U.S. General Services Administration
FY 2026 Annual Evaluation Plan

August 2025

Introduction

The U.S. General Services Administration (GSA) is committed to its mission of streamlining Federal operations, consolidating resources, and efficiently providing essential services that allow Government agencies to focus on their core missions.

Program evaluation is a key activity to assist GSA with achieving its mission by systematically assessing program or organizational effectiveness and efficiency to answer priority questions and provide rapid feedback for timely, data-driven decisions. The Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act) requires agencies to develop and implement evidence-building plans and evaluations, with the goal of improving the availability of evidence that supports efficient and effective Government programs and policies.

A key statutory requirement of the Evidence Act is an Annual Evaluation Plan (AEP). GSA's draft FY 2026 AEP identifies evaluations that GSA will conduct to provide evidence and recommendations for program and operational decision-making in alignment with Administration priorities.

The GSA Office of Evaluation Sciences (OES) is responsible for developing, coordinating, and implementing the AEP along with other statutory requirements of the Evidence Act. OES works closely with GSA partner program offices to plan and implement evaluations and produce evidence to support strategic decision-making to improve program effectiveness and efficiency.

GSA leadership has identified the following priorities, which guided the development of the AEP:

1. Optimize our Federal Buildings Portfolio
2. Streamline and centralize procurement
3. Rationalize our IT infrastructure and Software as a Shared Service
4. Embrace GSA's model (of efficiency) for ourselves

The evaluation topics below are the key areas for GSA's FY 2026 AEP evaluations and are aligned to the agency priorities.

- **Evaluation Topic 1:** Evaluate the effectiveness of a workspace optimization pilot to increase building occupancy
- **Evaluation Topic 2:** Evaluate procurement protests to minimize avoidable protests and streamline procurement processes
- **Evaluation Topic 3:** Evaluate the use of artificial intelligence (AI) chat tools to increase the value of GSA's AI-powered shared services

Evaluation Topic 1

Evaluate the effectiveness of a workspace optimization pilot to increase building occupancy

Agency Priority

1. Optimize our Federal buildings portfolio

Context

GSA is a champion and innovator of efficient work environments that enable Federal agencies to achieve their missions. GSA's Public Buildings Service (PBS) oversees the nationwide asset management, design, construction, leasing, building management, and disposal of over 360 million square feet of Government-owned and -leased space across the United States and six territories. To achieve this Administration priority, GSA will:

- Eliminate years of accumulated deferred maintenance liabilities,
- Increase office occupancy north of 80 percent, and
- Support greater collaboration and sharing between agencies.

Expected FY 2026 Evaluation

GSA is using technology to optimize Federal workspace. The short-term goal is to trial innovative ways to increase occupancy in Federal offices. The long-term goal is to dynamically match Federal agency workforce needs with real-time space availability, enabling colocation opportunities and reducing underutilized real estate.

GSA will conduct an evaluation to understand the feasibility, efficiency, and relative costs of the initial trial platform(s) and other similar tools deployed to optimize the Federal real estate portfolio and increase occupancy in owned and leased buildings. The optimization pilot is planned for the GSA headquarters building in Washington, D.C. to consolidate employees and desks in new combinations that results in fewer empty buildings and spaces.

Evaluation question(s)

- To what extent do the technologies address agency space needs?
- Is the pilot being implemented as intended?
- To what extent does the pilot meet implementation, adoption, use, and efficiency (time/cost) expectations?

Approach

GSA will conduct a formative evaluation that identifies key evaluation metrics and benchmarks to inform future PBS decisions relating to adoption and scalability.

Data Sources

The following are potential data sources for this evaluation:

- Administrative data from the pilot platform: booking and occupancy data
- Building occupancy data: agency-submitted occupancy counts, counts of laptops in GSA-controlled buildings, key card data
- Employee surveys: satisfaction with products and services
- Interviews with space managers: understanding logistics and workflows to enable use of platforms
- Leasing and operations data: estimation of cost avoidance,% optimization changes implemented within 30 days, reduction in sq. ft. per person, number of share space agreements

Use of findings

The evaluation will provide feedback for learning and improvement during early pilot implementation and provide GSA with timely data that can inform programmatic decisions to increase occupancy across the Federal real estate portfolio.

Evaluation Topic 2

Evaluate procurement protests to minimize avoidable protests and streamline procurement processes

Agency Priority

2. Streamline and centralize procurement

Context

In support of the Administration's Executive Order for "Eliminating Waste and Saving Taxpayer Dollars by Consolidating Procurement," GSA will centralize Government procurement for common goods and services in order to negotiate the best prices for the taxpayer. To achieve this Administration priority, GSA will:

- Maximize the negotiating power of volume buying,
- Streamline the procurement process and reduce the compliance burden to increase competition, and
- Deliver better technology tools for the contracting workforce.

Expected FY 2026 Evaluation

When vendors are unsuccessful at winning a GSA contract or joining a GSA contracting vehicle, they can protest the decision. Reducing vendor protests can save taxpayer money by decreasing agency resources devoted to resolving protests and improve efficiency by speeding up procurement timelines.

GSA will use AI to conduct an evaluation of (1) key characteristics of vendor protests that may help identify potentially avoidable cases and (2) the cost and time savings of using an AI-based tool to automatically collect, aggregate, and summarize protest data for timely, data-driven decisions.

Evaluation question(s)

- What protest characteristics are associated with avoidable protests, and how can processes be improved to reduce costly protests?
    - Which Federal procurement regulations are most commonly cited in protests, and how do they vary between sustained and denied protests?
    - What are the most frequent corrective actions (overall, volunteered preemptively, and prescribed by ruling)?
    - Are there patterns of business types/characteristics and how they approach protesting (e.g., which venue to target first)?

    - How often do incumbent contractors file protests and are there serial protesters?
- What are the time and cost savings of using an AI-based tool to automatically collect, aggregate, and summarize protest data for timely, data-driven decisions?

Approach

GSA will conduct a descriptive evaluation to better understand how avoidable protests originate. The evaluation will use AI to aggregate all non-protected protests filed against the agency over the past 10 years into a dataset and will identify optimal prompts for using AI to summarize each protest. GSA will estimate time and cost savings in using an AI-based tool across multiple systems and venues for a more comprehensive view of the protest process, a process that would be difficult to conduct without AI due to the labor-intensive level of effort required.

Having more timely and more complete data will facilitate timely, data-driven decisions and allow for the development of more targeted interventions to minimize avoidable protests, which will ultimately save GSA time and money.

Data Source

- Documents: AI review of non-protected protest decisions (PDFs) to discover vendor characteristics and other data-driven patterns related to protests. Protest data within documents includes: Grounds cited for filing, protest venue, filing date, phase targeted for protest, decision, withdraw status, dismiss status, legal representation, evidence of debriefing

Use of findings

The evaluation will provide data for understanding characteristics of non-protected vendor protests to help identify potentially avoidable protests. Findings will also be used to improve protest data aggregation and synthesis with limited resources, which will enable targeted interventions to avoid costly protests.

Evaluation Topic 3

Evaluate the use of AI chat tools to increase the value of GSA's AI-powered shared services

Agency Priority

3. Rationalize IT infrastructure and software as a shared service

Context

GSA plays a pivotal role in modernizing and streamlining Federal Government IT infrastructure by serving as a strategic enabler of technological innovation and efficiency. Through initiatives like FedRAMP 2025, centralizing data access, and developing standardized technology solutions, GSA helps Federal agencies adopt secure, cost-effective information technologies and software platforms. To achieve this Administration priority, GSA will:

- Consolidate the number of systems for each job,
- Centralize our data to be accessible across teams,
- Invest in shared services & cybersecurity, and
- Optimize GSA's cloud and software spending.

Expected FY 2026 Evaluation

GSA recently launched a new AI chat tool (USAi) designed to support staff in their daily work, reinforcing the agency's long-standing commitment to technological innovation and secure digital solutions. GSA is actively working toward offering the tool as a shared service to other federal agencies in the near future.

In accordance with the Administration's goal of promoting responsible AI innovation, GSA will track and evaluate performance of our internal tool to ensure its quality and value as a shared service.

Evaluation question(s)

- What barriers exist to responsible AI adoption and application, and how can they be mitigated?
- What is the effect of training and other resources on improving efficient use of AI tools and increasing AI adoption?
- What is the effect of AI tools on employee effectiveness and productivity (e.g., time saved on repetitive or manual tasks)?

Approach

GSA will conduct a formative evaluation to understand early use patterns to identify the most promising use cases and to identify opportunities to increase the adoption of AI tools. An outcome evaluation will be conducted to estimate the benefits of AI tools on employee effectiveness and productivity.

Data Sources

The following are potential data sources for this evaluation:

- Interviews: Targeted employee interviews will allow for an in-depth understanding of employees' experience using the tool and its potential value for workflows and products.
- Surveys: Surveys of early users will help allow for both broader coverage than interviews and will allow for asking specific questions that cannot be supported through administrative data (e.g., estimating productivity in work products). The surveys will contain a mix of quantitative and qualitative items on topics such as task use, barriers to use, training and support needs, confidence in using AI effectively, and perceived benefits of using AI in work-related tasks.
    - While self-reported responses have limitations (e.g., social desirability bias) they will provide valuable preliminary data given the limited information about how employees use the AI tools. The results will guide development of future activities to provide additional insight into potential changes in productivity and will help to triangulate among different data sources.
- Administrative data: Data containing prompts will inform general usage patterns across the entire organization and will allow for analysis of the topics and tasks people are using the tools for.

Use of findings

Since many features of the tools are still in development, timely evidence from the formative evaluation can inform prioritization of new features and resources that will result in more effective usage of AI chat tools.

The outcome evaluation will provide evidence about the effectiveness of specific features and resources, which can inform development decisions that enable real-world productivity gains and provide initial evidence of the value of GSA’s AI-powered shared services.