RESPONSIBLE AI BUILDS TRUST IN GOVERNMENT

By Steve Mills, Elias Baltassis, Nadim Abillama, and Matthew Mendelsohn

WHEN GOVERNMENTS INCORPORATE artificial intelligence into decision making, the results can be good. Governments have, for instance, deployed AI-based systems effectively to manage hospital capacity, regulate traffic, and determine the best way to distribute social benefits and deliver services.

However, some government AI systems have unfairly or unwittingly targeted or disadvantaged specific portions of their constituencies—and worse.

A major difference between systems that work and those that fail is the way in which they are created and overseen. An approach that leads to successful outcomes includes proper governance, thoughtfully conceived processes based on input from affected stakeholders, and transparency about AI's role in decision making.

We call such a comprehensive approach Responsible AI. To achieve it, governments must empower leadership and use AI to enhance human decision making—not re-place it. A Responsible AI approach should include regular reviews, integration with standard tools and data models, and a plan for potential lapses.

Governments have another powerful tool at their disposal. They can use procurement to promote widespread adoption of Responsible AI. By making adherence to Responsible AI principles a prerequisite for bidding on public-sector AI contracts, agencies can ensure that the systems they create are ethical and transparent, while gradually integrating AI into public-sector decision making.

Unintended Consequences of Government AI Systems

In the recent past, several well-publicized lapses have illustrated the unintentional harm that can befall individuals or society when government AI systems aren't designed, built, or implemented in a responsible manner.

- After college entrance exams were canceled because of the pandemic, the

UK government used an algorithm that determined grades based on students' past performance. The system reduced the grades of nearly 40% of students and led to accusations that it was biased against test takers from challenging socioeconomic backgrounds.

- A Dutch court ordered the government to stop using a system based on an undisclosed algorithm—which was intended to predict whether people would commit benefits or tax fraud—after it was determined that the system breached human rights laws and the EU's General Data Protection Regulation. Citizens groups initiated the legal action after discovering that the system had been used to target neighborhoods with mostly low-income and minority residents.

- A local NGO sued Buenos Aires for violating child rights after the city adopted a facial-recognition system to find criminal suspects. The system relied on a database that contained information about children as young as four.

- A US border control agency deployed a biometric scanning application built on machine-learning systems from multiple vendors. When it was unable to explain failure rates because of the proprietary nature of the technology, concerns arose about the agency's accountability and procurement practices.

These lapses and others like them have led to widespread criticism and legal action. In the US, concerns over algorithmic-based decision systems in health care, criminal justice, education, employment, and other areas have resulted in lawsuits. Lapses have also led to organized efforts to track "bad actor" vendors and enact policies or laws to limit the use of public-sector AI systems or, at a minimum, make them more transparent. A 2020 UK parliamentary committee report on AI and public standards found that government and public-sector agencies are failing to be as open as

they should be about AI use. The report called on the government to uphold new transparency standards and create effective oversight processes to mitigate risks.

When lapses occur in government AI systems, trust in public institutions is eroded. Over time, such a disconnect can damage government legitimacy and citizens' belief in and support for governmental authority in general. According to BCG research, when people have a positive experience with online public-service delivery, it strengthens their trust in government. But when they have a negative experience, it has a disproportionately detrimental effect on trust. If governments' deployment of AI is to earn citizens' trust, they need to get it right. Responsible AI will foster both trust and the ongoing license that governments need in order to use machine learning to improve their own performance.

Responsible AI's Value to Government

As governments contend with the rules that control the use of emerging technologies, those that cultivate citizens' trust will be better situated to use AI to improve societal well-being. In addition to providing the direct benefits of AI systems, Responsible AI initiatives can foster government legitimacy, increase support for AI use, and help attract and retain people with digital skills.

Fostering Government Legitimacy. AI-based applications make up a fraction of all government systems and processes, but they have an outsized impact on people's perception of government institutions. Creating them in a responsible manner can encourage the belief that government is serving the best interests of individuals and society. As governmental use of AI increases and its impact becomes more widespread, the benefits of taking a responsible approach will grow.

A hallmark of Responsible AI is inclusive development. Considering multiple perspectives ensures that an AI-based assessment or model will treat all groups fairly

and not inadvertently harm any one individual or group. By seeking input from a broad spectrum of stakeholders—citizens groups, academics, the private sector, and others—governments also bring transparency to the process, which can build trust and encourage productive dialogue about AI and other digital technologies.

Increasing Support for Government Use of AI.
In a previous BCG study, more than 30% of those surveyed said they had serious concerns about the ethical issues related to government use of AI. Such concerns make people more likely to advocate against AI. If governments want to realize AI's potential benefits, they must take steps to show that the necessary expertise, processes, and transparency are in place to address those ethical issues.

A Responsible AI program can ease public resistance and smooth the way for governments to pursue long-term adoption of AI tools and services. The program should include a well-designed communications strategy that details the steps being taken to ensure that AI is delivered responsibly.

More and more, people recognize the benefits and potential harm brought about by emerging technologies. If government institutions are believed to be acting responsibly, they will have more license to deploy AI to improve their own performance.

Attracting and Retaining Digital Workers.
Amid a global shortage of digital talent, workers with AI skills are in especially high demand. Because private-sector companies can pay more, they have a distinct advantage over governments when it comes to recruitment and retention. So governments must find a different way to compete.

One option is to become a beacon of Responsible AI. Tech workers are particularly concerned about the ethical implications of the systems they create. In a study of UK tech personnel, one in six AI workers said they had left a job rather than participate in developing products that they believed could be harmful to society. Offering the opportunity to work in an environment where the concerns of employees and citizens are taken into account, and where AI is clearly used for the public good, could attract more people with the needed skills to public-sector jobs.

Achieving Responsible AI

Pursuing AI in a responsible manner is not an either-or proposition. Governments can realize AI's benefits *and* deliver technology in a responsible, ethical manner. Many have taken the initial step of creating and publicizing AI principles. A 2019 article in the journal *Nature Machine Intelligence* reported that 26 government agencies and multinational organizations worldwide have published AI principles. However, the vast majority lack any guidance on how to put these principles into practice, leaving a gap that could deepen negative perceptions of government AI use.

Governments can create a roadmap for building a Responsible AI program by empowering leadership, developing principles and policies, establishing a human-plus-AI governance mechanism, conducting use case reviews, integrating tools and methods (see "Six Steps to Bridge the Responsible AI Gap" for more on this step, which we do not discuss below), and building and testing a response plan in the event of lapses. (See the exhibit.)

Empower leadership.
Government agencies don't operate in a vacuum. They work together to develop and deliver public services, and they must work across departmental silos to develop Responsible AI. To manage AI responsibly, government can create a central leadership unit that adopts consistent, holistic approaches and standards. This unit can sit at the center of government or within an agency, department, or regulatory body. The location isn't as important as ensuring that the unit has a diverse membership and is the recognized authority in designing and leading all Responsible AI efforts. It can be a newly formed dedicated team, a task force with personnel from across government, or a combination of both. To ensure good communication, each department or

A Roadmap to Responsible AI in Government

Empower Responsible AI leadership
Appoint a leader and a diverse team to design and lead the Responsible AI program and run initiatives
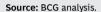
Develop principles and policies
Build, communicate, and disseminate Responsible AI principles and policies to all members of the AI team, including leaders

Establish human-plus-AI governance
Establish roles and responsibilities, a mechanism for review and adherence, escalation paths to raise concerns, and accountability for outcomes

Conduct Responsible AI reviews
Conduct end-to-end use case reviews and ensure that reviews are done at scale with independent oversight

Integrate tools and methods
Update standard data, technology, and model building to include Responsible AI considerations

Build and test a response plan
Create the roles and responsibilities, processes, and procedures to respond when a Responsible AI lapse occurs, and test and refine response plans periodically

Source: BCG analysis.

agency should appoint a liaison to the Responsible AI leadership unit, which should itself remain independent.

Regardless of how the Responsible AI unit is structured, it may take time to set up because of legal requirements or the need for policymaker approval. To avoid delays, a chief analytics officer or chief digital officer can be empowered to take temporary ownership of the program. Annual budgets may also have to be adjusted to account for its cost, although it may be possible to repurpose existing budgets for initial program expenses.

Develop principles and policies. It can take considerable time and effort for any organization to adopt new principles and policies. In government, the process is especially challenging. Implementing a new policy can involve consultations with multiple departments, agencies, and regulators, as well as feedback from academic experts, citizens groups, industry, and other external stakeholders. For example, when the Defense Innovation Board proposed AI ethics principles for the US Department of Defense, the group collected input from more than 100 experts and received close to 200 pages of public comments, a process that took more than 15 months. These steps are necessary, however, to ensure transparency and stakeholder buy-in. But the process must not prevent Responsible AI from being

implemented alongside early AI system adoption—or discourage it entirely.

To avoid delays, governments can use existing ethical principles laid out by international organizations while focusing on how Responsible AI should be tailored to their own context. For example, more than 50 countries have adopted AI ethical principles created by the Organisation for Economic Co-operation and Development. These principles can provide near-term guidance for nations beginning to develop Responsible AI programs.

Establish human-plus-AI governance.
Successful and ethical AI systems combine AI with human judgment and experience. People's role in creating and deploying Responsible AI should be elevated over the technology to ensure that the technology doesn't operate unchecked. Processes that accommodate people's expertise and understanding of context lead to better outcomes. In an example from the private sector, after redesigning its sales-forecasting model so that human buyers could use their experience and knowledge of upcoming fashion trends to adjust the algorithms, a retailer reduced its forecasting error rate by 50%, resulting in more than $100 million in savings a year.

Governments must adopt a similar human-plus-AI philosophy. AI development should include comprehensive end-to-end risk as-

sessments that incorporate human reviews and processes that allow anyone in the organization, or any end user, to raise concerns about a potential problem. Roles, responsibilities, and accountabilities need to be clear.

Conduct use case reviews. To ensure that program reviews are conducted at scale and with independent oversight, they should be led by the unit's leadership team. Reviews should occur regularly, with the participation of individual departments if a specific use case falls under their authority. Likewise, citizens should be part of the process when AI affects a specific constituency or society as a whole; this is analogous to companies asking for customer feedback on products in development. In such cases, the review should address potential unintended consequences and suggestions for potential remedies. Engaging citizens in reviews engenders trust in the government's use of AI, while the feedback gathered can be used to improve service quality.

Build and test a response plan. Mistakes happen. Roles and responsibilities, processes, and procedures should be established to address them. A communications plan to alert the public should be a key part of any Responsible AI response process, since transparency will help build and maintain trust over the long term.

Some recent government AI lapses were exacerbated by limited or poor communication. Communication problems underscore the need for AI use to be governed by common standards and oversight so that any issues can be addressed in a coordinated way. These also help spread expertise and knowledge across government, allowing every department to learn from AI's successes and failures. Responsible AI units can borrow from the response plans that governments have instituted in other areas, such as those adopted by many jurisdictions to inform citizens about cybersecurity breaches.

Using Procurement to Promote Responsible AI

When it comes to promoting Responsible AI, governments have a key advantage: the considerable amount of money that they already spend on AI products and services. AI spending in the Middle East and Africa was expected to jump 20% in 2020, to approximately $374 million, led in part by purchases from federal and central governments. In fiscal 2019, the US government spent an estimated $1.1 billion on AI. In 2020, the UK government requested proposals for AI services projects valued at approximately $266 million. The request follows the UK government's 2018 announcement of nearly $1.34 billion in spending on AI for government, industry, and academia.

When governments spend at that scale, they can shape market practices. They can use their spending power to set and require high ethical standards of third-party vendors, helping to drive the entire market toward Responsible AI.

To start with, they can make a vendor's access to public-sector procurement opportunities contingent on adopting Responsible AI principles. Agencies can require that a vendor provide evidence that it has instituted its own Responsible AI program, or they can include Responsible AI principles in their vendor evaluation criteria. The UK government has instituted AI procurement guidelines and vendor selection criteria that explicitly include such principles. The Canadian government has established a list of preapproved vendors that provide Responsible AI services. These examples, and guidelines such as those created by the World Economic Forum, can serve as models for other governments.

Responsible AI procurement practices require governments to balance transparency with the need to protect trade secrets. Visibility into software, algorithms, and data is required so that governments and the public can identify and mitigate potential risks. But companies want to keep proprietary information private to maintain a competitive advantage. Canada confronted this challenge by requiring commercial vendors to allow the government to examine proprietary source code. In exchange, the government assumed liability for maintaining

the confidentiality of the source code and underlying methodologies.

AS GOVERNMENTS EMBRACE more advanced technologies such as AI, they must continue to earn the trust of their citizens. By adopting an ethical and responsible approach, they can strengthen their legitimacy in the eyes of their constituents, earning the license to deploy AI in ways that deliver more effective, efficient, and fair solutions.

About the Authors

Steve Mills is a managing director and partner the Washington, DC, office of Boston Consulting Group. You may contact him by email at mills.steven@bcgfed.com.

Elias Baltassis is a partner and director in the firm's Dubai office. You may contact him by email at baltassis.elias@bcg.com.

Nadim Abillama is a senior knowledge analyst in BCG's London office. You may contact him by email at abillama.nadim@bcg.com.

Matthew Mendelsohn is a senior advisor in the firm's Toronto office. A former deputy secretary to the cabinet in Canada's Privy Council Office, he led the Canadian government's data strategy. You may contact him by email at mendelsohn.matthew@advisor.bcg.com.