

# Site Scanning

*Providing real-time data about federal government websites*



Office of Solutions | Platforms and Services  
Technology Transformation Services

# The data is the product

## The numbers...

- **28,965** websites
- **1,440** domains
- **148** agencies
- **90** data points per website

## The analysis...

- Website/Service Inventory
- US Web Design System
- Digital Analytics Program
- Third Party Services
- 21<sup>st</sup> Century IDEA compliance
- IPv6, HTTPS, cloud

# How the scanning works



## The Process

---

- Hypothesis
- Research
- MVP
- Feedback/iterate
- Validate
- Reliability



## The Scan Itself

---

- Combine .gov registry + open datasets to find all known websites
- Run daily automated scans of the target list
- Makes scan data available via API and bulk (CSV) download



## Principles

---

- Open sources
- What is public on the internet
- Open data
- Real user need

# Use Cases in Action

- **Digital Analytics Program** - User adoption; Analysis of implementation
- **OMB OFCIO - 21st Century IDEA** - Track implementation of OMB-23-22 requirements
- **TTS Digital Council** - Utilization of TTS programs by federal agencies, Market research
- **GSA Customer Experience Team** - Inventory of GSA websites, DAP/USWDS compliance
- **GSA IT** - Adoption of IPv6
- **Search.gov** - Preparatory work for new implementations; State of .gov SEO
- **US Web Design System** - User adoption; Analysis of implementation
- **DigitalDashboard.gov** - Inventory of Federal Websites; USWDS compliance

# Data Available Now

- Program website [digital.gov/site-scanning](https://digital.gov/site-scanning)
- Data downloads/API documentation [open.gsa.gov/api/site-scanning](https://open.gsa.gov/api/site-scanning)
- Contact Gray Brooks or Lauren Ancona [site-scanning@gsa.gov](mailto:site-scanning@gsa.gov)