
GS Federal AI Procurement Safeguards Index

CODEX-REVIEWED RELEASE

31.5 / 100

Equal-weight mean of ten reviewed acquisition safeguard estimates

Study period: January 1, 2024 - August 28, 2026

GS Consulting

Executive summary

GS Consulting analyzed 1,532 candidate federal AI solicitation notices and their available public packages. Across ten equally weighted acquisition safeguards, the reviewed GS Federal AI Procurement Safeguards Index is **31.5 out of 100**.

The most frequently documented safeguard was **Performance Metrics**, estimated in 59.7% of scoreable public packages (95% CI 50.8%-68.5%). The least frequently documented was **High Impact Disclosure**, estimated at 3.3% (95% CI 0.0%-6.9%).

Reviewed category estimates

The index is the equal-weight mean of the ten category prevalence estimates. Category estimates use the stratified Codex semantic review sample; confidence intervals include finite-population corrections.

Labels were produced by one versioned Codex semantic review over public notice text and extracted packages at the user's direction. This release does not claim independent human review, inter-rater agreement, or Cohen's kappa.

	Prevalence	95% CI	Score	Weighted Score
Performance Metrics	59.7%	50.8%-68.5%	98.0%	55.7%
Data And Ip Rights	43.3%	34.0%-52.5%	86.0%	52.3%
Portability And Exit	41.9%	31.3%-52.5%	96.2%	5.5%
Pricing Transparency	38.8%	27.5%-50.0%	90.0%	27.2%
Test And Evaluation	38.8%	27.2%-50.3%	78.0%	21.9%
Interoperability	32.9%	26.1%-39.7%	88.0%	73.6%
Ongoing Monitoring	25.9%	17.8%-34.0%	96.0%	53.8%
Sunset And Closeout	23.5%	15.8%-31.1%	96.0%	55.9%
Feature Change Notification	7.2%	2.3%-12.0%	100.0%	1.3%
High Impact Disclosure	3.3%	0.0%-6.9%	18.0%	44.8%

Descriptive comparisons

Subgroup comparisons use the deterministic retrieval rules and are not substituted for the reviewed population estimates above. Among agencies with at least ten candidate notices, **NATIONAL AERONAUTICS AND SPACE ADMINISTRATION** had the highest average raw rule-detection index (16.7 of 100; n=54). This is descriptive and may reflect procurement mix, package availability, or rule performance.

The largest absolute pre/post rule-detection change was for **Performance Metrics**: 27.8% before M-25-22 and 20.7% on or after April 3, 2025 (-7.1 percentage points). This comparison is not a causal estimate.

Coverage, validation, and awards

Public text was scoreable for 1,076 notices; 456 lacked scoreable public text. The AI classifier validation estimated precision at 95.6% (95% CI 92.6%-98.7%) and recall at 96.3% (95% CI 93.2%- 99.5%).

All 310 grouped award candidates were adjudicated; 307 groups covering 277 unique notices were accepted. Award values are not summed across multi-award vehicles, ceilings, or modifications without a separate aggregation rule.

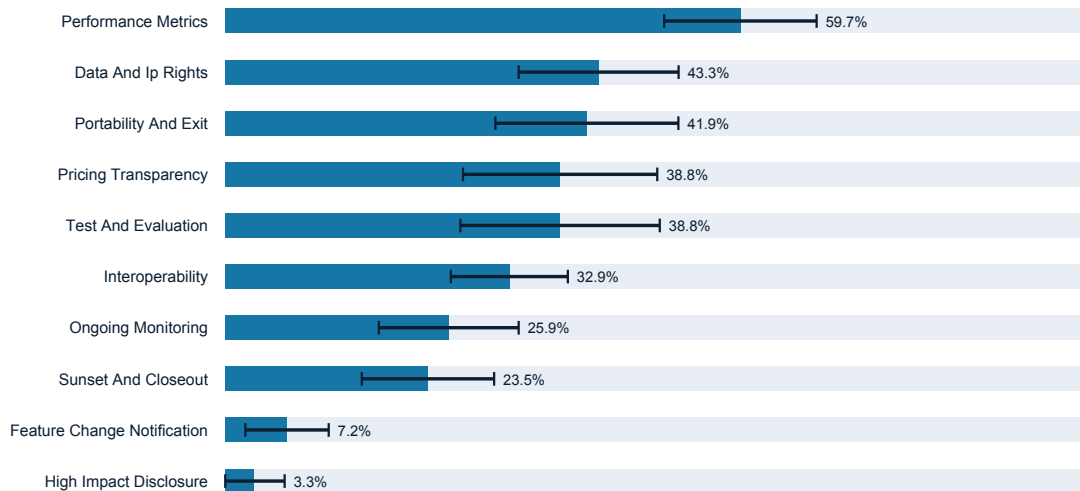
Interpretation and limitations

The study measures language observable in public solicitation packages. A non-match does not prove that an agency lacks a safeguard internally. Public packages may omit clauses stored in restricted systems or incorporated elsewhere. Agency and policy-period comparisons are descriptive. See [methodology.md](#) and [data-dictionary.md](#).

Figures

Codex-reviewed prevalence of AI procurement safeguards

Weighted estimates among scoreable public packages; lines show 95% confidence intervals



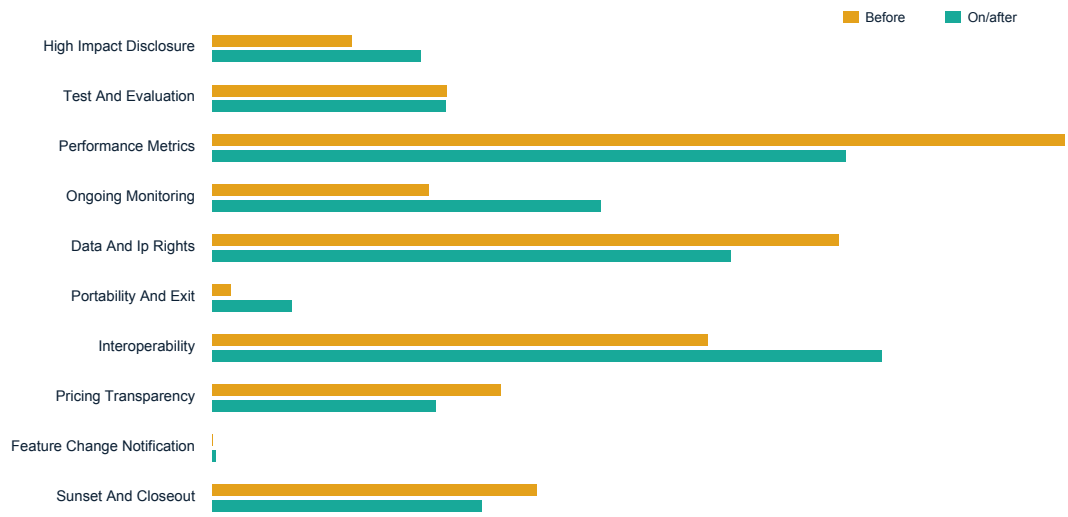
Source: SAM.gov public packages; stratified Codex semantic review

GS Federal AI Procurement Safeguards Index

Figure 1. 01 Overall Safeguards

Public safeguard language before and after M-25-22

Deterministic rule detection; April 3, 2025 cutoff — descriptive, not causal



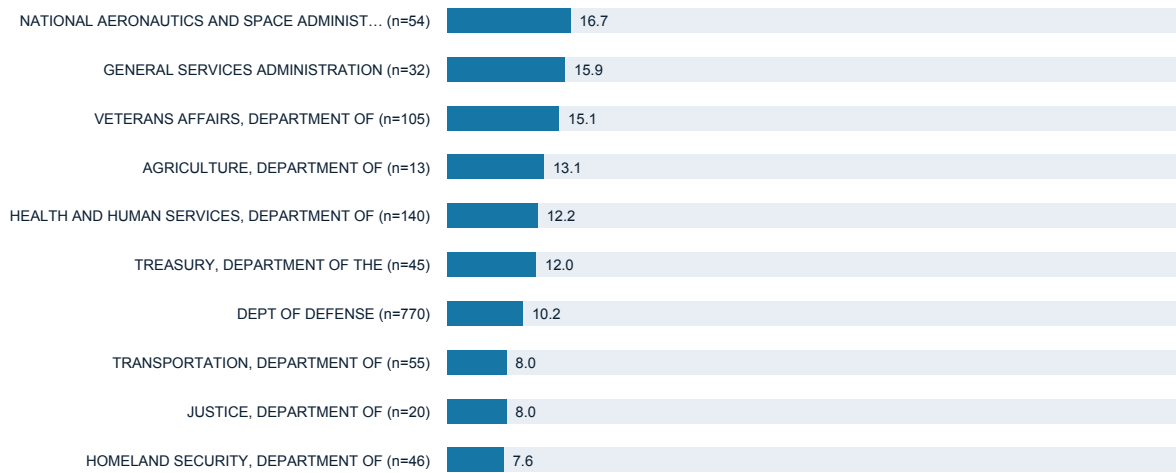
Source: SAM.gov; M-25-22 issued April 3, 2025

GS Federal AI Procurement Safeguards Index

Figure 2. 02 Before After M 25 22

Average raw safeguards index by agency

Agencies with at least 10 candidate notices; raw index ranges from 0 to 100



Source: SAM.gov; agency mix and missing public text vary

GS Federal AI Procurement Safeguards Index

Figure 3. 03 Agency Index

Distribution of safeguards found per notice

Count of deterministic category matches across the 1,532-notice candidate frame

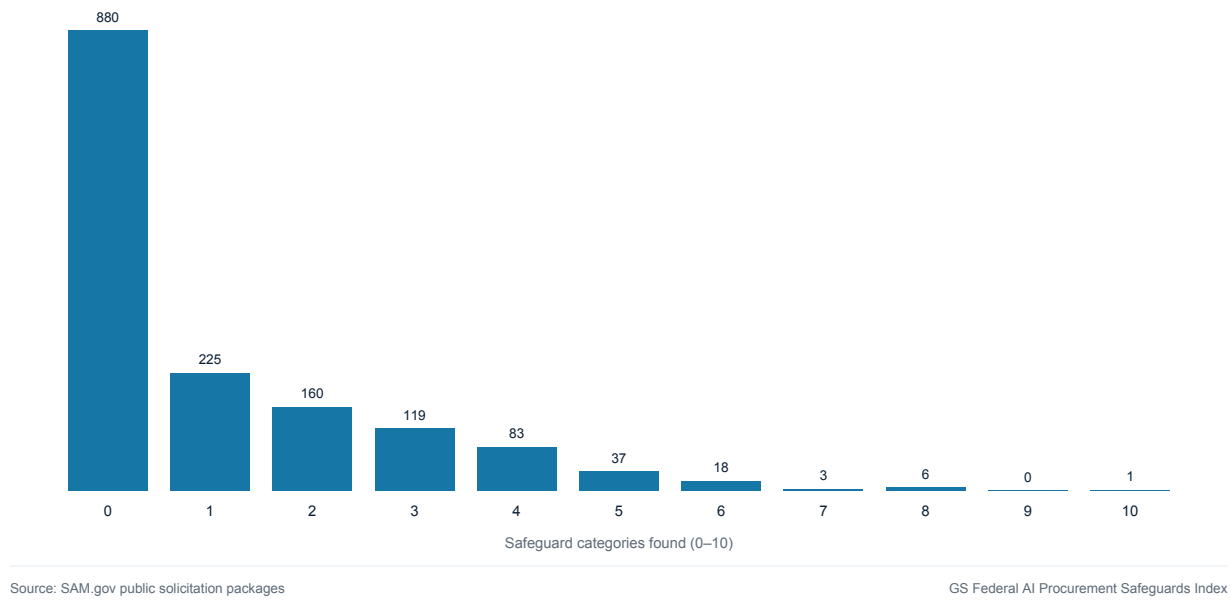
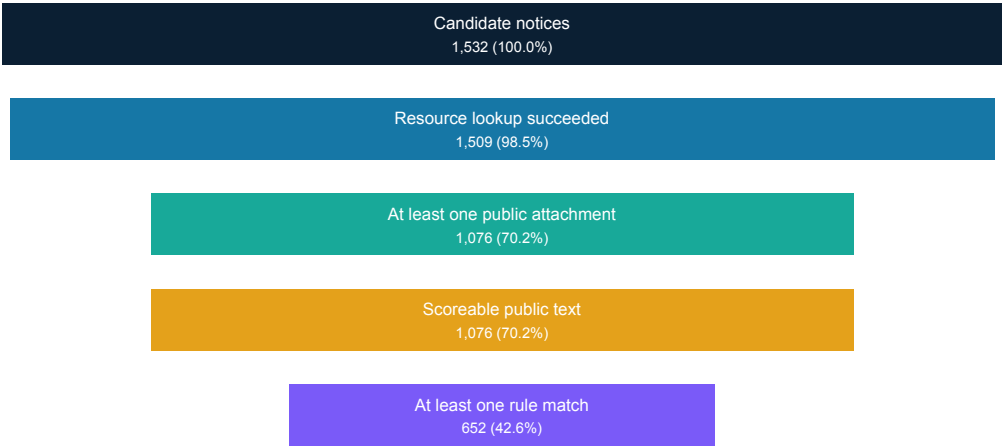


Figure 4. 04 Safeguard Count Distribution

From candidate notice to scoreable public package

Acquisition and observability funnel



Source: GS Consulting acquisition manifest; SAM.gov

GS Federal AI Procurement Safeguards Index

Figure 5. 05 Public Package Coverage