

Overview of Data.gov

January 2017

@usdatagov





The home of the U.S. Government's open data

Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and more.

GET STARTED

SEARCH OVER 104,942 DATASETS



Federal Student Loan Program Data

BROWSE TOPICS



Agriculture



Climate



Education



Energy



Finance



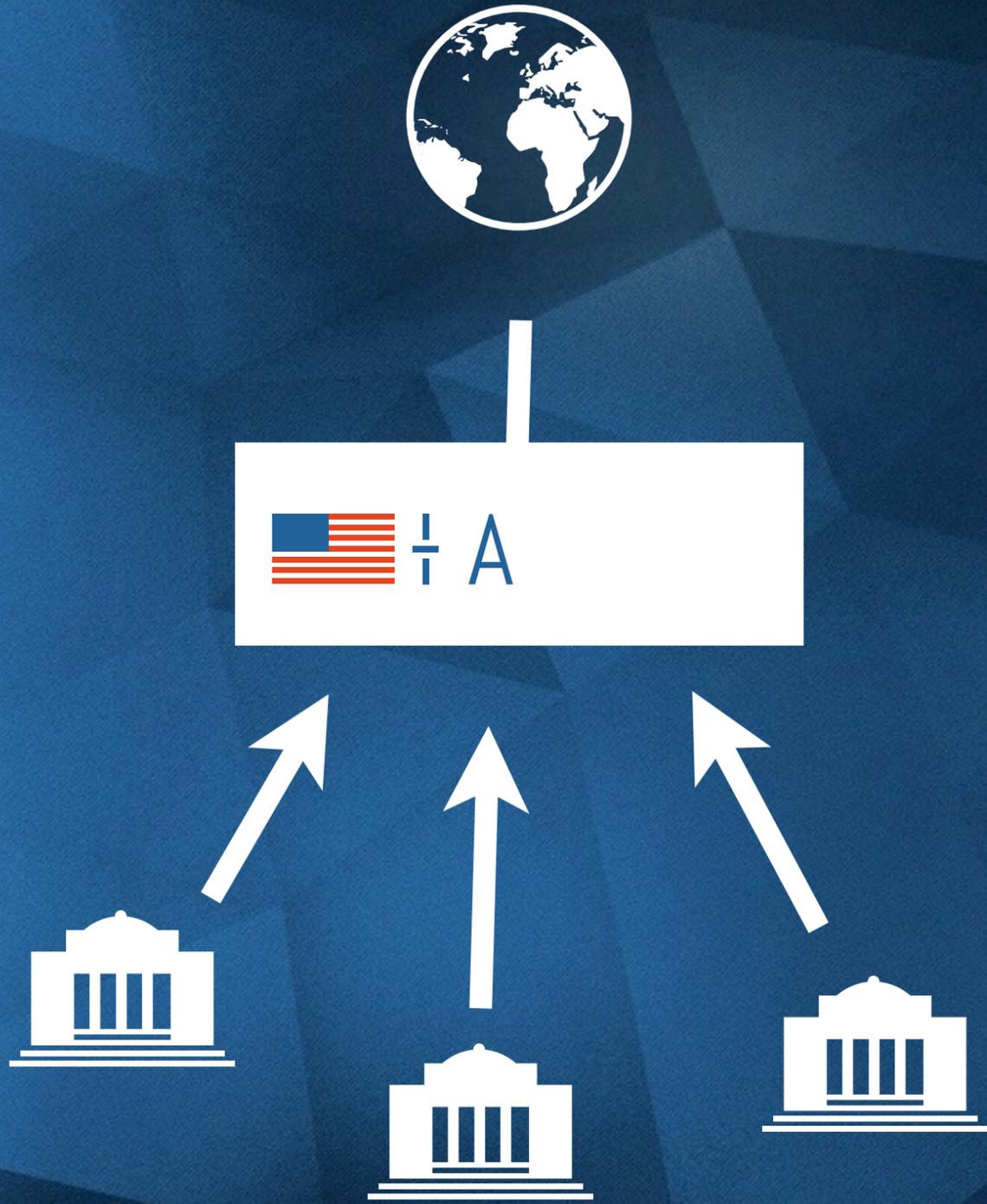
Geospatial



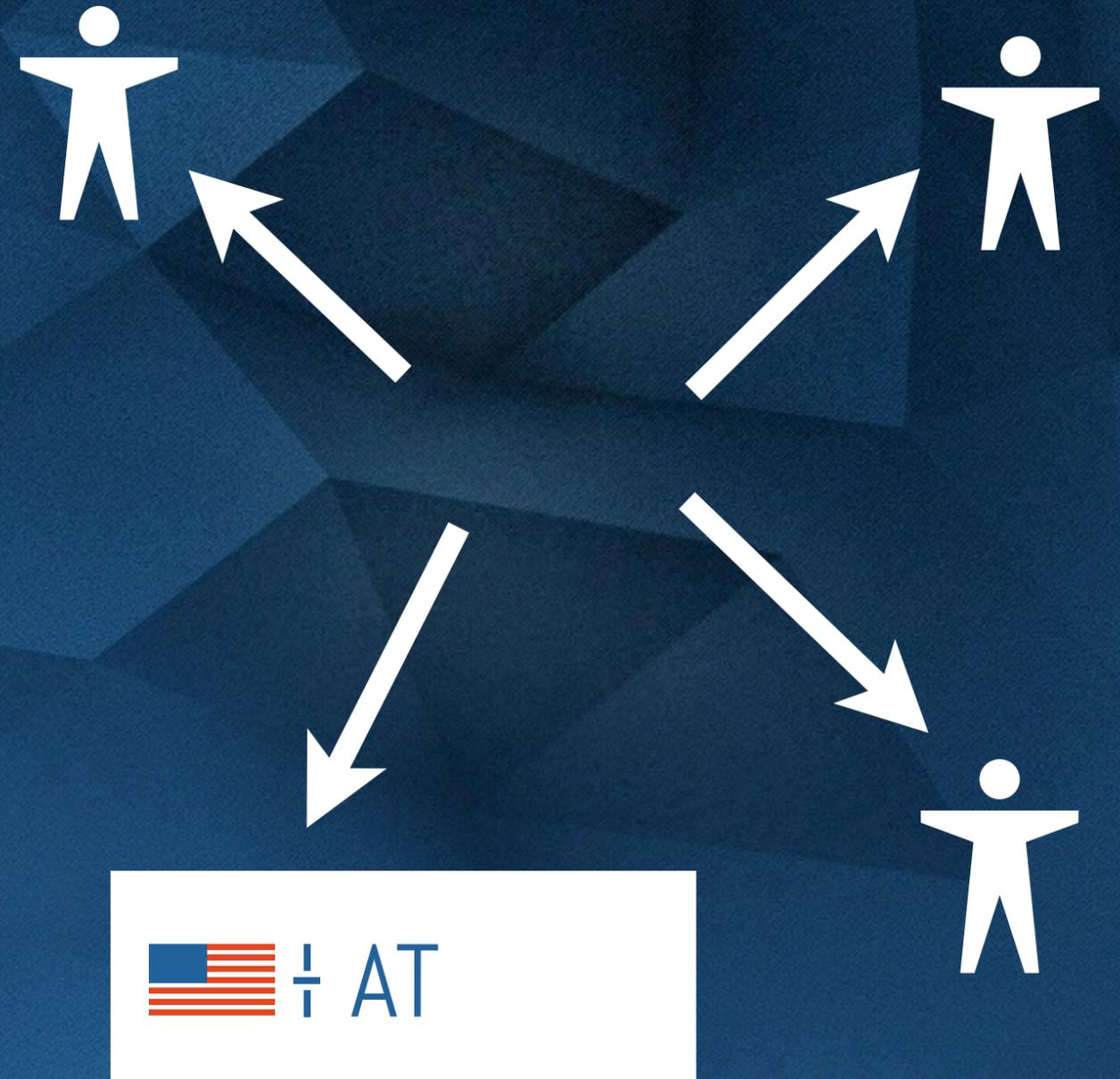
Executive Order on Open Data (2013)

Project Open Data

Before



After



PROJECT OPEN DATA

Open Data Policy – Managing Information as an Asset

1. Background

Data is a valuable national resource and a strategic asset to the U.S. Government, its partners, and the public. Managing this data as an asset and making it discoverable, and usable – in a word, *open* – not only strengthens our democracy and promotes efficiency and effectiveness in government, but also has the potential to create economic opportunity and improve citizens' quality of life.

For example, when the U.S. Government released weather and GPS data to the public, it fueled an industry that today is valued at tens of billions of dollars per year. Now, weather and mapping tools are ubiquitous and help everyday Americans [navigate their lives](#).

The ultimate value of data can often not be predicted. That's why the U.S. Government released a [policy](#) that instructs agencies to manage their data, and information more generally, as an asset from the start and, wherever possible, release it to the public in a way that makes it open, discoverable, and usable.

The White House developed Project Open Data – this collection of code, tools, and case studies – to help agencies adopt the Open Data Policy and unlock the potential of government data. Project Open Data will evolve over time as a community resource to facilitate broader adoption of open data practices in government. Anyone – government employees, contractors, developers, the general public – can view and contribute. [Learn more about Project Open Data Governance](#) and dive right in and help to build a better world through the power of open data.

2. Definitions

This section is a list of definitions and principles used to guide the project.

- 2-1 [Open Data Principles](#) - The set of open data principles.
- 2-2 [Standards, Specifications, and Formats](#) - Standards, specifications, and formats supporting open data objectives.
- 2-3 [Open Data Glossary](#) - The glossary of open data terms.
- 2-4 [Open Licenses](#) - The definition for open licenses.
- 2-5 [Project Open Data Metadata Schema](#) - The schema used to describe datasets, APIs, and published data at [agency.gov/data](#).

3. Implementation Guidance

Implementation guidance for open data practices.

Catalog Fields

These fields describe the entire Public Data Listing catalog file. Publishers can also use the `describedBy` field to reference the default [JSON Schema](#) file used to define the schema (<https://project-open-data.cio.gov/v1.1/schema/catalog.json>) or they may refer to their own JSON Schema file if they have extended the schema with additional schema definitions. Similarly, `@context` can be used to reference the default [JSON-LD Context](#) used to define the schema (<https://project-open-data.cio.gov/v1.1/schema/catalog.jsonld>) or publishers can refer to their own if they have extended the schema with additional linked data vocabularies. See the [Catalog section](#) under *Further Metadata Field Guidance* for more details.

Field	Label	Definition	Required
<code>@context</code>	Metadata Context	URL or JSON object for the JSON-LD Context that defines the schema used.	No
<code>@id</code>	Metadata Catalog ID	IRI for the JSON-LD Node Identifier of the Catalog. This should be the URL of the data.json file itself.	No
<code>@type</code>	Metadata Type	IRI for the JSON-LD data type . This should be <code>dcat:Catalog</code> for the Catalog.	No
<code>conformsTo</code>	Schema Version	URI that identifies the version of the Project Open Data schema being used.	Always
<code>describedBy</code>	Data Dictionary	URL for the JSON Schema file that defines the schema used.	No
<code>dataset</code>	Dataset	A container for the array of Dataset objects. See Dataset Fields below for details.	Always

Dataset Fields

See the [Further Metadata Field Guidance](#) section to learn more about the use of each element, including the range of valid entries where appropriate. Consult the [field mappings](#) to find the equivalent v1.0, DCAT, Schema.org, and CKAN fields.

Field	Label	Definition	Required
<code>@type</code>	Metadata Type	IRI for the JSON-LD data type . This should be <code>dcat:Dataset</code> for each Dataset.	No
<code>title</code>	Title	Human-readable name of the asset. Should be in plain English and include sufficient detail to facilitate search and discovery.	Always

Browse Issues Milestones

New Issue

- Everyone's Issues 17
- Assigned to you 0
- Created by you 8
- Mentioning you 6

Milestone: Next Version of Co. 17 open issues

- Labels
- GIS 2
 - Non-Federal 3
 - deprecate 1
 - guidance 2
 - new-field 6
 - rename 2
 - schema 13

Clear milestone and label filters
17 Open 1 Closed Sort: Newest

- Close Label Assignee Milestone
- Clarify how to specify which version of the schema is being used #309
schema schema structure
Opened by philipashlock on May 8
- Ability to specify primary or alternate metadata source #308
new-field schema
Opened by philipashlock on May 8 5 comments
- updating publisher guidance Non-Federal #306
Opened by gbinal on May 5
- Provide guidance for how agencies should designate Geospatial data #303
GIS guidance
Opened by gbinal on May 1
- Add label/value option under format #293
new-field schema
Opened by justgrimes on Mar 13 10 comments
- Further accrualPeriodicity values #292
deprecate guidance
Opened by gbinal on Mar 13 12 comments
- Improve metadata structure for listing web APIs #291

Non-public



-  Publisher
- OCSIT/18F
-  Contact
- Amanda Robinson
-  Share on Social Sites
-  Google+

 This metadata record is available for the public, but the data itself is not public for privacy or security reasons. [See details](#)

18F Internal Team API

 Metadata Updated: Apr 15, 2016

An API for information on what 18F is working on.

Access & Use Information

-  **Non-public:** This dataset is not for public access or use. Specific details are provided below.
-  License: Creative Commons CCZero

Downloads & Resources

 No file downloads are available because this is not a public dataset. See Access & Use Information for details.

Non-Federal



Order by:
Select an option

Datasets ordered by Popular

Topics:
Local Government

Filter by location Clear
Enter location...



10,002 datasets found

Crimes - 2001 to present [593 recent views](#)

City of Chicago - This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001...

CSV RDF JSON XML

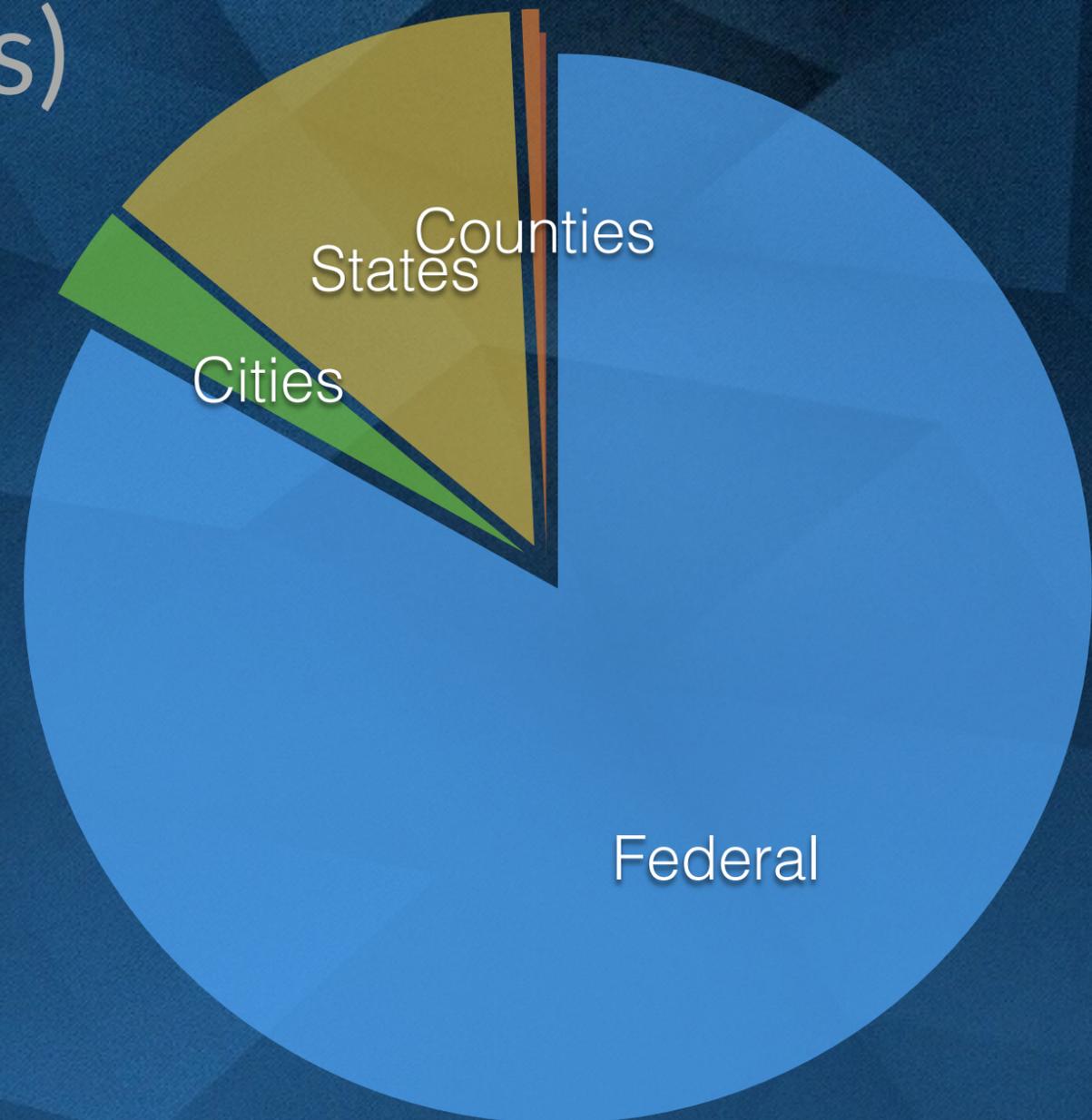
NYC Social Media Usage [149 recent views](#)



- 60+ Federal agency sources (145,192 datasets)
- 32 States (9170 datasets) + NSGIC (14,102 datasets)
- 4 Counties (926 datasets)
- 16 Cities (5024 datasets)

80% Federal & 20% Local

(As of June 2016)





*Measuring
Metadata Quality*

This is a public dashboard showing how Federal agencies are performing on the Open Data Policy. [Learn more](#)

Previous Milestone: The milestone selected is the most recently complete one. The status of each field won't be final until a few weeks after the milestone has passed

Selected: Milestone 6 - February 28th 2015

CFO Act Agencies	Leading Indicators Strategy						Public Data Listing Metrics	
	Enterprise Data Inventory	Public Data Listing	Public Engagement	Privacy & Security	Human Capital	Use & Impact	Public Datasets	Valid Metadata
Department of Agriculture	☑	☑	★	☑	☑	☑	505	100%
Department of Commerce	○	○	▲	☑	☑	★	313	100%
Department of Defense	○	○	○	▲	☑	○	372	100%
Department of Education	○	☑	☑	☑	☑	★	275	100%
Department of Energy	☑	☑	○	☑	☑	★	1026	100%
Department of Health and Human Services	▲	▲	☑	☑	☑	☑	710	100%

This is a public dashboard showing how Federal agencies are performing on the Open Data Policy. [Learn more](#)

Previous Milestone: The milestone selected is the most recently complete one. The status of each field won't be final until a few weeks after the milestone has passed

Selected: Milestone 6 - February 28th 2015 ▾

	Public Datasets	Valid Metadata	Programs	Bureaus	Public Datasets	Restricted Datasets	Non-public Datasets	Datasets with downloads	Total Download URLs	Working Download URLs	Correct Format	HTML Downloads	PDF Downloads
Department of Agriculture	505	100%	28	18	99.4%	0.0%	0.6%	100%	630	83.7%	75.7%	56.7%	15.4%
Department of Commerce	313	100%	15	7	99.7%	0.0%	0.3%	100%	344	80.2%	49.6%	72.5%	1.1%
Department of Defense	372	100%	10	9	100%	0.0%	0.0%	100%	380	45.8%	86.8%	42.5%	33.9%
Department of Education	275	100%	1	10	92.7%	0.0%	7.3%	76.7%	722	97.6%	64.4%	7.1%	3.0%
Department of Energy	1040	100%	26	5	96.7%	0.0%	3.3%	96.5%	2307	96.9%	87.7%	7.0%	11.2%
Department of Health and Human Services	710	100%	11	11	97.3%	0.0%	2.7%	36.6%	295	53.9%	1.3%	81.8%	1.9%

The fields below serve as quality assurance to verify that the download links included within the metadata are functioning properly

To see a detailed breakdown of these issues, download the [full error log as a CSV](#)

Server Not Found	0.0% (0 of 206)
Working links (HTTP 2xx)	77.2% (159 of 206)
Broken links (HTTP 4xx)	3.9% (8 of 206)
Error Links (HTTP 5xx)	0.0% (0 of 206)
Redirected Links (HTTP 3xx)	18.0% (37 of 206)
Correct format	76.7% (122 of 159)
PDF for raw data	3.8% (6 of 159)
HTML for raw data	2.5% (4 of 159)
Bureaus Represented	5
Programs Represented	18
License Specified	100% (217 of 217)
Datasets with Redactions	0.0% (0 of 217)
Redactions without explanation (rights field)	0.0% (0 of 217)
File Size	327.58KB
Last modified	Wednesday, 07-Dec-2016 20:30:19 EST

Feedback Mechanisms

Search Data.Gov



DATA TOPICS - IMPACT APPLICATIONS DEVELOPERS CONTACT

DATA CATALOG / Datasets Organizations

Department of Education / Office of Planning, ...

Submit Data Report Data Issue



College Scorecard

Metadata Updated: Nov 25, 2016

College Scorecards make it easier for students to search for a college that is a good fit for them. They can use the College Scorecard to find out more about a college's affordability and value so they can make more informed decisions about which college to attend.

Access & Use Information

- Public: This dataset is intended for public access and use.
- License: Creative Commons CCZero

Downloads & Resources

College Scorecard **5059 views**
College Scorecard tool at the College Affordability and Transparency Center

Visit page

- Topics
- Education
- Publisher
- Office of Planning, Evaluation, and Policy Development
- Contact
- Brian Fu
- Share on Social Sites
- Google+
- Twitter

DATA REQUEST

Please use the following form to submit a new request for data that is not currently available on Data.gov. If you would like submit an issue about existing data on Data.gov, please use the [Data Issue](#) page. This is part of the Data.gov Help Desk which is designed to route requests to the appropriate person and allow you to track the status of requests. Existing requests can be tracked on the [Data Requests](#) page.

All information provided other than your name and email address will be publicly visible. If you would like to keep your request private, please use the main [contact form](#) instead. For additional details, please consult our [Privacy Policy](#).

New Data Request

Dataset Name (Required)

Dataset Description (Required)

Related Agency (Optional)

Philip

Data Issues - Data.gov

https://www.data.gov/issues/

Search Data.gov

 DATA TOPICS - IMPACT APPLICATIONS DEVELOPERS CONTACT

DATA ISSUES

The details and current status of data issues submitted through the [Data Issue Help Desk](#) page can be seen here.

Military Installations, Ranges, and Training Areas
closed Submitted on Saturday December 3, 2016 at 9:52 pm

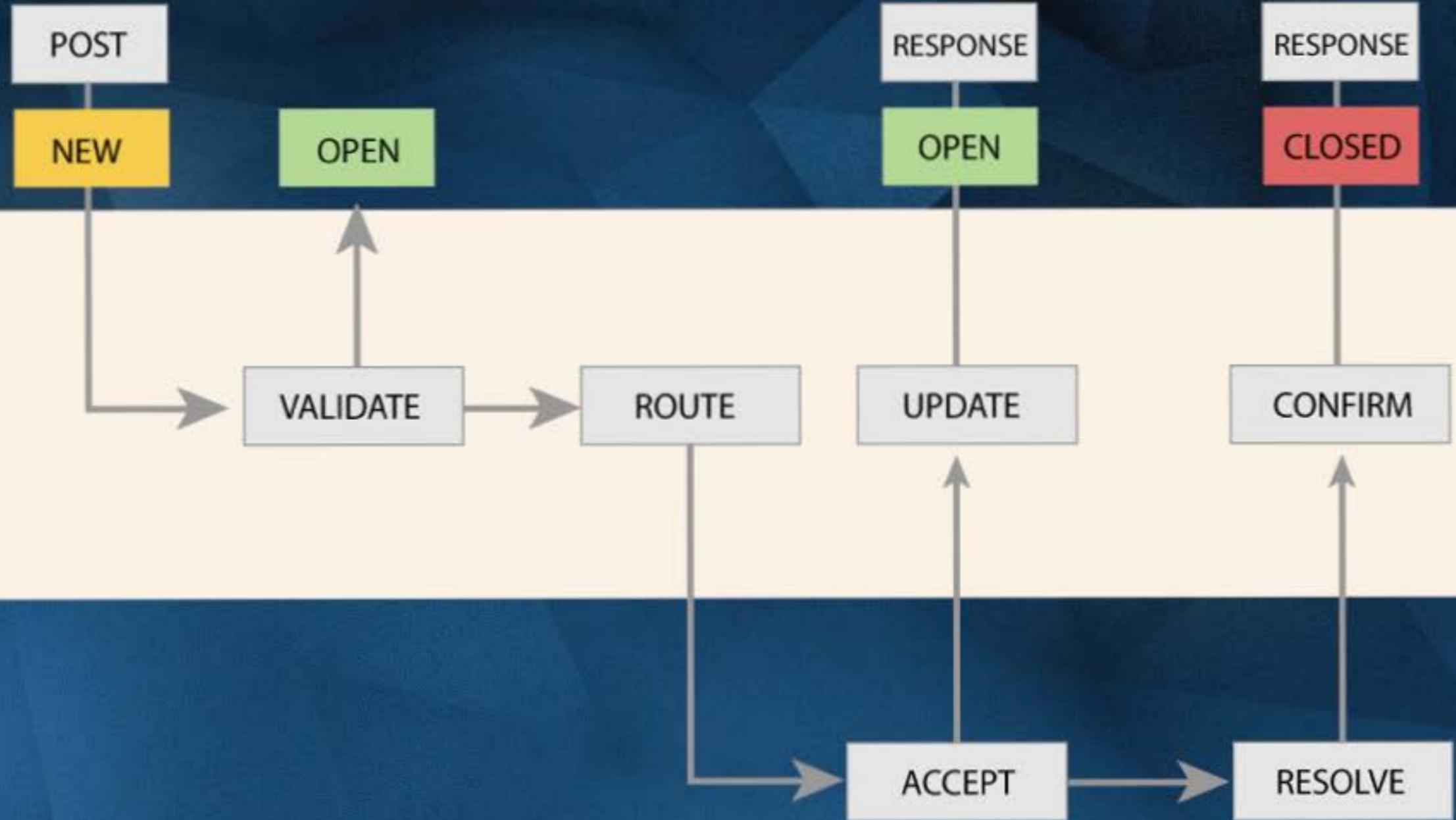
Page Not Found for data
open Submitted on Monday November 28, 2016 at 11:06 pm

File not found error
open Submitted on Monday November 28, 2016 at 8:02 pm

No access to the Safer company snapshot



 DATA.GOV





Demand-Driven Open Data



HHS IDEA LAB

ddod.healthdata.gov

***Data Documentation
&
Data Dictionaries***

Make field level metadata searchable and link common fields across the catalog #640

Edit New issue

Open philipashlock opened this issue on May 21, 2015 · 3 comments

 **philipashlock** commented on May 21, 2015 Member

We should be able to expose machine readable data dictionaries describing each field or variable in a dataset and index this as part of the catalog in order to more easily discover and understand each variable and provide linkages across disparate datasets.

This could also source metadata from resources linked from the `describedBy` field such as Swagger API specs and associated JSON Schema files. See <https://project-open-data.cio.gov/v1.1/apl/>

Some precedent with CKAN examples like <http://catalogue.data.gov.bc.ca/dataset/mta-mineral-placer-and-coal-titles> and related work with Data Packager <http://ckan.org/2014/06/09/the-open-knowledge-data-packager/> but ideally this wouldn't require data to be hosted within CKAN (DataStore) in order to provide the field level metadata.

This also relates to [project-open-data/project-open-data.github.io#465](https://github.com/project-open-data/project-open-data.github.io/issues/465) and [uscensusbureau/citysdk#100](https://github.com/uscensusbureau/citysdk/issues/100)

See also <http://www.w3.org/blog/data/2015/04/16/csv-on-the-web-seeking-comments-and-implementations/>

Assignees ⚙️

No one—assign yourself

Labels ⚙️

search

Projects

None yet

Milestone ⚙️

No milestone

Notifications

⏏ Unsubscribe



Object Description

Object Name: WHSE_MINERAL_TENURE.MTA_ACQUIRED_TENURE_POLY

Short Name: MTA_AT_PLY

Comments: The acquired tenure poly defines the spatial location of the acquired tenures. This layer contains multipart polygons. A tenure tenure is a form of ownership of mineral rights (mineral, placer or coal) over a parcel of land which has been located by a method set out in the legislation. The acquired tenure shape may be different from the located tenure shape due to previously located tenures or alienations.

Column Name	Short Name	Data Type	Data Precision	Comments
GEOMETRY	GEOMETRY	SDO_GEOMETRY	0	GEOMETRY is the column used to reference the spatial coordinates defining the feature.
FEATURE_AREA_SQM	null	NUMBER	19	FEATURE_AREA_SQM is the system calculated area of a two-dimensional polygon in square meters
FEATURE_LENGTH_M	null	NUMBER	19	FEATURE_LENGTH_M is the system calculated length or perimeter of a geometry in meters
TENURE_NUMBER_ID	TEN_NO_ID	NUMBER	10	Unique number of mineral tenure
MTA_TENURE_TYPE_CODE	TEN_TYP_CD	VARCHAR2	0	Tenure type code for a Mineral Tenure: C=Coal; M=Mineral; P=Placer.
REVISION_NUMBER	REVSNO	NUMBER	2	Revision Number to track the number of



Metadata Schema for Data Dictionaries

SCOPE Metadata Group

Jan 16, 2015 draft 1.01

Background:

The SCOPE (Statistical Community of Practice and Engagement) project was initiated in 2009 as an inter-agency effort to consolidate resources and practices across the major US federal statistical agencies. A team focused on metadata was initiated in 2013. The effort described here is part of the team's deliverables.

Parallel with SCOPE, the White House began an effort to make federal government data available through a single resource on the Web known as Data.Gov. As part of the effort to build a unified Data.gov, the federal agencies have been asked to standardize the format of certain metadata about the datasets they offer online, to enable public users and computer programs to find and use the data.

One element in the set of descriptors for a data set is a pointer to a data dictionary, if applicable. This specification is an effort to define the elements needed to describe the contents of a data dictionary.

*Search Engines
&
Schema.org*

Person

[Thing](#) > [Person](#)

A person (alive, dead, undead, or fictional).

Usage: Over 1,000,000 domains

[\[more...\]](#)

Property	Expected Type	Description
Properties from Person		
additionalName	Text	An additional name for a Person, can be used for a middle name.
address	PostalAddress or Text	Physical address of the item.
affiliation	Organization	An organization that this person is affiliated with. For example, a school/university, a club, or a team.
alumniOf	EducationalOrganization or Organization	An organization that the person is an alumni of. Inverse property: alumni .
award	Text	An award won by or for this item. Supersedes awards .
birthDate	Date	Date of birth.
birthPlace	Place	The place where the person was born.
brand	Brand or	The brand(s) associated with a product or service, or the brand(s) maintained by an organization or

About 474,000 results (0.40 seconds)

Congressman Derek Kilmer | Representing the 6th District of ...
kilmer.house.gov/ United States House of Representatives
Official site of U.S. Representative **Derek Kilmer** (WA-6), Democratic party.
Contact · Full Biography · About · Email Me

Email Me | Congressman Derek Kilmer
<https://kilmer.house.gov/contact/email-me> United States House of Representatives
Step One - Zip Authentication. * marks required fields of data. Your Zip Code. Regrettably, I am unable to reply to any email from constituents outside of ...

Contact | Congressman Derek Kilmer
<https://kilmer.house.gov/contact> United States House of Representatives
I value all feedback from residents of the 6th Congressional District of Washington. I am here to serve you, and so is my staff. Below please find the best means ...

Rep. Derek Kilmer (@RepDerekKilmer) | Twitter
<https://twitter.com/RepDerekKilmer>
3 days ago - View on Twitter
During a dark week it's important for LGBT Americans to know #We'rewithyou: bit.ly/1Ue8cq3
3 days ago - View on Twitter
Continuing my work to make sure that the Navy continues to be a better neighbor: bit.ly/1XtnU6m



Derek Kilmer
U.S. Representative

kilmer.house.gov

Derek Christian Kilmer is an American politician who has been the U.S. Representative for Washington's 6th congressional district since 2013. [Wikipedia](#)

Born: January 1, 1974 (age 42), Port Angeles, WA
Spouse: Jennifer Kilmer
Office: Representative (D-WA 6th District) since 2013
Residence: Artondale, WA
Education: University of Oxford (1999), Woodrow Wilson School of

Request

```
GET https://kgsearch.googleapis.com/v1/entities:search?query=Derek+Kilmer&key={YOUR_API_KEY}
```

Response

200

- Show headers -

```
- {
  "@context": {
    "@vocab": "http://schema.org/",
    "goog": "http://schema.googleapis.com/",
    "EntitySearchResult": "goog:EntitySearchResult",
    "detailedDescription": "goog:detailedDescription",
    "resultScore": "goog:resultScore",
    "kg": "http://q.co/kg"
  },
  "@type": "ItemList",
  "itemListElement": [
    - {
      "@type": "EntitySearchResult",
      - "result": {
        "@id": "kg:/m/03w9yd6",
        "name": "Derek Kilmer",
        - "@type": [
          "Person",
          "Thing"
        ],
        "description": "U.S. Representative",
        - "image": {
          "contentUrl": "http://t1.gstatic.com/images?q=tbn:ANd9GcSzLr1I3AMIPWwAoSthaaXcbUM75VvKTH7jcgToYOxKh8Q37AQN",
          "url": "https://en.wikipedia.org/wiki/Derek_Kilmer"
        },
        - "detailedDescription": {
          "articleBody": "Derek Christian Kilmer is an American politician who has been the U.S. Representative for Washington's 6th
```

Dataset

[Thing](#) > [CreativeWork](#) > [Dataset](#)

A body of structured information describing some topic(s) of interest.

Usage: Between 100 and 1000 domains

[\[more...\]](#)

Property	Expected Type	Description
Properties from Dataset		
datasetTimeInterval	DateTime	The range of temporal applicability of a dataset, e.g. for a 2011 census dataset, the year 2011 (in ISO 8601 time interval format). Supersedes temporal .
distribution	DataDownload	A downloadable form of this dataset, at a specific location, in a specific format.
includedInDataCatalog	DataCatalog	A data catalog which contains this dataset. Supersedes catalog , includedDataCatalog . Inverse property: dataset .
spatial	Place	The range of spatial applicability of a dataset, e.g. for a dataset of New York weather, the state of New York.
Properties from CreativeWork		
about	Thing	The subject matter of the content.
accessibilityAPI	Text	Indicates that the resource is compatible with the referenced accessibility API (WebSchemas wiki lists possible values).
accessibilityControl	Text	Identifies input methods that are sufficient to fully control the described resource (WebSchemas wiki lists



Google Research Blog

The latest news from Research at Google

Facilitating the discovery of public datasets

Tuesday, January 24, 2017

Posted by Natasha Noy, Google Research and Dan Brickley, Open Source Programs Office

There are many hundreds of data repositories on the Web, providing access to tens of thousands—or millions—of datasets. National and regional governments, scientific publishers and consortia, commercial data providers, and others [publish data](#) for fields ranging from [social science](#) to [life science](#) to [high-energy physics](#) to [climate science](#) and more. Access to this data is critical to facilitating reproducibility of research results, enabling scientists to build on others' work, and providing data journalists easier access to information and its provenance. For these reasons, many publishers and funding agencies now require that scientists make their research data available publicly.

However, due to the volume of data repositories available on the Web, it can be extremely difficult to determine not only where is the dataset that has the information that you are looking for, but also the veracity or provenance of that information. Yet, there is no reason why searching for datasets shouldn't be as easy as searching for recipes, or jobs, or movies. These types of searches are often

 Labels 

 Archive 

 Feed

Google on 

 Follow @googleresearch

Give us feedback in our [Product Forums](#).

Science Datasets | Search | X Philip

Secure https://developers.google.com/search/docs/data-types/datasets

Google Search Home Docs Tools Partner Access Search All Products

Documentation GUIDES DATA TYPES SEND FEEDBACK

Science Datasets ☆☆☆☆☆

The web contains specialized repositories for datasets in many scientific domains: life sciences, earth sciences, material sciences, and more. Similarly, many governments maintain repositories of civic and government data. However, much of that structured data is not readily available to search engines, which must extract the data from HTML pages in order to provide search services to users. When webmasters provide [structured markup](#), they enable search engines to “understand” this metadata, which in turn improves data discovery, leading scientists to the information they need for their work.

For example, consider this dataset that describes [historical snow levels in the Northern Hemisphere](#). This page contains basic information about the data, like spatial coverage and units. Other pages on the site contain additional metadata: who produces the dataset, how to download it, and the license for using the data. With structured data markup, these pages can be more easily discovered by other scientists searching for climate data in that subject area.

 **Dataset markup is available for you to experiment with** before it's released to general availability. When you implement the markup, you'll be able to test it in the Structured Data Testing Tool. You won't, however, see your datasets appear in Search.

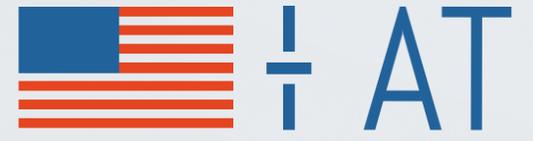
Contents

- What qualifies as a dataset?
- Mark up your dataset descriptions
 - Basic dataset properties
 - Data catalog properties
 - Download information properties
 - Temporal coverage
 - Spatial coverage
- Citations and publications
- Provenance and license information
- Complete Example: Datasets markup in JSON-LD
- Source and Provenance Guidelines

Authorized Presence
Corporate Contacts
Logos
Social Profile Links

Creative Works
Articles
Books
Courses
Music
Recipes
Reviews
TV & Movies
Videos

Commerce
Local Businesses
Events
Products



data.gov/contact

[@usdatagov](https://twitter.com/usdatagov)