



“Build to Share”

DATA.GOV

Open Government Data
One Year Later

13 October 2010

U.S. Federal Data Architecture Subcommittee (DAS)



... DAS ...

Agenda



- ◆ Origins
- ◆ Features of Data.gov
- ◆ Applications, Community, Semantic Web Highlights
- ◆ What's Next
- ◆ Summary



MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

SUBJECT: Transparency and Open Government

Government should be transparent. Transparency promotes accountability and provides information for citizens about what their Government is doing. Information maintained by the Federal Government is a national asset. My Administration will take appropriate action, consistent with law and policy, to disclose information rapidly in forms that the public can readily find and use. Executive departments and agencies should harness new technologies to put information about their operations and decisions online and readily available to the public. Executive departments and agencies should also solicit public feedback to identify information of greatest use to the public.

**“Information
maintained by the
Federal Government is
a national asset.”**

*President Obama
January 21, 2009*



... DAS ...

The Issues



There are over 24,000 .gov web sites – difficult to navigate

Web sites organized under the federal structure
– difficult to find datasets

Data exists in portals – are not “raw” machine downloadable
or data locked in legacy systems

Web sites not designed for new generation
of web technologies



When Data is Democratized



... DAS ...

When the Department of Defense released its satellite data... the market created affordable GPS devices!





... DAS ...

Tapping the Ingenuity of the American Public



◆ The Federal Government does not have a monopoly on information

- Democracy unleashes creativity
- Innovation drives sustainability





... DAS ...

Current Scope of Data.gov



◆ Datasets that:

- Are already in or ready for public domain
- Are high value
- Are hosted by the Executive Branch of the Federal Government
 - No Legislative or Judicial Branch datasets
 - No direct links to State, Local, or Tribal datasets
 - No non-governmental datasets
- Do not have copyright limitations
- Do not impose charges for access
- Do not require registration to access



... DAS ...

Preferred Formats for Raw Data



◆ Preferred formats for 'raw' data:

- XML
- CSV/TXT
- XLS
- Keyhole Markup Language (KML/KMZ)
- ESRI Shapefile

◆ Avoiding datasets in the following formats:

- PDF
- HTML (e.g., HTML tables containing data)



What Agencies Contribute



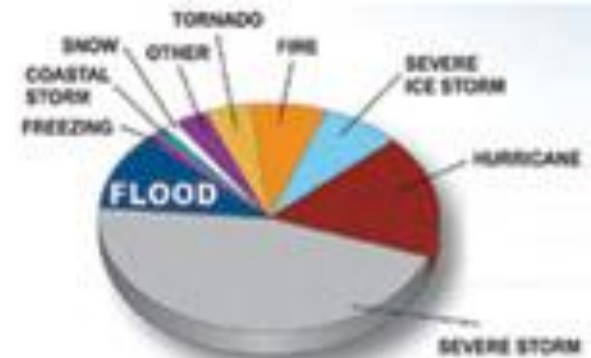
... DAS ...

Datasets that can be used to:

- increase agency accountability and responsiveness;
- improve public knowledge of the agency and its operations;
- further the core mission of the agency;
- create economic opportunity; or
- respond to need and demand as identified through public consultation



Veterans Affairs
Hospital Report Card



FEMA Disaster and
Emergency Declarations



MyPyramid Food Data



... DAS ...

The First Year



- ◆ Data.gov is a Federal Inter-Agency open government initiative that was chartered by the Federal CIO and is operated by GSA.
- ◆ Launched May 21, 2009: 47 Datasets

The screenshot shows the Data.gov homepage with the following elements:

- Header:** "TUESDAY, DECEMBER 15, 2009" and "Text A+ A- A" on the right.
- Navigation:** "DATA.GOV" logo and a menu with links: HOME, CATALOGS, STATE/LOCAL, FAQ, SUGGEST DATASETS, AGENCY PARTICIPATION, ABOUT, BLOG.
- Main Content:**
 - Left Column:** "DISCOVER. PARTICIPATE. ENGAGE." followed by "Search the following Data.gov catalogs:" and three icons for "RAW" DATA CATALOG, TOOL CATALOG, and GEODATA CATALOG.
 - Right Column:** "FEATURED DATASETS: VETERANS HEALTH ADMINISTRATION HOSPITAL REPORT CARD - QUALITY OF CARE - POPULATIONS". It includes a diagram with labels like "Patient Satisfaction", "Availability of Services", "Medical Center Staffing", "Patient Safety", "Wait Times", and "Physician Staffing". Below the diagram is a "VIEW THESE DATASETS" link.
- Footer:** "Welcome to Data.gov" with a paragraph about the site's purpose, and a "NEW!" banner for "Open Government Directive Mandates Open Data".



... DAS ...

Finding the Innovative Path



In just a few hours the public created an FBI most wanted app

In just a few weeks the public created several more

Public innovation translates to development \$'s saved



SUBMITTED APPS

MAICgregator
MAICgregator is a Firefox extension that aggregates information about colleges and universities embedded in the military-academic-industrial complex (MAIC)
Submitted June 22, 2009 by [Nicholas Knouf](#)

Top Dangerous Mines
A website showing a real-time list of the most dangerous mines in the US based on recent earthquake data.
Submitted June 18, 2009 by Benjamin Hardin

FlyOnTime.us
Find the most on-time flight from one city to another.
Submitted June 11, 2009 by Josh Sulkin

FBI Fugitive Concentration
Play Concentration with photos of FBI Fugitives
Submitted May 22, 2009 by Bertrand Fan

F.B.I. FUGITIVE CONCENTRATION

How it works

The FBI provides Fugitive RSS Feeds for the following categories:

- Top Ten Fugitives
- Most Wanted, Tennessee
- Wanted for Violent Crimes, Offender
- Wanted for Violent Crimes
- Wanted for Cyber Crimes
- Wanted for Crimes Against Children
- Wanted on Criminal Enterprise Investigations
- Wanted for White Collar Crimes

F.B.I. Fugitive Concentration passes those feeds and extracts an image of each fugitive and then randomly selects 12 fugitives for you to play with.

Using data from the [Federal Government](#), we can tell you how late your flight is on average or help you find the most on-time flight from one city to another.

FlyOnTime.us

Airline:

Flight #: (optional)

From: (city or airport)

To: (city or airport)

OR

[Terms of Use](#)

Data.gov Features One Year Later 273,011 Datasets



Search Raw Datasets, Tools, or Geodata Catalogs

Join the national dialogue and contribute ideas

See datasets most popular to date

See latest statistics on agency contributions, dataset suggestions and site visitors

Most Popular Datasets

1. U.S. Overseas Loans and Grants (Greenbook)
2. 2008 Early Release Toxics Release Inventory...
3. Worldwide M1+ Earthquakes, Past 7 Days
4. Latest Volumes of Foreign Relations of the...
5. Travel Warnings

See a list of apps developed using government datasets.

See states, cities, and countries that are forming community

See the latest in Data.gov semantic technology activity

APPS

COMMUNITY

Data.gov is leading the public sector data and data is being surfaced making the Government researchers to perform Developers are finding good uses for the datasets, providing interesting and useful applications that allow ordinary citizens and public

SEMANTIC WEB

Search Raw Datasets, Tools, or Geodata Catalogs



Participate



... DAS ...

An Official Web Site of the United States Government Sunday, July 11, 2010 Text: A+ A- A Share

DATA.GOV
EMPOWERING PEOPLE

Search our catalogs.. SEARCH

HOME DATA TOOLS COMMUNITY METRICS DIALOGUE

Toxics Release Inventory National data file of all US States and Territories

DATASET SUMMARY

Agency	Environmental Protection Agency
Category	Geography and Environment
Released	3/19/2009
Updated	6/10/2009
Period	Calendar Year 2007
Frequency	Annual
Description	The Toxics Release Inventory (TRI) is a publicly available EPA database that contains information on toxic chemical releases and waste management activities reported annually by certain industries as well as federal facilities.

DOWNLOAD INFORMATION

XML	CSV/TXT
XLS	KML/KMZ
Shapefile	Maps
RDF	PDF
36 MB	

Cannot find data you are looking for? Suggest other datasets!

DATASET RATINGS

	Current	Your Rating [?]
Overall	★★★★★ (3 votes)	★★★★★
Data Utility	★★★★★ (3 votes)	★★★★★
Usefulness	★★★★★ (3 votes)	★★★★★
Ease of Access	★★★★★ (3 votes)	★★★★★
Comments (optional)	<input type="text"/>	

SUBMIT (Privacy Policy)

DATASET METRICS

Number of Downloads	766
---------------------	-----

DATASET INFORMATION

Rate A Dataset and Provide feedback

Suggest A Dataset



Data.gov Features Apps



... DAS ...

People from across the country and around the world are using government datasets to create applications and mashups: 237 and counting!

A screenshot of the Data.gov website's "APPS SHOWCASE" page. The page features a navigation bar with "HOME", "DATA", "TOOLS", "COMMUNITY", "METRICS", and "DIALOGUE". Below the navigation bar is a search bar and a "SEARCH" button. The main content area is titled "APPS SHOWCASE" and includes a descriptive paragraph about government data applications. Below this, there are several application thumbnails, each with a title and a small preview image. The applications shown are: FlyOnTime, National Obesity Comparison Tool, Fix My City DC, Employment Market Explorer, Check it and See (Superfunds), DataMasher, Visualizing Community Health Data, and Clean Air Status—Ozone.

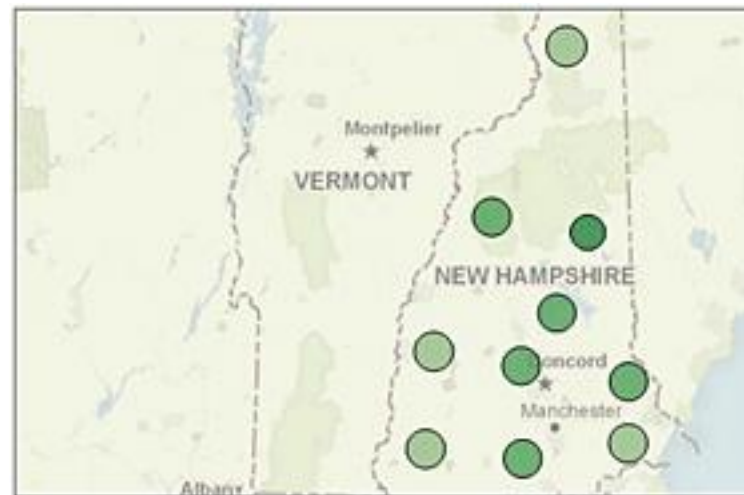
National Obesity: Eat Your Vegetables

DATA.GOV



... DAS ...

- ◆ Uses HHS community health data and shows obesity rates by county
- ◆ Supports the First Lady's fights against childhood obesity



<http://public.tableausoftware.com/views/contributorstoobesity/Eatyourvegetables?:embed=yes&:toolbar=yes>



Community



... DAS ...

Data.gov is leading the way in democratizing public sector data and driving innovation.

- 6 Other nations establishing open data
- 10 States now offering Open data sites
- 10 Cities in America with open data sites



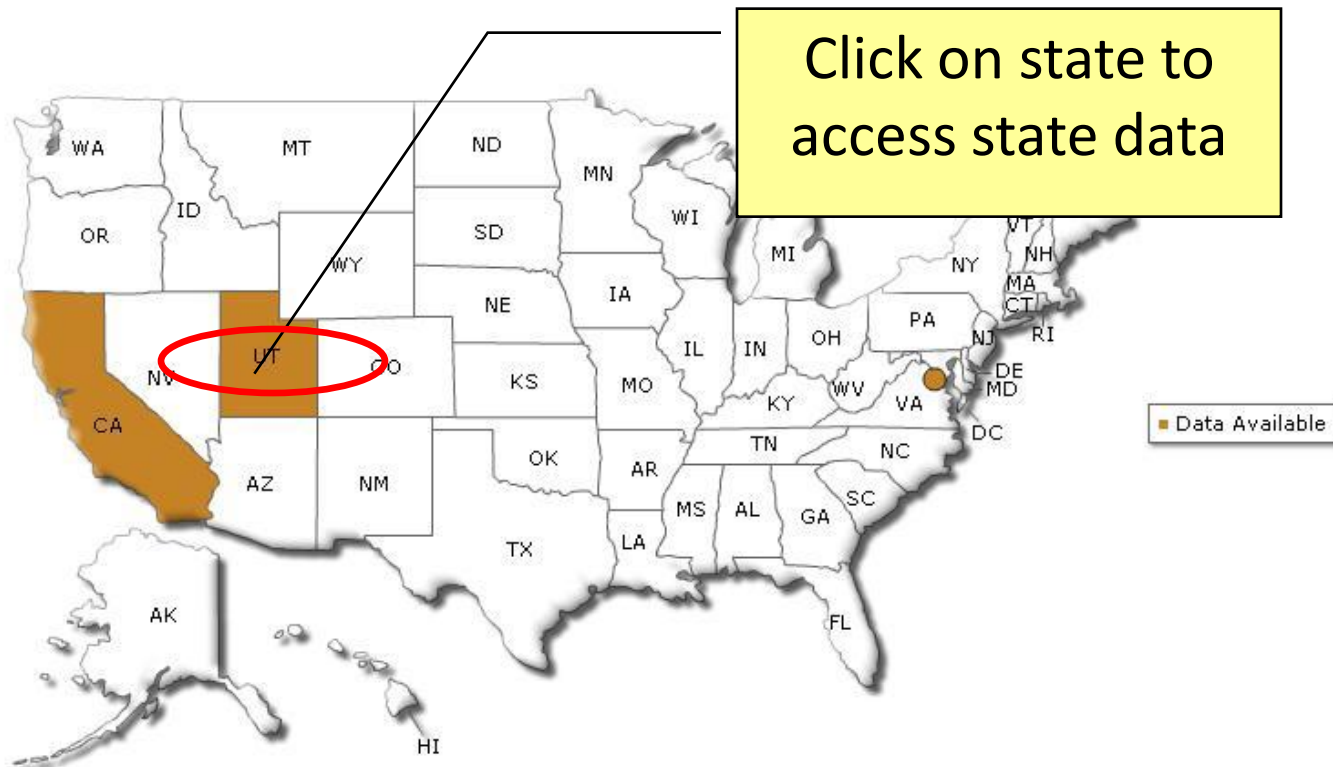
Access State and Local Specific Data

DATA.GOV

HOME | CATALOGS | **STATE/LOCAL** | ABOUT | FAQ | CONTACT US | SUGGEST OTHER DATASETS

State and Local Data Sites

We are encouraging and working with states and local governments to launch their own sites with access to machine-readable data. The map below provides easy access to the state and local data sites that have been launched.



Access State and Local Specific Data

Code	Title	Inexperienced	Average	Media	Middle Range From	Middle Range To
13-2011	Accountants and Auditors	\$18.13	\$29.26	\$26.03	\$19.98	\$34.95
27-2011	Actors	7.94	13.77	13.59	9.00	15.74
15-2011	Actuaries	25.53	42.11	34.66	27.69	49.66
23-1021	Administrative Law Judges, Adjudicators, and Hearing Officers	19.48	33.08	29.45	23.03	41.18
11-3011	Administrative Services Managers	21.68	32.65	30.49	24.12	38.94
25-3011	Adult Literacy, Remedial Education, and GED Teachers and Instructors	11.28	19.02	18.29	13.10	23.65
11-2011	Advertising and Promotions Managers	21.30	39.58	35.08	23.98	54.04
41-3011	Advertising Sales Agents	12.53	24.26	19.26	14.54	33.84
17-3021	Aerospace Engineering and Operations Technicians	17.25	23.89	23.82	20.03	28.06
17-2011	Aerospace Engineers	30.32	38.64	38.65	32.64	45.08
19-4011	Agricultural and Food Science Technicians	12.33	16.65	16.69	13.71	18.83
45-2091	Agricultural Equipment Operators	10.22	12.37	11.32	10.22	14.36
45-2011	Agricultural Inspectors	13.18	17.68	17.10	14.14	20.81
53-1011	Aircraft Cargo Handling Supervisors	10.66	14.73	13.14	11.30	16.80
49-3011	Aircraft Mechanics and Service Technicians	19.08	23.85	23.30	20.36	27.71
51-2011	Aircraft Structure, Surfaces, Rigging, and Systems Assemblers	12.64	17.98	18.61	14.32	21.41
53-2022	Airfield Operations Specialists	12.36	17.40	16.96	13.76	19.61
39-3091	Amusement and Recreation Attendants	6.42	7.97	7.47	6.43	9.06

THE OFFICIAL WEBSITE FOR THE STATE OF UTAH

Subscribers: 24/7 Help Font Size: A A A

GOVERNMENT RESIDENTS VISITORS BUSINESS EMPLOYMENT EDUCATION ABOUT ONLINE SERVICES

RELATED LINKS & RESOURCES

Utah Data

- [State Data Files](#)
- [State Data Sources](#)
- [Local Data Sources](#)
- [Data Visualizations](#)
- [Geographic Data Sources](#)
- [External Data Sources](#)
- [Swivel Utah Data Group](#)

Follow Data.Utah.gov

For immediate notification when new data becomes available, follow UtahData on Twitter.

More Data Sources

- [Federal Data Sources](#)
- [Other State and Local Data Sources](#)

Your Feedback

Was the information on this page helpful?

Yes No

If not, what were you looking for?

Please note, your feedback will be reviewed, but not responded to. If you would like

State Data Sources

Larger sources of state data which allow you to access, query and download in dozens of categories.

[View State Data Sources](#)

Local Data Sources

Data.Utah.gov will help you find and access data from local government sources in Utah.

[View Local Data Sources](#)

Geographic Data Sources


KML files can be used to create maps with popular mapping tools. See some examples.

[View Geographic Data Sources](#)

External Data Sources

Utah data is also available from external sources like the Federal government.

[View External Data Sources](#)





... DAS ...

Semantic Web RPI and Linked Data



Running on Data.gov are a set of government-related demonstrations at RPI showing how these RDF datasets can be used to build interactive visualizations using various popular visualization API's



Innovations at RPI



Running on Data.gov are a set of government-related demonstrations, originally created by a **team of developers** at the **Tetherless World Constellation** at **Rensselaer Polytechnic Institute (RPI)**, which show how these RDF datasets can be used to build interactive visualizations using various popular visualization API's. You can run the demonstrations here, served by our **SPARQL endpoint** that is powered by the **Virtuoso** linked data server. These demonstrations are also linked to pages at RPI that explain how they function and provide examples that you can cut and paste to build your own demos.



Clean Air Status and Trends Network (CASTNET)

This map shows the ozone levels across the U.S. Clicking on a dot brings up detailed information about the ozone measurements for that area over time. The ozone values shown are taken from **dataset 8**, "Clean Air Status and Trends Network (CASTNET); Ozone" released by the Environmental Protection Agency. The faceted browsing in this demo uses the **Exhibit framework** developed by the Simile Project at MIT.



Public Company Bankruptcy Cases Opened and Monitored for Fiscal Year 2009

This demonstration shows public company bankruptcies for fiscal year 2009 based on **dataset 1580** released by the Securities and Exchange Commission. The map shows the ratio of total debt over assets from all bankrupt cases at each state level. Clicking on a state displays details of those public companies that filed bankruptcies within this state, e.g. the company name, assets, and liabilities.



Metrics



... DAS ...

- Federal Agency Participation at Data.gov
- Suggested Dataset Stats
- Visitor Stats

An Official Web Site of the United States Government

Thursday, June 10, 2010 Text: A+ A- A Share



Search our catalogs..

SEARCH

HOME DATA TOOLS COMMUNITY METRICS DIALOGUE

FEDERAL AGENCY

SUGGESTED DATASETS

VISITOR STATS

Federal Agency Participation at Data.gov

This report is also available for download in the following formats: CSV XLS

* These numbers represent the number of times a user has clicked on the "XML" or "CSV" (for example) links in the Raw Data Catalogs to download datasets and user downloads of tools in the Tool Catalog available in these categories.

Data last updated on 06/10/2010 12:02 AM EDT.

Departments/Agencies

Agency/Sub-Agency	Raw Datasets (high-value)	Tools (high-value)	Geodata	Total	Latest Entry	# of times downloaded within the last week*
Department of Agriculture (USDA)	6 (3)	12	1	19	05/19/2010	298
Department/Agency Level (USDA)	1	4	0	5	12/10/2009	107
Agricultural Research Service (ARS)	1 (1)	2	0	3	05/19/2010	68
Economic Research Service (ERS)	1 (1)	4	0	5	05/19/2010	48
Food and Nutrition Service (FNS)	2	0	0	2	05/19/2010	17



... DAS ...

Visitors Statistics (As of June 2010)



Top 10 U.S. States

1. California
2. Georgia
3. Washington
4. Washington, DC
5. Texas
6. Virginia
7. New York
8. Illinois
9. Florida
10. Pennsylvania

Top 10 Nations

1. China
2. United Kingdom
3. Germany
4. Japan
5. Canada
6. France
7. India
8. Brazil
9. Spain
10. Taiwan

Top 5 Datasets

- Earthquakes – DOI/USGS
- U.S. Overseas Loans - USAID
- Latest Foreign Relations - State
- MyPyramid Food - USDA
- Environmental Regulation – EPA

Year to Date # of Visitors: 1,869,756

Visitors in January 2010: 159,656





... DAS ...

National Dialogue



- Encouraging the community to share ideas for Data.gov

Evolving Data.gov with You - by IdeaScale

new idea +

Categories

- all (142)
- Show Us Your Mashups! (5)
- Strategic Intent (18)
- Key Partnerships (12)
- Solution Architecture (36)
- Agency Next Steps (12)
- Developer's Corner (18)
- General (41)

Subscribe

142 Ideas, 339 Comments, 2240 Votes x

POWERED BY IdeaScale

Join the Dialogue

DATA.GOV
Evolving Data.gov With You

Login | Signup

Evolving Data.gov with You

We opened this discussion to encourage the community to share creative ideas and help us evolve Data.gov. We are getting some great ideas and discussion. We are working hard to incorporate many of your ideas, so stay tuned as Data.gov transforms throughout 2010.

If you are just visiting the site for the first time, we urge you to take a look at our Plan, the [Draft Concept of Operations](#), read through the ideas already posted to the site, contribute new ideas and comments, and to vote for your favorite ideas. Topic areas, including the Developer's Corner, are on the left hand tool bar and you can dive right in with the Getting Started section, which appears on the right hand tool bar. Let your voice be heard to make Data.gov a richer, more user-friendly resource for everyone!

Thank you very much for your contributions, they will be invaluable as we take the next steps to build future versions of Data.gov.

Linda A. Travers is CIO Council Data.gov Co-lead and Deputy CIO, US Environmental Protection Agency

Search Ideas

Search Ideas

Getting Started:

Add Your Idea by clicking on the "new idea" button to the left.

Here is a discussion question to help get you started:
What would you like Data.gov to look like by July 2010?

While you're here:

- Review Our Plans
Take a look at our ideas to evolve Data.gov. You can download our draft Data.gov *Concept of Operations* in Word or Acrobat formats.
- Word Document [49 pp, 6.5MB]



... DAS ...

National Dialogue



- Vote on your favorite ideas
- Some of the higher rated ideas have been implemented

Vote Up
48
Vote Down

Provide the Opportunity to Reduce Duplication

learon.dalby

Data.gov should become a nationally managed "access point" that provides a mechanism for all levels of government to participate or integrate with, thus creating a single location for citizens to access government data.
[Comments \(14\)](#) | [Strategic Intent, nsgic, key partnerships, coordination](#)

Vote Up
43
Vote Down

Create a Data.gov Evangelist Position

sressler

Most of the big APIs like Twitter, Facebook, and now LinkedIn have a dedicated core team that spends a lot of time evangelizing the API and working with developers to get what they need.

 Data.gov needs a Data.gov evangelist who can be the community manager out on the road listening and talking and generally spreading the word on data.gov and its value.

 I think this is a separate role from leadership like Vivek Kundra and Aneesh Chopra (who have a ton of priorities so hard to manage bandwidth).
[Comments \(10\)](#) | [Key Partnerships, evangelist](#)

Vote Up
43
Vote Down

Custom Notifications/RSS Feeds for New Datasets

David Smith

An API to provide customizable RSS feeds should be considered, to allow users to subscribe to specific thematic areas, geographic areas and so on. This way, if for example someone was interested in data on aquatic resources



... DAS ...

What's Next



- ◆ Year one focused on discovery of government datasets
- ◆ In addition to providing more datasets, year two will focus on:
 - ◆ Evolving the application of semantic technologies
 - ◆ Showcase more apps
 - ◆ Improving communication with suggested datasets
 - ◆ Geo visualizations
 - ◆ Building communities with:
 - Developers
 - Educators
 - Students (K-12 and universities)
 - Private industry
 - International

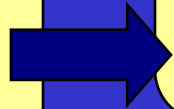


... DAS ...

Summary



Challenge :
Making Data Available



Solution:
Democratize Data using
DATA.GOV



Creating:
Transparent Open Government

*Tap the Ingenuity of the
American People*





Questions



... DAS ...



U.S. Federal Data Architecture Subcommittee (DAS)

Contact info:

Suzanne Acar:

**Co-Chair, Federal Data Architecture Subcommittee, and
FBI Senior Data Architect**

E-mail: suzanne.acar@ic.fbi.gov