whether policies achieve intended goals, examining both intended and unintended impacts. As part of the policy lifecycle, it informs redesign, continuation, or termination, and involves scientific, administrative, and political evaluations (Jann & Wegrich, 2017).

> Artificial intelligence can strengthen policy evaluation by widening the evidence base and **accelerating analysis**.

Natural language processing helps mine large volumes of consultation responses, monitoring reports and open data to detect patterns, such as emerging needs, sentiment shifts or recurring implementation barriers, far faster than any manual review.

For example, the European Commission's Text Mining and Analysis Competence Centre has operationalised such methods to support better regulation work and policy evaluations (European Commission, 2025b). Similarly, an AI tool has been used in Estonia to evaluate unemployed individuals who receive welfare services, based on data related to the labour market situation within the relevant segment for unemployed individuals, considering factors such as training, residence, and education (Vihalemm et al., 2025).

When policy evaluation is carried out by including the input of stakeholders other than public agencies (e.g., citizens, businesses, experts), AI can be used for analysing participatory input to the evaluation activities. In particular, AI agents could prove very helpful. After a policy is launched, an AI agent can support monitoring metrics, feedback, and contextual indicators to detect inequities, applies causal checks, explains drivers, and recommend targeted adjustments.

However, risks of using AI for policy evaluation are also significant. **Automation bias** is the tendency of decision-makers to over-rely on computer systems or algorithmic outputs, even when they are flawed (Rooy, 2025). For instance, as AI systems cannot provide outputs that go beyond their training, they cannot be expected to find new avenues for existing problems. In policy evaluation, automation bias occurs when evaluators accept AI-generated findings without sufficient critical scrutiny, sidelining contradictory evidence from qualitative insights or local expertise (*errors of commission*); or when they fail to notice problems or seek alternatives because they assume the automated system must be right (*errors of omission*). For example, an algorithm used in the UK to replace cancelled A-level exams was found to disproportionately downgrade disadvantaged students, but only after officials repeatedly relied on its "objectivity", due to an automation bias (Kippin & Cairney, 2022).

Lastly, the use of AI in policy evaluation can also contribute to reducing institutional accountability, when the algorithms used to evaluate policy outcomes are inscrutable and when their outputs cannot be explained by evaluators.

2.5 COMMON RISKS AND LIMITATIONS ACROSS THE POLICY LIFECYCLE

Besides the area-specific risks of AI applications in policymaking, there are four risks and limitations to AI that cut across all domains of use.

> Risks include the deskilling of public servants and vendor lock-in. Limitations, inherent to AI systems, include world-model deficiency and environmental impacts.

These transversal risks and limitations highlight systemic vulnerabilities that governments must address when integrating AI into policymaking.

First, the **deskilling of public servants** can arise when repetitive or complex tasks are increasingly automated, potentially undermining both individual expertise and the institutional memory essential to policymaking (Parra-Moyano et al., 2025). Overreliance on AI systems can diminish professional judgment, especially in areas such as legal interpretation, case assessment, or data analysis. In Denmark, for instance, municipalities experimented with AI tools for child welfare case management. While they promised efficiency, critics argued that they risked reducing social workers' discretion and contextual understanding (Jørgensen, 2023). Moreover, when entire job categories are displaced by AI, public employees may face unemployment or redeployment challenges, raising ethical and political concerns about workforce resilience.

It is to be noted, however, that evidence pointing towards deskilling and job displacement impacts has to be contextualized within a wider and more complex view of job shift, where some tasks are automated, while others are augmented with the use of AI (World

Economic Forum, 2023). AI is reshaping work often not by eliminating jobs outright but by unbundling tasks, augmenting some cognitive activities while automating routine ones, which creates both productivity gains and distributional risks, and undermines apprenticeship-style learning (Feijóo et al., 2026).

Second, considerations around long-term flexibility arise when using AI systems across different policy stages. Much of today's AI infrastructure is developed and maintained by a relatively small number of commercial providers, which can shape how public organizations configure and evolve their digital capabilities. When governments rely on specialized or proprietary solutions for key tasks, it may become more challenging—not impossible, but more complex—to modify, diversify, or transition these systems over time.

A frequently cited example concerns the extensive use of cloud and AI services offered by large global providers, many of which are headquartered outside Europe. This market structure can, at times, expose public administrations to factors such as pricing adjustments, contractual limitations, or technology roadmaps that evolve independently of government needs. These dynamics do not imply inherent problems, but they highlight the value of maintaining flexibility and optionality in procurement strategies.

Besides these risks, there are inherent limitations to AI systems that must be considered. AI applications in policy making, particularly large language models (LLMs) and agentic AI, do not learn through self-directed engagement with the world in any meaningful sense. This **world-model deficienc**y is engrained in their training process that consists of statistical inference over vast corpora of human-generated text: these systems construct an internal model of linguistic regularities (i.e., a model of what a human would likely say next) rather than a model of how the world itself operates or responds to actions. As LLMs lack embodied feedback and mechanisms to update their representations in light of new interactions or consequences, genuine continual and realistic learning will thus require fundamentally new architectures that integrate perception, action, and feedback.

Lastly, AI applications for policy making carry important consequences in terms of potential **environmental degradation.** While AI can contribute to sustainability by optimizing resource allocation, enabling climate modelling (Medaglia et al., 2021) or reducing CO_2 emissions if AI is fully integrated in the energy system and rebound effects are minimal (IEA, 2025), the deployment of large-scale foundational models is energy- and water-intensive (Marabelli & Davison, 2025). For example, generating an image with AI can use the energy equivalent of half a smartphone charge (Luccioni et al., 2024).

Recent studies show that data centers in key AI infrastructure hubs in Europe, such as Ireland and the Netherlands, consume vast amounts of electricity and water (Kamiya & Bertoldi, 2024), potentially straining local resources and challenging the EU's twin green and digital transition commitments.

For example, in a country like Spain, already facing increasing concerns about the water usage in areas traditionally afflicted by drought, the government has recently launched a public consultation on a decree to impose strict energy efficiency and sustainability requirements on new data centers (Ministerio para la Transición Ecológica y el Reto Demográfico, 2025).

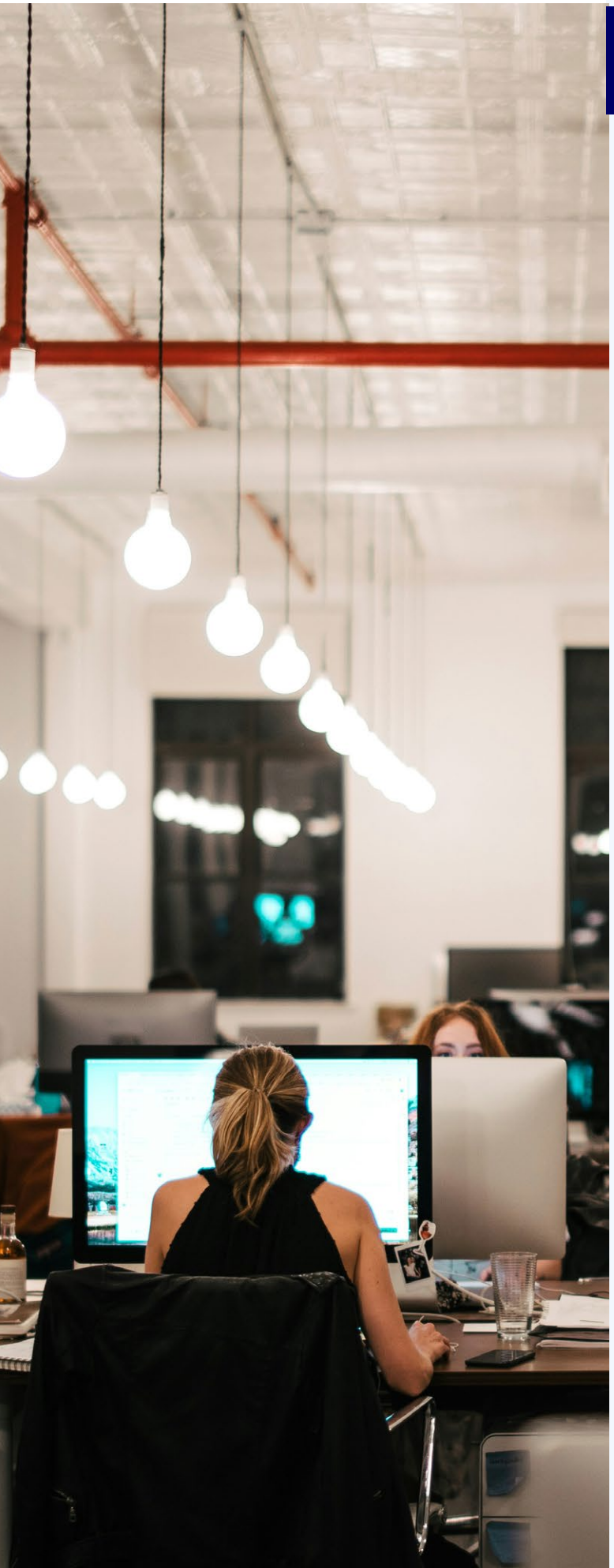
III. ACTIONABLE RECOMMENDATIONS FOR PUBLIC AND PRIVATE ACTORS

The discussion of potentials and risks across the agenda-setting, policy formulation, policy implementation, and policy evaluation stages makes clear that AI in policymaking embodies a duality: it can expand inclusiveness, transparency, and efficiency, but it can just as easily amplify exclusion, opacity, and institutional fragility.

Addressing this duality requires a set of actionable pathways that directly link opportunities to risks, ensuring that AI enhances democratic prosperity, also considering its inherent limitations.

The following pathways outline such strategies, with targeted recommendations for both public and private actors.

Figure 2. Actionable recommendations for public and private actors



3.1 PATHWAY 1: BUILDING AI LITERACY AS A FOUNDATION FOR POLICYMAKING

Opportunities such as AI-supported deliberation (Tessler et al., 2024) and automated administrative tasks (Young et al., 2022) cannot be realized without policymakers who understand both technical affordances and governance implications. Risks such as deskilling (Parra-Moyano et al., 2025) and automation bias (Kippin & Cairney, 2022) stem directly from gaps in AI literacy. Moreover, the unbundling of tasks ignited by AI, augmenting some cognitive activities while automating routine ones and providing distributional risks together with productivity gains (Feijóo et al., 2026), calls for actions that support literacy building.

A recent survey by the European Centre for the Development of Vocational Training indicates that 6 in 10 adult workers believe that, in the next 5 years, most of the impact of AI on their work will materialize in the form of new skill needs (Cedefop, 2025). At the same time, another recent European survey of nearly 5,000 employees indicates that only 18% of public sector workers feel they received sufficient AI training, the lowest among sectors (EY, 2025).

Although in this paper we are aware of different semantic distinctions, for clarity's sake we here refer to literacy as a synonym of competence, defined as 'the ability to apply learning outcomes adequately in a defined context (education, work, personal or professional development)' (Cedefop, 2008, p. 47). AI literacy should not only include technical competences, but also managerial, and policy / legal / ethical ones (Medaglia et al., 2024). AI-literate stakeholders (public and private) should bridge AI system designs into policymaking implications and vice versa (Misuraca et al., 2025).

ACTIONABLE RECOMMENDATIONS

01 **Incorporate AI literacy across existing training programs:** Integrate modules on AI ethics, legal frameworks, governance and technical basics into national civil service academies. Examples of successful AI education for policymakers remain rare but promising. The Master in Artificial Intelligence for Public Services (AI4Gov)[1], for example, shows that structured learning can strengthen cross-disciplinary fluency (Misuraca et al., 2025).

02 **Establish dedicated academies and platforms and leverage existing ones:** Scale up initiatives such as the recent AI4Gov-X[2] initiative—a joint platform for learning, experimentation, and knowledge sharing across governments in Europe to foster shared innovation and reduce risk—into a European-wide "AI for Policy Academy," pooling resources to reduce duplication and create peer-learning networks across administrations.

03 **Develop AI talent pipelines:** Launch fellowships and targeted recruitment campaigns at national and local level. Governments and private actors should commit to building long-term AI talent pipelines that address different demographic and professional groups. This means designing initiatives that target both current civil servants but also younger generations entering the workforce, mid-career professionals seeking to reskill, and underrepresented populations. At the same time, fellowship programs between government, academia, and industry can bring technical experts into the policy domain, while public sector professionals should be offered opportunities to rotate into AI-intensive roles to strengthen cross-sectoral exchange.

04 **Mandate program evaluation:** Introduce systematic impact evaluations for literacy programs to avoid symbolic training and ensure measurable improvements in critical capacity. With AI technologies evolving rapidly, evaluation needs to be dynamic and continuous, rather than a one-off assessment at the end of a funding cycle. Governments should require real-time or near real-time monitoring of program outcomes, using both quantitative indicators (e.g., number of participants trained, changes in competency levels) and qualitative assessments (e.g., participant confidence, ability to apply AI responsibly in policy contexts). Evaluation should also capture whether programs meet the needs of diverse target populations, such as local administrators, senior policymakers, or frontline public servants, who face different challenges in AI adoption.

By narrowing the literacy gap, governments can better scrutinize AI vendors, avoid overreliance on "black box" systems, and preserve public managers' judgment.

3.2 PATHWAY 2: EMBEDDING ACCOUNTABILITY AND RULE-OF-LAW PRINCIPLES

While AI can enhance legal drafting and regulatory experimentation (Fitsilis & Mikros, 2024), it also threatens to reduce institutional accountability and reproduce bias (Amnesty International, 2024). To ensure legitimacy, AI for policymaking must be embedded in democratic guardrails.

ACTIONABLE RECOMMENDATIONS

01 **Mandate pre-deployment rights and legal audits:** Go beyond technical checks to include discrimination testing, environmental footprint analysis, and explainability. Empirical studies of welfare algorithms in France and the Netherlands show discriminatory impacts when such safeguards are absent (Amnesty International, 2024; Rachovitsa & Johann, 2022).

02 **Institutionalize cross-functional procurement teams:** Require that procurement boards include legal experts, ethicists, technologists, and civil society observers. For example, Estonia's AI sandbox "Accelerate Estonia" demonstrates the value of multi-stakeholder experimentation in mitigating legal uncertainty (Accelerate Estonia, 2025). Contracting frameworks should also specify liability distribution between vendors and agencies.

03 **Operationalize rights impact assessments:** Provide templates and capacity-building resources for agencies to conduct assessments, ensuring that rights protection is practical. These templates should include frameworks, for example, to reduce AI bias in policy making (Rooy, 2025).

Framing this pathway as AI within democratic guardrails reinforces a legitimacy narrative: AI governance is not only about compliance, but about safeguarding liberal democratic values.

3.3 PATHWAY 3:
INVESTING IN OPEN, SUSTAINABLE, AND SOVEREIGN AI INFRASTRUCTURE

Opportunities lie in strengthening open repositories and interoperable infrastructures that broaden the range of technological options available to public administrations. Such approaches can complement existing partnerships with leading commercial providers. Some analyses also note that limited provider diversity can reduce flexibility over time, particularly when services are highly specialized or deeply integrated into public workflows (European Parliament, 2025). Recent assessments indicate that a substantial share of advanced digital and cloud services used by European governments and private organizations currently originates from outside the EU (European Commission, 2025). While these providers offer high-performing and reliable solutions, the figures also point to an opportunity for Europe to further develop its own capabilities and increase the diversity of the ecosystem.

While openness in AI infrastructures and models can enable transparency, reproducibility, collaboration, and the democratization of innovation, appropriate governance is needed to mitigate its potential risks. For example, recent developments such as the release of Meta's LLaMA, DeepSeek in China, and open-access variants of generative pre-trained transformer (GPT) models, illustrate a strong global trend towards openness that, paradoxically, may place Europe at risk of falling behind. U.S. and Chinese actors are scaling open solutions at massive speed, backed by computing resources and funding levels that far exceed European investments. Moreover, a proliferation of open variants without governance frameworks can lead to fragmented ecosystems, duplication of efforts, and uneven quality standards.

In addition, emerging evidence suggests that the rapid expansion of AI-related infrastructure may carry notable environmental implications. Some data indicate that energy consumption by data centers has risen significantly in several European locations. For instance, studies report that data centers in Ireland and the Netherlands can account for over 15% of national electricity use during peak periods (Kamiya & Bertoldi, 2024), illustrating the importance of integrating sustainability considerations into future planning. Managing infrastructure growth with environmental sustainability is a key area to focus on in using AI for policymaking.

ACTIONABLE RECOMMENDATIONS

01 **Tie public funding to open-source conditions**: Require that code, models, or datasets developed with public money be published under open licenses, so that governments can reduce dependency on single vendors and foster transparency and reproducibility. This also encourages collaboration between researchers, public administrations, and civic groups (Theben et al., 2021), helping ensure that AI systems used for policymaking are not only technically sound but also socially accountable.

02 **Create public repositories of models and data:** Similar to the European Data Portal[3], such repositories could provide trained models and benchmark datasets for reuse across administrations. Shared access would reduce duplication of effort, lower entry barriers for smaller public agencies, and establish common evaluation standards, ultimately accelerating responsible and cost-effective AI adoption in government.

03 **Strengthen technological sovereignty:** By directing funds toward computing capacity and large-scale model training within national and regional scope, governments can increase strategic agency. Such increase should happen gradually, also by designing AI applications that are as much as possible future-proof in avoiding lock-in situations. This approach, while geared towards the medium- and long-term, is still compatible with competitiveness, while supporting resilience and the ability to align AI development with democratic values.

04 **Invest in green AI infrastructures:** Requiring environmental impact disclosures for high-compute projects, and incentivizing the use of energy-efficient architectures, model compression, and optimisation techniques, would encourage the growth of environmentally sustainable infrastructures. This ensures that the push for digital innovation in policymaking also supports its environmental sustainability.

4. CONCLUSIONS

This paper has examined how artificial intelligence can shape policymaking across the entire policy cycle: from agenda-setting and policy formulation to implementation and evaluation. AI offers significant opportunities to improve inclusiveness, efficiency, and evidence use in public decision-making. At the same time, it features key limitations that range from world-model deficiency, to environmental degradation. The risks of using AI in policymaking thus include introducing distorting citizens representation, policy discrimination, increasing divides, and reducing automation accountability. These dynamics highlight the dual nature of AI: it can serve as an engine of democratic prosperity, or as a source of exclusion and institutional fragility.

The analysis demonstrates that AI governance cannot rely on narrow tools alone. The apparent intuitive nature of, and the low access barriers to, powerful AI applications also means that, without appropriate literacy and governance models, using AI to erode democracy became easier than using it for democratic prosperity. Rights assessments, cross-functional procurement teams, and pre-deployment audits are vital, but they are insufficient to anchor AI in democratic and sustainable trajectories. What is required is investment in digital and AI infrastructures that are value-based by design, not retrofitted with safeguards after deployment. In some cases, such infrastructures may need to be publicly owned and operated to guarantee accountability, resilience, and long-term sustainability. Regulation and procurement frameworks matter, but public investment is a decisive factor.

The recommendations offered in this paper point to three complementary pathways:

1. expanding AI literacy to enable policymakers and civil servants to exercise informed judgment;
2. embedding accountability and rule-of-law principles into the design and use of AI systems; and
3. developing open, sustainable, and sovereign AI models and infrastructures to reduce dependency on a small number of dominant providers.

Taken together, these pathways chart a course toward AI use in policymaking that is legitimate, inclusive, and trustworthy.

Looking ahead, mitigation efforts must focus where the risks are most acute. Policy formulation and implementation are particularly sensitive stages: biased predictive systems or opaque service delivery tools can entrench inequality and undermine public trust at scale. By contrast, agenda-setting and evaluation offer more experimental ground, where AI can support deliberation, broaden participation, and accelerate evidence use, provided safeguards remain in place. Prioritizing these risk hotspots allows governments to balance the advantages of AI while protecting against its most damaging effects.

The trajectory of AI in policymaking is still open. Whether AI reinforces democratic prosperity or erodes it depends on choices made today about governance, infrastructure, and investment.

REFERENCES

- Accelerate Estonia. (2025). *Making illegal things legal*. Accelerate Estonia. https://accelerate.ee/
- Alishani, A., Homburg, V., & Velsberg, O. (2025). Public Encounters and Government Chatbots: When Servers Talk to Citizens. *Public Administration Review, 0*(1), 1–10. https://doi.org/10.1111/puar.70005
- Amnesty International. (2024). France: Discriminatory algorithm used by the social security agency must be stopped. Amnesty International. https://www.amnesty.org/en/latest/news/2024/10/france-discriminatory-algorithm-used-by-the-social-security-agency-must-be-stopped
- Arana-Catania, M., Lier, F.-A. V., Procter, R., Tkachenko, N., He, Y., Zubiaga, A., & Liakata, M. (2021). Citizen participation and machine learning for a better democracy. *Digital Government: Research and Practice, 2*(3), [1-22]. https://doi.org/10.1145/3452118
- Asatiani, A., Malo, P., Nagbøl, P., Penttinen, E., Rinta-Kahila, T., & Salovaara, A. (2021). Sociotechnical Envelopment of Artificial Intelligence: An Approach to Organizational Deployment of Inscrutable Artificial Intelligence Systems. *Journal of the Association for Information Systems, 22*(2), 325–352. https://doi.org/10.17705/1jais.00664
- Barnett, J., Kieslich, K., & Diakopoulos, N. (2024). Simulating Policy Impacts: Developing a Generative Scenario Writing Method to Evaluate the Perceived Effects of Regulation. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 7*(1), 82–93. https://doi.org/10.1609/aies.v7i1.31620
- Bono Rossello, N., Simonofski, A., & Castiaux, A. (2025). Artificial intelligence for digital citizen participation: Design principles for a collective intelligence architecture. *Government Information Quarterly, 42*(2), 102020. https://doi.org/10.1016/j.giq.2025.102020
- Carter, L., Liu, D., & Cantrell, C. (2020). Exploring the Intersection of the Digital Divide and Artificial Intelligence: A Hermeneutic Literature Review. *AIS Transactions on Human-Computer Interaction, 12*(4), 253–275. https://doi.org/10.17705/1thci.00138
- Cedefop (Ed.). (2008). *Terminology of European education and training policy: A selection of 100 key terms*. Office for Official Publications of the European Communities.
- Cedefop. (2025). *Skills empower workers in the AI revolution: First findings from Cedefop's AI skills survey* [Policy brief]. Publication Office of the European Union. https://www.cedefop.europa.eu/en/publications/9201
- CrowdLaw for Congress. (2025). *vTaiwan: Using digital technology to write digital laws*. https://congress.crowd.law/files/vtaiwan-case-study.pdf
- European Commission. (2025a). *New AI services to simplify your texts*. https://commission.europa.eu/news-and-media/news/new-ai-services-simplify-your-texts-2025-03-24_en
- European Commission. (2025b). *Text Mining and Analysis Competence Centre*. https://knowledge4policy.ec.europa.eu/text-mining_en
- European Commission (Ed.). (2025). *The future of European competitiveness: Part A: A competitiveness strategy for Europe*. Publications Office. https://doi.org/10.2872/1823372
- European Parliament (2025). *Report on European technological sovereignty and digital infrastructure* (Nos. A10-0107/2025). https://www.europarl.europa.eu/doceo/document/A-10-2025-0107_EN.html
- EY. (2025). *European AI Barometer 2025*. https://www.ey.com/en_lu/insights/ai/ey-european-ai-barometer-2025
- Fatima, S., Desouza, K. C., & Dawson, G. S. (2020). National strategic artificial intelligence plans: A multi-dimensional analysis. *Economic Analysis and Policy,* 67, 178–194. https://doi.org/10.1016/j.eap.2020.07.008
- Feijóo, C., Blazquez, I., Luca de Tena, C., & Ugarelli, G. (2026). *AI for prosperity—A time of transition*.
- Fitsilis, F., & Mikros, G. (2024). *AI-based solutions for legislative drafting in the EU: Summary report*. Publications Office of the European Union. https://data.europa.eu/doi/10.2799/6892931
- Glickman, M., & Sharot, T. (2025). How human—AI feedback loops alter human perceptual, emotional and social judgements. *Nature Human Behaviour, 9*(2), 345–359. https://doi.org/10.1038/s41562-024-02077-2
- Habermas, J. (1996). *Between Facts and Norms. Contribution to a Discourse Theory of Law and Democracy*. MIT Press.
- Hannigan, T. R., McCarthy, I. P., & Spicer, A. (2024). Beware of botshit: How to manage the epistemic risks of generative chatbots. *Business Horizons, 67*(5), 471–486. https://doi.org/10.1016/j.bushor.2024.03.001
- Hashem, Y., Bright, J., Chakraborty, S., Onslow, K., Francis, J., Poletav, A., & Esnaashari, S. (2025). *Mapping the Potential of Generative AI and Public Sector Work: Using time use data to identify opportunities for AI adoption in Great Britain's public sector*. The Alan Turing Institute. https://www.turing.ac.uk/news/publications/mapping-potential-generative-ai-and-public-sector-work-using-time-use-data

- Hine, E., & Floridi, L. (2024). Artificial intelligence with American values and Chinese characteristics: A comparative analysis of American and Chinese governmental AI policies. *AI & SOCIETY, 39*(1), 257–278. https://doi.org/10.1007/s00146-022-01499-8
- IEA (2025), *AI and climate change.* https://www.iea.org/reports/energy-and-ai/ai-and-climate-change
- Ilves, L., Kilian, M., Peixoto, T., & Velsberg, O. (2025). *An Introduction to Agentic AI for Governments—Berlin Global GovTech Centre*. Global Government Technology Centre. https://www.globalgovtechcentre.org/an-introduction-to-agentic-ai-for-governments
- Jann, W., & Wegrich, K. (2017). Theories of the policy cycle. In *Handbook of public policy analysis* (pp. 69–88). Routledge.
- Jørgensen, A. M. (2023). *Digitalisation of social services for children, young people and families in Denmark: Country report 2022*. https://vbn.aau.dk/files/544755316/J_rgensen_2023_Webseite_DK_digitalisation_and_social_services_for_children_Jorgensen.pdf
- Kamiya, G., & Bertoldi, P. (2024). *Energy Consumption in Data Centres and Broadband Communication Networks in the EU.* JRC Publications Repository. https://doi.org/10.2760/706491
- Kippin, S., & Cairney, P. (2022). The COVID-19 exams fiasco across the UK: Four nations and two windows of opportunity. *British Politics, 17*(1), 1–23. https://doi.org/10.1057/s41293-021-00162-y
- Kopponen, A., Hahto, A., Villman, T., Kettunen, P., Mikkonen, T., & Rossi, M. (2024). Personalised public services powered by AI: the citizen digital twin approach. In *Research Handbook on Public Management and Artificial Intelligence*. Edward Elgar. https://www.elgaronline.com/edcollchap/book/9781802207347/book-part-9781802207347-20.xml
- Luccioni, S., Jernite, Y., & Strubell, E. (2024). Power Hungry *Processing: Watts Driving the Cost of AI Deployment? Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 85–99. https://doi.org/10.1145/3630106.3658542
- Marabelli, M., & Davison, R. M. (2025). Artificial intelligence and the environment: Ethical challenges and strategic opportunities for organizations. *The Journal of Strategic Information Systems, 34*(3), 101921. https://doi.org/10.1016/j.jsis.2025.101921
- Medaglia, R., Mikalef, P., & Tangi, L. (2024). *Competencies and governance practices for AI in the public sector*. Publications Office of the European Union. https://data.europa.eu/doi/10.2760/7895569
- Medaglia, R., Misuraca, G., & Aquaro, V. (2021). Digital Government and the United Nations' Sustainable Development Goals: Towards an analytical framework. *DG.O2021: The 22nd Annual International Conference on Digital Government Research*, 473–478. https://doi.org/10.1145/3463677.3463736
- Ministerio para la Transición Ecológica y el Reto Demográfico. (2025). *Audiencia e información pública sobre el proyecto de Real Decreto por el que se regula la eficiencia energética y la sostenibilidad para los centros de datos*. https://www.miteco.gob.es/es/energia/participacion/2025/detalle-participacion-publica-k-775.html
- Misuraca, G. C., Medaglia, R., Feijoo, C., Mortati, M., & Noordt, C. van. (2025). Developing competences for Artificial Intelligence in the Public Sector: The AI4Gov Canvas. *Annual International Conference on Digital Government Research, 26*. Annual International Conference on Digital Government Research. https://doi.org/10.59490/dgo.2025.951
- Parra-Moyano, J., Reinmoeller, P., & Schmedders, K. (2025). Research: Executives Who Used Gen AI Made Worse Predictions. *Harvard Business Review*. https://hbr.org/2025/07/research-executives-who-used-gen-ai-made-worse-predictions?tpcc=orgsocial_edit&utm_campaign=hbr&utm_medium=social&utm_source=linkedin
- Rachovitsa, A., & Johann, N. (2022). The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case. *Human Rights Law Review, 22*(2), 1–15. https://doi.org/10.1093/hrlr/ngac010
- Rooy, D. V. (2025). A cognitive science framework could prevent harmful public policy decisions involving AI. *Behavioral Science & Policy, 11*(1), 33–43. https://doi.org/10.1177/23794607251336099
- Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly, 36*(2), 368–383. https://doi.org/10.1016/j.giq.2018.09.008
- Tangi, L., Rodriguez Müller, A. P., & Combetto, M. (2025). A Silent Partner: The Shadow Presence of Generative Artificial Intelligence in Public Administrations. In S. Hofmann, L. Danneels, R. Dobbe, A.-S. Novak, P. Parycek, G. Schwabe, V. Spitzer, & J. Ubacht (Eds.), *Electronic Participation* (pp. 69–86). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-02515-9_5
- Tessler, M. H., Bakker, M. A., Jarrett, D., Sheahan, H., Chadwick, M. J., Koster, R., Evans, G., Campbell-Gillingham, L., Collins, T., Parkes, D. C., Botvinick, M., & Summerfield, C. (2024). AI can help humans find common ground in democratic deliberation. *Science, 386*(6719), eadq2852. https://doi.org/10.1126/science.adq2852

- The Ministry for Digital Transformation and the Civil Service. (2025). *Consenso por una administraciòn abierta*. https://digital.gob.es/consenso-por-una-administracion-abierta.html
- Theben, A., Gunderson, L., & López-Forés, L. (2021). *Challenges and limits of an open source approach to Artificial Intelligence*. European Parliament's Special Committee on Artificial Intelligence in a Digital Age (AIDA).
- UK Policy Lab. (2024). *Generative AI in subsurface science and policy*. https://openpolicy.blog.gov.uk/2024/07/30/generative-ai-in-subsurface-science-and-policy/
- Ünver, H. A. (2024). *Artificial intelligence (AI) and human rights: Using AI as a weapon of repression and its impact on human rights* | Think Tank | European Parliament. European Parliament. https://www.europarl.europa.eu/thinktank/en/document/EXPO_IDA(2024)754450
- Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly, 37*(4), 101509. https://doi.org/10.1016/j.giq.2020.101509
- van Noordt, C., Medaglia, R., & Tangi, L. (2025). Policy initiatives for Artificial Intelligence-enabled government: An analysis of national strategies in Europe. *Public Policy and Administration, 40*(2), 215–253. https://doi.org/10.1177/09520767231198411
- Veale, M., & Brass, I. (2019). *Administration by Algorithm? Public Management Meets Public Sector Machine Learning* (SSRN Scholarly Paper No. 3375391). Social Science Research Network. https://papers.ssrn.com/abstract=3375391
- Vihalemm, T., Männiste, M., Trumm, A., & Solvak, M. (2025). Specialists and Algorithms: Implementation of AI in the Delivery of Unemployment Services in Estonia. In P. Ahrweiler (Ed.), *Participatory Artificial Intelligence in Public Social Services: From Bias to Fairness in Assessing Beneficiaries* (pp. 97–117). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-71678-2_5
- Viscusi, G., Rusu, A., & Florin, M.-V. (2020). Public Strategies for Artificial Intelligence: Which Value Drivers? *Computer, 53*(10), 38–46. https://doi.org/10.1109/MC.2020.2995517
- Wilson, C. (2022). Public engagement and AI: A values analysis of national strategies. *Government Information Quarterly, 39*(1), 101652. https://doi.org/10.1016/j.giq.2021.101652
- World Economic Forum. (2023). *Jobs of Tomorrow: Large Language Models and Jobs*. https://www3.weforum.org/docs/WEF_Jobs_of_Tomorrow_Generative_AI_2023.pdf
- Young, M. M., Himmelreich, J., Honcharov, D., & Soundarajan, S. (2022). Using artificial intelligence to identify administrative errors in unemployment insurance. *Government Information Quarterly, 39*(4), 101758. https://doi.org/10.1016/j.giq.2022.101758
- Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly, 38*(3), 101577. https://doi.org/10.1016/j.giq.2021.101577

ENDNOTES

1 https://www.polidesign.net/en/formazione/business-design/master--artificial-intelligence-for-public-services

2 https://re-imagine.eu/ai4gov-x-project-article

3 https://data.europa.eu/en

IE UNIVERSITY
CENTER FOR THE GOVERNANCE OF CHANGE

WRITTEN BY:
Rony Medaglia

RECOMMENDED CITATION:
Medaglia, Rony. *Better policymaking in the age of AI*.
IE CGC, October 2025

Photos: Unsplash, Shutterstock
Design: epoqstudio.com