Mitigating the Risks of Generative AI in Government Through Algorithmic Governance

Mark Esposito, Daphne Halkias, Terence Tse, and Tatiana Harkiolakis

Abstract The launch of the generative artificial intelligence (gen AI) application ChatGPT by OpenAI launched artificial intelligence into public discourse and led to a wave of mass uptake of this technology in organizations in the private sector. At the same time, AI is increasingly incorporated into government functions and the public sector. We propose that governments and the public sector can set an example for the responsible use of AI technologies by following the principles of algorithmic governance traditionally recommended to the private sector. Algorithmic governance has historically been defined in the literature as governance by algorithms, or how artificial intelligence is used to make governance decisions and affect social ordering. However, we take an alternative approach; instead, we conceptualize algorithmic governance as the governance of algorithms. We summarize the risks of generative AI use in governments and the public sector, then outline algorithmic governance principles, a step-by-step approach to implementing algorithmic governance into government or public sector projects, opportunities for inter-sector collaboration, and conclusions.

M. Esposito
Berkman Klein Center for Internet & Society at Harvard University, Boston, MA, USA
e-mail: markesposito@fas.harvard.edu

D. Halkias
Ecole des Ponts ParisTech Business School, Paris, France
e-mail: daphne.halkias@gmail.com

T. Tse
HULT International Business School, London, UK
e-mail: terence.tse@faculty.hult.edu

T. Harkiolakis (✉)
ISC Paris Business School, Center for Policy and Competitiveness, Ecole des Ponts ParisTech Business School, Paris, France
e-mail: tatiana.harkiolakis@gmail.com

A. Skeja et al. (eds.), *Navigating AI in Business*, Contributions to Management Science, https://doi.org/10.1007/978-3-032-02526-5_5

1 Introduction

The release of OpenAI's ChatGPT marked a significant turning point, propelling generative artificial intelligence (gen AI) into widespread public discourse and accelerating its adoption within private sector organizations. While this surge in interest is notable, it represents an acceleration of a trend evident over the preceding decade. Key milestones include Google's 2014 acquisition of DeepMind (Shu, 2014) and subsequent substantial investments in large language models (LLMs) by major technology companies, including Microsoft and Google. The shift in terminology from LLMs to "gen AI," coupled with the capacity to generate diverse outputs, including language, visual, and audio content from user prompts, has further increased user engagement. This accessibility has democratized access to AI and amplified concerns about its potential misuse.

The governance of generative AI (gen AI) is a growing concern (Esposito & Tse, 2018). In the United States, regulatory discourse centers on the appropriate level of intervention, considering the nation's prominent role in gen AI development. Conversely, European Union legislators are debating the applicability of existing AI legislation to this emerging technology (Entsminger et al., 2023). Initial governance efforts have prioritized labeling AI-generated video, audio, text, and images to mitigate the risk of public deception (Eastwood, 2024). These measures are, in part, a response to the potential for computational propaganda, wherein state actors leverage AI to manipulate public opinion and disseminate disinformation (Harkiolakis & Esposito, 2023). The ease with which gen AI can create realistic fake content necessitates robust authentication and provenance tracking mechanisms.

The rising prominence of gen AI has amplified anxieties regarding its potential adverse effects, prompting increased scholarly and industry focus on risk identification and mitigation (Bird et al., 2023; Hacker et al., 2023) in the private sector as well as in entrepreneurial ecosystems (Fergnani et al., 2020; Groth et al., 2015). Simultaneously, AI is being progressively integrated into governmental operations and the public sector (Engstrom & Ho, 2020), supporting decision-making, public service delivery, policy formulation, and other functions. Applications span diverse domains, including urban planning and transport, environmental regulation, welfare planning, healthcare resource allocation, industry compliance, and energy management (Henman, 2020). While government regulation of the AI industry has garnered considerable attention, the internal governance of governmental AI usage to minimize risks remains comparatively under-explored (Wirtz et al., 2020). This internal governance is crucial, as governments must regulate AI in the private sector and ensure its responsible use within their operations. A lack of internal governance can lead to inconsistencies in policy and erode public trust. Furthermore, governments are uniquely accountable for protecting citizen rights and ensuring equitable outcomes when deploying AI in public services.

2 The Risks of Gen AI in the Government

Artificial intelligence (AI) involves the capacity of machines to execute tasks typically requiring human intelligence (Copeland, 2024). While AI can mimic certain cognitive functions, such as problem-solving and pattern recognition, it currently lacks the capacity for human-level creativity, critical thinking, and empathy (Montemayor et al., 2022; Preece & Çelik, 2023). A key concern regarding AI implementation in government is the potential for biased outputs. Because AI systems are trained on existing data, they can inherit and amplify biases and stereotypes present within that data. This can lead to increased discrimination against marginalized groups, exacerbating existing inequalities with significant economic, social, and political ramifications (Skaug Sætra, 2020).

A common example from the tech industry illustrates this issue: social networks often employ algorithms that curate content based on user preferences, creating "echo chambers" where individuals are primarily exposed to information reinforcing their existing beliefs, thus contributing to polarization and potentially influencing political outcomes (Boulianne et al., 2020; Syeda, 2024). Within governments, algorithmic bias has been identified as a potential driver of racial, gender, and socioeconomic discrimination (Henman, 2020). Predictive policing, which utilizes algorithms and data analysis to forecast criminal activity and allocate law enforcement resources, is a particularly contentious application. Critics argue that this technology perpetuates racial bias by disproportionately targeting minority groups and neighborhoods based on historical crime data (Díaz, 2021). While predictive policing has been primarily piloted in the U.S.A., several jurisdictions have recently discontinued their programs due to these concerns (Lau, 2024).

A further prominent risk associated with generative AI is its potential to facilitate the spread of propaganda and misinformation. The technology lowers the cost and effort required to create and disseminate persuasive content that manipulates public opinion (Yang & Roberts, 2023). Experts suggest that advancements in AI may lead to the development of chatbots and other interfaces trained on private knowledge bases explicitly designed to generate tailored narratives for specific target groups (Sedova et al., 2021). Chatbots, in particular, pose a significant threat due to their potential for disseminating persuasive misinformation. Utilizing AI for this purpose presents lower risks compared to other methods while simultaneously ensuring the source (the AI) remains impervious to persuasion or manipulation (Lombardo, 2021).

Despite these documented risks, the development and deployment of AI across the private, public, and third sectors continue to accelerate at an unprecedented pace (Jonk & Iren, 2021). Generative AI projects attract substantial investment and resource allocation from government and private entities. However, this rapid development necessitates careful guidance informed by ethical principles. Addressing the abovementioned risks and mitigating potential worst-case scenarios of this technology requires careful consideration (Henman, 2020). Algorithmic governance offers a promising framework in this context. This rapid expansion of AI adoption and the increasing sophistication of generative AI models create a sense of urgency

for establishing robust governance frameworks. Without such frameworks, the potential for unintended consequences, from algorithmic bias to the spread of misinformation, becomes significantly amplified.

The substantial investment in generative AI underscores the perceived potential of this technology to transform various sectors. Governments are exploring its use in public service delivery and policy analysis, while private companies see opportunities for increased efficiency and innovation. However, this investment should not be solely driven by technological possibilities but also by a thorough understanding of the ethical and societal implications.

A balanced approach is needed, one that fosters innovation while simultaneously safeguarding against potential harms. This requires proactive measures, including risk assessments, ethical guidelines, and ongoing monitoring of deployed AI systems. Furthermore, it necessitates a multi-stakeholder approach involving technologists, policymakers, ethicists, legal experts, and public members to ensure that AI development and deployment align with societal values and public interest. The development of clear ethical guidelines and standards for AI development is crucial to ensure responsible innovation and prevent the misuse of this powerful technology.

3 Algorithmic Governance for Governments and the Public Sector

As we conceptualize it, algorithmic governance entails effectively implementing practices and rules governing algorithm design, construction, use, and deployment within AI systems (Esposito & Tse, 2024). We propose that algorithmic governance focuses on establishing measures to ensure two fundamental conditions: (a) that algorithms function as intended and (b) that they are protected from errors and inherent social, economic, and political risks, including legal noncompliance and the perpetuation of bias and discrimination. Any algorithmic governance program must adhere to relevant local and international laws and regulations concerning AI. While ethical AI guidelines and charters developed by the private sector can serve as a starting point, they often require further development to adequately address AI's political and economic implications (Attard-Frost et al., 2023). These guidelines, while valuable, usually lack the teeth of legal enforceability and may not fully address the specific needs and contexts of the public sector.

Effective algorithmic governance necessitates the integration of diverse perspectives from academia and industry, including fields such as anthropology, sociology, political theory, and data science, to ensure a comprehensive and robust approach (Larsson, 2020). Government and public sector organizations must carefully consider their duties, obligations, and rights concerning the generative AI they intend to design or implement and end-users' needs. Furthermore, other stakeholders, such as different public sector agencies, other governmental bodies involved in the same project, and any private partners contributing to the technology's development or

deployment, must be considered (Wang & Cui, 2022). This multifaceted approach acknowledges that AI systems operate within complex social and political contexts, and their development and deployment must be sensitive to these contexts.

Data collection and analysis are essential components of algorithmic governance and should be integrated into every phase of an AI project (Ernst & Young, 2023). Quantitative and qualitative evidence are crucial for informing AI systems' design, deployment, and refinement to maximize benefits and minimize risks. Data and evidence are necessary to understand the real-world problem the AI aims to address; to guide effective technical development (including representativeness and sampling tests, algorithmic training, computer programming, and algorithm validation); to assess the effectiveness of AI deployment and citizen usage; and to facilitate iterative improvements through agile methodologies. This data-driven approach ensures that AI systems are developed and deployed based on evidence rather than assumptions and that their performance is continuously monitored and improved.

Government and public sector agencies should adopt a combined top-down and bottom-up approach. A top-down approach involves establishing a steering committee to provide project oversight, with project managers implementing specific practices down the organizational hierarchy to achieve project goals (Pastor-Escuredo & Treleaven, 2021). Conversely, a bottom-up approach empowers all staff members and end-users to provide feedback and actively participate in the project from development to deployment (Bruce et al., 2024). A common pitfall in algorithmic governance projects is over-reliance on one approach at the expense of the other. A balanced approach is essential to ensure that the project minimizes risks and maximizes benefits for all stakeholders. The top-down approach provides structure and direction, while the bottom-up approach fosters inclusivity and ensures that the AI system meets the needs of its users.

While a hierarchical approach is essential to AI governance, transparency is even more critical. Regardless of whether a project involves explicitly generative AI, all AI initiatives must incorporate checks and mechanisms to ensure transparency throughout their lifecycle (Coglianese & Lehr, 2019). The opacity surrounding algorithmic operations is a well-documented issue (Katzenbach & Ulbricht, 2019). Governments and the public sector can demonstrate best practices in embedding transparency within algorithmic projects while fostering inter-stakeholder collaboration through transparent knowledge sharing (Campion et al., 2022). Transparency should encompass the datasets used for training, the source code, and the criteria employed in algorithmic training. Making these elements open source and publicly accessible can bolster citizen trust in government generative-AI projects, provided that doing so does not compromise national or international security (Coglianese & Lehr, 2019). This level of transparency allows for public scrutiny and helps to build trust in the government's use of AI.

4 Practical Steps Towards Algorithmic Governance

Initially, the government or public sector agency should establish an ethics committee, either internal or external, to oversee the algorithmic governance of the project. A dedicated committee facilitates a more streamlined process, allowing project team members to focus on their respective specialized tasks (Janssen et al., 2020). This committee should develop a structured plan for integrating algorithmic governance into each stage of the generative AI project. All phases must be considered, from initial planning and design/technical specification through testing/beta to final deployment.

A risk assessment, encompassing ethical, reputational, technical, and scientific risks, should be conducted and tracked throughout the project lifecycle. Following establishing the project's fundamental structure, the committee should develop a charter or framework outlining the project's guiding principles and highlighting any potential areas of concern or contention requiring ongoing attention. This charter should be a living document, subject to review and revision as the project evolves and new risks or ethical considerations emerge. It should also clearly define roles and responsibilities related to AI governance.

One prominent example of a successful ethics steering committee in the public sector is the UK Centre for Data Ethics and Innovation (CDEI), which was established to guide the UK government on issues related to the responsible deployment of AI. This Centre played a crucial role in shaping policies related to bias mitigation, fairness, and algorithmic transparency. The CDEI operates as an independent entity, which allows it to engage with projects across government departments and ensures that its recommendations do not serve changing political agendas but rather the common public interest. In 2020, their landmark report on algorithmic bias in policing and recruitment spurred the widespread implementation of bias-detection measures in public sector agencies planning to deploy AI models (Centre for Data Ethics and Innovation [CDEI], 2020).

The CDEI is a prime example of how an independent ethics committee or body can ensure sector-wide ethical adherence, spreading beyond advisory functions to more advanced bias auditing, impact assessment, and general oversight of algorithmic decision-making. The CDEI's success highlights the importance of independence and a broad mandate for such bodies. They should not be limited to providing advice but should have the authority to conduct audits, assess impacts, and enforce ethical guidelines. Furthermore, their work should be transparent and publicly accessible to ensure accountability. Other governments are looking to models like the CDEI as they consider how to establish similar bodies to guide their own AI governance efforts.

4.1 Planning Phase

In the initial planning phase, the government or public sector agency must clearly define the problem that the AI system will address. This problem should be thoroughly investigated from multiple perspectives, including social, economic, political, security, and international relations, and supported by robust data and evidence. Critically, the agency must assess whether generative AI is the most appropriate solution or if a less expensive, safer, or more accessible alternative could achieve comparable results (Bruce et al., 2024). Developing technology for its own sake can misallocate public resources and create ineffective or irrelevant programs. This assessment should be documented and publicly available to ensure transparency and accountability.

This initial assessment should be overseen by an independent ethics body, similar to the CDEI, which plays a key role in guiding government AI deployment and shaping policies on algorithmic transparency, fairness, and bias mitigation (CDEI, 2020). These bodies, operating independently to ensure objectivity and public interest focus, should not only advise but also conduct impact assessments, bias auditing, and oversee algorithmic decision-making in public services, much like the CDEI's work on algorithmic bias in recruitment and policing, which prompted public sector agencies to implement bias-detection measures (CDEI, 2020). Establishing such independent bodies is crucial for responsible AI governance, ensuring ethical considerations are at the forefront of AI development and deployment (Zuiderveen Borgesius et al., 2016). These bodies should be able to access all relevant data and information necessary to conduct their work effectively.

Suppose generative AI is deemed the optimal solution. In that case, the steering or ethics committee should identify a comparable use case, preferably from the private or third sector, and analyze its governance framework, paying particular attention to how issues such as transparency, ethics, and risk mitigation were addressed (Ernst & Young, 2023). This exemplary case can serve as a best-practice model for implementing effective governance. Conversely, examining the use case can highlight areas where governance was deficient or lacking, providing valuable insights for improving the government or public sector project and achieving higher standards. This comparative analysis should be documented and used to inform the development of the government's AI project.

A preliminary risk assessment is crucial to ensure any ethical concerns are thoroughly investigated and accounted for before complex AI applications are developed. Canada has been leading the way in this area with its Directive on Automated Decision-Making (DADM), introduced in 2019, to guide public sector agencies in analyzing the risks and impact of AI projects before they are deployed. The DADM mandates that any public sector agency wishing to deploy an AI-based project must first conduct an Algorithmic Impact Assessment, which categorizes the project according to risk level (Treasury Board of Canada Secretariat, 2023). Stricter oversight is required on higher-risk projects, such as law enforcement, social benefits, or immigration. AI is compared against human-led approaches; alternative solutions

are prioritized if the latter outperforms the former. This approach aligns with the growing recognition of the importance of algorithmic impact assessments in ensuring responsible AI development (Hagendorff, 2020). The results of these assessments should be made public to foster transparency and allow for public input.

4.2 Business and Technical Specification Phase and Testing/ Beta Phase

During the business and technical specification and testing/beta phases, government and public sector agencies must carefully evaluate the AI system's risks, benefits, and overall impact (Bruce et al., 2024). This comprehensive evaluation extends beyond financial costs to encompass environmental impact, stakeholder literacy, and algorithmic explainability. For instance, while public technology projects are often assessed regarding financial costs, their ecological impact is frequently underanalyzed. This includes considering the extraction and import of raw materials for hardware and the computing power and energy consumption associated with software, particularly during algorithm training and operation (Esposito et al., 2023). For example, the carbon footprint of training large language models can be substantial, requiring careful consideration of energy sources and optimization strategies (Strubell et al., 2020).

Agencies should also assess and enhance the technical literacy of their stakeholders. Public sector employees who will be involved in the project or utilize the final product may require reskilling and upskilling initiatives (Diaz & Halkias, 2021; Diaz et al., 2022; Halkias et al., 2020), while citizen users may benefit from public awareness campaigns in media and educational venues (Bruce et al., 2024). Consider the rollout of a new AI-powered benefits system; staff would need training on how to use the system effectively, while citizens would need to understand how their data is used and how to appeal decisions. This also includes accessibility considerations for users with disabilities.

Explainability calculations are crucial during this stage for mitigating the risk of bugs, errors, and bias in the developing algorithm, thereby enhancing transparency and managing the variability of its responses (Rabiul Islam et al., 2020). Explainability calculations employ statistical methods to analyze and control the operational logic of an algorithm. These methods should be systematically applied before, during, and after algorithmic training, beginning with the training datasets (Alikhademi et al., 2021). This tool can help ensure that algorithms do not exhibit biased responses based on factors such as race and gender, a problem frequently observed in facial recognition algorithms (Alikhademi et al., 2021). For example, the COMPAS system used in the U.S. criminal justice system to predict recidivism has been shown to exhibit racial bias, even when race is not explicitly included as a variable in the algorithm (Angwin et al., 2016).

Explainability calculations could have been used to identify and mitigate this bias during the development and testing phases. Similarly, Amazon abandoned its AI recruiting tool after discovering it discriminated against women. Again, rigorous testing and explainability analysis could have revealed and addressed this issue before deployment. These examples highlight the critical importance of thorough evaluation and explainability throughout the AI development lifecycle. The development of robust explainability methods remains an active area of research (Doshi-Velez & Kim, 2017). Furthermore, explainability should be tailored to the audience. Technical users may require detailed information about the algorithm's inner workings, while nontechnical users may need more straightforward explanations of how the AI system arrives at its decisions.

4.3 Deployment Phase

Finally, during the deployment phase, the AI system and public perception of it must be carefully monitored and analyzed, with necessary adjustments made to address any technical issues, usability problems, or ethical concerns that arise (Bruce et al., 2024). An agile software development approach is particularly valuable in this context. Agile methodologies emphasize continuous feedback from end-users to ensure that the software meets their needs and to identify any problems with the final product (Alsaadi & Saeedi, 2022). This approach also involves iterative refinement, starting with a minimal viable product and adding features incrementally based on user feedback. Public generative AI projects can be piloted or released in beta versions, allowing the government or public sector agency to assess the AI's effectiveness and make necessary adjustments before the final release (Bruce et al., 2024). Post-deployment features can be added, and corrections can be implemented through updated versions. This iterative process allows for continuous improvement and adaptation to real-world conditions. This also enables rapid responses to unexpected issues or changes in user needs.

For example, the initial rollout of the UK government's online system for visa applications encountered significant usability problems, causing frustration and delays for applicants. An agile approach, incorporating user feedback and iterative improvements, would have been beneficial in addressing these issues more effectively (Garth Coates Solicitors, n.d.). Similarly, implementing predictive policing algorithms in several U.S. cities has raised concerns about bias and discriminatory targeting. Continuous monitoring and analysis of the algorithm's performance and community feedback are crucial to identifying and mitigating these ethical concerns (Lum & Isaac, 2016). A static, "launch and leave" approach to deployment is not appropriate for complex AI systems that can have a significant social impact. These examples highlight the need for ongoing evaluation and adaptation, particularly in high-stakes applications of AI.

Furthermore, post-deployment monitoring should include technical performance metrics and public perception and acceptance of the AI system. Public trust is

essential for adopting AI in government services (Organisation for Economic Co-operation and Development [OECD], 2019). Surveys, focus groups, and social media analysis can provide valuable insights into how the public perceives the AI system and identify any concerns or misunderstandings. This information can then inform communication strategies and make necessary adjustments to the system itself.

For instance, if a facial recognition system is perceived as invasive and a threat to privacy, the government might need to revise its data collection and usage policies or implement stronger transparency measures to address public concerns. The ongoing evaluation and adaptation of AI systems in the public sector are essential for ensuring their effectiveness, fairness, and societal acceptance. This includes technical adjustments and addressing AI deployment's social and ethical implications. Responsible AI development requires ongoing dialogue with the public to understand their concerns and build trust (Bateman, 2020). This dialogue should be proactive and inclusive, seeking diverse community and stakeholder input.

5 Collaboration and Partnerships

Algorithmic governance initiatives within government and the public sector offer significant opportunities for collaboration and partnerships with the private and third sectors (Janssen et al., 2020). Such alliances can enhance reputation and alleviate administrative burdens and workloads for agencies and departments facing staffing or resource constraints [30]. For example, data collection and analysis can be outsourced to independent research institutions, private technology companies can provide initial software for development, and specialized consulting firms can offer tailored guidance. Inter-sector partnerships can also promote innovation and knowledge sharing (Carbonara & Pellegrino, 2020). This collaborative approach allows public sector agencies to leverage the expertise and resources of external partners, accelerating development and improving the quality of AI systems.

One example of successful public-private partnerships in AI governance is the collaboration between the City of Amsterdam and private tech companies to develop algorithms for social good. The city partnered with organizations like the Amsterdam Data Science Initiative to develop AI tools for addressing traffic congestion and social inequality. This collaboration allowed the city to access cutting-edge technology and expertise while ensuring that the development and deployment of these AI systems aligned with ethical principles and public values. This type of collaboration is increasingly recognized as crucial for effective AI governance (Hagendorff, 2020). Another example is government agencies' use of private-sector cloud computing platforms. By leveraging existing infrastructure and expertise, governments can reduce costs and accelerate the deployment of AI-powered services. However, such partnerships also require careful consideration of data security, privacy, and vendor lock-in. As these partnerships become more common, best practices for managing risks and ensuring accountability are being developed (OECD, 2019).

Several international AI governance initiatives provide research guidelines and facilitate inter-sector collaboration to achieve the highest governance standards. A prominent example is the World Economic Forum's AI Governance Alliance. This initiative convenes government, third-sector organizations, industry, and academia experts to research, advocate for, and develop guidelines for designing and implementing responsible generative AI systems (World Economic Forum, n.d.). The alliance focuses on three workstreams: responsible applications and transformation, resilient governance and regulation, and safe systems and technologies (World Economic Forum, 2024). Its open-access resources include a briefing paper series and numerous reports, articles, and videos. The Organisation for Economic Co-operation and Development's (OECD's) Principles on Artificial Intelligence (OECD, 2019) are another essential example, providing a set of guidelines for the responsible development and use of AI. These principles emphasize human-centered values, fairness, transparency, and accountability.

Beyond these high-level initiatives, sector-specific collaborations are also emerging. For instance, partnerships between hospitals, research institutions, and technology companies in the healthcare sector drive innovation in AI-powered diagnostics and treatment. These collaborations often involve sharing data and expertise to develop and validate AI algorithms. However, ensuring that these partnerships adhere to strict ethical guidelines and protect patient privacy is crucial. Data sharing in healthcare requires robust privacy safeguards and ethical oversight (Meskó et al., 2018). Similarly, in the financial sector, collaborations between banks and fintech companies explore using AI for fraud detection and risk assessment. These partnerships can lead to more efficient and effective financial services, but they also raise concerns about algorithmic bias and the potential for discriminatory outcomes. Therefore, robust governance frameworks are essential to guide these collaborations and ensure they serve the public interest (OECD, 2019).

6 The Role of Polycentric Governance

Polycentric governance offers a valuable framework for navigating the complexities of governing generative AI-based and algorithm-based projects in government. As noted by the Digital Commons Governance Project at Harvard University (n.d.), polycentric governance recognizes that multiple centers of authority, both public and private, operate at different levels and interact to govern a particular domain. In the context of AI, this means acknowledging that governance is not solely the responsibility of government agencies but also involves private sector actors, civil society organizations, technical communities, and individual citizens. This distributed approach is particularly relevant for AI due to its rapid evolution, its diverse applications across sectors, and the inherent uncertainties and ethical dilemmas it poses.

This approach draws heavily from the work of Elinor Ostrom (1990), a Nobel laureate who studied how communities manage common-pool resources, like

forests or fisheries. Ostrom demonstrated that centralized, top-down control is not always the most effective way to govern shared resources. Instead, she found that complex systems can be successfully managed through polycentric governance, where multiple, overlapping jurisdictions and organizations interact to create rules and enforce them.

Ostrom's work on the commons provides a valuable lens for understanding AI governance. Just like a physical commons, AI can be seen as a shared resource that requires careful management to prevent overuse, degradation, and inequitable distribution of benefits. The "digital commons" (Hess & Ostrom, 2006) further extends this idea to the digital realm, encompassing data, algorithms, software, and other digital resources that are collectively owned or managed. AI, particularly generative AI, often relies on large datasets, many drawn from the digital commons. Therefore, effectively governing AI requires considering its impact on the digital commons and ensuring its development and use are consistent with open access, shared benefit, and democratic participation. A polycentric approach to AI governance recognizes the interconnectedness of different actors and interests involved in the digital commons. It seeks to create compelling and legitimate governance mechanisms that foster collaboration and communication among diverse stakeholders, promote transparency and accountability, and ensure that AI systems are developed and used to benefit society.

While the concept of the digital commons is still evolving in the context of AI governance, several real-world examples illustrate its potential. One example is the development of open-source datasets and AI models. Communities of researchers and developers contribute to shared repositories, making these resources available for others to use and build upon. This fosters innovation and democratizes access to AI technology (Richardson, 2025). For instance, large language models like those used in generative AI are often trained on massive datasets scraped from the web, raising questions about ownership and usage rights (Chai, n.d.). However, some efforts are underway to create curated, ethically sourced datasets that could be considered part of the digital commons.

Another example is the development of open standards and protocols for AI. These shared technical specifications facilitate interoperability and prevent vendor lock-in, promoting a more open and competitive AI ecosystem. Organizations like the Partnership on AI are working to develop such standards, fostering collaboration among different stakeholders (Moltzau, 2020). Furthermore, initiatives promoting data transparency and algorithmic explainability can be seen as contributing to the digital commons. By making data and algorithms more accessible and understandable, these initiatives empower citizens and civil society organizations to scrutinize AI systems and hold them accountable. This increased transparency strengthens democratic participation in AI governance and helps ensure that AI systems are developed and used in the public interest (Dittmar, n.d.). These examples, while still nascent, point towards a future where the principles of the digital commons play a more prominent role in shaping the governance of AI.

A polycentric approach to AI governance in government can involve several key elements. First, it necessitates the establishment of clear roles and responsibilities

for different actors involved in the AI lifecycle, from data collection and algorithm development to deployment and monitoring. This can include government agencies setting standards and regulations, private sector companies developing and implementing AI solutions, and civil society organizations providing oversight and advocacy. Second, polycentric governance emphasizes collaboration and coordination among these diverse actors (Yadav et al., 2024) and is achieved through multi-stakeholder platforms, public-private partnerships, and open forums for dialogue and knowledge sharing. As Ginsburg (2024) argues in Forbes, effective AI governance requires a "telegraph revolution," a rapid and widespread communication and coordination among stakeholders.

Third, polycentric governance promotes experimentation and learning (Yadav et al., 2024). Different approaches to AI governance can be tested and evaluated, allowing for adaptation and improvement over time. This is particularly important in the rapidly evolving field of AI, where best practices are still emerging. Fourth, polycentric governance recognizes the importance of local context. AI governance frameworks should be flexible enough to accommodate different communities and sectors' specific needs and values. As the Laguna Health roundtable discussion (Laguna Health, 2024) highlights, the healthcare sector, for example, requires specific AI governance considerations related to patient privacy, data security, and clinical safety. Finally, polycentric governance emphasizes accountability (Yadav et al., 2024). Mechanisms should be in place to ensure that all actors involved in AI development and deployment are held responsible for their actions, including independent audits, public reporting requirements, and legal recourse for individuals harmed by AI systems. By embracing a polycentric approach, governments can create more robust, adaptable, and legitimate AI governance frameworks that effectively address the challenges and opportunities presented by this transformative technology.

Several examples illustrate the practical application of polycentric governance in AI. One is **the Partnership on AI, a** multi-stakeholder organization that brings together a diverse range of actors, including technology companies (e.g., Google, Microsoft, Meta), civil society groups (e.g., ACLU, Access Now), and academic institutions, to address the ethical and societal implications of AI. The Partnership on AI serves as a platform for collaboration, knowledge sharing, and the development of best practices for responsible AI development and deployment (Partnership on AI, n.d.). It exemplifies the polycentric approach by involving actors from different sectors and levels of authority in shaping the future of AI governance (Thiel, 2023). While not a government-led initiative, its influence on government policy and industry practices is significant. It provides a forum for diverse perspectives to be heard and contributes to developing shared norms and standards.

Another successful example is **the Montreal Declaration for Responsible AI.** This declaration, developed through a broad public consultation process involving academics, civil society organizations, and citizens, articulates ethical principles for AI development and use (Montreal Declaration Responsible AI, 2025), emphasizing human dignity, fairness, transparency, and accountability. The Montreal Declaration is not a legally binding document, but it serves as a moral compass for

AI development and has influenced policy discussions at both national and international levels. This example demonstrates the polycentric nature of AI governance by involving a wide range of stakeholders in developing ethical guidelines. It also highlights the importance of public engagement in shaping the future of AI. The declaration's influence on policy underscores the power of collaborative, multi-stakeholder initiatives to shape the AI landscape and is a testament that governance can emerge from various centers of authority, not just government (Zhang, 2024).

7 Conclusion

Generative AI offers transformative potential for government and public sector operations, but like any emerging technology, it presents inherent challenges. Generative AI carries significant social, ethical, economic, and legal implications that must be carefully considered during development and deployment in any sector (Hacker et al., 2023). While it can enhance efficiency and drive innovation within government and the public sector, it also demands stringent governance and ethical oversight (Campion et al., 2022). This disruptive technology can automate administrative processes, generate predictive models, and perform other valuable functions. Still, it also carries the potential to sow societal discord, perpetuate bias and discrimination, and infringe upon citizens' privacy (Skaug Sætra, 2020). For example, generative AI for creating public service announcements could be misused to spread misinformation or propaganda, undermining public trust in government communications. Similarly, AI-powered predictive policing tools could exacerbate existing biases in the criminal justice system, disproportionately targeting marginalized communities and eroding trust in law enforcement. The potential for misuse also extends to other areas, such as welfare distribution, where biased algorithms could deny individuals access to essential services.

The principles of algorithmic governance provide a valuable framework for governments and the public sector to ensure that AI projects adhere to norms of transparency and accountability. A practical, phased approach to algorithmic governance should emphasize evidence-based decision-making, robust risk management, multi-stakeholder engagement, and inter-sector collaboration. This comprehensive strategy can ensure that public generative AI projects enhance, rather than erode, citizen trust in public institutions while maximizing societal benefits and minimizing harm. Transparency is crucial; citizens should understand how AI is being used in government and have avenues for redress if they believe an AI-driven decision has unfairly impacted them. Understanding the data used to train the AI, the logic behind its decision-making processes, and the potential for bias in the system. Without this transparency, citizens cannot hold governments accountable for using AI.

Accountability mechanisms are also essential, ensuring that individuals and organizations are responsible for the outcomes of AI systems, which require clear lines of responsibility and established procedures for investigating and addressing harms caused by AI. It also necessitates ongoing monitoring and evaluation of AI

systems to identify and mitigate potential risks. Furthermore, ethical guidelines and regulations are needed to govern the development and deployment of AI in the public sector, ensuring that it aligns with democratic values and protects fundamental rights. The development of these guidelines should involve broad public consultation to ensure that they reflect societal values and concerns. Finally, public education and awareness campaigns are crucial to inform citizens about the use of AI in government and empower them to participate in the governance process. Building public trust in a government's use of AI requires a proactive and transparent approach that prioritizes ethical considerations and ensures accountability.

8 Recommendations for Future Research

However, further research is needed to explore algorithmic governance in real-world government and public sector settings. Empirical studies can offer valuable insights into the practical application of algorithmic governance principles, particularly in specific domains such as citizen services, resource allocation, and international relations (Jonk & Iren, 2021). Future research could also investigate how national and geographic contexts, funding and resource allocation, and citizen attitudes and reception influence the governance of public generative AI projects (Kuziemski & Misuraca, 2020). For instance, comparative studies could examine how different countries approach the regulation of AI in healthcare or education, considering variations in cultural values, legal frameworks, and political systems. Such research could identify best practices and inform the development of tailored governance models. This includes exploring the role of different regulatory approaches, from self-regulation to binding legislation.

Research could also explore the impact of public awareness campaigns on citizen acceptance of AI-driven government services. Understanding how different communication strategies affect public perception and trust in AI is crucial for fostering responsible AI adoption. This could involve investigating the effectiveness of various messaging approaches, the role of media coverage, and the influence of community engagement initiatives. Understanding how different demographics perceive and interact with AI-driven services is also essential. Furthermore, research should investigate the long-term societal impacts of generative AI, including its effects on labor markets, democratic processes, and social cohesion. For example, studies could examine how generative AI might automate jobs currently performed by public sector employees and what retraining or upskilling programs might be necessary. Research should also investigate the potential for generative AI to be used for malicious purposes, such as creating deepfakes or spreading disinformation, and how governments can mitigate these risks. This includes exploring the development of detection technologies and legal frameworks for addressing harmful uses of AI.

Understanding these complex issues is essential for developing effective and responsible AI governance frameworks. Technical considerations, such as

algorithm design and data privacy, and social and ethical dimensions, such as fairness, accountability, and transparency, are foundational to responsible AI governance. Interdisciplinary research, bringing together computer science, law, ethics, and social sciences experts, is crucial for addressing these multifaceted challenges. Finally, research should explore the development of specific tools and methodologies for auditing AI systems and ensuring their compliance with ethical guidelines and legal regulations. This could include developing standardized metrics for evaluating algorithmic fairness, creating frameworks for assessing AI systems' societal impact, and designing mechanisms for public participation in AI governance. Further research should also explore the role of international cooperation in AI governance, given the transboundary nature of many AI-related challenges. This includes examining the potential for international agreements and developing shared standards and best practices.

Acknowledgments The authors take full responsibility for the contents of this paper. The authors acknowledge using Grammarly (grammarly.com) ONLY to proofread parts of this paper for correct use of the English language and academic writing style.

References

Alikhademi, K., Richardson, B., Drobina, E., & Gilbert, J. E. (2021). Can explainable AI explain unfairness? A framework for evaluating explainable AI. *ArXiv*. https://doi.org/10.48550/arXiv.2106.07483

Alsaadi, B., & Saeedi, K. (2022). Data-driven effort estimation techniques of agile user stories: A systematic literature review. *Artificial Intelligence Review, 55*(7), 5485–5516. https://doi.org/10.1007/s10462-021-10132-x

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias. *ProPublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Attard-Frost, B., De los Ríos, A., & Walters, D. R. (2023). The ethics of AI business practices: A review of 47 AI ethics guidelines. *AI Ethics, 3*(2), 389–406. https://doi.org/10.1007/s43681-022-00156-6

Bateman, W. (2020). Algorithmic decision-making and legality: Public law dimensions. *Australian Law Journal, 94*(7), 520–530. https://ssrn.com/abstract=3496386

Bird, C., Ungless, E., & Kasirzadeh, A. (2023). Typology of risks of generative text-to-image models. In F. Rossi, S. Das, J. Davis, K. Firth-Butterfield, & A. John (Eds.), *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 396–410). ACM Digital Library. https://doi.org/10.1145/3600211.3604722

Boulianne, S., Koc-Michalska, K., & Bimber, B. (2020). Right-wing populism, social media and echo chambers in western democracies. *New Media & Society, 22*(4), 683–699. https://doi.org/10.1177/1461444819893983

Bruce, D., Fadia, A., Isherwood, T., Marcati, C., Mitchell, A., Münstermann, B., Shenai, G., Vuppala, H., & Weber, T. (2024). *How generative AI can help governments maximize their economic potential*. McKinsey & Company. https://www.mckinsey.com/industries/public-sector/our-insights/unlocking-thepotential-of-generative-ai-three-key-questions-for-government-agencies

Campion, A., Gasco-Hernandez, M., Jankin Mikhaylov, S., & Esteve, M. (2022). Overcoming the challenges of collaboratively adopting artificial intelligence in the public sector. *Social Science Computer Review, 40*(2), 462–477. https://doi.org/10.1177/0894439320979953

Carbonara, N., & Pellegrino, R. (2020). The role of public-private partnerships in fostering innovation. *Construction Management and Economics, 38*(2), 140–156. https://doi.org/10.1080/01446193.2019.1610184

Centre for Data Ethics and Innovation. (2020). *Review into bias in algorithmic decision making.* https://assets.publishing.service.gov.uk/media/60142096d3bf7f70ba377b20/Review_into_bias_in_algorithmic_decision-making.pdf

Chai, J. (n.d.). *What is generative AI? What are large language models (LLMs)?* [Video]. University of Michigan Online. https://online.umich.edu/collections/artificial-intelligence/short/what-is-generative-ai-what-are-llm/

Coglianese, C., & Lehr, D. (2019). Transparency and algorithmic governance. *Administrative Law Review, 71*(1), 1–56. https://www.jstor.org/stable/27170531

Copeland, B. J. (2024). Artificial intelligence (AI). In *Encyclopedia britannica*. https://www.britannica.com/technology/artificial-intelligence

Díaz, A. (2021). *Data-driven policing's threat to our constitutional rights*. Brookings. https://www.brookings.edu/articles/datadriven-policings-threat-to-our-constitutional-rights/

Diaz, J., & Halkias, D. (2021). Reskilling and upskilling 4IR leaders in business schools through an innovative executive education ecosystem: An integrative literature review. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3897059

Diaz, J., Halkias, D., & Thurman, P. W. (2022). *The innovative management education ecosystem: Reskilling and upskilling the future workforce*. Routledge.

Digital Commons Governance Project. (n.d.). *About*. Retrieved February 25, 2025, from https://www.iq.harvard.edu/digital-commons-governance-project

Dittmar, L. (n.d.). *What does transparency really mean in the context of AI governance?* OECG. https://www.oceg.org/what-does-transparency-really-mean-in-the-context-of-ai-governance/

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *ArXiv*. https://doi.org/10.48550/arXiv.1702.08608

Eastwood, B. (2024). How should AI-generated content be labeled? *MIT Sloan Management Review*. https://mitsloan.mit.edu/ideas-made-to-matter/how-should-ai-generated-content-be-labeled

Engstrom, D. F., & Ho, D. E. (2020). Algorithmic accountability in the administrative state. *Yale Journal on Regulation, 37*(3), 800–854. https://heinonline.org/HOL/P?h=hein.journals/yjor37&i=812

Entsminger, J., Esposito, M., Tse, T., & Jean, A. (2023, June 26). The dark side of generative AI: Automating inequality by design. *California Management Review.* https://cmr.berkeley.edu/2023/06/the-dark-side-of-generative-ai-automating-inequality-by-design/

Ernst & Young. (2023). *The AIdea of India: Generative AI's potential to accelerate India's digital transformation*. https://www.ey.com/content/dam/ey-unified-site/ey-com/en-in/newsroom/2023/12/documents/ey-the-aidea-of-india-generative-ai-s-potential-to-accelerate-india-s-digital-transformation-v4.pdf

Esposito, M., & Tse, T. (2018). DRIVE: The five megatrends that underpin the future business, social, and economic landscapes. *Thunderbird International Business Review, 60*(1), 121–129. https://doi.org/10.1002/tie.21889

Esposito, M., & Tse, T. (2024). Mitigating the risks of generative AI in government through algorithmic governance. In H.-C. Liao, D. Duenas Cid, M. A. Macadar, & F. Bernardini (Eds.), *Proceedings of the 25th Annual International Conference on Digital Government Research* (pp. 605–609). ACM Digital Library. https://doi.org/10.1145/3657054.36571

Esposito, M., Halkias, D., Tse, T., & Harkiolakis, T. (2023). Environmental and climate impacts of the metaverse. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4616695

Fergnani, A., Hines, A., Lanteri, A., & Esposito, M. (2020). Corporate foresight in an ever-turbulent era. *European Business Review, 25*, 26–33. https://www.europeanbusinessreview.com/corporate-foresight-in-an-ever-turbulent-era/

Garth Coates Solicitors. (n.d.). *The challenges of the UK's immigration system: Digitalisation delays and legislative complexity*. https://garthcoates.com/news/the-challenges-of-the-uks-immigration-system-digitalisation-delays-and-legislative-complexity/

Ginsburg, R. (2024, December 18). *AI governance and the telegraph revolution*. Forbes. https://www.forbes.com/sites/robertginsburg/2024/12/18/ai-governance-and-the-telegraph-revolution/

Groth, O. J., Esposito, M., & Tse, T. (2015). What Europe needs is an innovation-driven entrepreneurship ecosystem: Introducing EDIE. *Thunderbird International Business Review, 57*(4), 263–269. https://doi.org/10.1002/tie.21709

Hacker, P., Engel, A., & Mauer, M. (2023, June). Regulating ChatGPT and other large generative AI models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1112–1123). ACM Digital Library. https://doi.org/10.1145/3593013.3594067

Hagendorff, T. (2020). The ethics of algorithms: An evaluation of guidelines. *Minds and Machines, 30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Halkias, D., Neubert, M., Thurman, P. W., Adendorff, C., & Abadir, S. (2020). *The innovative business school*. Routledge.

Harkiolakis, T., & Esposito, M. (2023). Populism, democratic participation, and media polarization in the digital age: The case of vaccine hesitancy in Greece. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4599236

Henman, P. (2020). Improving public services using artificial intelligence: Possibilities, pitfalls, governance. *Asia Pacific Journal of Public Administration, 42*(4), 209–221. https://doi.org/10.1080/23276665.2020.1816188

Hess, C., & Ostrom, E. (2006). *Understanding knowledge as a commons: From theory to practice*. MIT Press.

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy artificial intelligence. *Government Information Quarterly, 37*(3), 101493. https://doi.org/10.1016/j.giq.2020.101493

Jonk, E., & Iren, D. (2021). Governance and communication of algorithmic decision making: A case study on public sector. In *2021 IEEE 23rd Conference on Business Informatics (CBI)* (pp. 151–160). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/CBI52690.2021.00026

Katzenbach, C., & Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review, 8*(4), 1–18. https://doi.org/10.14763/2019.4.1424

Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy, 44*(6), 101976. https://doi.org/10.1016/j.telpol.2020.101976

Laguna Health. (2024, December 6). *Exploring the future of AI governance in healthcare: Insights from our executive roundtable with Mark Esposito*. https://www.lagunahealth.com/blog/exploring-the-future-of-ai-governance-in-healthcare-insights-from-our-executive-roundtable-with-mark-esposito

Larsson, S. (2020). On the governance of artificial intelligence through ethics guidelines. *Asian Journal of Law and Society, 7*(3), 437–451. https://doi.org/10.1017/als.2020.19

Lau, T. (2024, January 21). *Predictive policing explained*. Brennan Center for Justice. https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained?ref$=$axion.zone

Lombardo, S. (2021). The bad, the good, and the rebellious bots: World's first in artificial intelligence. In T. H. Musiolik & A. D. Cheok (Eds.), *Analyzing future applications of AI, sensors, and robotics in society* (pp. 221–237). IGI Global. https://doi.org/10.4018/978-1-7998-3499-1.ch013

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance Magazine, 13*(5), 14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Meskó, B., Hetényi, G., & Győrffy, Z. (2018). Will artificial intelligence solve the human resource crisis in healthcare? *BMC Health Services Research, 18*, 545. https://doi.org/10.1186/s12913-018-3359-4

Moltzau, A. (2020, June 17). *The global partnership for artificial intelligence*. Medium. https://alexmoltzau.medium.com/exploring-the-partnership-on-ai

Montemayor, C., Halpern, J., & Fairweather, A. (2022). In principle obstacles for empathic AI: Why we can't replace human empathy in healthcare. *AI & Society, 37*(4), 1353–1359. https://doi.org/10.1007/s00146-021-01230-z

Montreal Declaration Responsible AI. (2025). *The Montréal declaration for a responsible development of artificial intelligence*. Retrieved February 25, 2025, from https://montrealdeclaration-responsibleai.com/the-declaration/

Organisation for Economic Co-operation and Development. (2019). *OECD principles on artificial intelligence* (OECD/LEGAL/0449). OECD Legal Instruments. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Partnership on AI. (n.d.). *About us*. Retrieved February 25, 2025, from https://partnershiponai.org/about/

Pastor-Escuredo, D., & Treleaven, P. (2021). Multiscale governance. *ArXiv*. http://arxiv.org/abs/2104.02752

Preece, C., & Çelik, H. (2023). *AI is a powerful tool, but it's not a replacement for human creativity*. World Economic Forum. https://www.weforum.org/agenda/2023/06/ai-cannot-replace-humancreativity/

Rabiul Islam, S., Eberle, W., & Ghafoor, S. K. (2020). Towards quantification of explainability in explainable artificial intelligence methods. In *The Thirty-Third International Flairs Conference* (pp. 75–81). FLAIRS. https://cdn.aaai.org/ocs/18410/18410-79310-1-PB.pdf

Richardson, D. (2025, January 21). *Why open source is critical to the future of AI*. Red Hat. www.redhat.com/en/blog/why-open-source-critical-future-ai

Sedova, K., McNeill, C., Johnson, A., Joshi, A., & Wulkan, I. (2021). *AI and the future of disinformation campaigns: Part 2: A threat model*. Center for Security and Emerging Technology. https://cset.georgetown.edu/publication/ai-and-the-future-of-disinformation-campaigns-2/

Shu, C. (2014, January 26). *Google acquires artificial intelligence startup DeepMind for more than $500M*. TechCrunch. https://techcrunch.com/2014/01/26/google-deepmind/?guccounter$=$1&guce_referrer$=$aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig$=$AQAAAIfAI12Lw2IZzXRyWktGu7PvuHy8WZqVG7cBOgR-ITIq2RDmenqUFSnVjHfxzpLfyflxUvQ1Jic6-0C9mH3OQbCJDM8DPYFB2emTtppN8ypVwlrueCfm4MOK57EevZECZtkQ83SZf0NsN8VF7tdV8Bnfa7DJbWMXlxRqWwj8xAz

Skaug Sætra, H. (2020). A shallow defence of a technocracy of artificial intelligence: Examining the political harms of algorithmic governance in the domain of government. *Technology in Society, 62*, 101283. https://doi.org/10.1016/j.techsoc.2020.101283

Strubell, E., Ganesh, A., & McCallum, A. (2020). Energy and policy considerations for modern deep learning research. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(9), 13693–13696. https://doi.org/10.1609/aaai.v34i09.7123

Syeda, N. (2024). Algorithms and populism: Practices in microtargeting. In J. Chacko Chennattuserry, M. Deshpande, & P. Hong (Eds.), *Encyclopedia of new populism and responses in the 21st century* (pp. 1–7). https://doi.org/10.1007/978-981-16-9859-0_135-1

Thiel, A. (2023). Polycentric governing and polycentric governance. In F. Gadinger & J. Aart Scholte (Eds.), *Polycentrism: How governing works today* (pp. 98–120). Oxford University Press. https://doi.org/10.1093/oso/9780192866837.003.0005

Treasury Board of Canada Secretariat. (2023). *Directive on automated decision-making*. Retrieved February 25, 2025, from https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592

Wang, X., & Cui, X. (2022). PPP financing model in the infrastructure construction of the park integrating artificial intelligence technology. *Computational Intelligence and Neuroscience, 2022*(1), e6154885. https://doi.org/10.1155/2022/6154885

Wirtz, B. W., Weyerer, J. C., & Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration, 43*(9), 818–829. https://doi.org/10.1080/01900692.2020.1749851

World Economic Forum. (2024). *Our workstreams*. Retrieved February 25, 2025 from https://initiatives.weforum.org/ai-governance-alliance/home
World Economic Forum. (n.d.). *AI governance alliance*. Retrieved February 25, 2025 from https://initiatives.weforum.org/ai-governancealliance/home
Yadav, A., Anwer, N., Mahapatra, K., Shrivastava, M. K., & Khatiwada, D. (2024). Analyzing the role of polycentric governance in institutional innovations: Insights from urban climate governance in India. *Sustainability, 16*(23), 10736. https://doi.org/10.3390/su162310736
Yang, E., & Roberts, M. E. (2023). The authoritarian data problem. *Journal of Democracy, 34*(4), 141–150. https://muse.jhu.edu/pub/1/article/907695
Zhang, B. (2024, February 5). *Public participation is essential to decide the future of AI*. Tech Policy. https://www.techpolicy.press/public-participation-is-essential-to-decide-the-future-of-ai/
Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, B. (2016). Should we worry about filter bubbles? *Internet Policy Review, 9*(1). https://doi.org/10.14763/2016.1.401

Mark Esposito, PhD, is a professor of economics and public policy with appointments at Hult International Business School and Harvard University. He equally serves as an Adjunct Professor of Public Policy at Georgetown University's McDonough School of Business. He co-founded the Machine Learning research firm, Nexus FrontierTech and The Chart ThinkTank and The AI Native Foundation. He was ranked by Thinkers50 in 2016 as 1 of the 30 rising business thinkers in the world and was shortlisted for the Breakthrough Award in 2019 and the Strategy Award in 2023.

Daphne Halkias, PhD, is a Visiting Fellow at the Futures Impact Lab at Hult International Business School, a Distinguished Research Fellow at the Center for Policy and Competitiveness at École des Ponts ParisTech Business School, a Fellow at The Institute of Coaching, McLean Hospital at Harvard Medical School; and Research Affiliate at Institute for Social Sciences, Cornell University. Dr. Halkias is CEO of Executive Coaching Consultants and Editor in Chief of *International Journal of Environment, Workplace and Employment*, *International Journal of Teaching and Case Studies*, and *International Journal of Technology Enhanced Learning.* She has authored 14 academic books and over 100 peer-reviewed papers.

Terence Tse, PhD, is a Professor of Finance at Hult International Business School and a Visiting Professor at ESCP Business School and Cotrugli Business School. He is the Co-founder and Executive Director of the AI company Nexus FrontierTech. A highly sought speaker, he is co-author of *The AI Republic* (Lioncrest, 2019), *Understanding How the Future Unfolds* (Lioncrest, 2017), *The Great Remobilization: Strategies and Designs for a Smarter Global Future* (MIT Press, 2023). Additionally, he is the author of *Corporate Finance: The Basics* (Routledge, 2017), now in its second edition (2023).

Tatiana Harkiolakis, MSc., is a Research Fellow at the Center for Policy and Competitiveness, Ecole des Ponts ParisTech Business School, and a DBA candidate at ISC Paris Business School. She holds a Master's in Media and Communications from the London School of Economics and Political Science and is Director of Communication at Executive Coaching Consultants. She is a published journalist and Managing Editor of the *International Journal of Environment, Workplace and Employment*. Her research interests include the future of work, circular business models and shared value creation, disruptive innovation ecosystems, digital sustainable entrepreneurship, and regional entrepreneurship ecosystems.