# McKinsey & Company

Public Sector Practice

Unlocking the potential of generative AI: Three key questions for government agencies

Government organizations may seek to jump on the gen AI bandwagon, but the technology's complexities could sideline their efforts. Our framework addresses some critical implementation questions.

This article is a collaborative effort by Damien Bruce, Ankit Fadia, Tom Isherwood, Chiara Marcati, Aldous Mitchell, Björn Münstermann, Gayatri Shenai, Hrishika Vuppala, and Thomas Weber, representing views from McKinsey's Public Sector Practice.
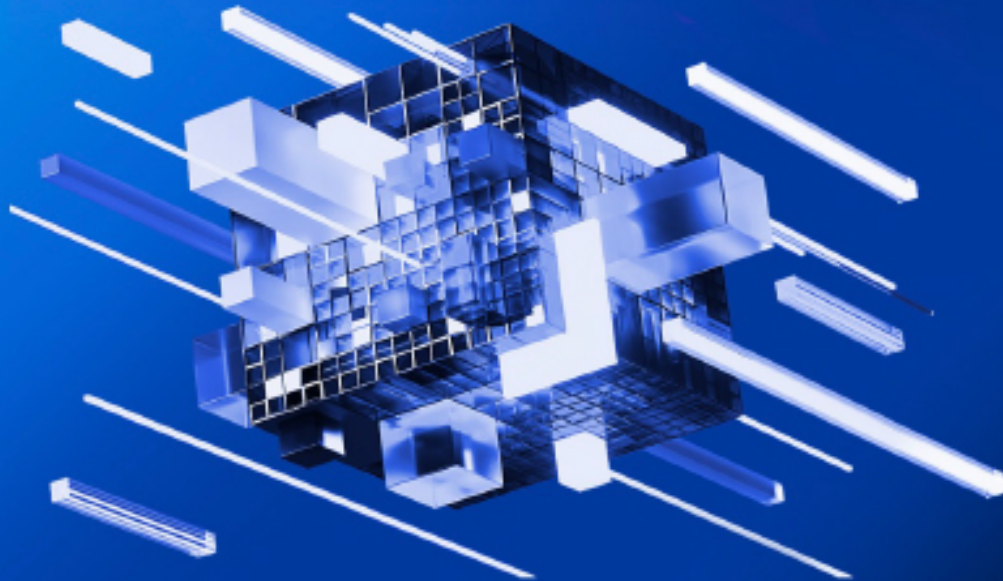
It's been just a year since generative AI (gen AI) tools first captured public attention worldwide. But already the economic value of gen AI is estimated to reach trillions of dollars annually—even as its risks begin to worry businesses and governments across the globe. Gen AI offers government leaders unique opportunities to steer national economic development (Exhibit 1). At the same time, they face the heavy burden of monitoring the technology's downsides and establishing robust guidelines and regulations for its use.

Many government agencies have started investing in transformations made possible by gen AI, but the technology's rapid evolution means that predicting where it can contribute the most value is difficult. In this article, we discuss three important questions that public sector organizations may need to consider before choosing areas for investment:

— How can government agencies address the potential risks of gen AI?

— How can public sector entities begin to transform their own service delivery?

— Should governments develop national gen AI foundation models (core models on which gen AI applications are built)?

We conclude with a suggested eight-step plan for government organizations that are just beginning to implement gen AI use cases.

1. How can government agencies address the potential risks of gen AI?

By now, the risks of gen AI—such as its tendencies toward unpredictability, inaccuracy, and bias—are widely known. Government agencies face different risks than do private sector companies. For

example, the technology can be misused to spread political propaganda or compromise national security. Confidential government data can be leaked or stolen if government employees inadvertently introduce that information into foundation models through prompts.

Some outputs from gen AI models might contain inaccurate information—also called "hallucinations"—that could erode public trust in government services that leverage these technologies. Like many private sector organizations, government agencies face challenges with gen AI's transparency and with the difficulty of explaining the conceptual underpinnings of gen AI, as well as the logic of the models' decisions and output. Consequences might include low public acceptance of gen-AI-powered government services and unclear liability when unintended effects occur. And like all organizations, government entities run the risk that criminals may misuse gen AI to carry out powerful cybersecurity attacks.

To address those risks, many countries—such as the United States, Australia, and China—have launched initiatives to create frameworks of regulations and policies for AI, and some have expanded their existing AI regulations to explicitly include gen AI, too. The European Union is leading a global effort to build safeguards for any product or service that uses an AI system. Many state government agencies in the US have also released AI-related legislation, executive actions, and policies focused on mitigating the potential risks of AI systems—by highlighting the negative aspects of AI, transparently communicating where AI is used in government, and addressing the ethical aspects of AI usage.
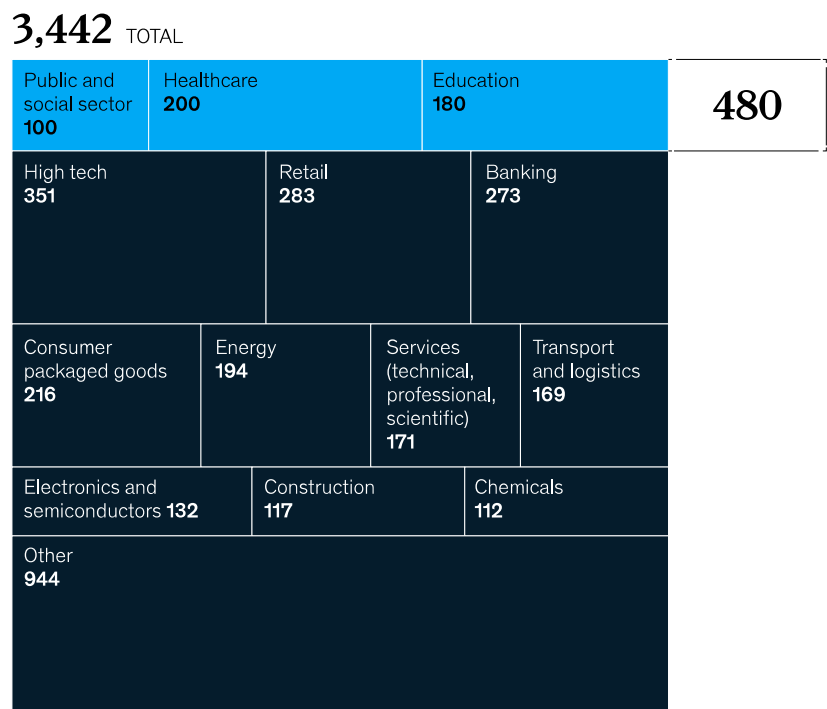
However, those mitigation efforts are still in their early stages in most parts of the world, and gen AI is evolving fast, which means that governments must revise their regulations continually to keep pace. Some government organizations have started ongoing awareness programs among stakeholders—especially end users—about gen AI's risks and how to address them. For example, the United Kingdom's Central Digital and Data Office has released a guide for civil servants on safe and informed use of gen AI tools. Similarly, Australia's Digital Transformation Agency and its Department of Industry, Science and Resources provide interim guidance to government agencies on responsibly using publicly available gen AI platforms, with emphasis on ethical AI usage, security, and human oversight.

Exhibit 1

Generative AI could have an estimated $480 billion productivity effect on the public sector and adjacent industries.

Generative AI productivity effect, by industry,[1] $ billion

3,442 TOTAL

Public and social sector **100**	Healthcare **200**	Education **180**	**480**

High tech **351**	Retail **283**	Banking **273**

Consumer packaged goods **216**	Energy **194**	Services (technical, professional, scientific) **171**	Transport and logistics **169**

Electronics and semiconductors **132**	Construction **117**	Chemicals **112**

Other **944**

[1]Excluding implementation costs (eg, training, licenses).
Source: McKinsey analysis

McKinsey & Company

2. How can public sector entities begin to transform their own service delivery?

As key providers of services to the public, government agencies are likely to prioritize the delivery of those services as a critical area for AI-driven improvements. A good place to start may be our "4Cs" framework, comprising four cross-industry categories: content summarization and synthesis, coding and software, customer engagement, and content generation (Exhibit 2). Most gen AI implementations we have seen fall into one of those four categories, which could apply to both private and public sector enterprises.

— *Content summarization and synthesis.* This category involves culling the most relevant insights from a large knowledge repository. For example, Singapore's GovTech has developed the Pair app, which summarizes text and generates reports for internal use.

— *Coding and software.* Software development could gain speed and increase productivity by using gen AI to write code and automate testing. Use cases will then need to be prioritized according to their potential impact, feasibility, and susceptibility to risk. For example, the United Kingdom's HM Treasury (economic and finance ministry) is testing GitHub Copilot (an AI pair programmer that offers coding suggestions) to accelerate software development.

— *Customer engagement.* Customer and client services could get a boost from gen AI apps—for example, in government agencies, chatbots could answer questions from or customize services for residents. The city of Heidelberg, in Germany, has launched the Lumi chatbot, the country's first digital citizen assistant. The tool enables people to easily navigate government services such as applying for a new identity card, getting a driving license, and registering a place of residence.

— *Content generation.* Gen AI can help produce a vast variety of content, including emails, social media posts, contracts, and proposals. For example, the US Department of Defense has developed an AI-powered contract-writing capability, called Acqbot, to speed up procurement.

Gen AI implementations could streamline a broad range of services that governments typically provide, in areas such as education, healthcare, defense and intelligence, and urban development (see sidebar "Potential applications of gen AI in government functions and services"). Across all of those areas, we have seen government agencies implement gen AI use cases in both external and internal operations that fall within the categories of our framework (see Exhibits 3 and 4). For example, in customer-facing applications, gen AI can help the public navigate government services and get access to real-time language translation. Internally, gen AI can draft creative content such as speeches and official correspondence, simplify complex official documents, and consistently generate financial reports and KPIs on schedule.

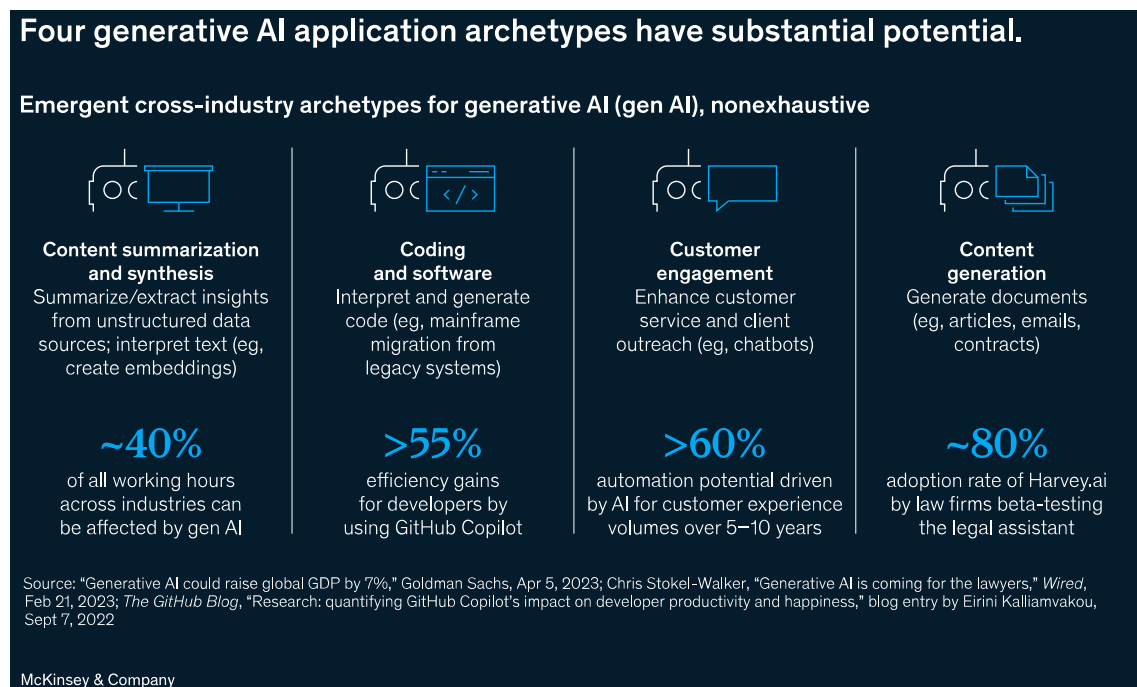## 3. Should governments develop national gen AI foundation models?

Some governments may aspire to develop foundation models—the core models on which gen AI applications are built. But leaders of government agencies must be aware that this endeavor requires considerable investment of time and resources. The many barriers to entry include the availability of

talent to build, train, and maintain gen AI models; the necessary computing power; and experience in addressing potential risks inherent in building and serving gen AI foundation models. Almost all current work in these models is led by a few large private sector tech companies (Cohere, Google, Meta, and others) and by open-source initiatives that are quickly becoming popular (such as Hugging Face, Stability AI, and Alpaca).

Unlike global private-sector tech players, government organizations simply lack the capabilities to develop foundation models while managing their risks. For example, violations of intellectual property and copyright laws can expose government agencies that own foundation models to litigation; gen AI's occasional lack of proper

source attribution makes it even harder to detect potential copyright infringement in its responses. Legal implications also apply to manipulated content—including text, images, audio, and video—that malicious actors may use to harass, intimidate, or undermine individuals and organizations. Users could act unscrupulously or illegally by exploiting inherent biases in the data that a specific foundation model was trained on. As a result, some governments—such as those of Iceland and Finland—have chosen to partner with global large language model (LLM) providers to get access to their existing models and augment and customize them to suit their own needs, by adding proprietary data and insights.

Exhibit 2

Four generative AI application archetypes have substantial potential.

Emergent cross-industry archetypes for generative AI (gen AI), nonexhaustive

Content summarization and synthesis	Coding and software	Customer engagement	Content generation
Summarize/extract insights from unstructured data sources; interpret text (eg, create embeddings)	Interpret and generate code (eg, mainframe migration from legacy systems)	Enhance customer service and client outreach (eg, chatbots)	Generate documents (eg, articles, emails, contracts)
~40%	**>55%**	**>60%**	**~80%**
of all working hours across industries can be affected by gen AI	efficiency gains for developers by using GitHub Copilot	automation potential driven by AI for customer experience volumes over 5–10 years	adoption rate of Harvey.ai by law firms beta-testing the legal assistant

Source: "Generative AI could raise global GDP by 7%," Goldman Sachs, Apr 5, 2023; Chris Stokel-Walker, "Generative AI is coming for the lawyers," *Wired*, Feb 21, 2023; *The GitHub Blog*, "Research: quantifying GitHub Copilot's impact on developer productivity and happiness," blog entry by Eirini Kalliamvakou, Sept 7, 2022

McKinsey & Company

Eight steps for getting started

For public sector agencies just beginning to venture into gen AI, we suggest this eight-step plan:

1. *Define your organization's risk posture.* After identifying your agency's risk parameters, devise a plan to mitigate the risks of using gen AI—with a mix of internal policies, guidelines, and awareness sessions.

2. *Identify and prioritize use cases.* Not everything needs gen AI technology to power it. Government agencies may find our 4Cs framework helpful in developing a list of potential use cases—and then prioritizing them according to potential impact and feasibility—while avoiding implementations with high potential for risk or limited tolerance for errors.

3. *Select the underlying model; upgrade technical infrastructure as needed.* Most public sector agencies begin with an off-the-shelf LLM and fine-tune it with proprietary data and integration with internal systems to deliver customized results. In very rare cases have we seen government agencies develop and train a new model from scratch. When that happens, it is driven primarily by aspirations to develop a national asset, manage data-sovereignty issues, or reduce dependence on private sector tech companies.

4. *Ensure that the necessary skills and roles are available.* "Head of AI" is one of the hottest jobs around, and governments will need to hire for it—only a senior executive can coordinate all gen AI–related activities and ensure that risks are addressed effectively. Traditionally, governments haven't had AI engineers, AI

Exhibit 3

Generative AI can help improve customer-facing operations in governments.

Potential use cases for generative AI in customer-facing government operations, by archetype

	Content summarization and synthesis	Coding and software	Customer engagement	Content generation
Programs and benefits	Eligibility and benefit distribution requirements	Transform legacy code (eg, COBOL to Python); generate code requirements	Citizen interactions via chatbot assistants	Personalized program content
Policies, licensing, and regulation	Policy design research; regulation summaries		License renewals	Legal, permitting, and/or regulatory guidance for new programs
Services	Standard operating procedure syntheses; economic-development incentives		Expert advisers (eg, to improve mental health responses)	Targeted emergency instructions
Infrastructure	Licensing infrastructure projects (eg, multicounty permits)		Auditing of regulatory and/or plan compliance	Requests for proposal and contracts to improve supply chain resilience

McKinsey & Company

ethics officers, or prompt engineers, but such roles must now be created and filled.

5. *Develop gen AI apps jointly with end users.* Gen AI is a fast-evolving technology, so early involvement of end users is critical not only for educating them on privacy and safety but also for collecting their feedback to improve the accuracy and performance of LLM responses. For example, users can provide a quantitative score for the quality of each response.

6. *Keep humans in the loop, at least for now.* Until gen AI technologies mature and enforceable regulations are in place, it may be prudent for government agencies to keep human managers accountable and use gen AI implementations only to execute models and not to monitor or assess them.

7. *Design a comprehensive communication plan.* Embed necessary disclaimers in all communication efforts to clarify the limitations of gen AI use cases and ensure safe adoption.

8. *Start small and scale up.* Our research shows that 72 percent of leading organizations find managing data to be one of the top impediments to scaling AI use cases. In our article on scaling gen AI programs, we identify seven actions that data leaders should consider as they move from experimentation to scale.

Exhibit 4

Generative AI can help improve internal operations in governments.

Potential use cases for generative AI in internal government operations, by archetype

	Content summarization and synthesis	Coding and software	Customer engagement	Content generation
Operations and delivery	Document summaries	Synthetic data generation	Virtual assistants for suggested responses/ diagnoses	
Technology	Current state mapping; data cleaning	Transform legacy code (eg, COBOL to Python); generate code requirements		Cybersecurity test case generation
Talent	Résumé summaries; feedback synthesis		Personalized recruiting and onboarding materials	Job descriptions; interview guides; training materials
Fiscal, asset, and performance management	Budget summaries; transaction syntheses		Citizen budget navigators	Financial reports; standard performance reports; KPIs

McKinsey & Company

Potential applications of gen AI in government functions and services

We see gen AI uses across different public sector domains. Here's a sampling.
In *education*, gen AI can do the following:

— serve as a virtual tutor to review and grade essays

— act as a student assistant chatbot, providing well-sourced answers

— help students in workshops and labs, guiding them through experiments

— generate preliminary drafts of teaching aids, such as lecture scripts and quizzes

— draft initial versions of university applications based on student profiles

In *urban development*, gen AI can do the following:

— act as a design-compliance assistant for urban administrators

— summarize citizen feedback from hotlines for city planners

— draft initial urban planning layouts to enhance city designs

— design optimal public transport routes to reduce traffic congestion

— simulate urban design impacts, such as traffic flow or sunlight exposure

In *tax and customs*, gen AI can do the following:

— serve as a virtual assistant for complex import-guidance synthesis

— be a real-time document verifier for customs and tax officials

— generate risk-assessment summaries from diverse data sources

— prepare preliminary audit reports for potential noncompliance areas

— act as a public assistant for real-time tax-filing guidance

— prepare first drafts of personalized notifications to taxpayers and traders

In *agriculture, forestry, and fishing*, gen AI can do the following:

— summarize and categorize farmer helpline calls for timely departmental assistance

— act as a virtual assistant to guide farmers on government schemes and benefits

— provide personalized crop-management advisories for updated farming techniques

— analyze crop and soil images to suggest restoration techniques for improved yield

In *justice, courts, and legal systems*, gen AI can do the following:

— act as a real-time legal assistant to judges and attorneys

— simplify and explain intricate legislation and case law

— keep legal professionals updated with summarized regulatory changes

— counsel repeat offenders with tailored legal education

— suggest potential outcomes for judicial sentencing for judges' consideration

— automate the drafting of various legal documents to ease administrative burdens

In *defense and intelligence*, gen AI can do the following:

— act as a real-time translator for intelligence agencies

— draft initial intelligence-operation reports

— tailor training content to specific military missions

— simulate potential conflict scenarios for strategic planning

— generate synthetic misinformation data to enhance surveillance AI

In *healthcare*, gen AI can do the following:

— simplify the descriptions of complex diseases to improve patients' understanding of them

— summarize post–acute-care instructions for patients

— generate concise discharge instructions to optimize clinical operations

— act as a co-pilot to summarize imaging diagnostics for clinicians

— summarize patients' medical histories for streamlined review

START YOUR FINANCE

起点财经，网罗天下报告