twosix
TECHNOLOGIES

AI BUYING GUIDE FOR GOVERNMENT AGENCIES:

WHAT TO KNOW BEFORE YOU BUY

Cutting Through the Hype:

What to Ask Before You Buy

The impact of artificial intelligence (AI) is hard to ignore. The widespread use of this technology and its success in the private sector have captured the interest of government, with several agencies already reaping tangible benefits. In fact, over 700 federal AI use cases were recently disclosed across major AI and machine learning (ML) categories such as computer vision, natural language processing (NLP), and generative AI. It's clear that AI holds a great deal of potential for the public sector, but federal decision-makers still need to exercise caution. Because AI continues to change, develop, and grow at an exponential rate, many government agencies might feel they're navigating uncharted territory, and the unique challenges of AI are not always evident. As out-lined in the Executive Order (EO) 14110 on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, federal entities still have much work to do when it comes to AI education, staffing talent, and improving procurement processes.

> We face a genuine inflection point in history, one of those moments where the decisions we make in the very near term are going to set the course for the next [several] decades."
>
> – *President Biden, Remarks on the AI Executive Order*

While further guidance from executive departments and offices is pending, there are steps that can be taken now to alleviate stress and confusion when evaluating AI solutions for adoption. Chief among these concerns is the ability to discern real value from marketing hype in a time when so many tools are branded as "AI-enabled," or "powered by machine learning."

Government buyers and end users need to think critically and ask questions to identify legitimate AI tools that can solve pressing problems and use cases. To help separate the wheat from the chaff, we will arm you with the knowledge required to understand the current state of AI for government use, challenges of adopting this technology, and guidance for choosing the right solution.

Questions To Ask an AI Vendor Before You Buy

When vetting vendors and their solutions, you should ask pointed questions to uncover key details about what each tool is designed to do and how AI specifically helps achieve the desired results. We recommend using these essential questions as a foundation when communicating with vendors and preparing your AI evaluation process:

1. Can you share a few examples of practical use cases for your technology?
2. What types of expertise do we need (data science, programming, prompt engineering, etc.) to enable AI and adopt your solution?
3. What data are required for your solution, and who is responsible for providing it?
4. Where is the solution/data hosted, and who is responsible for security?
5. How do you train and test your models for accuracy?
6. Why is AI the best choice for the tasks at hand?
7. How does your solution result in measurable efficiency, quality, or effectiveness gains?
8. How do you ensure that the right data inputs are used when training your models?
9. How can I differentiate the outputs that are solely the result of AI vs. the outputs that have been verified by a human?
10. How do you ensure the AI generates accurate, unbaised information?
11. Is the solution entirely your own intellectual property, or does it leverage third-party technologies (e.g., open source, pass through)?
12. How is your AI/ML resilient against technology changes or outside efforts to mislead your solution?

The following sections offer additional context on why these questions are important and how to identify vendors selling solutions that actually use AI to provide value (vs. those jumping on the bandwagon).

The Reality of AI and ML in the Public Sector

The U.S. government is eager to harness the power of AI and ML for better resource utilization, enhanced decision-making, and a competitive advantage in the global arena. 83% of senior public sector leaders say they are willing and able to adopt intelligent technologies.

However, when answering the call to prioritize AI capabilities, it's important to be realistic about what AI can and can't do. The term "artificial intelligence" often conjures images of highly advanced applications that have been in development for years, like autonomous vehicles and target identification. But there are a multitude of valuable use cases that fly under the radar and deliver a faster time to value.

Looking at the inventory of disclosed government use cases, you will find practical examples of how agencies are using AI to better understand large and disparate datasets, automate manual, time-consuming tasks, and surface valuable insights that can help guide decision-making. These include:

- Document processing to identify sensitive information.
- Converting text from images into a machine-readable format.
- Building data analytics frameworks for process optimization.

Further examination reveals that many of these government projects were carried out under data scientists, machine learning engineers, and AI developers — among other specialized roles — who not only design these tools but also operate them and train others. In the absence of expert oversight, AI tools can run amuck, leading to problems such as:

- Information overload
- Oversimplification of complex tasks
- Bad intelligence presented as fact

Your team, like the 80% of government organizations still at the initial or developing digital maturity stages, may not have AI experts on staff to solve these challenges. It's advisable to seek out a vendor that can help you combine the right data, models, and experts to produce AI-driven outputs you can trust.

ASK A VENDOR:

- Can you share a few examples of practical use cases for your technology?
- What types of expertise do we need (e.g., data science, programming, prompt engineering, etc.) to enable AI and adopt your solution?

Items to Consider When Procuring AI Solutions

Federal AI guidance is iterative (see: the October 2022 Blueprint for an AI Bill of Rights and NIST's January 2023 AI Risk Management Framework [RMF]). Now, with EO 14110, we are seeing a push to supply tangible direction inthe area of AI procurement.

While individual agencies are expected to release more specific guidance in the next year, the Office of Management and Budget (OMB) has published Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence. This draft memorandum — a byproduct of EO 14110 — provides insight into the types of considerations you will need to bear in mind when evaluating AI solutions.

The memorandum's "recommendations for responsible Federal procurement of AI" include:

Aligning to National Values and Law

Ensure that procured AI is consistent with the Constitution and complies with all other applicable laws, regulations, and policies.

Transparency and Performance Improvement

Understand how procured AI produces the desired outcomes and request proof that vendor claims are accurate.

Promoting Competition in Procurement of AI

Take steps, such as emphasizing interoperability and open source tech, so as to not "improperly entrench incumbents."

Maximizing the Value of Data for AI

Treat relevant data and modifications (e.g., cleaning and labeling) as a critical asset. Ensure that contracts include sufficient rights to protect that data from nauthorized disclosure and use.

Responsibly Procuring Generative AI

Consider including tailored risk management conditions in generative AI contracts, such as requiring adequate testing and safeguards to prevent deceptive or unsafe outputs.

Items to Consider When Procuring AI Solutions

Proper data governance for AI solutions is vital, as flawed or mismanaged data can lead to misleading results, failure to support mission needs, and incidents that jeopardize national security.

If you are evaluating a solution that requires you to provide some or all of the data, make sure the vendor has a strong track record of data protection, including robust encryption policies and full transparency about how this information will be transmitted. If the vendor is responsible for sourcing the data inputs for its AI, you should investigate how they choose this data and what steps are taken to maintain data quality over time.

In both cases, you will also want to understand how the vendor uses the data to train their AI models for accuracy and usefulness. There is no guarantee that every model from every contractor is refined to fit your specifically desired use case. Additionally, some solutions marketed as AI may in fact require you to provide your own data models.

You need to partner with a vendor that has deep expertise in your unique domain and is capable of fine-tuning its models accordingly. Confirm that their team has specifically trained the tool for usability, tested it extensively to ensure it produces the intended results, and are confident that the outputs are beneficial and highly actionable.

ASK A VENDOR:

- What data is required for your solution, and who is responsible for providing it?
- Where is the solution and data hosted, and who is responsible for security?
- How do you train and test your models for accuracy?

What’s in a Name?

Confusion regarding what AI actually does and how it works can make it difficult to identify false advertising — many solutions are now marketed as “AI-enabled” or “powered by machine learning” solely to drive sales. Despite the fact that the Federal Trade Commission (FTC) has released several warnings about the consequences of false AI claims, this issue has been difficult to snuff out as new products labeled as AI are released every day.

Fortunately, buzzwords are not as enticing to savvy buyers who understand that adopting AI just for the sake of joining the trend is a losing proposition. The key is to focus on the desired end state instead of fixating on the means. Start with a specific problem or set of problems to be solved, then decide if AI is the right tool for the job. If you adopt a tool with no plans on how you will use it, the old “hammer, nail” adage will take effect.

> “AI is not always the best solution and in many cases is not viable. A common issue with emerging technologies, such as those emerging from the field of AI, is the risk that people start with solutions and then look for problems for the technology to solve.”
>
> – *OECD, Artificial intelligence and its use in the public sector*

Having a firm understanding of different categories of AI and application scenarios is necessary to determine if AI functionality is relevant to solving a given problem. A great deal of government use cases fall under a subset of AI called machine learning, “a branch of computational statistics that focuses on designing algorithms that can automatically and iteratively build analytical models from new data without explicitly programming the solution.”

While there are many different types of AI, some of which overlap, we are going to examine computer vision, natural language processing (NLP), and generative AI as three high-potential groups that have demonstrated value for government. These categories are particularly useful because they help government entities comprehend and make decisions based on the massive amounts of data (text, audio, statistical and visual, etc.) at their disposal.

ASK A VENDOR:

- Why is AI the best choice for the tasks we have at hand?
- How does your solution result in measurable efficiency, quality, or effectiveness gains?

Understanding Computer Vision

Computer vision is a type of machine learning that involves using machines to extract information and derive meaning from digital visual inputs. Instead of using many hours of human bandwidth, computer vision can help you scale image and video evaluation capabilities, in near real time.

Like other categories of machine learning, computer vision requires careful training and testing of its models to achieve accuracy. Factors that influence this include the quality, volume, and cleanliness of data inputs as well as the training process used for the model. Examples of computer vision use cases for government include:

- Fraud detection: Assess large amounts of documents, transactions, and other data points to identify discrepancies and falsified information, such as COVID relief fraud.
- Inspecting electronics: Examine sophisticated hardware, such as printed circuit boards, to review their composition and identify defects or maliciously inserted components.
- Military intelligence: Analyze footage from drones and other aircraft to identify noteworthy or hostile activity.
- Supporting NLP: Together, computer vision and NLP can target inventive new use cases (e.g., using computer vision to digitize paper documents and then NLP to enable smart searches for finding information within the data).

TWO SIX PRODUCT EXAMPLE:

circuitRE is a Reverse engineering automation platform for embedded systems that automates hardware enumeration and evidentiary risk management for supply chain, intellectual property, and security. CircuitRE is built on a number of models that are fine-tuned using curated data inputs, powering an automation platform that makes hardware reverse engineering more scalable and less resource-intensive. It enables use cases such as:

1. AI assists and automates identification and classification of electrical engineering components from photo images.
2. Computed tomography (CT) results are analyzed separating traces of connectivity from components and non-conductive material.
3. Text written on components, silkscreens of the board, or in documents are captured with optical character recognition.

Understanding Natural Language Processing

NLP is a type of machine learning that involves training computers to understand and manipulate human language. This allows you to more effectively surface and manage insights from the unfathomably large volumes of structured and unstructured text and audio at your disposal.

However, just because NLP enables computers to comprehend language doesn't mean that it knows what's happening in the real world or can discern fact from fiction. Without some degree of human oversight, AI systems such as NLP can produce faulty outputs that lead to issues of accountability. In addition to prioritizing models that have been trained for accuracy, you should choose a solution that emphasizes human-centric decision loops for use cases such as:

- Machine translation: Automatically translate text from one language to another, supported by subject matters experts, to conduct information intelligence analysis and understand how foreign actors are influencing the global information environment.
- Text summarization: Turn documents, online content, social media posts, etc. into short, precise summaries that capture key information. \
- Sentiment analysis: Identify trends in digital communications, such as social media posts, to gauge the feelings and views expressed by a certain group or groups.
- Named entity recognition (NER): Detect and categorize important information known as named entities (e.g., events, organizations, people, topics, values, etc.) so that computers can better understand their relationships and manipulate this data.

TWO SIX PRODUCT EXAMPLE:

IKE is an orchestration, automation, and analysis platform for all-domain mission command, control, and decision support. This capability equips teams with actionable insights for planning missions and making decisions in real time. IKE's data analysis tools include NLP technologies such as classifiers, regressors, and clustering algorithms.
IKE's uses of AI/ML includes:

1. Document analysis by performing NER on documents of various types (e.g., text files, PDFs, emails, etc.) and converts extracted entities into IKE-compatible data objects.
2. NLP for specific document upload tasks, such as mission upload, where, instead of creating an IKE mission manually, the user can upload a pdf document containing the mission details and interact with this extracted information.

Generative AI

AIGenerative AI is a type of machine learning that focuses on using large datasets to train a system so that it can create new content (e.g., text, images, video) based on specified inputs. Since the release of ChatGPT in November 2022, generative AI has garnered considerable attention for its potential to boost productivity as well as concerns about training data.

EO 14110 notes that “agencies are discouraged from imposing broad general bans or blocks on agency use of generative AI” and should adopt safeguards to mitigate potential risk. And, while the vast majority of government program executives and IT officials believe that generative AI offers significant benefits, they are not all confident in their ability to implement controls for responsible information generation or to verify that outputs are suitable for decision-making.

Because generative AI is comparatively new, there are few time-tested examples to showcase. That said, budding applications of generative AI for government include:

- Administrate task automation: Use generative AI to increase the speed of routine administrative tasks, such as drafting templates or writing emails.
- Citizen services: Make government information and services more accessible for citizens to interact with online, such as using AI chatbots to respond to common inquiries and help website users find what they’re looking for.
- Information and intelligence discovery: Ask questions and request information (think: ChatGPT) to quickly surface relevant details and insights for a given mission or project.
- Understanding the data domain: Leverage generative AI transformer models to better understand the data at hand, such as finding large objects that exist in satellite images.

TWO SIX PRODUCT EXAMPLE:

PULSE is an AI-powered information advantage platform that collects data, generates insights and enables direct engagement with hard-to-reach audiences around the world. Pulse uses generative AI to help users navigate and interact with the overwhelming amounts of data collected from dozens of social media platforms, social chat, traditional media, deep/dark web, and adtech. In Pulse, GenAI assists users by:

1. Quickly summarizing the most impactful insights across diverse media and social platforms, allowing users to dive deeper into the data driving each insight
2. Integrating pre-engineered prompts into the search experience to simplify interactions with the underlying data with knowledge of the user’s role and Two Six’s mission expertise.

ASK A VENDOR:

- How do you ensure that the right data inputs are used when training your models?
- How can I differentiate the outputs that are solely the result of AI vs. the outputs that have been verified by a human?
- How do you ensure the AI generates accurate and actionable information?

Potential Barriers to AI/ML Adoption

> AI is a long-term data competency grounded in high-quality training quality datasets (TQD) that are the pieces of information and associated labels used to build algorithmic models."
>
> – *U.S. Department of Defense (DoD) Data Strategy*

Federal agencies looking to adopt AI often encounter technical, organizational, and policy-related challenges. Fortunately, you can learn from the trials and tribulations of others to ensure that your team is well-prepared to navigate any obstacles that arise. According to the Federal AI Landscape 2023 report from Govwin, a Deltek company, top challenges of government AI adoption include:

- Poor data quality
- AI model auditability
- Lack of data culture
- Risk in AI responsibility

Data quality is absolutely critical for AI solutions because it directly impacts their accuracy, performance, and reliability. Model auditability is a related challenge of the "inner workings" of these systems, and basically means that stakeholders are able to understand how and why AI produced certain outputs, and can explain these details to outside parties.

A reliable software vendor will ensure that the data and models used for your AI application have been thoroughly tested and hold up against any level of scrutiny. They will also ensure AI Robustness, which is the concept of ensuring that AI can defend itself against attacks both during training and the inference process, as threat actors may wish to alter results of your solution depending on the data you are working with.

Some vendors require you to provide your own model, a significant lift that counteracts the convenience and security of choosing a third-party solution. Hosted models introduce another layer of attack vectors, given the outsourcing of infrastructure, model hosting, and lack of control over the true inputs and outputs of the system. Trustworthy vendors, on the other hand, have AI chain of custody for their models and use vetted, curated, high-quality data intended for a specific use case.

AI risk management is another factor that your vendor should account for, but it is also a shared responsibility and skill set that your agency will need to build independently. While federal entities are still actively working to develop comprehensive guidance in this area, we recommend studying the NIST AI RMF and the NIST AI RMF Playbook for guidance on mitigating the risk of unintended bias, inferences, outcomes, and predictions.

ASK A VENDOR:

- Is the solution entirely your own intellectual property or does it leverage third-party technologies (e.g., open source, pass through)?
- How is your AI/ML resilient against technology changes or outside efforts to mislead your solution?

A Reliable Partner for Sensitive AI Use Cases

Distinguishing tangible value from marketing hype is entirely possible when you embrace a healthy skepticism. Now that you understand why it's important to ask certain questions during the procurement process, you can choose solutions that will actually help you solve problems and achieve key objectives.

At Two Six Technologies, our team of experts has spent years working with and training AI so that it can be safely and effectively used for defense, security, and safety missions. We've successfully integrated this technology into our products, such as the IKE and Pulse, and possess a proven track record of helping federal agencies use AI to achieve the desired outcomes. We also have a tenured relationship with DARPA and are conducting ongoing research in the area of AI and ML assurance.

HAVE FOLLOW-UP QUESTIONS ON HOW AI CAN ASSIST YOUR DEFENSE, SECURITY, AND/OR SAFETY MISSION? REACH OUT TO US TODAY.
Solutions@twosixtech.com