

Berlin
Global
Government
Technology Centre



In collaboration with Capgemini

Making Agentic AI Work for Government: A Readiness Framework

INSIGHT REPORT

APRIL 2026

Contents

Foreword	3
Executive summary	5
1 The agentic opportunity	6
1.1 From process digitization to outcome orchestration	7
1.2 Why agentic AI matters for governments	9
1.3 The tactical challenge: where to begin	9
2 An innovative government readiness framework for agentic AI	10
2.1 A function-based assessment lens	11
2.2 The assessment of potential against complexity	13
2.3 A topography of government readiness	15
2.4 From global topography to regional roadmap	22
3 Learning from successful deployments	26
Conclusion	30
Appendices	31
A1 Methodology	31
A2 Comprehensive breakdowns of agentic AI assessment scores for all 70 core government functions	33
Contributors	37
Endnotes	39

Disclaimer

This document is published by the World Economic Forum as a contribution to a project, insight area or interaction. The findings, interpretations and conclusions expressed herein are a result of a collaborative process facilitated and endorsed by the World Economic Forum but whose results do not necessarily represent the views of the World Economic Forum, nor the entirety of its Members, Partners or other stakeholders.

© 2026 World Economic Forum. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, including photocopying and recording, or by any information storage and retrieval system.

Foreword



Martina Klement

Permanent Secretary for Digital Transformation and Administrative Modernization; Chief Digital Officer, State of Berlin



Fabian Mehring

State Minister for Digital Affairs and Chief Information Officer, Free State of Bavaria



Mohamed Bin Taliah

Assistant Minister, Cabinet Affairs for Government Experience Exchange Affairs, Ministry of Cabinet Affairs, United Arab Emirates

Over the past decade, governments around the world have invested heavily in digital transformation – and it has paid off. Public services have moved online, data-driven decision-making has gained real traction across many administrations, and the govtech market has matured from a niche into a global force. Technology is no longer a support function for the state. It is becoming central to how governments operate, deliver services and earn the trust of their citizens.

The drivers for this development vary. In some countries, citizen expectations have outpaced what traditional administration can deliver, creating urgency to rethink how services are designed and provided. In others, demographic decline means that doing more with fewer people is not a choice but a reality. For some, digital infrastructure is the backbone of economic competitiveness; for others, it is the path to rebuilding public trust after years of institutional underperformance. The starting points may be different, but the direction is the same: technology is moving to the core of government.

Agentic artificial intelligence (AI) accelerates this shift – and brings a new quality to it. Where earlier waves of digitization moved paper processes on to screens, agentic AI systems can plan, decide and act across entire workflows, coordinating steps that previously required manual handoffs, interpreting context and delivering outcomes rather than outputs. For governments already under pressure from multiple directions, this is not a theoretical prospect. It is a practical lever.

The opportunities are vast. Yet, opportunity has never been the bottleneck for government technology. The real challenge is strategic: understanding which operations benefit most, sequencing adoption so that early efforts build capability rather than drain it and maintaining public accountability throughout. This demands clarity, not haste.

Governments that approach agentic AI with discipline and strategic intent will not only improve their own operations. They will shape the norms and expectations for how this technology is governed in the public interest. That makes this not just an opportunity, but a shared responsibility.

Foreword



Ammar Alkassar
Member, Board, GovTech
Deutschland



Roshan Soorunsingh Gya
Chief Executive Officer, Northern
Europe, Capgemini



Stephan Mergenthaler
Managing Director; Chief Technology
Officer, World Economic Forum

Governments around the world are exploring the potential of artificial intelligence (AI), and agentic AI in particular is attracting growing attention as a way to move beyond isolated tools towards systems that can coordinate, decide and act across complex workflows. Yet for most, this remains difficult to translate into practice. What is often missing is a credible basis for deciding where to begin, what to prioritize and how to move from experimentation to operational reality.

This matters because the cost of poor implementation is real. Governments that pursue agentic AI without a realistic understanding of where the potential lies – and where it does not – risk investing in the wrong places, building solutions that do not deliver and losing institutional confidence in the technology along the way. Enthusiasm alone does not prevent this. What governments need is enthusiasm paired with strategic clarity: a grounded view of which government operations are genuinely suited for agentic AI and which require a different approach.

That is what this report sets out to do. The strategy and action of public administration are shaped by organizational boundaries – ministry by ministry, department by department. Agentic AI does not work that way. It operates across workflows that span organizational lines: “cybersecurity monitoring”, “document life cycle management”, “eligibility assessment”, “fraud detection” and “benefit calculation”. Harnessing agentic AI therefore requires a fundamentally different understanding of public administration itself.

This report develops that understanding. It maps 70 core government workflows against two dimensions – the opportunity for agentic AI to add public value, and the complexity of deploying it responsibly – showing where governments can act with confidence, where targeted preparation is needed, and where caution is warranted. We offer this as a shared starting point (adaptable to local conditions), not a final answer. Agentic AI will reshape how governments operate. Whether that transformation is guided by evidence and strategic intent, or left to chance, is a choice. This report is our contribution to ensuring it is the former.

Executive summary

A pioneering framework to help governments move agentic AI from experimentation to scalable public value.

Public expectations are rising, fiscal space is tightening, and many administrations are being asked to deliver more with less. Against this backdrop, agentic artificial intelligence (AI) represents a fundamental shift in capability, enabling systems to autonomously execute end-to-end multi-step workflows, with the potential to transform how governments serve citizens.




Realizing this opportunity requires strategic, evidence-based adoption – grounded in a clear assessment of where agentic AI can deliver the greatest public value, what risks must be managed and what capabilities and safeguards need to be in place before deployment at scale.

This report responds to that need by introducing the first systematic framework for assessing government readiness for agentic AI and evaluating core global government activities (or “functions”). The framework is complemented by real-world use cases that ground the assessment in practice, highlighting current agentic AI initiatives.

The starting point: assessing agentic AI readiness across government functions

By focusing on functions – recurring workflows that cut across organizational segregation rather than isolated tasks or departments – this framework provides governments with a new basis for prioritizing deployment and implementation at scale.

Each function is assessed across two dimensions (agentic AI potential and implementation complexity) and mapped into three readiness areas:

-  **High-readiness area:** suitable for early deployment, with robust safeguards
-  **Medium-readiness area:** suitable for phased implementation requiring additional enabling conditions, capability building or analysis
-  **Low-readiness area:** suitable for monitoring, iterative testing and longer-term preparation

The review of 70 core government functions indicates that 50% combine significant agentic AI potential with manageable implementation complexity, pointing to a substantial opportunity

for scaled adoption where institutional capacity and safeguards are in place.

Clear takeaways emerge from this analysis:

- **Think in functions, not departments:** Agentic AI operates in workflows that cut across organizational boundaries.
- **Balance ambition with feasibility:** High agentic AI potential should be weighed against implementation complexity before operationalizing at scale.
- **Start where the best odds exist:** Build capability and confidence with high-readiness functions before tackling more complex ones.
- **Local context determines success:** Global scores are baselines. Local infrastructure, regulatory environments and cultural norms determine what is possible.
- **Expect the topography to evolve:** Reassess regularly. Functions in the low-readiness area today may be in the medium- or high-readiness area tomorrow.

From insight to action

The report translates analysis into a practical decision-support framework for governments considering agentic AI adoption. It provides a structured approach to identifying relevant functions and prioritizing opportunities based on potential public value and implementation barriers, helping decision-makers focus on where agentic AI is most likely to make an impact. These insights are intended to inform subsequent choices on governance design, risk management, piloting and scaling. Local adaptation is essential: each jurisdiction must tailor the framework to its organizational priorities, digital maturity and regulatory context.

This framework is a starting point – a tool to inform strategic choices. It is designed to help policy-makers and industry move together – establishing where to begin, how to build capability and ultimately how to convert agentic AI opportunities into measurable public value.

1

The agentic opportunity

Agentic artificial intelligence can transform public institutions by orchestrating end-to-end workflows, shifting government operations from task automation to outcome delivery.



Government technology has become a foundational driver of competitiveness, institutional capacity and public trust. [The Global Public Impact of GovTech: A \\$9.8 Trillion Opportunity](#) estimates a \$9.8 trillion opportunity from public-sector digital transformation by 2034.¹ Converting that opportunity into sustained impact, however, depends on how governments design and deploy the next generation of capabilities.

Agentic artificial intelligence (AI) is one of the most important levers for realizing this potential.

Unlike earlier AI applications that focused on narrow tasks such as classification, prediction or pattern recognition, agentic systems can coordinate multi-step processes, integrate information across multiple sources and adapt their actions based on context and evolving conditions. These capabilities are well aligned with the structure of many government workflows, which typically involve sequential decisions, multiple stakeholders and the application of rules and judgement over time.

1.1 From process digitization to outcome orchestration

“ Agentic systems can deliver meaningful efficiency gains while maintaining the accountability, quality and trust on which public services depend.

Many government digital transformation efforts have focused on digitizing existing processes: moving paper applications online, digitizing records and automating data entry. These efforts have delivered real value, but they have largely preserved the underlying logic of public administration – the medium changes, while the process remains fundamentally the same. Agentic AI marks a meaningful shift because it does more than analyse data: it can plan, decide and act across multiple steps of a workflow (see Box 1).

Until now, public-sector digitization has operated under a familiar constraint: digitizing a flawed process largely preserved its flaws. Agentic AI can help alleviate this constraint by reasoning across information and executing goal-directed actions. While processes must still be redesigned to unlock its full potential, improvements to existing workflows can deliver meaningful gains in the near term.

- **Agentic workflow optimization:** Instead of requiring a complete process redesign before automation, agents can directly coordinate and improve existing workflows. Their ability

to orchestrate interdependent steps can enhance performance and consistency without a full re-engineering effort.

- **Agentic reinvention:** Agentic AI enables a structural rethink of legacy processes by shifting the focus from task execution to outcome orchestration, opening the door to new service models and operating patterns that were not possible under traditional, fragmented system architectures.

Crucially, agentic AI is best understood as a model of augmentation rather than replacement. By handling routine coordination, accelerating case processing and surfacing relevant information, these systems allow public servants to focus on meaningful work that requires human judgement: complex cases, policy interpretation and direct engagement with citizens. When deployed with appropriate oversight, governance and safeguards, agentic systems can deliver meaningful efficiency gains while maintaining the accountability, quality and trust on which public services depend.

BOX 1 What is agentic AI?

An AI agent is a software system that autonomously executes tasks based on goals, applying decision logic while operating within predefined constraints. Agentic AI refers to the coordinated use of one or more AI agents that operate with limited human intervention and bounded autonomy across workflows. Unlike tools that require constant prompting or follow only predetermined scripts, AI agents can operate in unstructured environments, make context-sensitive decisions in real time, learn from outcomes and take proactive action. To ensure accountability, agentic AI is typically deployed with human oversight. This may take the form of direct involvement in key decisions (human-in-the-loop)

or supervisory control with the ability to intervene when needed (human-on-the-loop).

Agentic AI is not the only pathway to improving public sector performance. Predictive analytics, robotic process automation (RPA) and generative AI tools already deliver measurable efficiency gains in many government contexts. In some cases, these approaches are more appropriate and offer a resource-efficient initial step before exploring more advanced agentic capabilities. The key is to apply the right level of intelligence to the right type of workflow. Table 1 contrasts traditional automated workflows, generative AI and agentic AI.

TABLE 1 | Comparison of RPA, generative AI and agentic AI

Category	Automated workflows (RPA)	Generative AI	Agentic AI
Focus	Automates repetitive tasks and processes	Primarily creates new content	Primarily executes actions
Output	Facilitates task execution across software systems	Generates text, images, code, audio, video, etc. based on patterns learned from vast training data	Generates not just content, but executes actions, decisions, and multi-step processes to reach a larger objective
Autonomy	Has no autonomy, relies entirely on predefined instructions; designed exclusively to reproduce tasks as directed by humans or previous RPA tasks	Has limited autonomy, needs specific prompts and guidance to produce output	Has high autonomy, capable of making decisions and executing tasks independently to pursue complex goals
Learning ability	Has no learning ability	Is pre-trained, with limited to no real-time learning; uses memory through mechanisms like conversation history but does not act on it dynamically	Is learning and adapting continuously; actively uses long-term memory; dynamically retrieves or acquires new information and has contextual awareness
Interactivity	Has no interactivity	Cannot perform real-time interactions; LLM model itself has no external tool access unless agentic features are added	Interacts with external and internal data, tools, and systems in real time; is better positioned to handle errors
Level of capabilities	 Low Medium High		

Source: Based on Capgemini. (2025). *Rise of agentic AI: How trust is the key to human-AI collaboration*.

BOX 2 **Further reading**

The following publications provide additional perspectives on the implementation and governance of agentic AI:

- [Ilves, I. et al. \(2025\). *The Agentic State: Rethinking Government in the Era of Agentic AI*. Global Government Technology Centre Berlin and The World Bank.](#)
- [Capgemini. \(2025\). *Rise of agentic AI: How trust is the key to human-AI collaboration*.](#)
- [World Economic Forum. \(2025\). *AI Agents in Action: Foundations for Evaluation and Governance*.](#)

1.2 Why agentic AI matters for governments

“ For agentic AI to deliver value without eroding transparency, entrenching bias, weakening accountability or undermining public trust, strategic sequencing and governance capacity are key.

Governments face a decisive moment shaped by both acceleration and constraint. Agentic AI has reached operational maturity for public sector use, supported by increasingly robust commercial ecosystems. Yet fiscal pressure and service demands continue to intensify. The mandate to do more with less has become a structural requirement.

A recent survey conducted by Capgemini² – spanning 350 public-sector organizations globally across six sectors (predominantly public administration, tax and customs) indicates that 90% of surveyed institutions plan to explore or deploy agentic AI within two to three years. This reflects more than technological optimism. It signals recognition that governments cannot meet current mandates with current operating models alone. Agentic AI offers a credible path to operational improvement at scale, and the conditions for successful deployment are stronger than they have ever been. At the same time, safeguarding sovereign control over critical elements of AI systems – across regulatory, operational, technological and geopolitical dimensions – will be essential to ensuring the transformation remains aligned with public-interest requirements.

This transition marks a broader shift towards what can be described as an **agentic state**: disciplined deployment of agentic AI across core government workflows. An agentic state represents a model of public administration in which AI-enabled systems autonomously coordinate and execute across institutional boundaries, reduce administrative latency and enable more proactive, outcome-oriented service delivery.

The benefits can be assessed through various value metrics, including greater consistency in decision-making, improved compliance, time savings, reduced errors, improved citizen satisfaction and equity impacts. Defining these early provides a clear frame for evaluating where agentic AI creates tangible administrative value. Yet this transformation carries real risks. For agentic AI to deliver value without eroding transparency, entrenching bias, weakening accountability or undermining public trust, strategic sequencing and governance capacity are key, and its autonomy must be intentionally limited. Agents should operate within clearly circumscribed mandates, escalate to human oversight when necessary and make their decisions transparent.

1.3 The tactical challenge: where to begin

Given ongoing technical progress, the central question for governments is now strategic rather than purely technical: which activities are appropriate for agentic automation, which offer sufficient public value returns – meaning economic returns as well as tangible social benefits – to justify investment and which involve risks or complexities that require caution.

The costs of both error and inaction are real. Delay can force dependence on externally developed solutions that do not reflect local needs or policy priorities; precipitous action can lead to fragmented pilots that drain resources, erode institutional confidence and impede future AI initiatives without delivering results.

While many public sector institutions have recognized the opportunity, Gartner® projects that “over 40% of agentic AI projects will be cancelled by the end of 2027, due to escalating costs, unclear business value or inadequate risk controls”.³ This provides a cautionary note, underscoring how enthusiasm can outpace strategic planning. A systematic, evidence-based approach is therefore essential to identifying where agentic AI can deliver long-term public value.

② An innovative government readiness framework for agentic AI

By evaluating the AI readiness of core government functions, the framework highlights where AI agents can be deployed and where risks outweigh benefits.



The report introduces a framework with four building blocks to guide decision-making on agentic AI in government (see Figure 1). It assesses functions against both its potential and its complexity, producing a clear picture of where

governments can act with confidence, where they should invest in preparation and where caution is warranted – helping them translate the broad promise of agentic AI into strategic decisions.

FIGURE 1 The government readiness framework for agentic AI

	1 →	2 →	3 →	4
What?	A function-based assessment lens The report specifies what unit governments need to evaluate for agentic AI decisions.	The assessment of potential against complexity The report defines how governments should evaluate and score agentic AI readiness.	A topography of government readiness The results are visualized as an agentic AI readiness map, showing where to act, prepare and monitor.	From global topography to regional roadmap The results must be adapted to local context to become actionable.
How?	Use of a domain-agnostic unit of analysis 70 recurring core functions in nine categories	Setting of assessment criteria in two dimensions <ul style="list-style-type: none"> 1 Potential (opportunity) 2 Complexity (barriers) 	Scoring of government readiness Potential – complexity = agentic AI readiness	Creation of topography with three prioritization areas <ul style="list-style-type: none"> 1 High-readiness area (deploy) 2 Medium-readiness area (prepare) 3 Low-readiness area (monitor)
				Adaption to local context in six steps <ul style="list-style-type: none"> 1 Assess baselines 2 Reassess functions 3 Sequence implementation 4 Test 5 Scale 6 Iterate

2.1 A function-based assessment lens

“ Earlier AI applications were designed for narrow tasks within a single context. Agentic AI, by contrast, can coordinate entire workflows that span multiple systems and decision points.

The framework maps government activities not by department but by function: the recurring workflows that cut across organizational boundaries. In this report, the term function is not used in the traditional sense of a department unit – e.g. human resources (HR), finance, logistics. Instead, it refers to outcome-oriented operational activities that encompass one or more end-to-end workflows.

Thinking agentic: a new lens for public administration

Agentic AI requires a different lens for thinking about how public administration is organized and evaluated. Traditional approaches to technology implementation in government tend to follow organizational boundaries: assessing digitalization department by department, or cataloguing AI use cases sector by sector. These approaches

made sense for earlier generations of technology, but not for adaptive systems that operate across workflows, institutions and decision cycles.

The distinction becomes clear when examining how agentic systems function. Earlier AI applications were designed for narrow tasks within a single context – classifying documents, predicting demand or recognizing patterns. Agentic AI, by contrast, can coordinate entire workflows that span multiple systems and decision points.

For example, a traditional AI system used in application processing might automate one step, such as extracting data from submitted documents or flagging incomplete forms for review. An agentic AI system, however, could manage the full end-to-end workflow: receiving and logging submissions, cross-checking information across databases, orchestrating required steps and routing cases to the appropriate authority.



“ Governments can identify where agentic AI is both feasible and valuable, and create opportunities for reuse rather than generating fragmented pilots shaped by institutional segregation or vendor offerings.

This shift has implications for how readiness should be assessed. A department-based approach defines the boundary by institution (for example, a national tax authority) without identifying which underlying processes should change. A technology-led approach starts with a tool, such as chatbots, and looks for applications, rather than diagnosing the underlying operational need. A use case approach can be equally limiting, often treating isolated scenarios, producing bespoke pilots that fail to generalize or capture repeatable workflows across agencies.

By contrast, a workflow-based analysis begins with the recurring functions governments perform, including processing applications, allocating benefits and issuing permits, rather than with departments or tools. It examines where these processes rely on shared data, require coordination across agencies or involve sequential decision-making over time. By focusing on these cross-cutting functions, governments can identify where agentic AI is both feasible and valuable, and create opportunities for reuse rather than generating fragmented pilots shaped by institutional segregation or vendor offerings. The result is guidance that is more scalable, more transferable across institutions and more grounded in how public administration actually operates.

The function-based methodology

Governments perform millions of workflows every day. Assessing them meaningfully requires a unit of analysis that reflects real workflow logic, without becoming too granular to compare or prioritize. For agentic AI, that unit is a function, a recurring government activity that delivers a clear result and

requires several connected workflows to move from initiation to final decision. A function is concrete enough to be assessed, yet general enough to recur across contexts – allowing its readiness to be evaluated once, rather than separately within every organization or policy domain where it appears. For instance, the “eligibility assessment” function incorporates several workflows (gathering information, verifying credentials, applying criteria, making determinations, documenting decisions) and spans diverse departments, such as social services, healthcare, education and housing.

The function-based approach offers clear advantages:

- First, it reflects operational reality – how government work gets done – rather than relying on organizational structures that vary across jurisdictions and are more prone to political and administrative reconfiguration.
- Second, the function-level abstraction provides an appropriate unit for focused pilots, with clearly defined scope and measurable outcomes.
- Third, it promotes cross-domain learning and reuse; insights from a pilot in one function can be adapted elsewhere, supporting a standardized, reusable operating model for agent deployment. This reduces duplication and lowers barriers to entry, making previously complex or costly initiatives more feasible. Shared components – such as authentication, front-end interfaces or data stores – form an agentic template that can be deployed across organizations, shortening initial development cycles and providing a ready-made foundation for subsequent pilots.

The inventory: 70 functions across nine categories




Existing classifications of government functions, such as the UN's COFOG (Classification of the Functions of Government),⁴ as well as skills or occupation-based frameworks like O*NET (developed by the US Department of Labor)⁵ provide a useful foundation. However, they organize government activity by expenditure categories, occupational roles or skills rather than the underlying workflows through which government operations are actually carried out, and therefore do not fully meet the needs of this framework.

This new catalogue of 70 core government functions was developed in consultation with government officials and subject-matter experts

across geographies, drawing on the experience of leading delivery and public-sector practitioners. The catalogue was created through a structured set of criteria. Each function represents at least one complete, recurring workflow with clear inputs and outputs, appears across diverse national government contexts independent of administrative structure and is performed regularly. Every function also meets at least one significance test: substantial citizen impact, a legal mandate or a prerequisite role in enabling other public functions. The resulting inventory includes 70 functions grouped into nine categories, which are further organized into three overarching themes (see Figure 2).

The complete list appears in the Appendix. The taxonomy intentionally focuses on core recurring workflows rather than attempting to create a universal classification of all government activities.

FIGURE 2 70 core government functions organized into nine categories

 Core service delivery	Public services	15 functions
	Citizen communication and interaction	08 functions
	Security and risk response	07 functions
	Compliance and oversight	03 functions
 Policy and governance	Policy execution	04 functions
	Policy planning and strategy	08 functions
 Support functions and enabling infrastructure	Operations and resources	08 functions
	Organization and workforce	08 functions
	IT infrastructure and data governance	09 functions

2.2 The assessment of potential against complexity

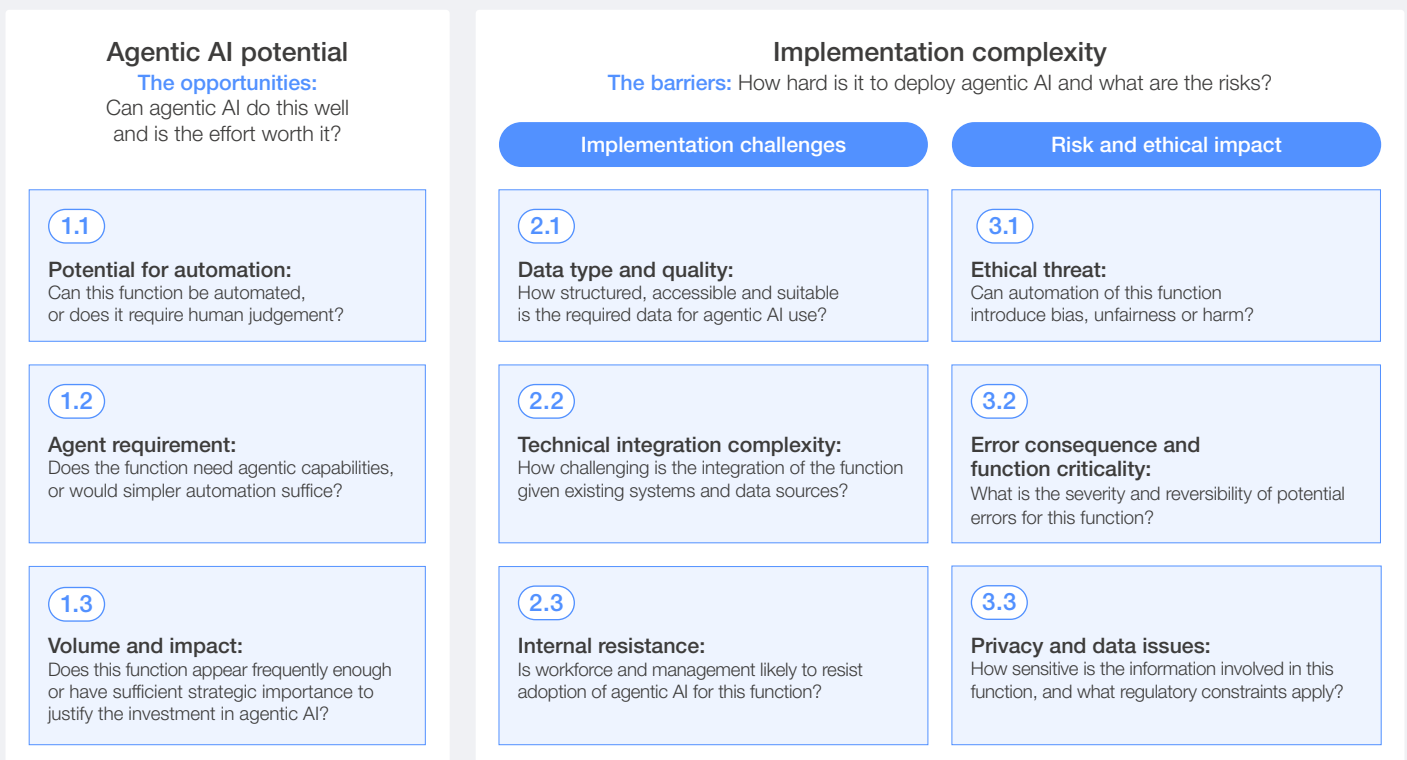
Cataloguing functions establishes scope; prioritization requires deeper analysis. Effective decision-making depends on understanding which functions are well-suited to agentic AI and which face meaningful implementation barriers. The assessment framework addresses this by evaluating each function across two dimensions: potential and complexity. Figure 3 outlines the detailed criteria underpinning this assessment.

Agentic AI potential evaluates whether AI agents can perform a given function effectively and generate meaningful public value. This dimension comprises three sub-criteria.

Implementation complexity assesses the practical barriers to deploying agentic AI for a given function. This dimension comprises six sub-criteria. The first three address implementation challenges, and the last three focus on risk and ethical impact.



FIGURE 3 | Assessment dimensions for the government readiness framework for agentic AI



Each sub-criterion is scored on a standardized scale, then aggregated into composite potential and complexity ratings. The scoring was conducted by a team of experts specializing in agentic AI and public-sector operations and reviewed with practitioners and global public-sector specialists to validate and reconcile differing perspectives.

The ratings in this report reflect global estimates based on typical conditions observed across public administrations and draw on shared patterns in how governments are structured. Individual countries will score higher or lower than these benchmarks depending on their data infrastructure maturity and

regulatory context. These deviations are analytically informative, revealing where national contexts create advantages or barriers.

Crucially, the two-axis framework requires readiness on both potential and complexity. High potential alone does not warrant prioritization when implementation complexity remains elevated. This guards against a common failure mode in technology adoption: allowing enthusiasm for capability to outpace institutional readiness for responsible deployment. Strategic assessment must precede operational planning. Governments need to know where to act before determining how to act.

2.3 A topography of government readiness

In the assessment, each of the 70 functions was scored on potential (vertical axis) and complexity (horizontal axis), illustrated in Figure 5. When plotted together, these scores form a visual map of readiness: a topography that reveals where agentic AI can deliver the greatest public value.

The functions are categorized based on their level of readiness. In this framework, readiness reflects how well a government function is suited for agentic AI – balancing the value it could create (potential

dimensions) with the practical, organizational and ethical considerations required for responsible implementation (complexity dimensions). Figure 4 illustrates this categorization across three readiness areas, ranging from high to low.

The analysis indicates that 50% of functions demonstrate medium to high readiness for agentic AI, suggesting that many core government activities are viable candidates provided appropriate safeguards are in place.

FIGURE 4 Overview of readiness areas

Calculation of agentic AI readiness score = potential score minus complexity score

10 functions (top 14%)
with highest agentic AI
readiness scores



High-readiness area

Suitable for early deployment,
with robust safeguards

25 functions (next 36%)
with medium agentic AI
readiness scores



Medium-readiness area

Suitable for phased
implementation, requiring
additional enabling conditions,
capability building or analysis

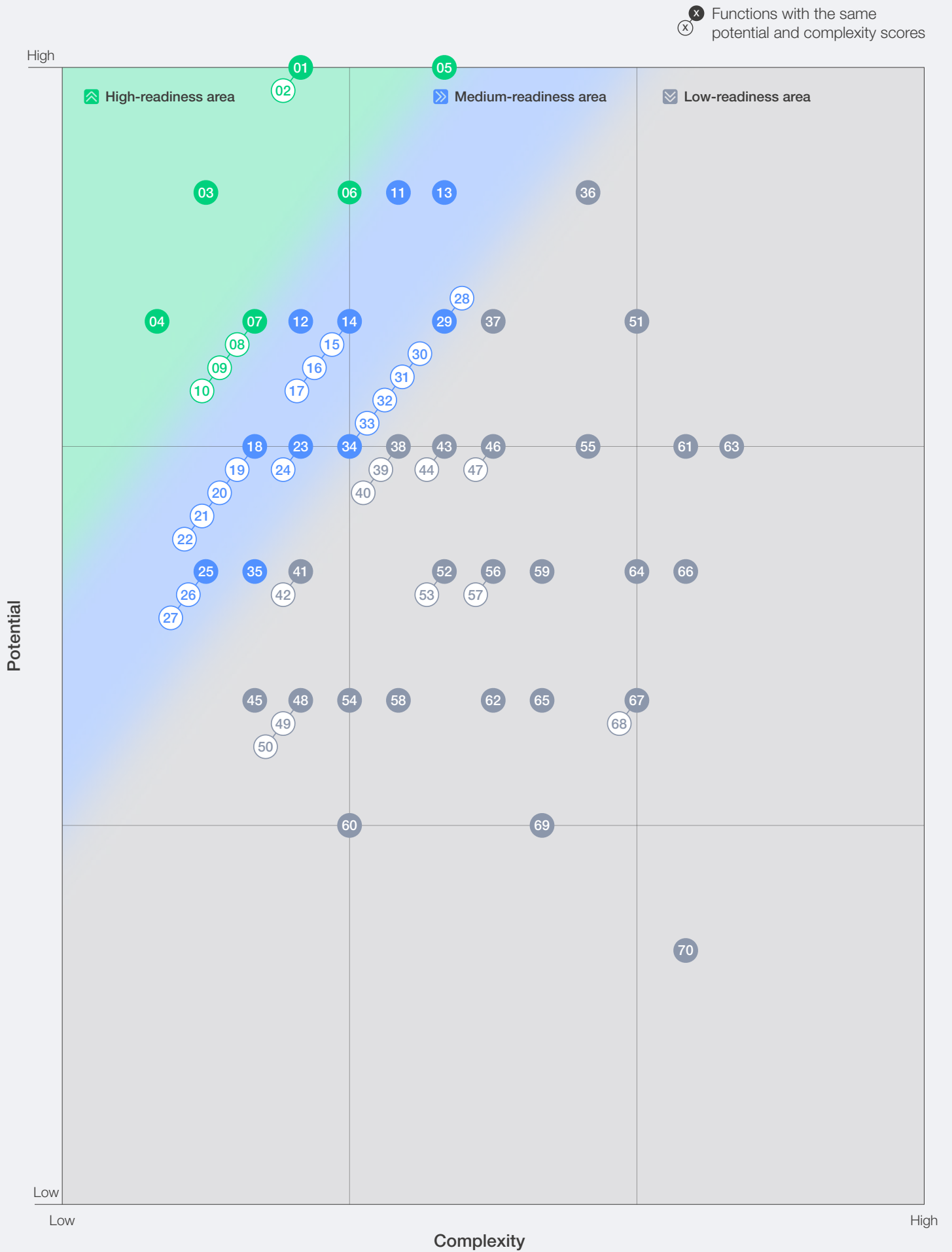
35 functions (bottom 50%)
with lowest agentic AI
readiness scores



Low-readiness area

Suitable for monitoring,
iterative testing and
longer-term preparation

FIGURE 5 | Assessment for 70 functions on agentic AI potential vs implementation complexity

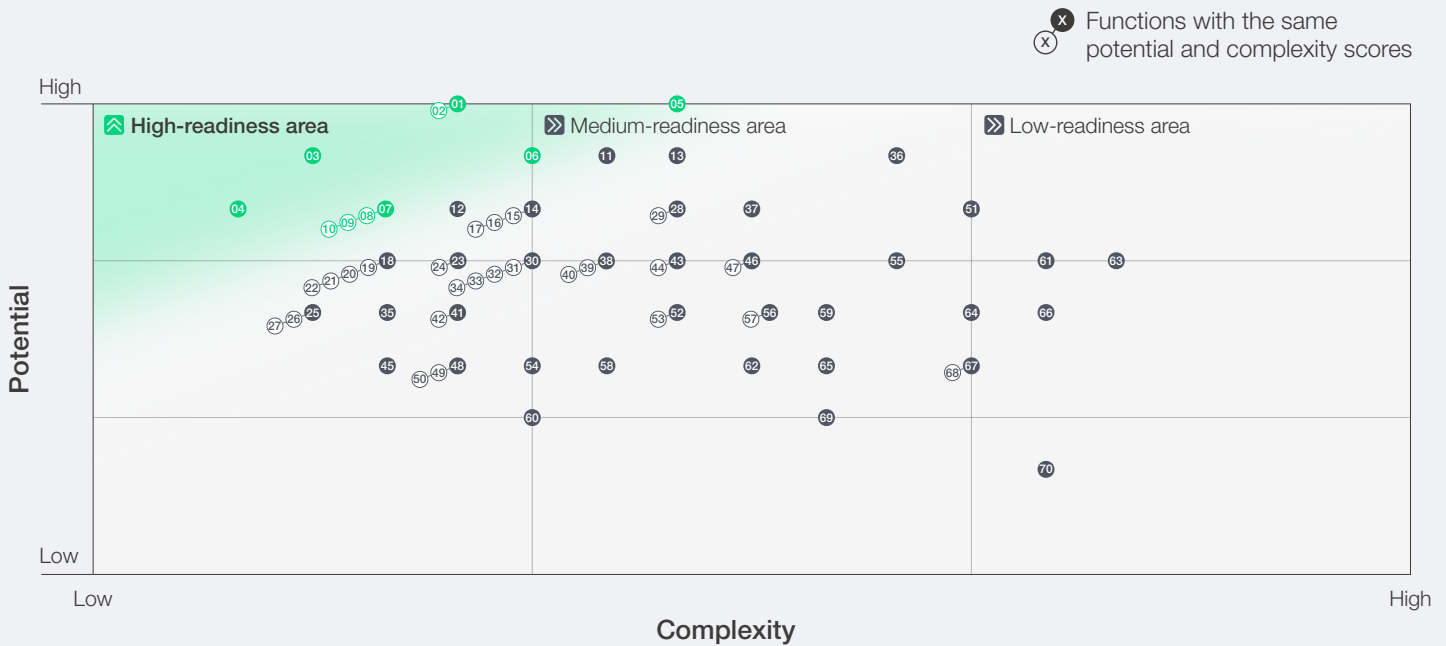


Core government functions by readiness area (sorted by readiness score)

High-readiness area 		Low-readiness area 	
01	Cybersecurity monitoring	36	Service communication channel operations
02	Public information and guidance provision	37	Interoperability and data exchange management
03	Systems performance monitoring	38	Emergency preparedness monitoring
04	Service appointment and queue management	39	Complaint intake and resolution
05	Threat intelligence and early warning	40	Legal and regulatory records management
06	Tender preparation and awarding	41	Contract and supplier management
07	Document validation and processing	42	Archiving and long-term preservation
08	Transparency reporting and disclosure	43	Policy forecasting and scenario modelling
09	Financial performance monitoring and compliance	44	Capital investment planning
10	Document life cycle management	45	Legal and regulatory framework analysis
		46	Risk mitigation
		47	Corporate compliance inspection
		48	Multilingual content adaptation
		49	Policy option formulation and drafting
		50	Infrastructure, asset and facility management
		51	Cybersecurity incident response and recovery
		52	Public warning and alert dissemination
		53	Budget planning and allocation
		54	Workforce development and training
		55	Policy impact prediction
		56	Public consultation facilitation
		57	Multi-level governance coordination
		58	Workforce planning and analytics
		59	Privacy protection and compliance management
		60	Program implementation planning
		61	Eligibility assessment and verification
		62	Stakeholder relationship mapping
		63	Crisis response coordination
		64	Inter-agency service orchestration
		65	Inter-governmental policy coordination
		66	Investigation and evidence management
		67	Grant and subsidy allocation
		68	Policy impact assessment and evaluation
		69	Organizational transformation implementation
		70	Employee performance management

The following sections describe each area and related implications for prioritization and implementation.

FIGURE 6 High-readiness area



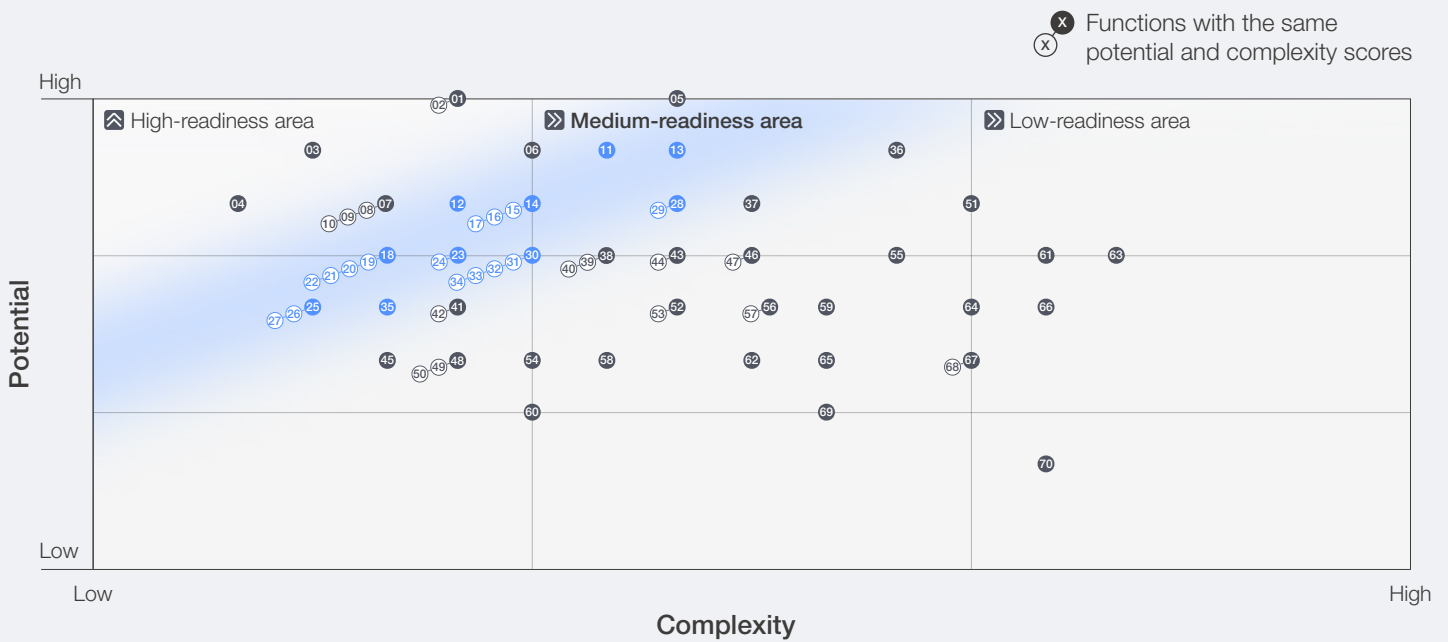
Functions in this area show the highest readiness for agentic AI and should be prioritized (see Figure 6). They typically rely on structured data, raise fewer ethical concerns, are well-suited to agentic coordination and present relatively low implementation barriers. While some carry meaningful error consequences (i.e. severe or hard-to-reverse mistakes), these can be managed through appropriate safeguards and controls, making them logical starting points for deployment.

“Cybersecurity monitoring”, “public information and guidance provision”, “systems performance monitoring”, and “service appointment and queue management” are the functions with the highest readiness for agentic AI implementation. Each function combines high potential with relatively low complexity, although none should be treated as low-stakes in practice.

Less publicly visible functions also feature prominently in this category. “Document life cycle management”, “data quality assurance and validation”, and “financial performance monitoring and compliance” may attract less attention than citizen-facing services. Yet they involve substantial volumes of recurring work on well-structured data and help build organizational capability and confidence before agentic systems are deployed in higher-risk contexts.

Five of the nine function categories appear in the high-readiness area, demonstrating that readiness is distributed across most government activities. “Public services” (see Figure 2) is the most represented category – it accounts for three of the top ten functions (it also contains the largest number of functions overall). This creates a strong foundation for improving citizen services across common life events. These functions handle high transaction volumes, deliver cost-efficiency gains and directly affect citizen experience and satisfaction.

FIGURE 7 | Medium-readiness area



These functions exhibit average to strong overall readiness (see Figure 7). Compared with high-readiness functions, they generally require less autonomous coordination or involve smaller transaction volumes or strategic impact, making them less likely to deliver transformative effects in the near term.

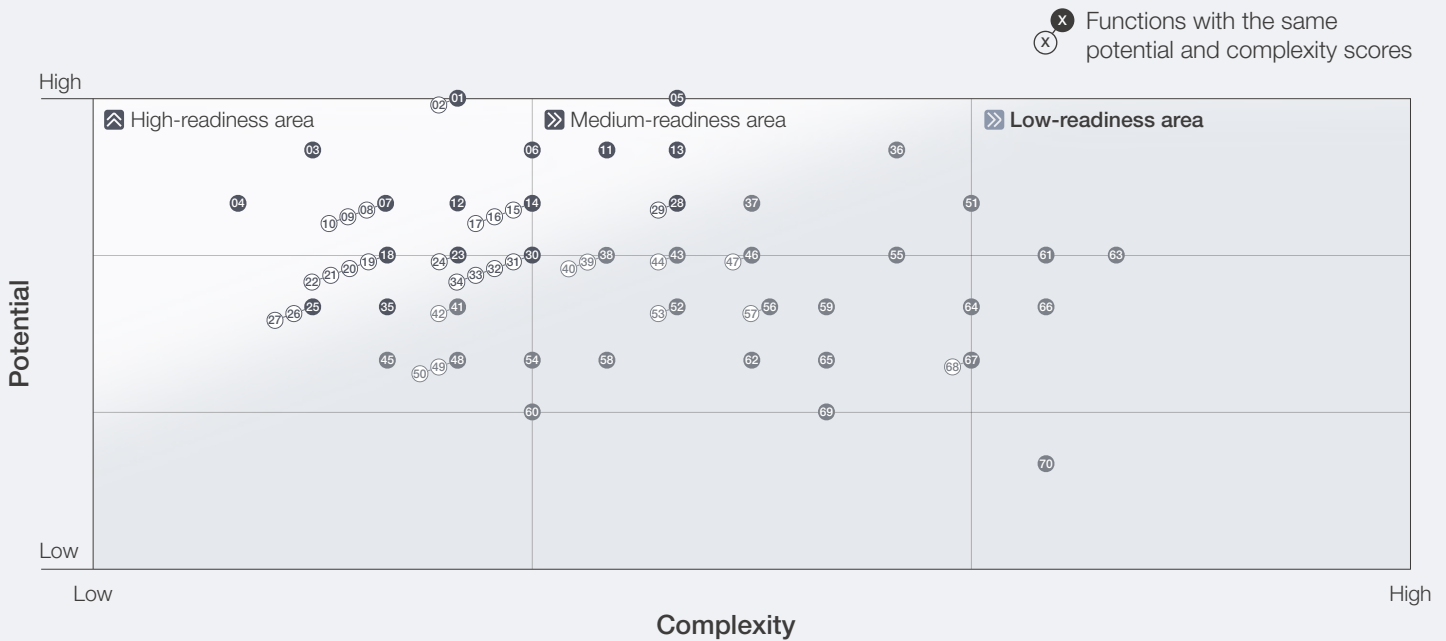
Within the area, two clusters emerge:

- **Medium potential and low complexity:** This cluster merits deeper subprocess-level analysis to identify fast, high-certainty opportunities for agentic enablement. A phased approach is recommended – i.e. establish data and integration baselines via traditional

automation and subsequently introduce agentic functionalities in incremental steps. Targeted, process-level pilots will generate learning and lower the risk of broader deployment.

- **High potential and medium complexity:** While these functions represent considerable opportunity, they will usually require distributed, multidimensional investment – e.g. data and integration work, risk mitigation measures and capability building – before scalable agentic deployment is viable. Eligibility determination workflows (granting or denying rights, benefits or access) illustrate this dynamic: they offer major potential but involve pronounced ethical and privacy risks.

FIGURE 8 | Low-readiness area



The lower half of functions fall into the low-readiness area (see Figure 8). Most of the functions in the categories “policy execution” and “policy planning and strategy” (see Figure 2) sit here because they require nuanced interpretation, strategic judgement and extensive stakeholder engagement; data is often fragmented or qualitative, limiting the near-term value of agentic automation.

This area reveals a structural constraint: high potential does not offset high complexity. Despite strong potential, functions such as “interoperability management” remain constrained by cross-agency coordination and data-sharing barriers, while “cybersecurity incident response and recovery” entails severe error consequences.

Most of the “security and risk response” functions fall into the low-readiness area. These functions are less suitable starting points, as they require a long-term, rigorously governed approach. They operate in high-stakes environments, where complex system integrations and severe error consequences limit the feasibility of near-term agentic deployment.

By contrast, the “organization and workforce” category (see Figure 2) highlights how widely readiness can vary within a single function category. Half of its functions fall into the low-readiness area,

yet real-life use cases – such as agent-supported onboarding – demonstrate the existing automation opportunities enabled by well-structured HR data. In contrast, low-readiness HR functions such as “employee evaluation” require a more cautious, longer-term adoption pathway due to their sensitivity and the associated bias risks.

However, all these functions should be actively monitored and periodically reassessed as legal, technical and organizational conditions evolve.




This analysis operates on a meso-level

It assesses functions using globally applicable dimensions, so its findings have limited direct applicability to specific local contexts. Moreover, functions with low overall readiness may contain subprocesses suitable for agentic AI, warranting targeted pilots. Classifications are not fixed. A function that ranks low today may become a priority as data infrastructure strengthens, governance frameworks mature or institutional acceptance increases. The framework is designed to be adaptive, providing a structured basis for ongoing reassessment as conditions evolve.

“Agentic AI priorities must reflect more than technical readiness, aligning efficiency gains with legal, ethical and societal expectations.”

Patterns across the topography

The aggregate analysis yields several strategic recommendations for governments developing an agentic AI strategy (see Figure 9). The following five are particularly noteworthy:

- 
Prioritize “public services” functions as prime targets for agentic automation:
 Two-thirds of “public services” functions are positioned in the high-readiness area, with repeatable, rule-based workflows that deliver visible citizen-facing improvements. These functions can serve as starting points for demonstrating tangible value and generating institutional buy-in.
- 
Build on mature data foundations to accelerate agentic IT operations:
 Governments with strong IT operations and mature data structures can find early wins in their own technology management functions and are well-positioned for agentic experimentation and rapid, scalable deployments.
- 
Focus on information-driven functions before tackling judgement-based ones:
 Functions that present, aggregate or monitor

information, e.g. “cybersecurity monitoring”, are lower-risk and easier to implement than functions like “eligibility assessment”, where agents would influence sensitive decision-making. As a result, they offer a practical entry point for adopting agentic AI. For governments, this highlights that agentic AI priorities must reflect more than technical readiness, aligning efficiency gains with legal, ethical and societal expectations.



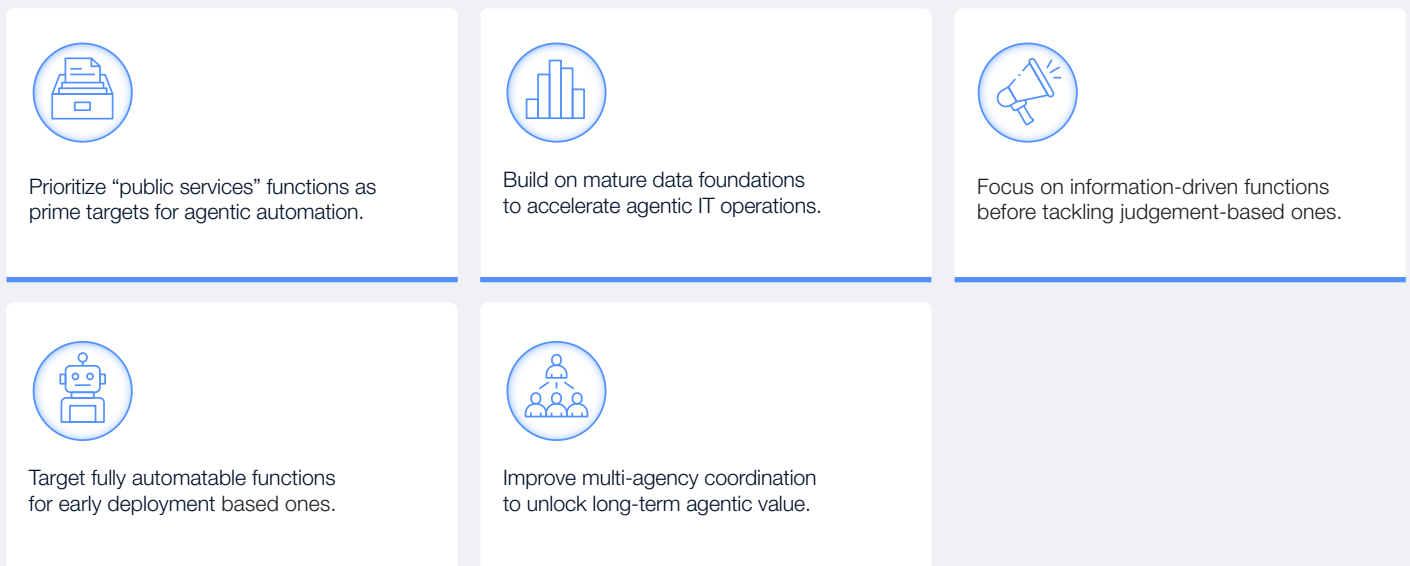
- 
Target fully automatable functions for early deployment: Functions that can be entirely automated face fewer organizational, ethical and error-related barriers and are strong initial candidates for agentic implementation. This is not coincidental: functions that have resisted automation often do so for reasons that also make agentic deployment challenging.
- 
Improve multi-agency coordination to unlock long-term agentic value: Functions spanning multiple agencies face a hidden tax from data-integration challenges and fragmented responsibilities, slowing down implementation – particularly in federal systems. While this constrains adoption in the near term, cross-agency automation should be seen as a longer-term goal where multi-agent systems may unlock compounding value.

FIGURE 9 Analytical patterns across the readiness analysis

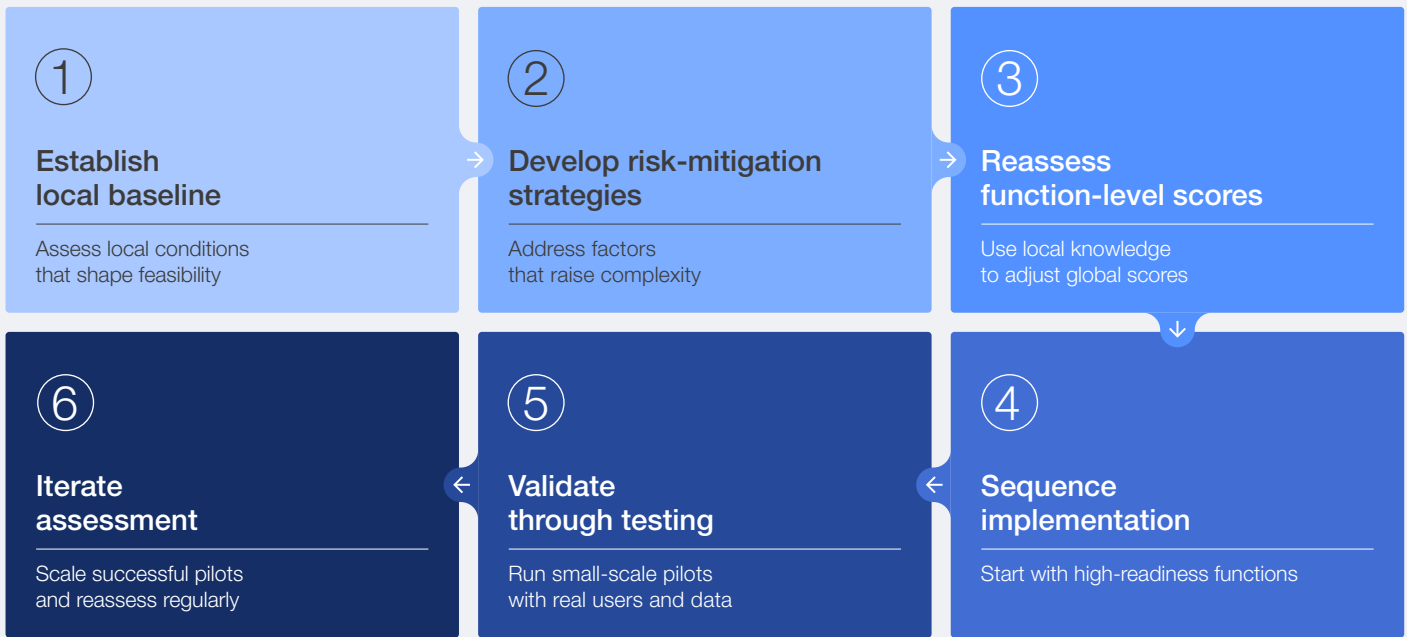


2.4 From global topography to regional roadmap

The scores reflect cross-country baseline estimates, not country-specific recommendations. They highlight general patterns, but feasibility depends on local context. A function scoring low in the global analysis might be a suitable

opportunity in a country with mature digital infrastructure and strong data governance. The same function might be effectively out of reach in a jurisdiction where legal constraints are binding, or legacy systems are deeply entrenched.

FIGURE 10 Steps to translate the global topography to local roadmaps



“ Governments earlier in their digital journey may lack foundational infrastructure but can also benefit from building on greenfield conditions without legacy constraints.

Variation from global scores is not a flaw – it is the framework’s value. When local assessments diverge from baselines, governments learn where they have advantages or face barriers. If high-complexity functions prove easier to implement than the global benchmark suggests, prior infrastructure investments will yield returns. If they prove more difficult, specific institutional or technical barriers require attention. The global assessment provides the reference point; the national assessment reveals relative readiness. Translating the framework into a national implementation roadmap requires six steps:

1 Establish local baselines

Before evaluating individual functions, governments should assess the cross-cutting conditions that determine feasibility. These systemic factors reshape the implementation landscape, establishing the contextual baseline for subsequent function-specific assessments.

Digital maturity encompasses the overall state of government digitization: whether core services are already online, whether internal processes

have been re-engineered for digital delivery and whether the organization has experience managing technology-driven change. The relationship to readiness for agentic AI is not straightforward. Governments with higher digital maturity bring valuable experience and established capabilities. Evidence from the GovTech Maturity Index (GTMI)⁶ shows that these digitally advanced governments continue to progress faster than less mature peers, widening the digital divide and demonstrating how strong foundational capabilities can accelerate further transformation. However, they may also carry complex legacy systems and technical debt that can complicate integration efforts. In contrast, governments earlier in their digital journey may lack foundational infrastructure but can also benefit from building on greenfield conditions without legacy constraints. Attention should shift from measuring levels of digitization to identifying the specific factors that enable progress and those that constrain it.

Data infrastructure refers to the availability, quality and accessibility of the data that agentic systems require. This includes not just whether data exists, but whether it is structured, standardized and accessible. Governments with mature data infrastructure and established data governance

practices start from a stronger position; those with fragmented, uncoordinated or poorly structured data will face higher barriers.

Cloud readiness affects the speed and cost of deployment. Agentic systems benefit from scalable compute resources and modern development environments. Governments with established cloud infrastructure and clear policies on cloud use can move faster; those still navigating procurement, security or sovereignty concerns around cloud adoption will face delays.

Workforce capabilities include both technical skills (data science, AI engineering, systems integration) and organizational capacity for managing AI-driven change. Governments with relevant in-house expertise or that are accessible through trusted partners are better positioned; those facing skills gaps will need to invest in recruitment, training or external support before ambitious deployments become feasible.

Funding availability determines what scale of investment is realistic. Agentic AI implementation requires upfront spending on technology, integration and change management, with returns materializing over time. Governments with dedicated digital transformation budgets or access to innovation funding can pursue larger initiatives; those operating under tighter fiscal constraints may need to seek and prioritize smaller-scale entry points.

Political mandate and leadership shape what is organizationally feasible. Agentic deployment often requires changes to processes, roles and decision-making authorities that encounter resistance without clear leadership support. Many stakeholders have difficulty accepting the occasional failure inherent to innovation;

decisive political leaders can mitigate this by articulating a clear mandate for structured experimentation, endorsing controlled pilots and establishing accountability frameworks that protect civil servants from undue risk. Governments with a strong commitment to digital transformation and AI adoption can thus overcome organizational inertia and translate technical feasibility into real-world deployment.

Legal and regulatory clarity is essential for defining accountability, liability and acceptable automation boundaries; jurisdictions with clear, enabling laws can move faster, while uncertainty about legal limits discourages civil servants from taking necessary operational risks. Because digitized systems will never be error-free, legal frameworks play a critical role in defining how residual risk is managed, and in shaping the conditions under which agentic systems can be deployed responsibly.

Citizen acceptance of AI-driven services and public trust strongly influences adoption speed and determines whether agentic AI is perceived as beneficial to service quality or not.

A mature local govtech ecosystem, comprising vendors and innovation partners, enables governments to access expertise and proven solutions. Where ecosystem maturity is low, governments may need to invest in partnerships or capacity-building before scaling.

These nine factors provide an overall readiness baseline. Governments that score strongly across them are likely to find a broad set of functions within reach. Those facing gaps across multiple dimensions should expect to strengthen foundational capabilities before expanding into more complex areas.



“ Agentic AI should be applied only where it demonstrably delivers substantial benefit, driven by high volumes or a clear requirement for agentic capabilities.

2 Develop clear strategies to mitigate risks and challenges

Many contextual factors that raise complexity can be addressed through targeted measures. Clear guardrails and risk frameworks are essential. As with other AI challenges, safe public-sector use of agentic systems depends on explicit, practical strategies to manage those risks.

Sustainability: The growing computing footprint of agentic systems is a critical challenge, with significant energy and water demands for training and operation. The assessment framework’s potential dimension supports the prioritization of functions where expected public value impact justifies (ecological) cost. Agentic AI should be applied only where it demonstrably delivers substantial benefit, driven by high volumes or a clear requirement for agentic capabilities. Deployments must be justified on that basis, and implementation choices should favour energy-efficient model architectures, optimized cloud provisioning and operational practices that minimize environmental and financial burden.

Workforce readiness: Two primary workforce risks emerge with agentic AI: excessive reliance on automation and resistance to change. Overreliance occurs when staff become too dependent on automated agents, reducing their ability to intervene effectively in cases where human judgement is needed or maintain critical judgement skills. Resistance to change can slow adoption if employees lack trust in new agentic systems or fear role displacement. Therefore, it is critical to invest in AI literacy, skills development and change management to drive adoption and responsible use.

Value alignment: Another risk is posed by agents acting contrary to public values, producing discriminatory outputs or compromising privacy. Mitigation includes ethical impact assessments, sandbox testing, continuous monitoring for bias and behavioural drift, human-in-the-loop validation, and definition of clear autonomy boundaries and rollback mechanisms for unintended actions.

Infrastructure resilience: Technical risks include exposure to adversarial attacks in environments with weak infrastructure or security, as well as vulnerabilities arising from external tool or application programming interface (API) integrations. Mitigation requires strong authentication and authorization for external tools and APIs, end-to-end encryption and hardened operational security for agent interactions.

Explainability: Explainability and auditability are non-negotiable in the public sector. It is therefore essential to embed explainability mechanisms that make decision rationales visible in user-facing interfaces, while maintaining detailed audit trails and observability tools to record agent actions and detect anomalies.

Taken together, these measures form practical risk-management guidelines allowing for safe implementation. Specific mitigations must be tailored to local legal frameworks, technical environments and organizational contexts.

3 Reassess function-level scores

Once the baseline and risk strategies are established, governments should reassess functions within their own national contexts, as the assessment assigns scores based on globally applicable estimates that may not reflect local realities. Jurisdictions should therefore reassess the dimensions using domestic knowledge and data. Key questions include: where are the biggest operational pain points (and thus the largest opportunities), what legal constraints apply, how does the public perceive AI, and do technical interfaces exist for data exchange across agencies?

Local reassessment will shift the position of some functions. Activities that appear complex at the global level, such as those requiring cross-agency coordination, may be more feasible in jurisdictions with centralized authority, streamlined governance or highly interoperable systems. The same holds true for countries that operate shared platforms and standardized data-exchange frameworks. Where local assessments diverge from global scores, the reasons matter: positive divergences highlight opportunities to move faster than the baseline suggests, while negative divergences reveal specific barriers that may require targeted investment or indicate functions to defer until conditions improve.

4 Sequence and prepare implementation

Not every high-readiness function should be pursued at once. Implementation capacity is limited, and initial choices influence future options. Therefore, careful sequencing of efforts is essential.

Start with high-readiness functions to build capability and confidence. Early deployments should target functions where complexity is low. Success in these initial efforts builds organizational confidence in agentic systems, develops internal expertise and demonstrates value to stakeholders whose support will be needed for more ambitious initiatives. Governments that move directly to complex deployments often encounter resistance that could have been avoided with a stronger foundation of demonstrated success.

Recognize dependencies between functions. Some functions enable others and act as prerequisites. Investments in data quality assurance, interoperability and identity verification create infrastructure that makes subsequent deployments easier. Governments should consider

“ Governments should treat early deployments as opportunities to develop and test governance approaches on the ground – the essential base for more complex implementations.

these dependencies when sequencing, prioritizing foundational functions that serve as multipliers, even if their standalone value may appear moderate.

Prepare workflows for agentic deployment.

Before deploying agentic AI into any function, governments should apply a structured process transformation lens such as the ESOAR (eliminate, standardize, optimize, automate, robotize) framework.⁷ This ensures that legacy inefficiencies are removed and that workflows are streamlined to create public value. Embedding ESOAR into the sequencing stage helps governments avoid automating inefficiency by automating old underlying function processes.

Use early deployments to develop governance frameworks.

High-readiness functions are not just opportunities to capture value – they are opportunities to test and learn. Implementing agentic systems, even in lower-stakes contexts, surfaces questions about oversight, accountability error handling, and human-machine interaction that must be answered before moving to higher-stakes functions. Governments should treat early deployments as opportunities to develop and test governance approaches on the ground – the essential base for more complex implementations.

Exchange learnings, successes and challenges.

Regular knowledge exchange is essential for implementing agentic AI successfully. Given the limited number of agentic AI pilots globally currently under way, international dialogue helps governments sequence implementation by learning from each other’s successes and avoiding repeated mistakes. Platforms such as the Global Government Technology Centre Berlin (GGTC) and the World Economic Forum’s GovTech Network can provide a structured forum for continuous sharing of insights and best practices, enabling public organizations to benefit from each other’s experiences.

5 Validate through testing

Assessment frameworks, however rigorous, cannot surface every barrier. Procurement rules, legacy system dependencies, vendor capabilities, workforce concerns and stakeholder politics often become visible only when implementation begins. Governments should validate priorities through small-scale pilots before committing to full deployment.

Pilots serve several critical purposes. They validate that a function performs as intended with real users and live data, and they expose integration issues that are not visible in abstract assessments. Pilots surface the organizational and political factors that shape implementation while generating practical evidence to inform scaling decisions.

Designing pilots on a function basis is particularly valuable, as it enables consistent comparison between human and agent performance and allows for clearer assessment of operational impact.

Effective pilots are scoped narrowly enough to be delivered quickly and at manageable cost, but broadly enough to encounter realistic conditions. A pilot that operates on sanitized data in a controlled environment will not reveal the challenges of production deployment. Governments should design pilots that engage real workflows, real data and real users, while containing the scope of any unintended consequences.

The output of a pilot is not just a working system – it is a decision. Does the evidence support scaling this pilot? What modifications are needed? What can be learned that affects the assessment of related functions? Governments should build deliberate decision points into their pilot processes rather than allowing pilots to drift indefinitely or expand without evaluation.

6 Iterate as conditions evolve

The topography is not static. Technology advances, infrastructure matures, regulations evolve, organizational capabilities develop and societal attitudes towards AI shift over time. A function that appears out of reach may become accessible after targeted investments. Governments should therefore use the framework as a dynamic diagnostic tool, revisiting assessments over time as capabilities mature and constraints evolve.

Regular reassessment serves several purposes. It captures the effects of investments already made (e.g. if a government has improved its data infrastructure, the scores for data-dependent functions should reflect that improvement). It also maintains strategic focus as the landscape shifts, ensuring that implementation priorities remain aligned with current conditions rather than outdated assessments.

Readiness will shift as conditions evolve. For example, HR automation currently faces trust and bias concerns, but these barriers may diminish once trust in agentic systems becomes more established and technical capabilities for bias-free models mature. Annual framework reviews and more frequent checks on active priorities are a reasonable starting point. Rather than viewing initial assessments as fixed, reassessment should be built into ongoing strategic processes.

The topography is a map, not a plan. It shows what’s possible – each government must chart its own path and adjust as conditions change.

3

Learning from successful deployments

A closer look into agentic AI use cases shows the tangible impact agents are already delivering.



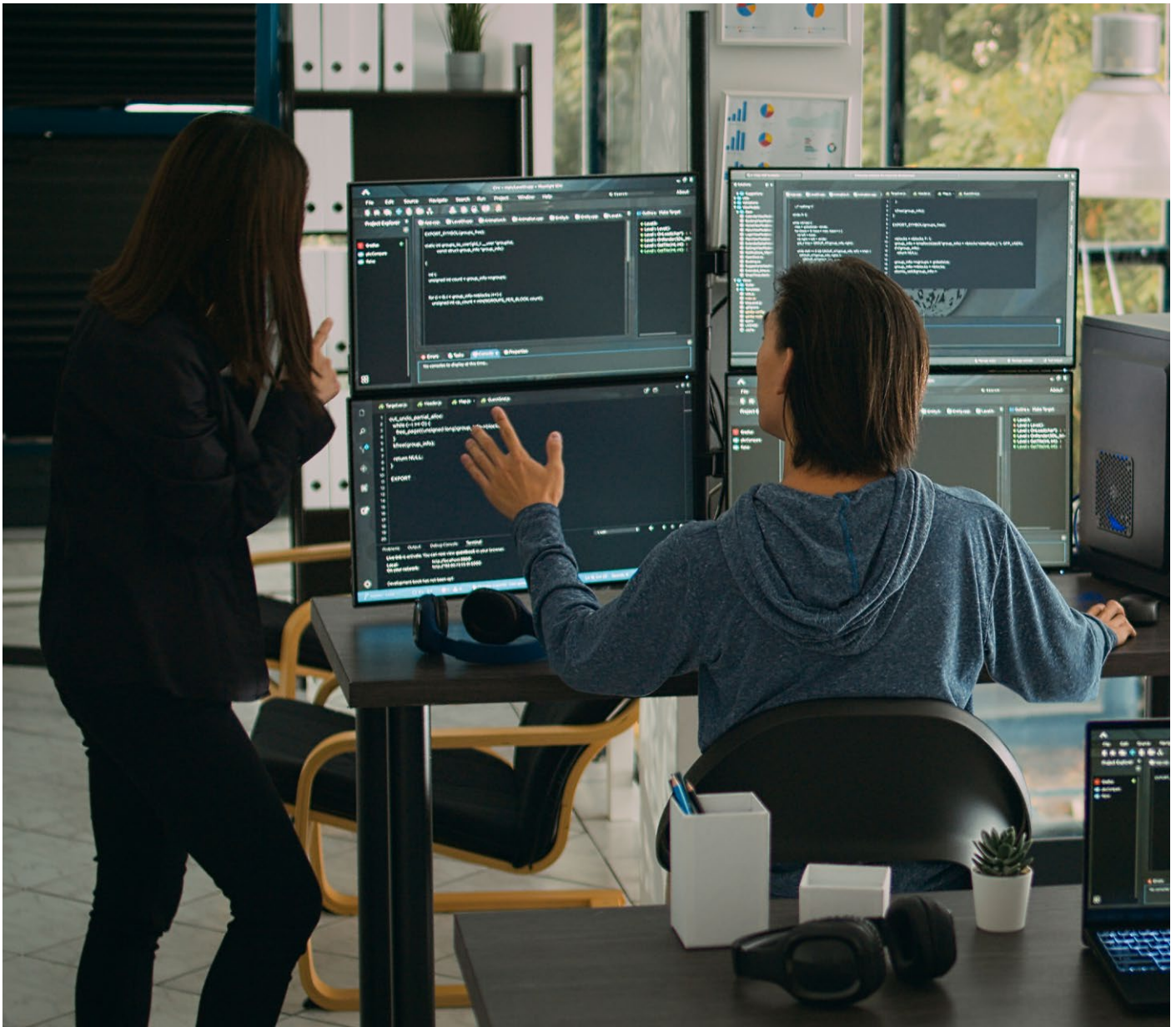
Detailed interviews with public- and private-sector stakeholders on early experiences with agentic AI and use cases have informed the framework in this report. Early implementations show that agentic AI can create tangible value in government operations when applied to clearly defined challenges.

The following examples from the public sector illustrate where agentic approaches are most effective and capture key learnings that enable solutions to progress from experimentation to scalable impact. Private-sector use cases are available on the GovTech Intelligence Hub, presented in a practical, real-world format and curated as an evolving collection to support ongoing cross-sector learning.

Explore all contributing use cases on the GovTech Intelligence Hub

[Organizations contributing use cases:](#)

- German Federal Employment Agency
- German Federal Ministry for Digital Transformation and Government Modernization (BMDS)
- HP
- Ministry of Digital Transformation of Ukraine
- Salesforce
- Scale AI
- ServiceNow
- United Arab Emirates Federal Authority for Human Resources



USE CASE 1

Ukraine: Diia.AI National AI Assistant – Ministry of Digital Transformation of Ukraine

Function category: Citizen communication and interaction

Over the past six years, Ukraine has transformed its public sector, rising from 102nd to fifth globally in the UN Online Services Index. Through the Diia ecosystem, citizens can access digital documents and more than 200 government services via a unified app and portal. Yet as services expanded, many citizens still struggled to identify the right service or navigate procedures efficiently. To address this, Ukraine is moving beyond static digital platforms towards an agentic model in which the state proactively delivers personalized services with minimal friction.

Diia.AI serves as a national AI assistant that enables citizens to access public services through conversational interaction. Unlike traditional chatbots, it not only provides guidance but executes services end-to-end. For example, a citizen can

request an income certificate in natural language, specify the time period and receive the official document within minutes. Human oversight remains central: Diia.AI follows a zero-trust approach, adheres to strict data protection standards and does not grant the model direct access to personal data.

Since its launch in September 2025, Diia.AI has been used by more than 290,000 citizens and issued over 7,000 income certificates. Continuous feedback mechanisms support system improvement, while resilience and bias mitigation were strengthened through adversarial testing. A phased rollout, beginning with the web portal before scaling to the full over 24 million-user app, ensured controlled deployment. In parallel, AI tools within the Diia support team now process 90% of the over one million citizen inquiries, further enhancing responsiveness and efficiency.

USE CASE 2

Germany: AI-based construction permit system – German Federal Ministry for Digital Transformation and Government Modernization

Function category: Public services

Large-scale infrastructure permit procedures in Germany are among the most documentation-intensive processes. Reviews sometimes take longer than the projects they are meant to authorize, delaying investments despite available mandates and funding. This erodes planning reliability, pushes strategic objectives beyond legislative cycles and reduces the effective use of public resources.

To address this, the BMDS developed a solution allowing permit submissions to be screened on the same day they arrive. The system analyses application materials, structures submissions, maps case-specific facts against applicable legal norms and consolidates relevant provisions across interconnected regulatory domains. On this basis, it can generate drafts for legally grounded permit decisions under human oversight to support the review process.

At this stage, the system has limited autonomy, allowing for gradual expansion once confidence and reliability increase. This reflects that scaling is not only a matter of technical feasibility – it also requires people to trust and confidently work with agentic systems.

While agentic approaches to automating legal reasoning are emerging across several fields, scaling to a truly agnostic solution is challenging due to computing capacity, cost and energy constraints. Therefore, rather than scanning the full body of administrative law, the system narrows the legal search to the relevant domains for each case – for example, a wind turbine permit only draws on specific environmental and regional regulations. This targeted approach enables scalable deployment while applying AI capacity only where it genuinely adds value.

USE CASE 3

United Arab Emirates: HR AI Agent – Federal Authority for Human Resources

Function category: Organization and workforce

HR management in large public-sector systems is often slowed by procedural complexity, fragmented information sources and heavy reliance on manual processing. In response, the United Arab Emirates' Federal Authority for Government Human Resources (FAHR) launched the HR AI Agent at the federal level, advancing its broader ambition to embed AI across government operations and reduce internal bureaucracy.

The HR AI Agent introduces a conversational, agentic interface within the federal HR ecosystem, supporting more than 130 digital HR services and automating a wide range of employee self-service transactions. Initial deployment focused on high-frequency, information-centric interactions where efficiency gains were measurable and adjudication risks remained limited (such as policy clarification, service eligibility and procedural guidance). Early results indicate that the system

autonomously resolves over 80% of HR legislation and policy inquiries, significantly improving response speed, consistency and accessibility for more than 50,000 federal employees.

Beyond employee experience, the agent restructures workflows end-to-end, reducing manual intervention and allowing HR professionals to shift towards higher-value activities such as workforce planning, capability development and performance optimization. The initiative underscores the importance of secure architecture, strong data governance and clear human oversight to ensure regulatory compliance and protect sensitive personnel data.

By starting with structured, high-volume use cases and expanding capabilities progressively, the HR AI Agent illustrates a pragmatic pathway towards an AI-enabled government workforce ecosystem that reduces bureaucracy while preserving human accountability and institutional integrity.

USE CASE 4

Germany: Jira ticket creation – German Federal Employment Agency

Function category: Information technology (IT) infrastructure and data governance

The German Federal Employment Agency faces a growing shortage of skilled professionals, particularly in software development. By 2032, an estimated 35% of its workforce is expected to retire or leave the organization. This increases pressure on the agency's ability to manage expanding workloads and sustain the pace of digital transformation.

One operational bottleneck was the manual conversion of requests for change and user stories into Jira tickets. This repetitive task consumed valuable developer time, while existing automation tools lacked the flexibility to handle diverse inputs.

To address this, the agency piloted an agentic AI solution to translate requests for change and user stories into structured,

compliant Jira workflows. The system operates within the agency's secure IT environment and combines autonomous processing with human oversight to ensure accountability and quality.

The pilot showed that ticket generation could be performed reliably and at scale, significantly reducing manual effort and freeing staff to focus on higher-value work. Assuming an average of 20 tickets per day and approximately 20 minutes to process, this solution could save around 150 hours a month for the team involved. The pilot underscored that deploying agentic systems into a secure on-premises architecture requires close coordination across operations, security and development teams, as unaligned priorities can complicate integration efforts. Additionally, managing model updates and handling large input documents requires extra attention but proved manageable with appropriate technical safeguards.

Conclusion

Governments are at a clear inflection point: agentic AI is no longer hypothetical – it is expected to reshape public administration. The central question is not whether to engage with agentic AI, but how to do so responsibly and in ways that produce measurable public value.

A function-level view reveals where automation delivers the greatest impact, regardless of organizational charts. The report presents a clear lesson: progress depends less on pursuing the most sophisticated applications and more on disciplined sequencing. Governments that begin with high-readiness functions and treat pilots as structured learning exercises are better positioned to scale responsibly. By contrast, prioritizing complex, high-stakes applications too early can strain resources and slow broader adoption.

Now, intentional leadership is required. Successful adoption depends less on any single technology than on clear mandates, aligned incentives and the willingness to invest in data, integration and capability building. This report does not prescribe a fixed roadmap. Instead, it provides a shared frame of reference to help governments decide where to start, what to invest in and how to sequence action.

Operationally, adoption should be governed by bounded autonomy – defined agent operating scopes, explicit human escalation mechanisms and transparency around decisions – so officials can deploy systems that are effective and accountable. They should define desired outcomes and constraints upfront, use pilots to test or validate assumptions and expand system scope incrementally as trust and capability grow.

The next step is practical: adapt the framework to local conditions, reassess assumptions against real constraints and move from analysis to carefully scoped action. Continued empirical learning from test deployments and follow-on research will be essential as both the technology and its governance evolve.

Approached with strategic clarity, agentic AI can be more than another wave of digitization. It can underpin a more resilient, responsive and outcome-focused public administration – strengthening public trust and helping governments fulfil their core mission.

Appendices

A1 Methodology

This report employs a structured methodology designed to address governments' most pressing questions and provide actionable guidance for implementing agentic AI. The approach introduces a new analytical lens and methodological approach to 1) identify core government functions as a unit of analysis, 2) define assessment dimensions for the potential and complexity of agentic AI in the public sector and score the functions on their readiness, and 3) create a visual topography with three prioritization areas. The methodology integrates a systematic review of existing frameworks, expert consultation and an iterative development process to ensure robust results.

1 Identification of core government functions:

Review of existing frameworks and research:

A systematic literature review provides a basis for understanding the current state of research and the conceptual foundations of public-sector functions. This review reveals that no existing framework, such as the UN's COFOG or the US Department of Labor's O*NET, sufficiently supports a cross-departmental approach, which is central to this report's methodology.

Development of a tailored function list: To address this gap, the methodology applies an expert-driven and iterative process to identify core government functions. The approach includes consultation with government officials, policy experts and domain specialists across multiple countries to understand which functions are consistently performed across ministries and agencies, as well as analysis of government organizational structures. Selection criteria for functions are based on cross-domain applicability, international relevance, recurring execution and legal mandate or prerequisite for other public services. This strict definition ensures that each of the 70 identified functions across nine categories represents a globally meaningful, assessable unit. The final list is validated through expert review (the complete function list appears in A2 – see Table 2).

2 Definition of the assessment framework to score agentic AI potential and implementation complexity across 70 government functions:

Development of a multidimensional assessment framework: To assess the functions' potential and complexity, a framework with globally applicable dimensions is established, providing an analytical structure for evaluating each function.

The development of these dimensions follows a qualitative, research-informed process, combining insights from the literature review with iterative expert consultations. Dimensions were refined through expert review to ensure each dimension is globally relevant, practically assessable and suited for scoring functions on a global level.

Agentic AI potential determines whether agentic AI can deliver meaningful value for a given function. It includes three sub-criteria:

- **Potential for automation** examines whether the function inherently requires human judgement, discretion or physical presence. Functions with clear rules and structured processes score higher because they can be reliably automated without compromising quality.
- **Agent requirement** focuses on whether agentic capabilities – such as multi-step coordination, contextual reasoning and adaptive decision-making – are essential to the function. If simpler automation, like machine learning, can achieve the desired outcome, the agentic requirement is lower.
- **Volume and impact** assesses how frequently the function occurs and/or how strategically important it is. High-volume functions represent recurring, operationally relevant activities. High-impact functions, even if performed less frequently, play a critical strategic role for government operations.

Implementation complexity measures the practical barriers to deploying agentic AI for a given function, and is split into two sub-dimensions:

1. **Implementation challenges** captures practical barriers to deploying agentic AI in real-world government settings. It includes three sub-criteria:
 - **Data type and quality** evaluates whether the function relies on structured, standardized and accessible data, which facilitates automation or (on fragmented, inconsistent or analogue data) which requires extensive preparation and investment before agentic AI can be applied effectively.

- **Technical integration complexity** examines how easily agentic AI can be integrated. Functions that span multiple systems or agencies face higher integration complexity.
 - **Internal resistance** estimates the likelihood of workforce resistance, which can stem from job displacement fears, cultural barriers or lack of trust in automation.
2. **Risk and ethical impact** covers fairness, transparency and accountability that must be safeguarded in public-sector adoption. It includes three sub-criteria:
- **Ethical threat** considers whether automation could introduce bias, reinforce disparities or compromise fairness. Functions involving decisions that affect vulnerable populations carry higher ethical risk.
 - **Error consequence and function criticality** assesses the severity of potential mistakes. Mistakes in some functions may cause minor administrative delays, while others could lead to irreversible harm, legal liability or safety risks.
 - **Privacy and data issues** evaluates the sensitivity of the information involved. Functions that handle highly sensitive information, such as health or biometric data, present greater complexity and require stringent controls.

Creation of detailed coding scheme: A detailed coding scheme for each sub-dimension is developed to enable a consistent, traceable and transparent assessment of all 70 functions. Each sub-dimension is designed to be scored on a one-to-four scale, with higher values indicating greater potential or complexity and assuming evenly spaced rating intervals.

Expert scoring of all 70 functions: The assessment of all functions is based on an iterative, expert-driven scoring process of the potential and complexity dimensions for each government function (a complete ranking with scores across all dimensions appears in Table 2). A team of practitioners and global public-sector specialists reviewed results to validate and reconcile differing scorings.

Composite scores are calculated as follows:

- **Potential score** = (potential for automation + agent requirement + volume and impact) ÷ 3
- **Complexity score** = (implementation challenges + risk and ethical impact) / 2

The two components are calculated as:

Implementation challenges = (data type and quality + technical integration complexity + internal resistance) / 3

Risk and ethical impact = (ethical threat + error consequence and function criticality + privacy and data issues) / 3

Agentic AI readiness score calculation: For each function, a readiness score is calculated, resulting in one score per function:

- **Readiness score** = potential score - complexity score

3 Deduction of readiness areas to create a visual readiness topography:

Categorization into three readiness areas: The functions are then visualized in a matrix – forming the “readiness topography” (see Figure 5) – and grouped into three agentic AI readiness areas, using threshold values of > 1.5, ≥ 1.0 and < 1.0 for the readiness score. These thresholds were set to reflect expert consensus on meaningful distinctions in operational readiness. The resulting areas are intended to support prioritization and sequencing rather than to imply sharply defined boundaries:

High-readiness area: Functions with readiness scores above 1.5. This threshold indicates a balance of potential and complexity that signals strong readiness for early deployment.

Medium-readiness area: Functions with readiness scores of 1.0 or higher. At this threshold, potential may still outweigh complexity, suggesting that targeted investment and experimentation can be viable.

Low-readiness area: Functions with readiness scores below 1.0. This threshold indicates that continued observation is most appropriate, particularly as contextual factors or implementation barriers evolve.

The scores presented in this report reflect global estimates across governments. These ratings naturally vary by country and organization; this variation is a defining feature of the methodology. The framework provides value by revealing general patterns and strategic priorities that hold across contexts, while also enabling local adaptation to support decision-makers in navigating the challenges of public-sector transformation.

A2 Comprehensive breakdowns of agentic AI assessment scores for all 70 core government functions

TABLE 2 Potential and complexity scores for all 70 core government functions

Category	Function name	Potential	Complexity	
		Agentic AI potential	Implementation challenges	Risk and ethical impact
Security and risk response	Cybersecurity monitoring	4.00	1.00	2.67
Public services	Public information and guidance provision	4.00	1.33	2.33
IT infrastructure and data governance	Systems performance monitoring	3.67	1.67	1.33
Public services	Service appointment and queue management	3.33	1.00	1.67
Security and risk response	Threat intelligence and early warning	4.00	2.33	2.33
Operations and resources	Tender preparation and awarding	3.67	2.00	2.00
Public services	Document validation and processing	3.33	1.33	2.00
Citizen communication and interaction	Transparency reporting and disclosure	3.33	2.00	1.33
Operations and resources	Financial performance monitoring and compliance	3.33	1.33	2.00
IT infrastructure and data governance	Document life cycle management	3.33	1.33	2.00
Public services	Case tracking and orchestration	3.67	2.67	1.67
Public services	Benefit calculation and distribution	3.33	1.33	2.33
Compliance and oversight	Fraud detection and prevention	3.67	1.67	3.00
Public services	Payment processing and reconciliation	3.33	1.33	2.67
Policy execution	Policy implementation monitoring	3.33	2.67	1.33
Organization and workforce	Talent acquisition	3.33	2.00	2.00
Organization and workforce	Project management and progress monitoring	3.33	2.33	1.67
Public services	Authorization, licence issuance and suspension	3.00	1.33	2.00

TABLE 2 Potential and complexity scores for all 70 core government functions (continued)

Category	Function name	Potential	Complexity	
		Agentic AI potential	Implementation challenges	Risk and ethical impact
Citizen communication and interaction	Media and discourse monitoring	3.00	2.00	1.33
Policy planning and strategy	Legislative process support and documentation	3.00	2.00	1.33
Operations and resources	Environmental, social and governance oversight and integration	3.00	2.00	1.33
Organization and workforce	Employee onboarding and offboarding	3.00	1.67	1.67
Citizen communication and interaction	Citizen satisfaction and sentiment analysis	3.00	2.00	1.67
Organization and workforce	Knowledge capture and institutional memory	3.00	2.00	1.67
Citizen communication and interaction	Communication effectiveness monitoring	2.67	1.67	1.33
Citizen communication and interaction	Service interface design and accessibility	2.67	1.67	1.33
IT infrastructure and data governance	IT infrastructure and application provisioning	2.67	1.33	1.67
Operations and resources	Revenue collection and reconciliation	3.33	1.67	3.00
Policy planning and strategy	Public needs identification and analysis	3.33	2.33	2.33
IT infrastructure and data governance	Data quality assurance and validation	3.00	2.33	1.67
IT infrastructure and data governance	Access control and rights management	3.00	1.33	2.67
Citizen communication and interaction	Communication and campaigning design	3.00	2.00	2.00
Public services	Identity verification and authentication	3.00	1.33	2.67
Public services	Application intake and initial processing	3.00	2.00	2.00
Public services	Service quality monitoring and assurance	2.67	2.33	1.00
Public services	Service communication channel operations	3.67	3.00	2.67
IT infrastructure and data governance	Interoperability and data exchange management	3.33	2.67	2.33

TABLE 2 Potential and complexity scores for all 70 core government functions (continued)

Category	Function name	Potential	Complexity	
		Agentic AI potential	Implementation challenges	Risk and ethical impact
Security and risk response	Emergency preparedness monitoring	3.00	2.67	1.67
Public services	Complaint intake and resolution	3.00	2.33	2.00
IT infrastructure and data governance	Legal and regulatory records management	3.00	2.33	2.00
Operations and resources	Contract and supplier management	2.67	2.00	1.67
IT infrastructure and data governance	Archiving and long-term preservation	2.67	1.67	2.00
Policy planning and strategy	Policy forecasting and scenario modelling	3.00	2.33	2.33
Operations and resources	Capital investment planning	3.00	2.33	2.33
Policy planning and strategy	Legal and regulatory framework analysis	2.33	1.67	1.67
Security and risk response	Risk mitigation	3.00	2.67	2.33
Compliance and oversight	Corporate compliance inspection	3.00	2.67	2.33
Citizen communication and interaction	Multilingual content adaptation	2.33	1.67	2.00
Policy planning and strategy	Policy option formulation and drafting	2.33	1.67	2.00
Operations and resources	Infrastructure, asset and facility management	2.33	2.33	1.33
Security and risk response	Cybersecurity incident response and recovery	3.33	3.00	3.00
Security and risk response	Public warning and alert dissemination	2.67	2.00	2.67
Operations and resources	Budget planning and allocation	2.67	2.00	2.67
Organization and workforce	Workforce development and training	2.33	2.00	2.00
Policy planning and strategy	Policy impact prediction	3.00	3.00	2.67
Citizen communication and interaction	Public consultation facilitation	2.67	2.67	2.33
Policy planning and strategy	Multi-level governance coordination	2.67	3.33	1.67

TABLE 2 | Potential and complexity scores for all 70 core government functions (continued)

Category	Function name	Potential	Complexity	
		Agentic AI potential	Implementation challenges	Risk and ethical impact
Organization and workforce	Workforce planning and analytics	2.33	2.00	2.33
IT infrastructure and data governance	Privacy protection and compliance management	2.67	2.00	3.33
Policy execution	Programme implementation planning	2.00	2.67	1.33
Public services	Eligibility assessment and verification	3.00	2.67	3.67
Policy planning and strategy	Stakeholder relationship mapping	2.33	2.67	2.33
Security and risk response	Crisis response coordination	3.00	3.33	3.33
Public services	Inter-agency service orchestration	2.67	3.67	2.33
Policy execution	Inter-governmental policy coordination	2.33	3.33	2.00
Compliance and oversight	Investigation and evidence management	2.67	2.67	3.67
Public services	Grant and subsidy allocation	2.33	2.67	3.33
Policy execution	Policy impact assessment and evaluation	2.33	3.33	2.67
Organization and workforce	Organizational transformation implementation	2.00	3.00	2.33
Organization and workforce	Employee performance management	1.67	3.00	3.33

Note: Sorted by readiness score (final readiness score not included in the table)

Contributors

World Economic Forum

Vidhi Bhatia

Communications Specialist

Claude Dyer

Lead, Digital Governance and Partnership

Leon Martinez

Thematic Lead, Government Technology

Stephan Mergenthaler

Managing Director; Chief Technology Officer

Kelly Ommundsen

Head, Digital Inclusion; Member,
Executive Committee

Global Government Technology Centre Berlin

Manuel Kilian

Managing Director

Capgemini

Timo Graf von Koenigsmarck

Managing Director, Capgemini Invent

Marc Reinhardt

Executive Vice-President, Public Sector Global
Industry Leader

Hannah Schweren

Consultant, Public Sector

Inga Stange

Manager, Connected Products & Services

Acknowledgements

The authors would like to extend their sincere thanks to the experts who contributed their valuable insights and perspectives in interviews and consultations.

Combiz Abdolrahimi

Global Head, Government Affairs & Public Policy,
ServiceNow

Achyut Chandra

GovTech and Open Innovation Lead,
HCLSoftware, HCLTech

Uwe Henkel

Head, Automation Center of Excellence, IT System
House, German Federal Employment Agency

Luukas Ilves

Advisor to the Deputy Prime Minister, Ministry of
Digital Transformation of Ukraine

Dalius Kazlauskas

Chief Technology Officer and Innovation Strategist,
City of Vilnius

Balasubramanian Kalyan Kumar

Chief Product Officer,
HCLSoftware, HCLTech

Aaron Maniam

Fellow of Practice and Director,
Digital Transformation Education, Blavatnik
School of Government, University of Oxford

Philip McGrath

Senior Governance Specialist, World Bank

Simone Parazzoli

Consultant, World Bank

Eldar Sultanow

Enterprise Architect Director, Capgemini

Raj Suri

Vice-President, Global Head of Enterprise AI, HP

Dan Tadross

Vice-President and General Manager,
Public Sector, Scale AI

Mohamed Bin Taliah

Assistant Minister, Cabinet Affairs for Government
Experience Exchange Affairs, Ministry of Cabinet
Affairs, United Arab Emirates

Paul Tatum

Executive Vice-President,
Global Public Sector Solutions, Salesforce

Danylo Tsvok

Chief AI Officer, Ministry of Digital Transformation,
Global Government Technology Centre Kyiv

Florian Tursky

Executive Fellow, Global Government Technology
Centre Berlin

Joao Ricardo Vasconcelos

Senior Governance Specialist, World Bank

The authors also wish to thank Capgemini
colleagues Katharina Reusch, Pavel Romanenko
and Oliver Stuke.

Production**Louis Chaplin**

Editor, Studio Miko

Laurence Denmark

Creative Director, Studio Miko

Jay Kelly

Designer, Studio Miko

Endnotes

1. World Economic Forum. (2025). *The Global Public Impact of GovTech: A \$9.8 Trillion Opportunity*. <https://www.weforum.org/publications/the-global-public-impact-of-govtech/>.
2. Capgemini. (2025). *Data foundations for government From AI ambition to execution*. https://www.capgemini.com/wp-content/uploads/2025/05/Capgemini-Research-Institute-report_Data-foundations-for-government_From-AI-ambition-to-execution-2.pdf.
3. Gartner. (2025) *Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027*. <https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027>. GARTNER is a trademark of Gartner, Inc. and/or its affiliates.
4. United Nations (UN) Statistics Division. (2000). *Classification of the Functions of Government*. <https://unstats.un.org/unsd/classifications/cofog/revision>.
5. US Department of Labor, Employment & Training Administration (ETA). (2025). *O*NET Resource Center*. <https://www.onetcenter.org>.
6. World Bank. (2025). *GovTech Maturity Index 2025: Tracking Public Sector Digital Transformation Worldwide*. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099121725193511608>.
7. Capgemini. (2020). *ESOAR: A unique and straightforward transformation methodology*. <https://www.capgemini.com/solutions/esoar/>.



COMMITTED TO
IMPROVING THE STATE
OF THE WORLD

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744
contact@weforum.org
www.weforum.org