Fiscal Patriots Dashboard

Follow the Money

Trust The Recipient

By: Khaled Alkurd, Nikita Chandrasing, Mariam Debas, Pranavi Doodala, Andy Yaro

George Mason University Costello College of Business

Meet The Fiscal Patriots

Khaled Alkurd
Team Lead
- Forensic Accounting
- Accounting Analytics
- Python • SQL • R
- Alteryx • Tableau

Pranavi Doodala
ML Lead
- Business Analytics
- Data Science
- Python • R
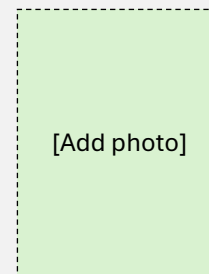

Mariam Debas
Visualization Lead
- Accounting
- Data Analytics
- Tableau • R

Nikita Chandrasing
Research Lead
- Business Analytics
- Python
- R
- SQL

[Add photo]

Andy Yaro
Documentation Lead
- Cybersecurity Engineering
- Python
- Git / Github
- Technical Writing

Government Data is Powerful

The Data

We Connected the Dots

- 74M USAspending records processed

- 57,448 audited entities analyzed

- 3 federal systems integrated

Citizens can finally see where tax dollars go

The Problem

Trust Can't Be Verified

- 3 systems, 1 identifier, 0 integration

- Audit failures don't reduce funding

- $1.09T to Entities with Severe Audit Failures

- 33,780 excluded entities traceable, none auto-blocked

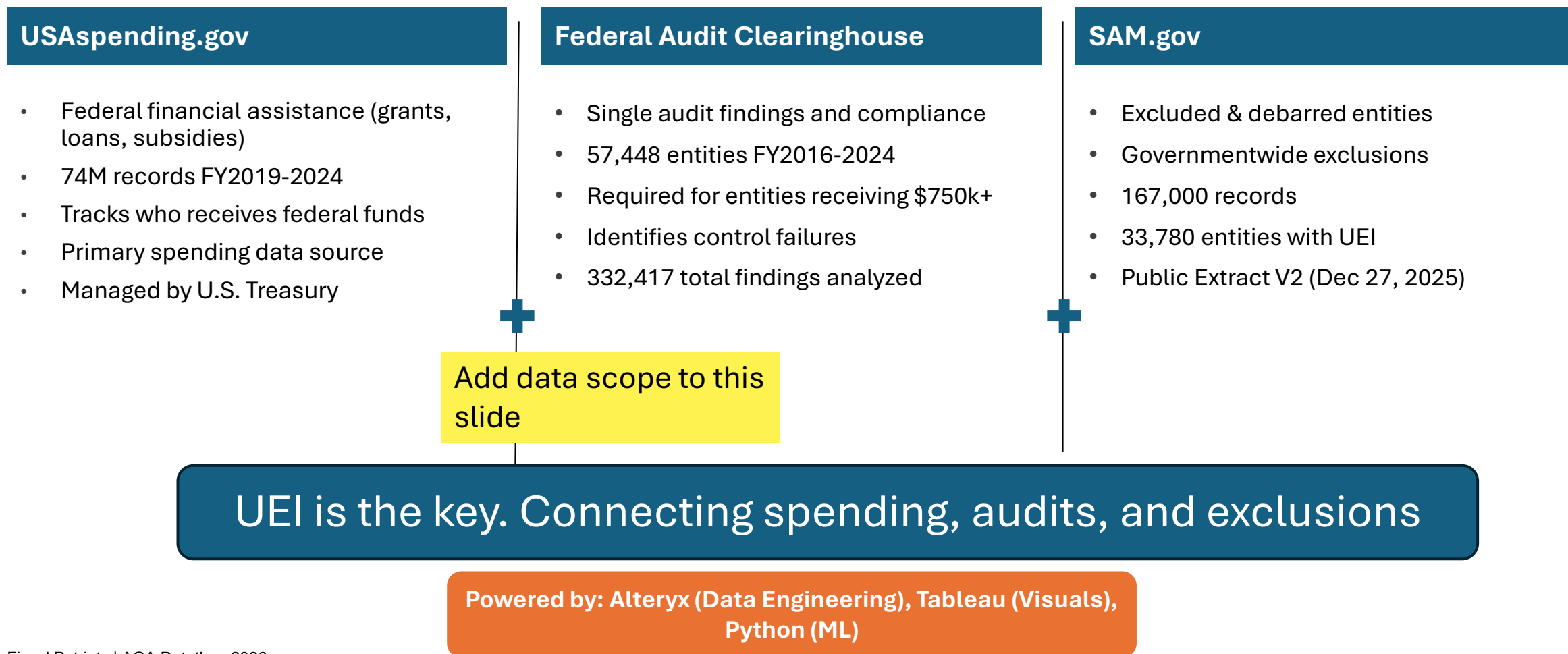**But they can't verify which entities to trust**

The Solution

We Built the Tools

- 42,233 entities matches & scored

- Taxpayer Protection Score (0-100)

- Predictive risk model

Until now

Source: USAspending.gov (FY2019-2024), Federal Audit Clearinghouse (FY2016-Present), SAM.gov Public Extract 25361 (Dec 27, 2025)
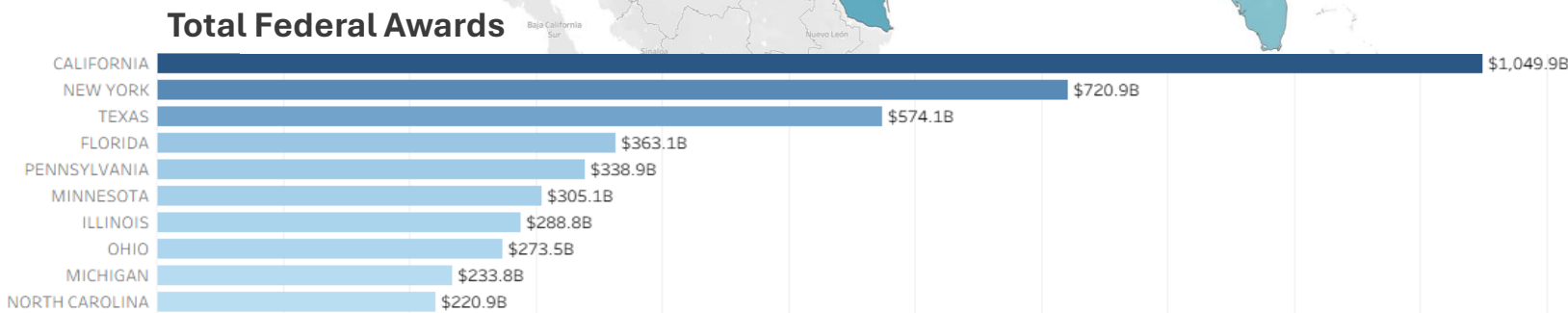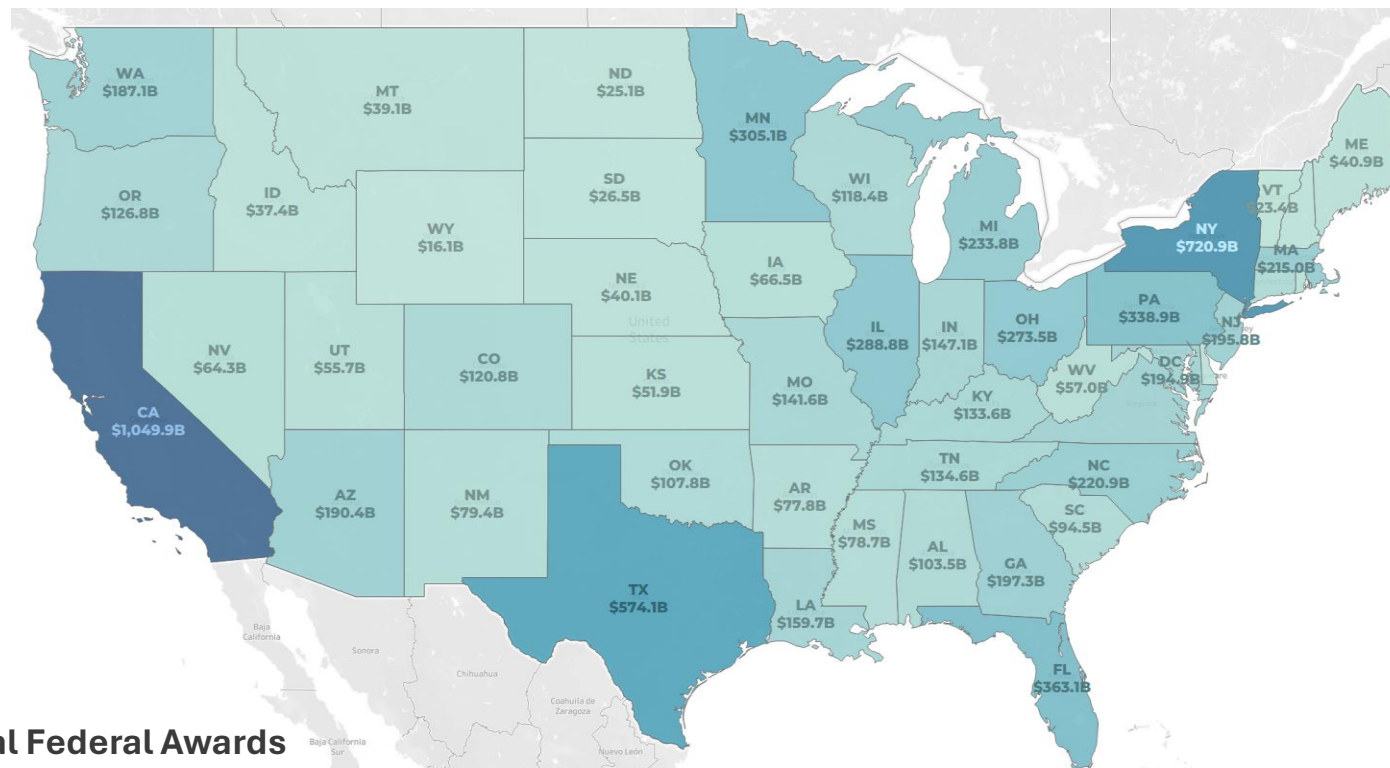
A Unified View of Federal Accountability

Three federal systems connected by one identifier. Separately, open data. Together, accountability.

USAspending.gov

- Federal financial assistance (grants, loans, subsidies)
- 74M records FY2019-2024
- Tracks who receives federal funds
- Primary spending data source
- Managed by U.S. Treasury

Federal Audit Clearinghouse

- Single audit findings and compliance
- 57,448 entities FY2016-2024
- Required for entities receiving $750k+
- Identifies control failures
- 332,417 total findings analyzed

SAM.gov

- Excluded & debarred entities
- Governmentwide exclusions
- 167,000 records
- 33,780 entities with UEI
- Public Extract V2 (Dec 27, 2025)

Add data scope to this slide

UEI is the key. Connecting spending, audits, and exclusions

Powered by: Alteryx (Data Engineering), Tableau (Visuals), Python (ML)

Fiscal Patriots | AGA Datathon 2026

Source: USAspending.gov (FY2019-2024), Federal Audit Clearinghouse (FY2016-2024), SAM.gov Public Extract 25361 (Dec 27, 2025)

4

Where Does Federal Funding Go?

Half of all federal assistance awards flow to just 10 states. This is where the money lands.

Total Federal Awards

State	Amount
CALIFORNIA	$1,049.9B
NEW YORK	$720.9B
TEXAS	$574.1B
FLORIDA	$363.1B
PENNSYLVANIA	$338.9B
MINNESOTA	$305.1B
ILLINOIS	$288.8B
OHIO	$273.5B
MICHIGAN	$233.8B
NORTH CAROLINA	$220.9B

Fiscal Patriots | AGA Datathon 2026

Source: USAspending.gov (FY2019-2024)

Funding Snapshot

$8.58 Trillion
Federal Financial Assistance
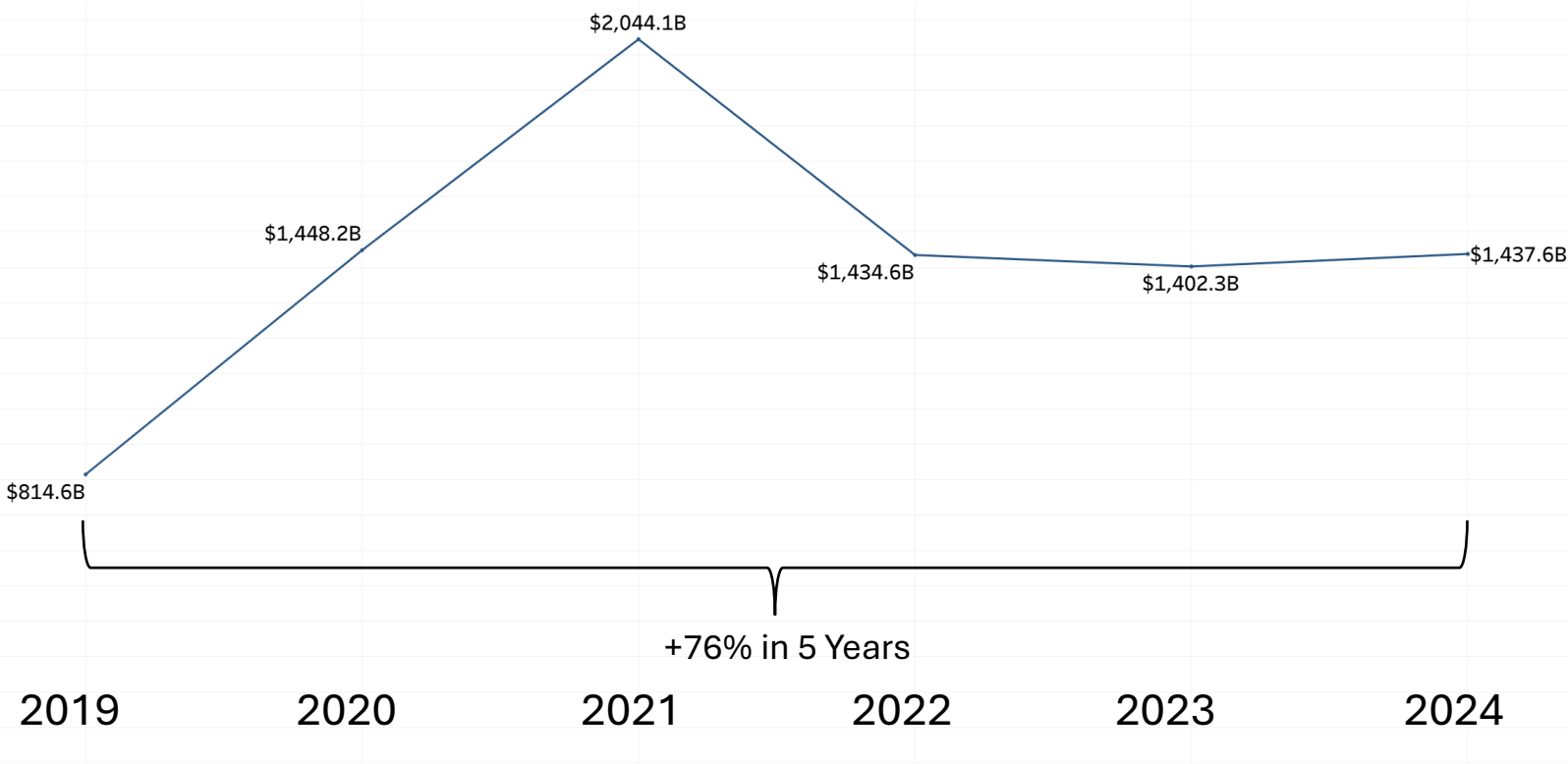
$4.37T to Top 10 States (51%)

$1.05T to California Alone (12%)

181,922 Recipients Tracked

You can see where the money goes. But can you trust who receives it?

5

Federal Financial Assistance Over Time

COVID drove funding to record highs. The baseline nearly doubled. Did accountability keep pace?

$2,044.1B

$1,448.2B

$1,434.6B

$1,402.3B

$1,437.6B

$814.6B

+76% in 5 Years

2019   2020   2021   2022   2023   2024

Trend Snapshot
$2.04 Trillion FY2021 Peak (COVID)
$1.4 Trillion Stabilized by FY2022
3 Years Flat FY2022-2024
$8.58T Total Over 6 Years

Spending up 76%. Oversight unchanged.

How Is Funding Concentrated?

Just 10 agencies are responsible for 92% of all federal financial assistance to recipients nationwide.

Top 10 Agencies by Federal Funding (FY2019-2024)

Agency	Funding
Department of Health and Human Services	$3,578.7B
Department of Agriculture	$2,160.5B
Department of Transportation	$641.5B
Department of Housing and Urban Development	$482.7B
Department of Homeland Security	$261.6B
Department of Education	$249.5B
Department of the Treasury	$170.3B
Agency for International Development	$155.4B
Department of Justice	$110.0B
Department of the Interior	$101.6B

33 Other Agencies: 8% ($670B)

Top 10 Agencies: 92% ($7.9T)
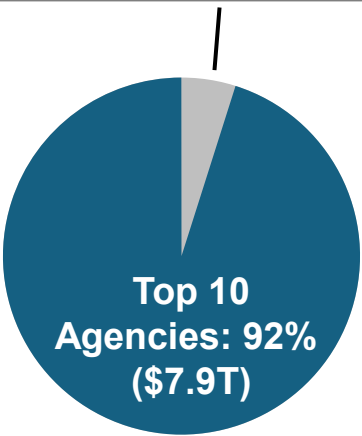
Concentration Snapshot

$3.58 Trillion to HHS Alone (42%) (Primarily Medicare/Medicaid)

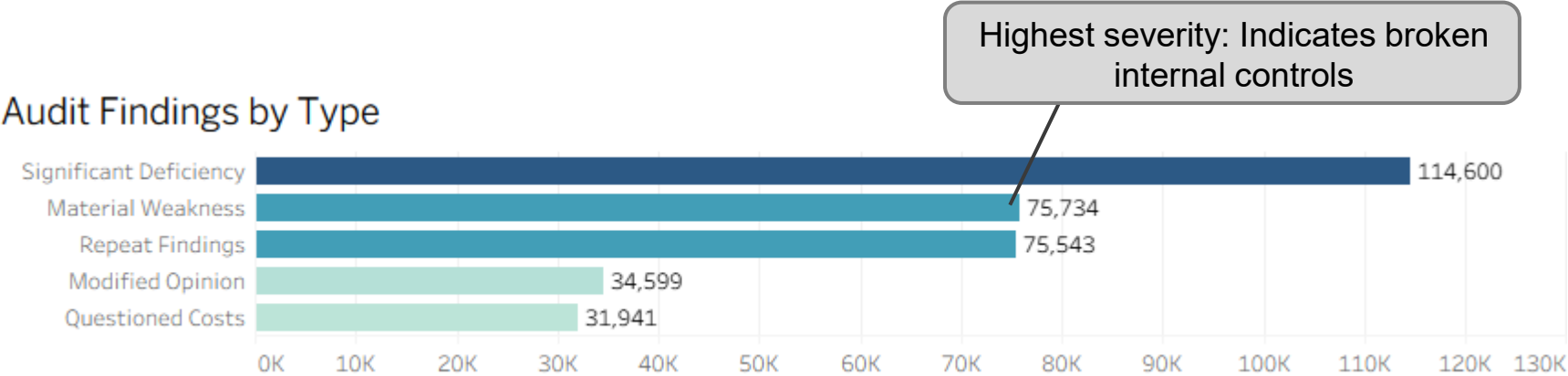$5.74T to Top 2 Agencies (67%)

$7.9T to Top 10 Agencies (92%)

$670B to 33 Other Agencies (8%)

**67% of federal dollars flow through just two agencies.
Who's watching where it goes?**

What the Audits Reveal

Cumulative Audit Findings (FY2016-2024) reveal a chronic pattern of over 332,417 findings of control failures and financial mismanagement.

Audit Findings by Type

Finding Type	Count
Significant Deficiency	114,600
Material Weakness	75,734
Repeat Findings	75,543
Modified Opinion	34,599
Questioned Costs	31,941

Highest severity: Indicates broken internal controls

Federal Audit Snapshot

57,448 Entities
Audited

332,417 Total Findings FY2016-2024

Average of 5.8 Findings per Entity

75,734 Material Weaknesses
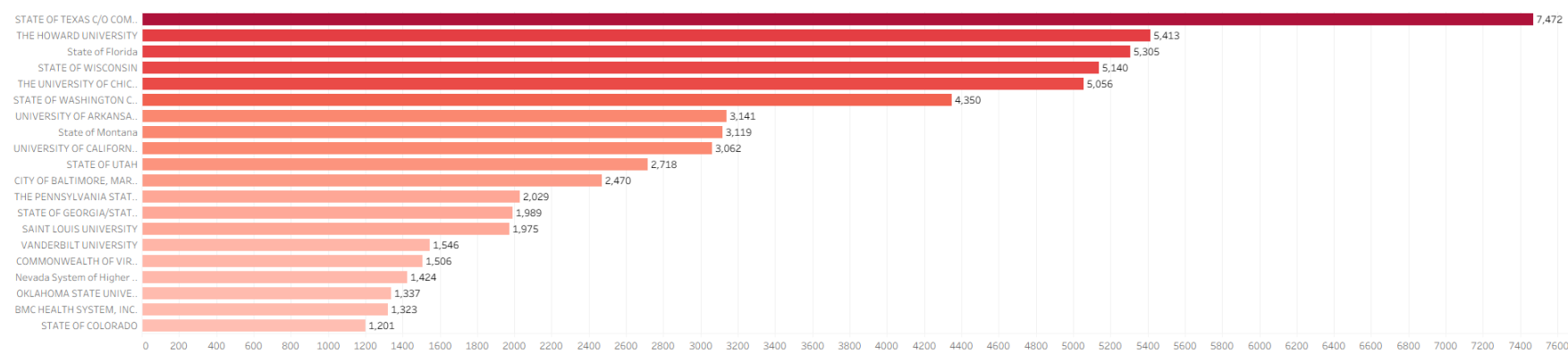
What These Findings Mean

Significant Deficiency	Material Weakness	Repeat Findings	Modified Opinion	Questioned Costs
• Notable control weakness	• Severe control failure	• Same issue year after year	• Auditor couldn't fully approve	• Spending flagged as improper

Audits surface the problems.
But how do we measure the risk?

Where Risk Concentrates

A small number of entities drive most audit failures.

Top 20 Entities by Audit Findings

Entity	Findings
STATE OF TEXAS C/O COM..	7,472
THE HOWARD UNIVERSITY	5,413
State of Florida	5,305
STATE OF WISCONSIN	5,140
THE UNIVERSITY OF CHIC..	5,056
STATE OF WASHINGTON C..	4,350
UNIVERSITY OF ARKANSA..	3,141
State of Montana	3,119
UNIVERSITY OF CALIFORN..	3,062
STATE OF UTAH	2,718
CITY OF BALTIMORE, MAR..	2,470
THE PENNSYLVANIA STAT..	2,029
STATE OF GEORGIA/STAT..	1,989
SAINT LOUIS UNIVERSITY	1,975
VANDERBILT UNIVERSITY	1,546
COMMONWEALTH OF VIR..	1,506
Nevada System of Higher ..	1,424
OKLAHOMA STATE UNIVE..	1,337
BMC HEALTH SYSTEM, INC.	1,323
STATE OF COLORADO	1,201

Concentration Snapshot

332,417 Findings

16,334 Entities with findings

Top 20 entities → 13% of findings

Top 20 average: 2,239 findings

All others average: 18 findings

> 20 entities. 13% of failures.
> 124x the average.

Fiscal Patriots | AGA Datathon 2026

Source: Federal Audit Clearinghouse (FY2016-2024)

The Audit Health Score

Turning complex audit data into a simple 0-100 score. Making federal accountability accessible to everyone.
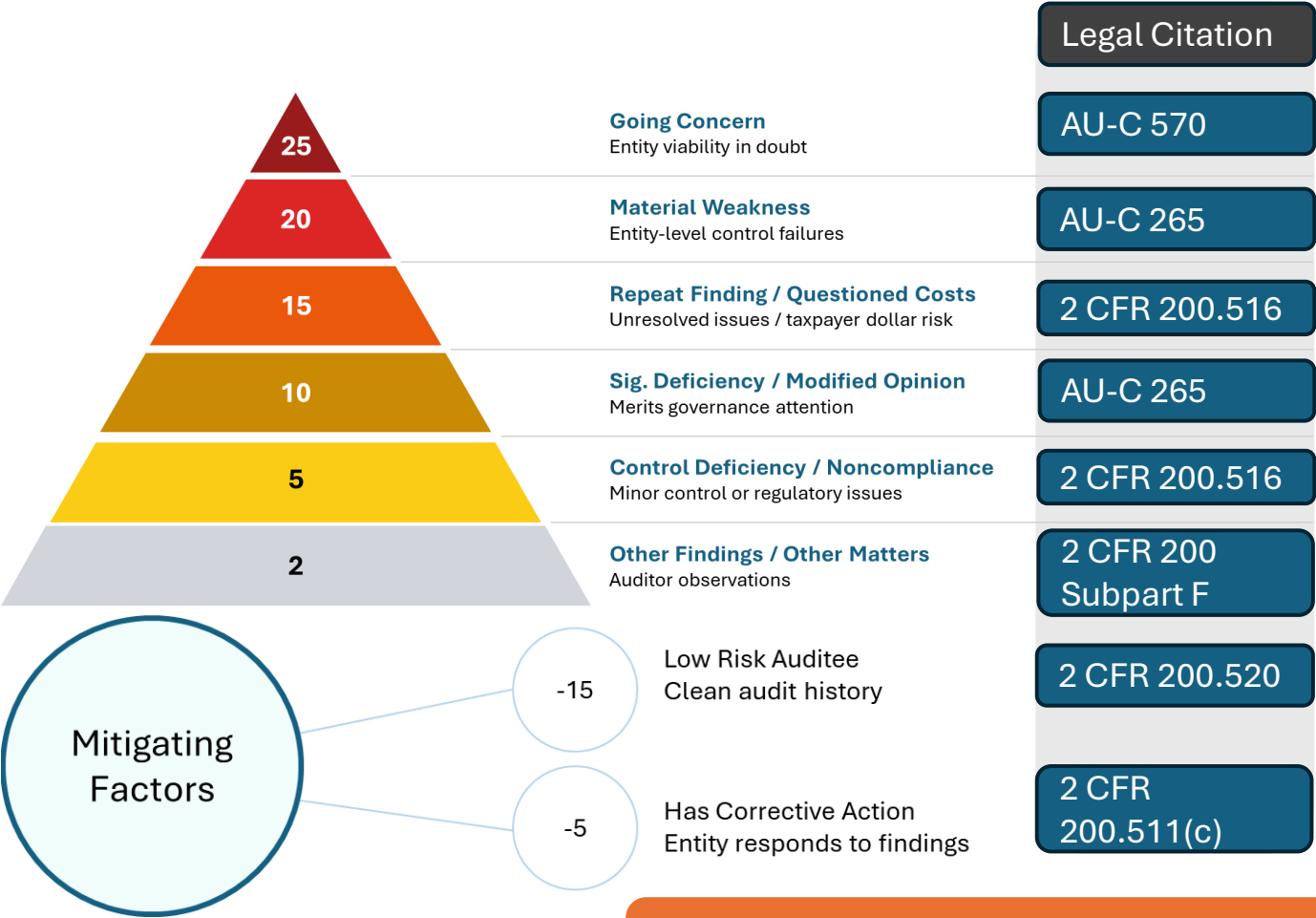
Score = 100 - Risk Points (capped 0-100)

Assessment parameter	Risk Factor Rationale
Key Risk Factors	• Going Concern (+25) | Auditor doubts entity can continue operating. • Material Weakness (+20) | Severe internal control failure. • Repeat Findings (+15) | Same problem year after year, the entity isn't improving. • Questioned Costs (+15) | Taxpayer dollars flagged as being used improperly. • 11 total risk factors weighted by severity (see appendix)
Mitigating Factors	• Low Risk Auditee (-15) | Clean audit history, proven track record. • Corrective Action Plan (-5) | Entity responded to findings with an action plan.
Risk Tiers	🔴 High Risk (0-59). Multiple audit failures 🟡 Moderate Risk (60-79). Some concerns identified 🟢 Low Risk (80-100). Minimal findings, trusted

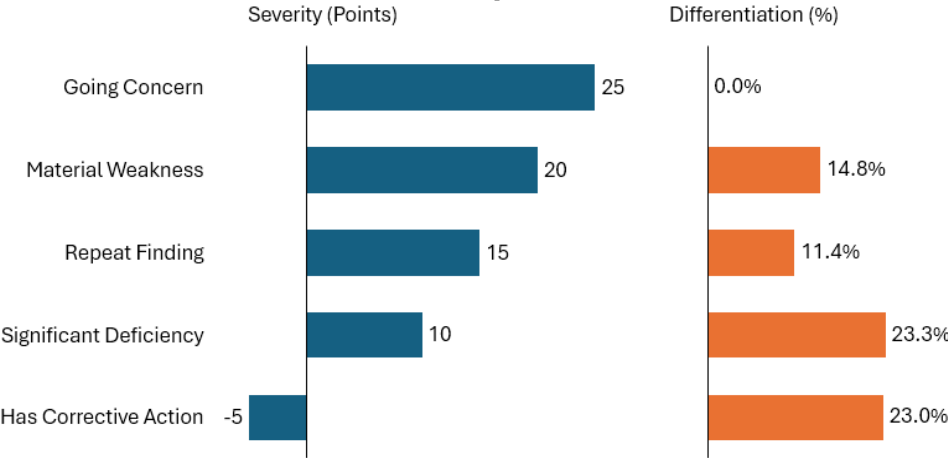Now let's apply this score to real federal funding data.

The Science Behind the Score

Federal audit standards define severity. Our model proves they predict risk.

Score = 100 - Risk Points (capped 0-100)

	Legal Citation
Going Concern — Entity viability in doubt	AU-C 570
Material Weakness — Entity-level control failures	AU-C 265
Repeat Finding / Questioned Costs — Unresolved issues / taxpayer dollar risk	2 CFR 200.516
Sig. Deficiency / Modified Opinion — Merits governance attention	AU-C 265
Control Deficiency / Noncompliance — Minor control or regulatory issues	2 CFR 200.516
Other Findings / Other Matters — Auditor observations	2 CFR 200 Subpart F
Low Risk Auditee — Clean audit history	2 CFR 200.520
Has Corrective Action — Entity responds to findings	2 CFR 200.511(c)

Pyramid values: 25, 20, 15, 10, 5, 2

Mitigating Factors: -15, -5

Random Forest feature importance

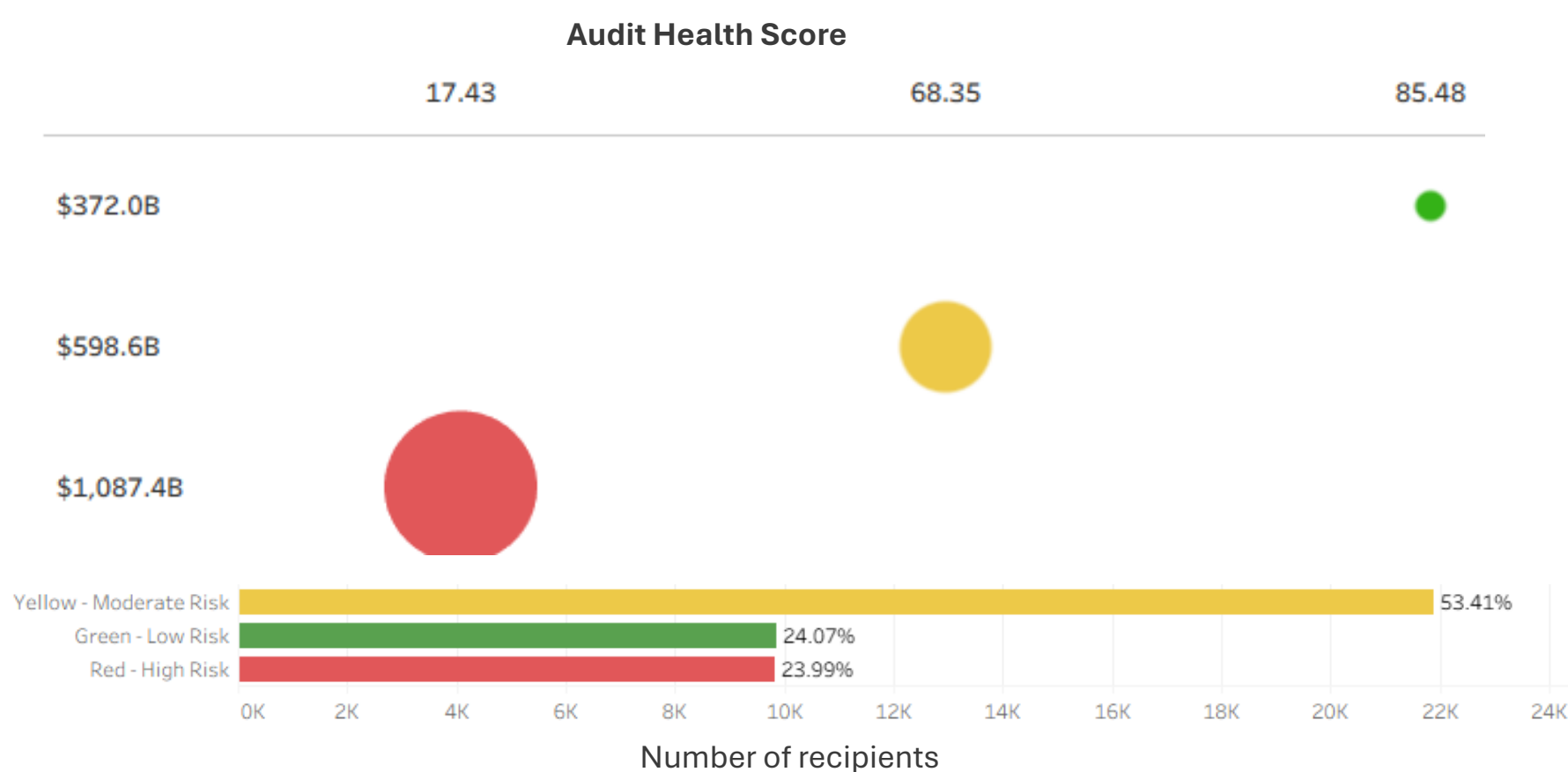|	Severity (Points)	Differentiation (%)
Going Concern	25	0.0%
Material Weakness	20	14.8%
Repeat Finding	15	11.4%
Significant Deficiency	10	23.3%
Has Corrective Action	-5	23.0%

Risk Tiers
- 🟢 80-100 Low Risk
- 🟡 60-79 Moderate Risk
- 🔴 0-59 High Risk

Built on Law, Proven by Math: A Transparent Standard for Federal Risk.

The Funding Paradox: High-Risk Entities Get 3x More

$1.09 trillion in taxpayer dollars went to recipients that auditors flagged for repeated failures.

Audit Health Score

17.43   68.35   85.48

$372.0B

$598.6B

$1,087.4B

Yellow - Moderate Risk — 53.41%
Green - Low Risk — 24.07%
Red - High Risk — 23.99%

Number of recipients

Funding Snapshot

$1.09 Trillion
to High-Risk Entities

3x More Than Low Risk

82% Funded to High & Moderate Risk Entities

42,222 audited entities receiving $2.06T in federal funding

Audit failures don't stop the money. Neither does being banned.

When the Score Hits Zero

The 10 largest recipients all have an Audit Health Score of 0.
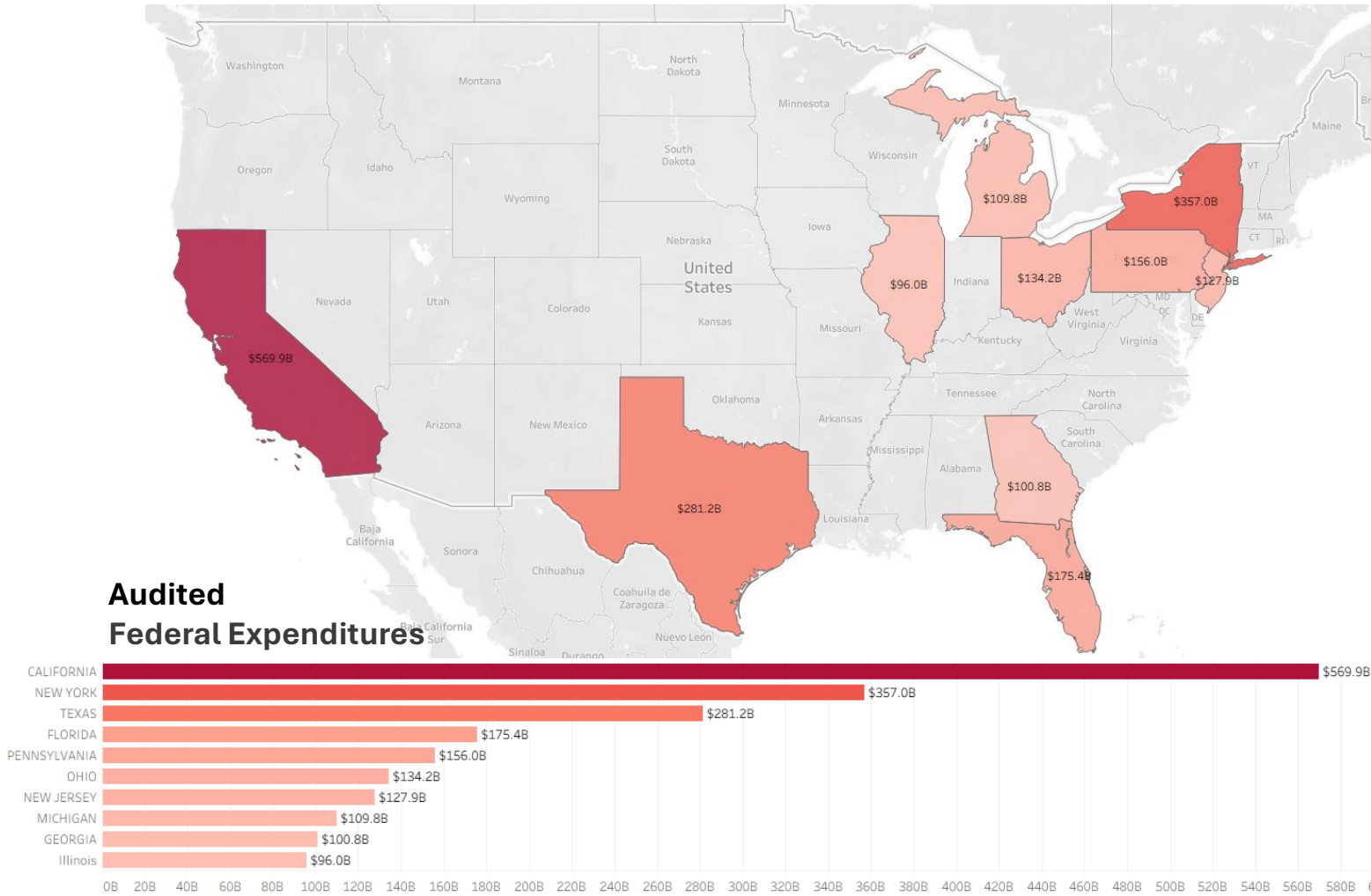
Risk Snapshot

10 state governments

#1 California: $570B

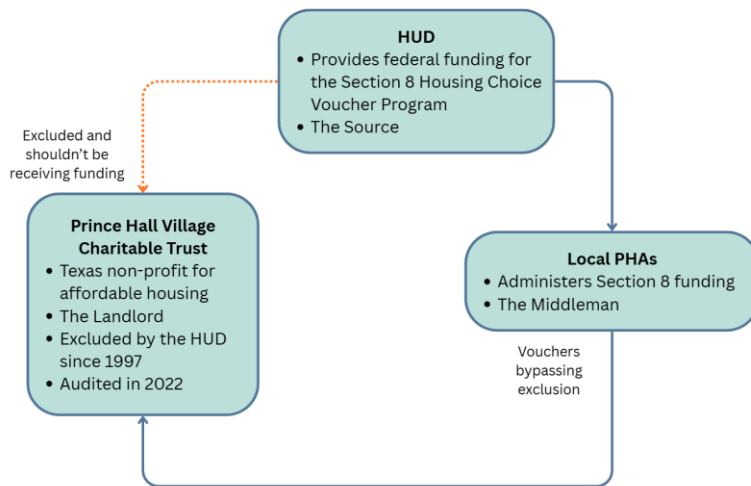Highest findings: 233 (MA)

$2.1T in expenditures

TPS = 0 (all)

Audited Federal Expenditures

State	Expenditure
CALIFORNIA	$569.9B
NEW YORK	$357.0B
TEXAS	$281.2B
FLORIDA	$175.4B
PENNSYLVANIA	$156.0B
OHIO	$134.2B
NEW JERSEY	$127.9B
MICHIGAN	$109.8B
GEORGIA	$100.8B
Illinois	$96.0B

Maximum risk. Maximum funding. Minimum accountability.

A Gap in the System

How HUD Paid an Excluded Entity for 28 Years

Key Players

HUD
- Provides federal funding for the Section 8 Housing Choice Voucher Program
- The Source

Excluded and shouldn't be receiving funding

Prince Hall Village Charitable Trust
- Texas non-profit for affordable housing
- The Landlord
- Excluded by the HUD since 1997
- Audited in 2022

Local PHAs
- Administers Section 8 funding
- The Middleman

Vouchers bypassing exclusion

What Went Wrong

2 CFR § 180.300: Requires checking exclusion status before covered transactions.

- Agencies are required to verify an entity isn't excluded before entering transactions.
- Prince Hall was excluded by HUD in 1997. Section 8 payments **were not screened**.
- PHAs relied on **manual checks**
- Exclusion enforcement **failed across decentralized systems**
- **~$2M paid** in Section 8 subsidies
- Payments confirmed through **2022 audit activity**
- Failure persisted for **28 years**

The Gap

- Without **automated, pre-payment exclusion screening**, permanent exclusions can be effectively ignored, even within a single agency.
- A glaring intra-agency oversight
- Clean audits don't reveal exclusion status. Data must be integrated.

- Undermines federal payment integrity and public trust.
- Creates legal, financial, and reputational risk.

Fiscal Patriots | AGA Datathon 2026

Source: SAM.gov Public Extract 25361 (Dec 27, 2025) | FAC (FY2016-2024) | USAspending.gov (FY2019-2024) | 2 CFR § 180.300

14

Banned by One Agency, Paid by Another

Cross-Agency Communication Failure

What Went Wrong

- **Entity:** Amerihost Services, LLC
- **Excluded by EPA:** September 8, 2021
- **Problem:** Despite the EPA ban, **HUD continued paying the entity** for housing-related awards in FY2022–FY2024.
- Entity had no Single Audit

> HUD paid a **legally excluded entity**, creating a **cross-agency compliance failure**.

Key Legal Principles

2 CFR § 180.100: Federal exclusions are **binding across all agencies**.
- An exclusion by **one agency (EPA)** automatically applies to **all other federal agencies**, including HUD.
- This is known as **government-wide reciprocal effect**.

2 CFR § 180.415(b)
- Agencies **cannot renew or extend transactions** with an excluded entity unless an exception is documented.
- 15 payments over FY2022–FY2024 (5 per year) **constitute renewals or extensions**, prohibited without an exception.
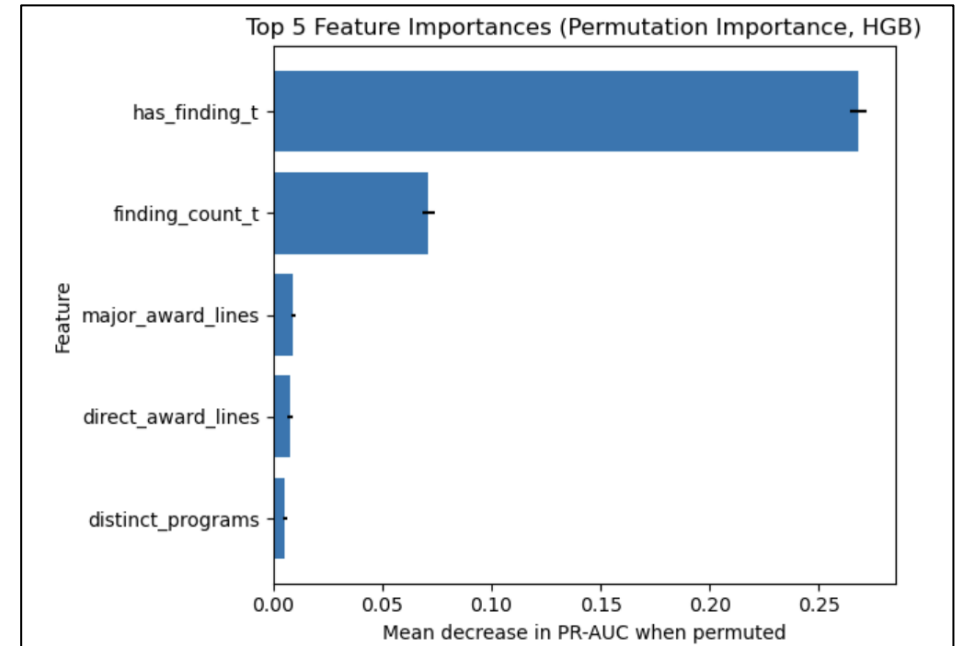
Why This Matters

- HUD paid $759,584 over 15 payments to a legally excluded entity.
- Cross-agency enforcement gaps
- Lack of automated screening for exclusions
- Direct legal violations of federal regulations

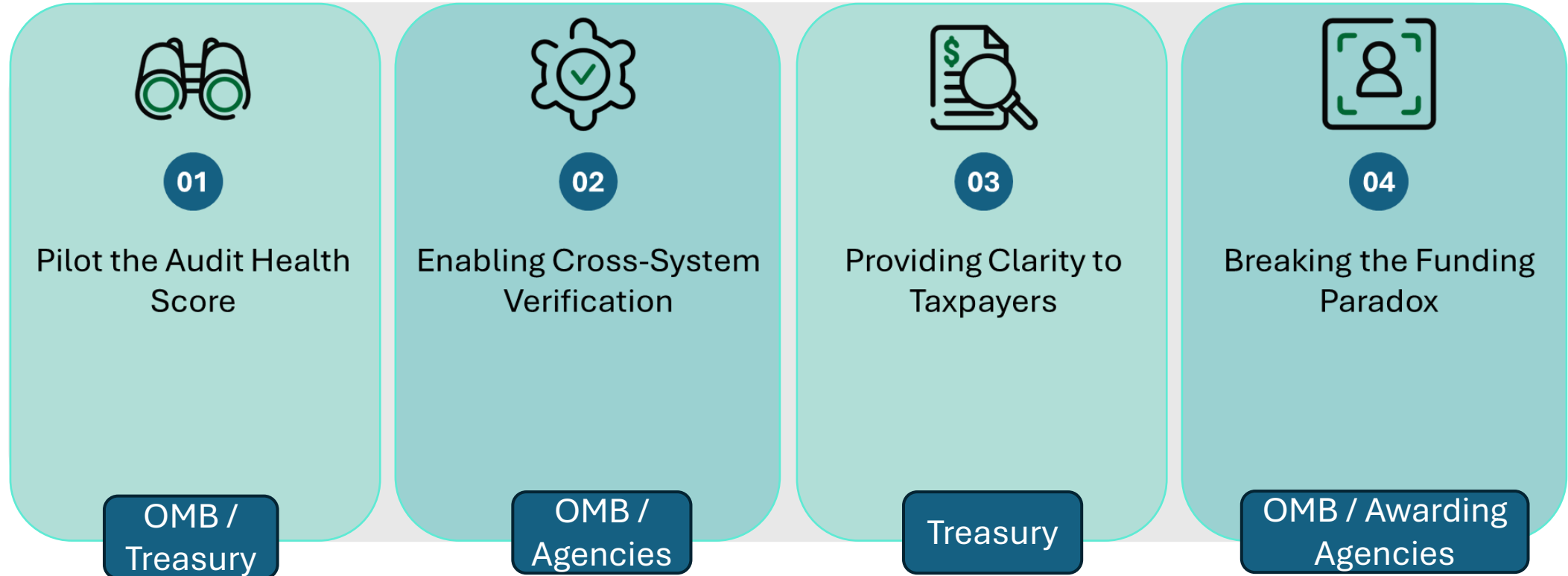# Predicting Audit Findings (Entity-Year Risk Model)

Predict whether a federal grant recipient will have **audit findings next year** using prior-year audit + funding characteristics to support risk oversight.

Model

Algorithm: HistGradientBoosting (HGB) classifier

Unit of Prediction: Entity-year (recipient in year t → predicts findings in t+1)

Inputs: prior findings indicators/counts, award line counts, funding totals/mix, program & agency diversity

Performance

ROC-AUC: 0.7656
PR-AUC: 0.5439

Key Signals

- Prior finding flag / finding count
- Award-line volume (# of direct award lines)
- Breadth of funding (distinct programs/agencies
- Scale of funding (program totals / major award lines)

Top 5 Feature Importances (Permutation Importance, HGB)

Closing the Gaps

From analysis to accountability.

01	**02**	**03**	**04**
Pilot the Audit Health Score	Enabling Cross-System Verification	Providing Clarity to Taxpayers	Breaking the Funding Paradox
OMB / Treasury	OMB / Agencies	Treasury	OMB / Awarding Agencies

---

The Bottom Line: The data exists. The tools exist. What's missing is the integration.

Appendix

DATA SOURCES

- USAspending.gov: API docs + data dictionary

- Federal Audit Clearinghouse (FAC): dataset overview + files used

- SAM.gov Exclusions: public extract documentation (UEI coverage limitation)

- "Allowable Data Sources " (competition rules screenshot / doc)

- Data dictionaries used in this repo (FAC, USAspending, Merged outputs)

TAXPAYER PROTECTION SCORE

- Variable Definitions (fields used in scoring)

- Weighting rationale (risk factors vs mitigating factors)

- Score formula and tier thresholds

- "Problem Entity" rule (logic and flags)

REGULATORY CITATIONS

- GAO-05-479 (2005) reference used

- GAO-09-174 (2009) reference used

- 2 CFR 200 Subpart F references(audit requirements)

METHODOLOGY

- Pipeline overview (Clean --> Score --> Merge) visual

- Join logic: UEI mapping (auditee_uei = recipient_uei) explanation

- Alteryx workflows (screenshots + the .yxmd files)

- Data acquisition notes (what we downloaded, and why)

ML MODEL

- ML training dataset build workflow

- Train/test split methodology

- Feature list + feature importance chart (or placeholder if not done

- Model validation metrics

- Final ML-ready outputs

CASE STUDIES (EXTENDED)

- Full audit findings for flagged entities

- Supporting documentation: screenshots/tables used in the write-up

Thank you