Quick Answer: What Are the Implications of LLM Applications Such as ChatGPT for Government CIOs?

Published 15 May 2023 - ID G00791640 - 8 min read

By Analyst(s): Ben Kaner, Dean Lacheca

Initiatives: Governmentwide Digital Innovation and Application Modernization

ChatGPT, GPT4, Bard and other large language models have entered into the mainstream conversation of government executives wanting to understand the technology, its opportunities and its risks. Government CIOs must be able to address their executives' questions and map a path forward.

More on This Topic

This is part of an in-depth collection of research. See the collection:

- Applying AI — Industries

Quick Answer

What are the Implications of large language model applications such as ChatGPT for government CIOs?

- Large language models (LLMs) will be used to generate value across a wide range of government use cases, both internal and facing citizens. But to achieve this, the platforms will need access to a wide range of government data, much of which may be sensitive.
- As for any emerging technology, significant limitations and risks exist. These must be managed as part of any deployment, including for embedded capability within other software. For government, ensuring the control of sensitive information and maintaining privacy are critical. Balancing the opportunity against these risks will require executives to formulate appropriate policies.
- Use of the growing number of publicly available interfaces by the government workforce, whether from their own or government devices, creates risks around disclosure, exposure of policy thinking, inadvertent bias and use of copyrighted content within government records or reports.

- The capability works both ways: Communication outbound to citizens and inbound from citizens or other parties may be influenced by these platforms as usage grows.
- While technology vendors and providers have been embedding LLM capabilities within their services for some time, their capabilities have grown fast. Government CIOs must understand where they are being incorporated and how they might be used.

More Detail

ChatGPT, by opening up interfaces into its LLM, has stimulated a large amount of interest, development, competition and concern. LLMs use significant quantities of data to create, in essence, a statistically based model of what is likely to be a usable or effective response to a prompt. In essence — "What is likely to be the next word?" While that sounds simple, it requires data harvested across the internet, significant research, and considerable compute power and mediation to achieve. Such models have been around for some time. But by harvesting data that is input into OpenAI's ChatGPT, OpenAI can, in effect, crowdsource additional data with which to train and enhance its model — further accelerating its capability and potentially exposing policy thinking within government. More constrained models are needed to retain control over sensitive data and are rapidly becoming available.

Potential Value for Government Service Delivery and Operations

Appropriately trained LLMs deployed with other automation tooling represent significant potential sources of value to government service delivery and operations. For example:

- **Text generation.** The ability to draft communications in multiple forms to address multiple target audiences — aiming at young audiences, marginal communities or different language communities represents opportunities for new levels of personalization in government services.
- **Text summarization.** Summarizing long or complex related cases for case managers to support improved decision making. Similarly summarizing complex or abstruse documents for laypeople or policymakers could improve productivity.
- **Question answering.** Significantly reducing the time needed to deliver responses to queries — both internal and external. Providing easy-to-understand guidance for employees to ramp them up into new roles, reducing the issues around loss of talent.

- **Text classification.** LLMs enable classification and collation of the large volume of unstructured text, improving the quality of the data used to support decision intelligence and policy development.
- **Sentiment extraction**. Text classification could also be used for sentiment analysis of citizen engagements and communications.

A useful way of considering the challenge is to think of an LLM as being an enthusiastic, highly articulate, but inexperienced graduate intern. The kinds of things you might ask an intern to work on would be any of the above use cases, but no government executive or CIO would assume that the intern's output was correct and to be published externally or used for internal decisions without experienced review.

To extend the analogy, the open versions of ChatGPT would be the equivalent of a work-experience intern who had not signed a nondisclosure agreement. The enterprise versions would be the equivalent of an intern who had signed such agreements — you would remain careful about sensitive information until their discretion had been proven.

Use cases that are likely to be early candidates are experimental early stages in research, the use of appropriately secured systems to summarize information and translating information that might be intended for publication for different audiences — whether to appeal to different language users or simply different communities (young or not, educated or not, marginalized communities, and so on). The intent must be to test where the system works and where the limits lie — and, therefore, develop policies that enable value to be obtained while respecting the strict risk limits, security/confidentiality and legal/regulation/copyright obligations of government.

Limitations and Risks That Currently Exist

As interest rises and the technology matures, delivering value through LLMs relies on gaining and maintaining the trust of the community and the government workforce alike. This requires strong governance and risk management, and a comprehensive understanding and management of the inherent limitations of the technology. There are four major sources of risk:

- **Accuracy.** LLMs are not cognitive models, but statistical ones. As the response is often returned in a manner that appears articulate, sometimes crass errors can be hard to spot. LLMs cannot directly update their model continuously with newly published information. Therefore, the data used to train the model cannot reflect the latest position and facts, thus impacting the accuracy of the result. Conversely, new updates may shift the output of the model, which means that answers may be significantly inconsistent over time.

- **Bias.** Like all machine learning implementations, the data used to train the LLM may be incomplete, poor quality or contain inherent biases. These biases will impede the accuracy of the model's output.

- **Copyright.** The potential for copyright violations exists, as the position of copyright in the data used to train GPT3 and GPT4 models, and in use of the results from them, is as yet unsettled in any major jurisdiction.

- **Privacy.** Publicly available LLMs, such as ChatGPT, are not secure, because they may use your input for further training. [1] They are, therefore, unlikely to meet privacy legislative obligations or community expectations. Government CIOs must take immediate steps to ensure no sensitive or private information is exposed. [2] All employees who use ChatGPT should be instructed to treat the questions they pose or information they post as if they were posting it on a public site.

Implications of External Use of LLMs

As the technologies evolve and are combined with other emerging AI capabilities, the communities that government supports will also use these technologies in both productive and malicious ways. For example:

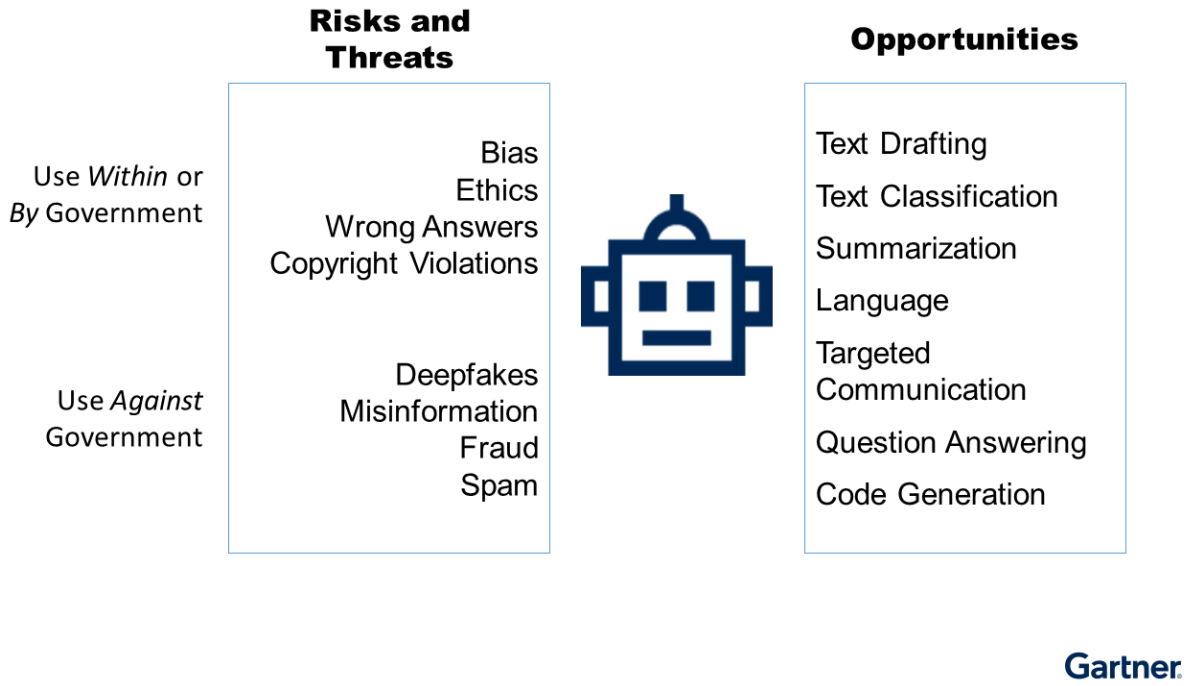
- Citizens with difficulties communicating or interacting with government could use LLMs to operate as an intermediary translating and "simplifying" complex government concepts. If done well, it will improve the inclusiveness of government services. But this also creates the risk of misrepresenting government policy or misleading individuals, leading to frustration within the community.

- Apparently, realistic communications can be created in volume. This may not only be aimed at corrupting the intended information flow from government, but also can overload government administration by flooding staff with queries to which they are legally obliged to respond, such as Freedom of Information requests.

- Using voice and image/video generative AI systems in conjunction with systems such as ChatGPT creates a real risk of deepfake attacks. Consider being faced with an apparently angry <faked> political leader or executive — correct voice, syntax and verbal style — demanding immediate action, such as permitting/licensing a questionable application or transferring government funds.

A high-level view can be seen in Figure 1.

Figure 1. Risks, Threats and Opportunities of Large Language Models



Establishing a Roadmap for LLM in Government

In commercial arenas, it may be possible to "take a view" on some of the risks incurred by using these systems — with potential gains from high-risk innovation outweighing the potential problems, particularly in high-growth areas. In a government, which can be considered a monopoly within its jurisdiction, this is unlikely to be a viable justification for the risk, and a more-measured approach is necessary.

CIOs should advocate for a policy that contains the risk of exposure of sensitive information — at the minimum, constraining the use of the open environment, except for low-impact experimentation, while allowing for a rapid, though measured, exploration of the limits of the technology and where each use case can deliver value that exceeds the residual risk.

Government CIOs involved in setting a roadmap for LLMs within their department should:

- Use ChatGPT as an example to highlight the importance of the organization taking a structured approach toward understanding, utilizing and implementing responsible artificial intelligence practices, perhaps through a center of excellence approach.
- Favor services with enterprise security and compliance controls, such as Azure OpenAI Service's ChatGPT (based on ChatGPT3.5 and ChatGPT4), where data is retained within the customer's infrastructure, over those that don't, such as the open version of OpenAI's ChatGPT.
- Monitor the chosen service's (such as Azure ChatGPT) security practices, once implemented, to ensure that your data is safe, as advertised.
- Ensure that acceptable use guidelines insist that humans should review the output to detect incorrect, misinformed or biased results until the systems are mature enough to be rigorously consistent.
- Implement a policy that prohibits employees from submitting any questions to OpenAI's ChatGPT that may include aspects of confidential or sensitive data.
- Ensure that the policy will allow for safe experimentation using appropriately constrained tools to allow innovation, particularly around the broad productivity enhancements that they may enable.
- Document and plan mitigations for the risks posed to government processes by malign activities exploiting LLM and generative capability, such as administrative flooding or deepfake attacks.

Recommended by the Authors

Quick Answer: What Is ChatGPT?

Tool: Enterprise Use Cases for ChatGPT

ChatGPT and GPT: A Board Reference Presentation

A Comprehensive Guide to Responsible AI

Quick Answer: How Can You Manage Trust, Risk and Security for ChatGPT Usage in Your Enterprise?

Evidence

[1] Quick Answer: How Can You Manage Trust, Risk and Security for ChatGPT Usage in Your Enterprise?

[2] Italy Bans OpenAI's ChatGPT Until OpenAI Makes the Hit A.I. Respect Europe's Privacy Laws, Fortune.