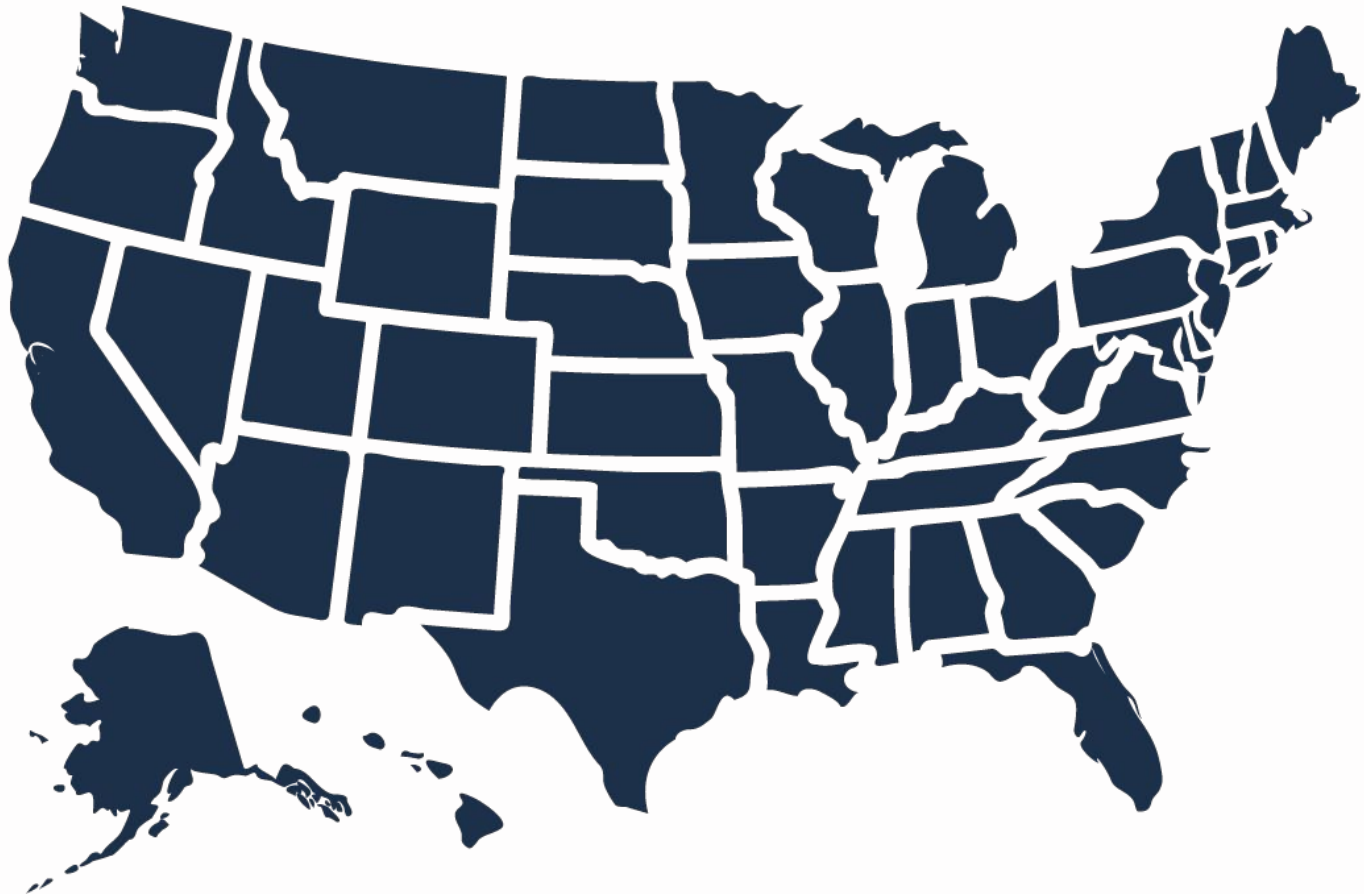


1/ Overview





Current U.S. Federal Statistical System



NOTE: OIRA/OMB is Office of Information and Regulatory Affairs/Office of Management and Budget.



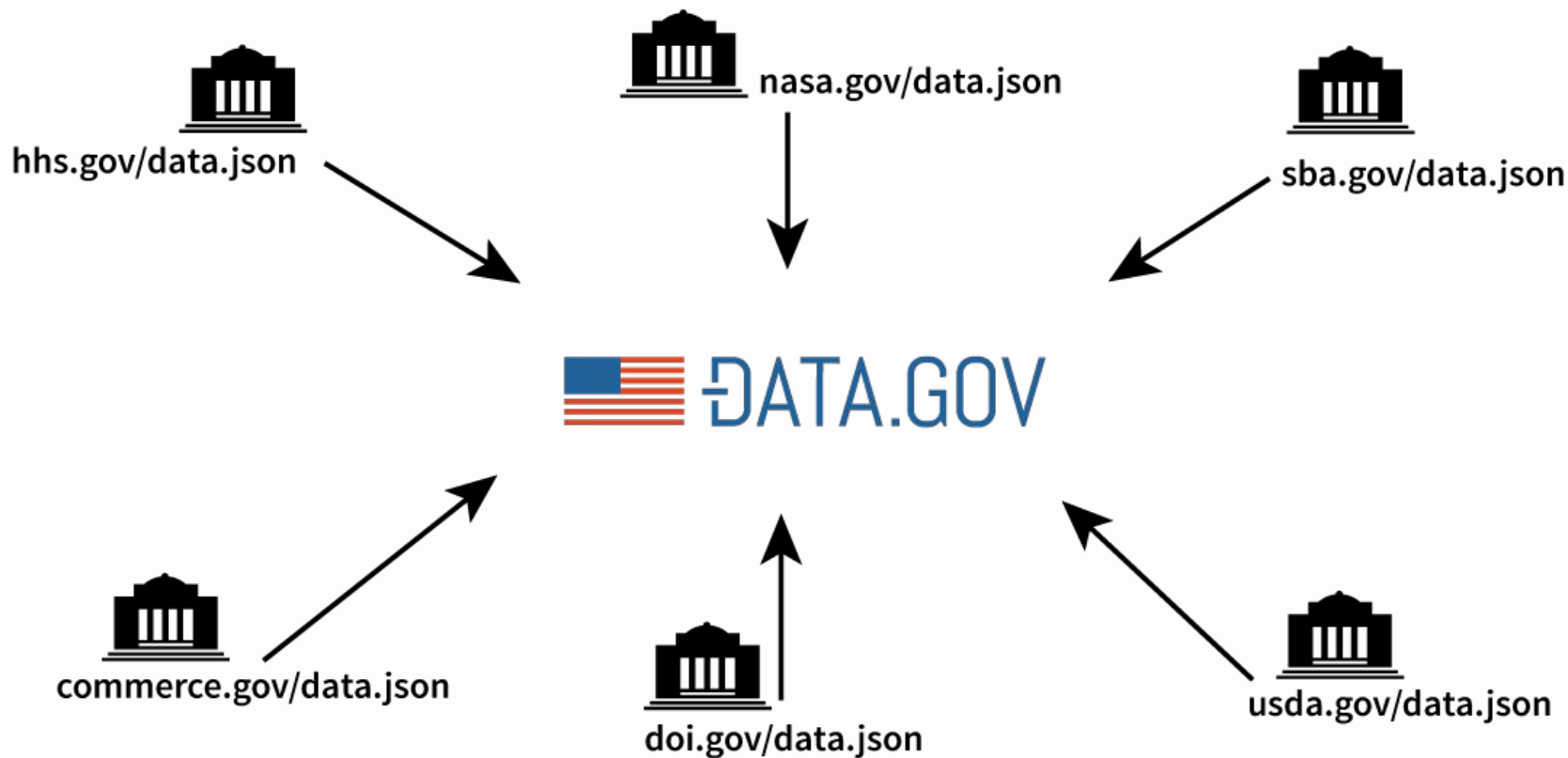
fed·er·a·tion

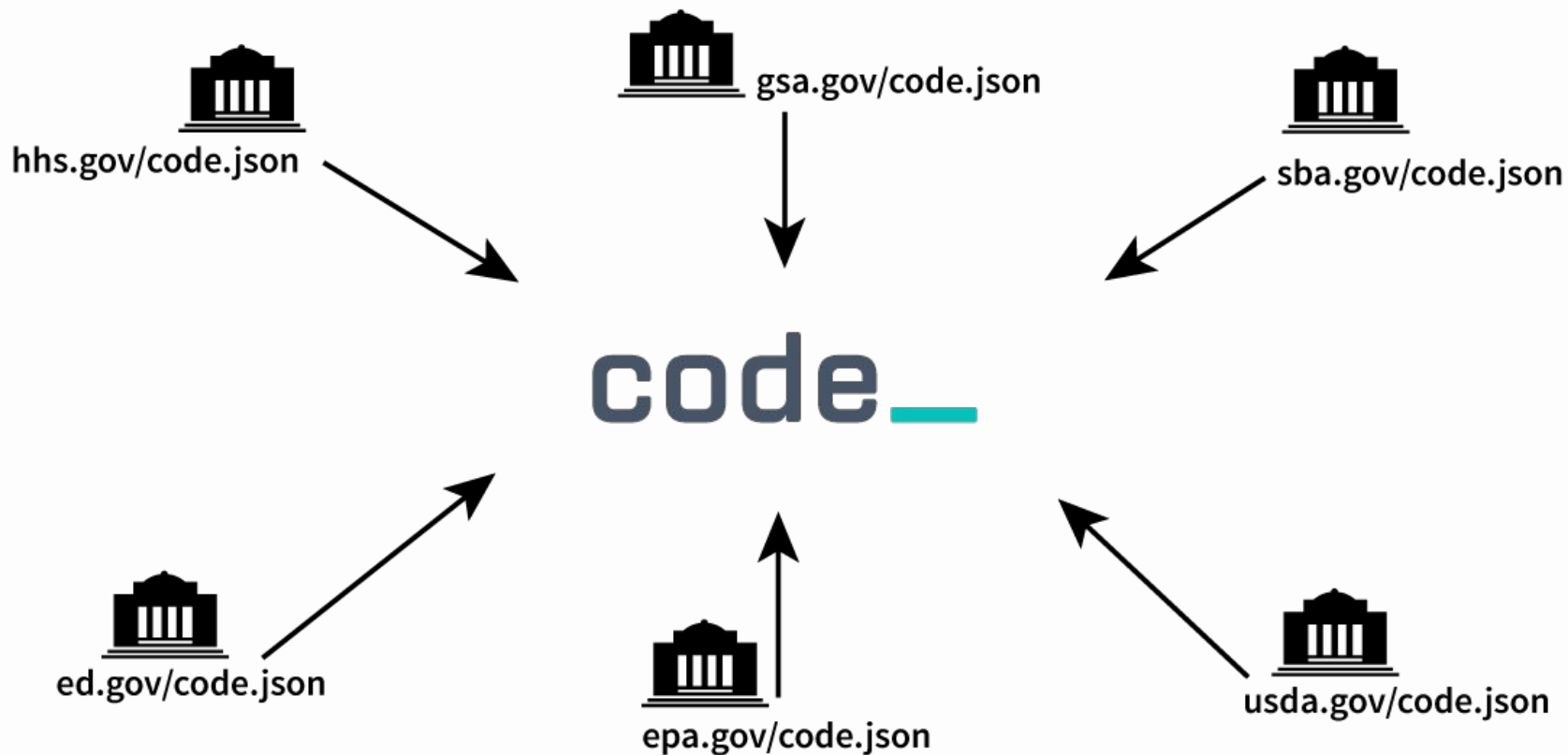
/ˌfedəˈrāSH(ə)n/

noun

1. a group of states with a central government but independence in internal affairs
2. a group of computing or network services agreeing upon standards of operation in a collective fashion







Project Open Data Metadata S x

Philip

Secure | https://project-open-data.cio.gov/v1.1/schema/

★

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](https://project-open-data.cio.gov/v1.1/schema/catalog.json) 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](https://project-open-data.cio.gov/v1.1/schema/catalog.jsonld) 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>dcatalog: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>dcatalog: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
<code>description</code>	Description	Human-readable description (e.g., an abstract) with sufficient detail to enable a user to quickly understand whether the asset is of interest.	Always
<code>keyword</code>	Tags	Tags (or keywords) help users discover your dataset; please include terms that would be used by technical and non-technical users.	Always
<code>modified</code>	Last Update	Most recent date on which the dataset was changed, updated or modified.	Always
<code>publisher</code>	Publisher	The publishing entity and optionally their parent organization(s).	Always
<code>contactPoint</code>	Contact	Contact person's name and email for the asset.	Always



Project Open Data Dashboard

Philip

Secure | https://labs.data.gov/dashboard/validate

Project Open Data DashboardAgenciesValidatorConverters ▾RubricHelp ▾About

Sign in with MAX

Validator

There are three ways you can validate data.json, either by validating a public URL, uploading a json file, or pasting the raw JSON into the form.

Validate data.json URL

SchemaFederal v1.1 ▾

☒ View in Browser☐ Output JSON

data.json URL

e.g. http://energy.gov/data.json

Validate URL

Validate data.json file upload

SchemaFederal v1.1 ▾

Upload a data.json file

Choose FileNo file chosen

Validate File

Validate raw JSON

data.json JSON



Creating your enterprise code x Philip

Secure | <https://code.gov/#/policy-guide/docs/compliance/inventory-code>

code_ ABOUT BROWSE PROJECTS FEDERAL AGENCIES DEVELOPERS COMMUNITY

Metadata Schema version 2.0.0

Version 2.0.0 of the [metadata schema file in JSON format can be found here](#).

File location and contents:

`code.json` must live in the root directory of your agency's website.
`code.json` must include a single object represented as JSON, with key-value pairs according to the list below.

Fields:

Required

version: [string] Code.gov's version of the metadata schema in use. Implements semantic versioning 2.0.0 rules as defined at <http://semver.org>. For example "2.0.0".

measurementType: [object] An object containing description of the open source measurement method.

method: [enum] An enumerated list of methods for measuring the open source requirement.

- cost**: Cost of software development.
- systems**: System certification and accreditation boundaries.
- projects**: A complete software solution / project.
- modules**: A self-contained module from a software solution.
- linesOfCode**: Source lines of code.
- other**: Another measurement method not referenced above.

ifOther: [string] A one- or two- sentence description of the measurement type used, if 'other' is selected as the value of 'method' field.

agency: [string] The agency acronym for Clinger Cohen Act (CCA) agencies can be found through the [CFO Act list](#), as CCA references the CFO Act. For example "GSA" or "DOD." Please use uppercase letters.

releases: [array] Contains objects representing each versioned source code release made available.

name: [string] The name of the release.

repositoryURL: [string] The URL of the public release repository for open source repositories. This field is not required for repositories that are only available as government-wide reuse or are closed (pursuant to one of the exemptions).

description: [string] A one- or two-sentence description of the release.

permissions: [object] An object containing description of the usage/restrictions regarding the release.

licenses: [null or array of objects] An object containing license details, if available. If not, **null** should be used.

URL: [string] The URL of the release license, if available.

name: [string] An abbreviation for the name of the license. For example, 'CC0' or 'MIT.'

usageType: [enum] A list of enumerated values which describes the usage permissions for the release.



code.gov

Secure | <https://code.gov/#/policy-guide/docs/compliance/inventory-code/tools/validate-schema>

code_

ABOUT BROWSE PROJECTS FEDERAL AGENCIES DEVELOPERS COMMUNITY

How To Read the Policy

Overview

Compliance

Agency Compliance

How to Procure Software

How to Inventory Code

Validate Schema

Upgrade Schema

Open Source Pilot

Codejson Validator

Please enter your code.json file below in order to validate it correctly meets the specification.

Codejson

Version 2.0.0

1

Problems (1)

[Expected a JSON object, array or literal \(1..1\)](#)



Case Studies & Interviews

DATA Act

data.gov

code.gov

Voting Information Project

Open Civic Data

OpenReferral

NIEM

FGDC

Rachel Bloom

Andrew Nicklin

James McKinney

Christian Hoehner

Mark Headd

Wo Chang

Tyler Kleykamp

Christina Bapst

Tools & Infrastructure

- Validation
- Aggregation

Guidance & Playbooks

- Documentation
- Community & Governance

Validation

Structural integrity, syntax and semantics, complex logic

Aggregation

Push, pull, automated, manual, etc

**Getting good,
reliable data is hard.**

**Data from different
sources* makes
this even harder.**

* otherwise known as “federated data”

Data providers are less likely to contribute to systems that are difficult and time consuming to use.

2/ Use case



US Department of Agriculture

USDA wants to start small and build out a federated data service with just one form, starting with data on school lunch programs collected at the school district. The tool will validate data as it is entered, starting with a pilot across a couple states.

Department of Agriculture, Food and Nutrition Service School Food Authority (SFA) Verification Collection Report			
State agencies must report the information on this form ANNUALLY for each SFA with schools operating the National School Lunch Program (NSLP) and/or the School Breakfast Program (SBP).			
All SFAs, including SFAs with all schools exempt from verification requirements, must complete applicable sections.			
According to the Paperwork Reduction Act of 1995, no person is required to respond to a collection of information unless it contains a valid OMB control number. The valid OMB number for this collection is 0548-0046. The time required to complete this information collection is 45 minutes per response, including the time to review instructions, search existing data resources, gather the data needed and complete and review the information collection.			
State Agency Name: SFA ID#		Type of SFA: <input type="checkbox"/> Public <input type="checkbox"/> Nonprofit/Private	School Year: From: 20 To: 20
SFA Name:		SFA City:	SFA Zip code: <input type="text"/> <input type="text"/> <input type="text"/>
Section 1	Total Schools, Resident Child Care Institutions (RCCI) and Excluded Students	A. Number of Schools OR Institutions	B. Number of Students
	1-1: Total schools (Do not include RCCIs)	1-1A	1-1B
	1-2: Total RCCIs (Do not include schools counted in 1-1)	1-2A	1-2B
	1-3: RCCIs (Do not include schools counted in 1-2B)	1-3A	1-3B
Section 2	SFA with alternate provisions must report Section 2	A. Number of Schools OR Institutions	B. Number of Students
	2-1: Operating Provision 23 in a BASE year for NSLP and SBP	2-1A	2-1B
	2-2: Operating Provision 23 in a NON BASE year for NSLP and SBP	2-2A	2-2B
	2-3a: Provision 23 students reported as FREE in a NON BASE year	2-3aB	2-3aC
	2-3b: Provision 23 students reported as REDUCED PRICE in a NON BASE year	2-3bB	2-3bC
	2-4: Operating the Community Eligibility Option	2-4A	2-4B
Section 3	Students approved as FREE eligible NOT subject to verification	A. Number of Schools OR Institutions	B. Number of FREE Students
	3-1: Students directly certified through Supplemental Nutrition Assistance Program (SNAP). Do not include students certified with SNAP through the letter method.	3-1A	3-1B
	3-2: Students directly certified through other programs: include those directly certified through Temporary Assistance for Needy Families (TANF), Food Distribution Program on Indian Reservations (FDPIR), or Medicaid (if applicable); those documented as homeless, runaway, foster, street, blind, free or even blind, or non-applicant but approved by local officials. DO NOT include SNAP students already reported in 2-3.	3-2A	3-2B
	3-4: Students certified categorically FREE eligible through SNAP letter method. Include students certified for free meals through the family providing a letter from the SNAP agency.	3-4A	3-4B
Section 4	Students approved as FREE or REDUCED PRICE eligible through household application	A. Number of Applications	B. Number of Students
	4-1: Approved as categorically FREE Eligible: (Based on those providing documentation (e.g. a case number for SNAP, TANF, FDPIR on an application)	4-1A	4-1B
	4-2: Approved as FREE eligible: (Based on household size and income information)	4-2A	4-2B
	4-3: Approved as REDUCED PRICE eligible: (Based on household size and income information)	4-3A	4-3B
T-1: Total FREE Eligible Students Reported:		T-1	T-2: Total REDUCED PRICE Eligible Students Reported:

Form FNS-742 (10/12) Previous Editions are Obsolete

SBU

Electronic Form Version Designed in Adobe 10.0 version

S-1: Check the box if ALL schools and/or RCCIs are exempt from verification (see instructions for list of exemptions). If S-1 is checked, no further reporting in Section 3 is required.			
S-2: Was verification performed and completed?		S-3: Type of Verification process used:	
<input type="checkbox"/> Yes, completed by November 15th, S-2 yes-by		<input type="checkbox"/> Standard (Lesser of 5% or 2,000 error-prone) S-3-1	
<input type="checkbox"/> Yes, completed after November 15th, S-2 yes-after		<input type="checkbox"/> Alternate one (Lesser of 3% or 1,000 selected randomly) S-3-2	
<input type="checkbox"/> No, verification was NOT performed or the process was not completed. S-2 no		<input type="checkbox"/> Alternate two (Lesser of 1% or 1,000 error-prone applications PLUS lesser of one-half of one percent of 500 applications with SNAP/TANF/FDPIR case numbers) S-3-3	
S-4: Total ERROR PRONE applications reported S-4	S-5: Number of applications selected for verification sample	S-6: A. Number of Applications	S-6: B. Number of Students
S-4: Total ERROR PRONE applications reported S-4	S-5: Number of applications selected for verification sample	S-6: A. Number of Applications	S-6: B. Number of Students
Section 5: Report FREE and/or REDUCED PRICE eligible as confirmed through direct verification with SNAP/TANF/FDPIR/MEDICAID as of November 15th.			
S-7: Confirmed through direct verification			
S-8: Results of Verification by Original Benefit Type			
For each original benefit type (A, B, & C), report the number of applications and students as of November 15th for each meal category (1, 2, 3, & 4). Do NOT include students and applications already reported in S-7A or S-7B.			
A. FREE-Categorically Eligible: Confirmed as FREE based on SNAP/TANF/FDPIR documentation (e.g. case number) or application		B. FREE-Income: Confirmed as FREE based on household size application	
Result Category		Result Category	
1. Responded, NO CHANGE	5-8-A1a	1. Responded, NO CHANGE	5-8-B1a
2. Responded, Changed to FREE	5-8-A2a	2. Responded, Changed to FREE	5-8-B2a
3. Responded, Changed to REDUCED PRICE	5-8-A3a	3. Responded, Changed to REDUCED PRICE	5-8-B3a
4. NOT Responded, Changed to FREE	5-8-A4a	4. NOT Responded, Changed to FREE	5-8-B4a
VC-1: Total questionable applications verified for cases (Enter "N/A" if not applicable)		VC-1	

certified for free meals through the family providing a letter from the SNAP agency.

Section 4	Students approved as FREE or REDUCED PRICE eligible through a household application	**ALL SFAs collecting applications must report Section 4**	
		A. Number of Applications	B. Number of Students
		4-1A	4-1B
		4-2A	4-2B
		4-3A	4-3B

T-1: Total FREE Eligible Students Reported:

T-1

T-2: Total REDUCED PRICE Eligible Students Reported:

T-2

Critical sources of data need to be accurate.

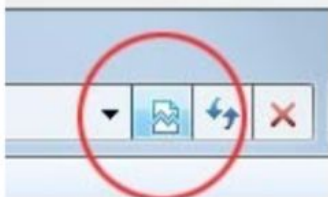
Policy

Legislation

Research

NOTE: If your SFA does not operate an alternate provision such as **Community Eligibility Provision (CEP) or Provision 2**, check this box and move on to section 3.

If you check this box, you will get a pop up message. The pop up must be answered for you to continue to the next section. If you do not see the pop up: Turn off your pop up blocker OR if you are using Internet Explorer 10, click on the “Compatibility View” icon.



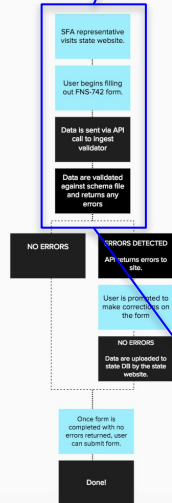
ONLY SFAs with alternate provisions must report Section2	A.Number of Schools AND Institutions	B.Number of Students
<input type="checkbox"/> Check if SFA does not operate under CEP or Provision 2		
2-1: Operating Provision 2/3 in a BASE year for NSLP and SBP:	<input type="text"/>	<input type="text"/>
2-2: Operating Provision 2/3 in a NON BASE year for NSLP and SBP:	<input type="text"/>	<input type="text"/>
2-2a: Provision 2/3 students reported as FREE in a NON BASE year:		<input type="text"/>
2-2b: Provision 2/3 students reported as REDUCED PRICE in a NON BASE year:		<input type="text"/>
-3: Operating the Community Eligibility Provision:	<input type="text"/>	<input type="text"/>
2-4: Operating Other alternatives for NSLP and SBP:	<input type="text"/>	<input type="text"/>
2-5: Operating an alternate provision(s) for only SBP or only NSLP:	<input type="text"/>	<input type="text"/>

**We can't expect 50 states
to code and maintain
100+ validation checks in
their web systems.**

**The data service
builds on what states
are already doing.**

USDA Case Study - FNS 742 form

USDA Case Study - FNS 742 form



SFA representative visits state website.

...

User begins filling out FNS-742 form.

...

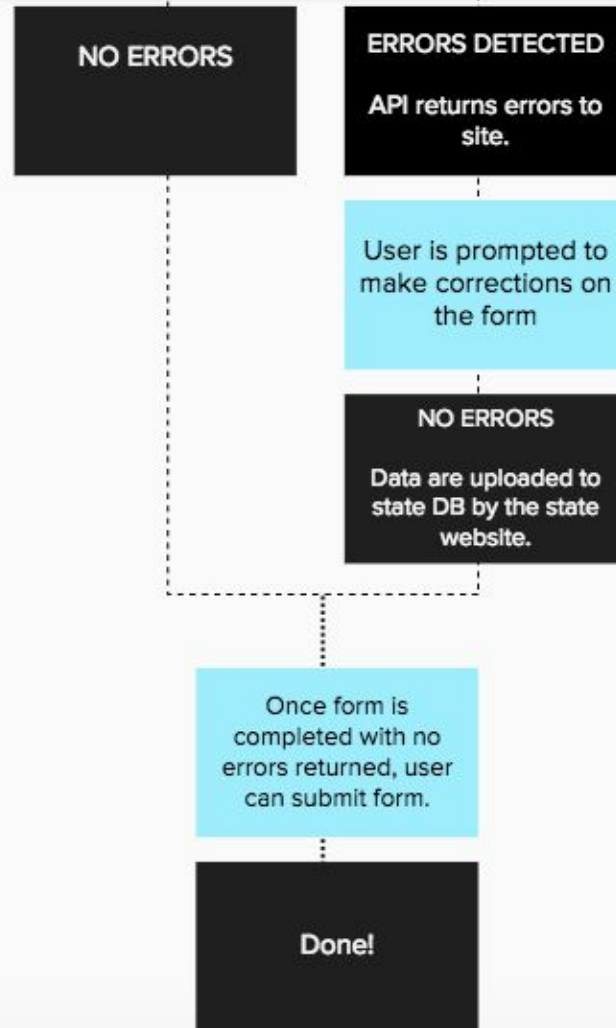
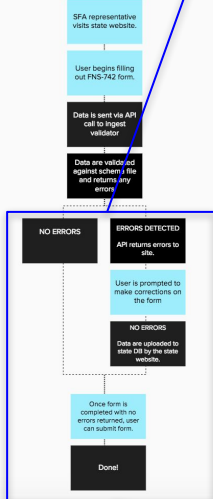
Data is sent via API call to ingest validator

...

Data are validated against schema file and returns any errors

...

USDA Case Study - FNS 742 form



3/ DEMO



4/ PHASE III



Plans for Phase III funding:

1

Design a service and software offering. Build out the tool to address other critical federated data efforts by focusing on the next use cases.

2

Integrate an option into data.gov / cloud.gov as a paid hosting service. This provides a mechanism for GSA to recover operating costs.

3

Use these tools to promote the federated data playbook and best practices throughout government. Focus first on early adopters. Scale going forward.

Phase III Milestones

Timeline	Product	Engineering
Week 1 - 4	Design a service and software offering. Build out the tool to address other critical federated data efforts by focusing on the next use cases. Set up hosting package option with the cloud.gov team.	Build out the API feature set. Building off of Phase II use case, further develop the API feature set in partnership with USDA and other use case partners interested in leveraging the API to improve existing data ingest systems. From a user perspective, these will continue to be systems that will still look the same but provide smarter data validation on top of existing applications.
Week 5 - 8	Develop the pricing and cost recoverability strategy. Develop a pricing model that prices both the service and software product for government partners. Offer as an option to existing data-centric engagements.	Build out the UI feature set. Building off of the Phase II use case, this work will include development of the user interface workflow use case. This will also involve preliminary steps for setting up the cloud.gov hosted solution for partners that want to deploy on our infrastructure.

Phase III Milestones, cont'd

Timeline	Product	Engineering
Week 9 - 12	Enter pre-sales and service agreements with two partners. By the end of week 12 the team will have two agreements in place with prototype partners to stand up service.	Focus on building MVPs for two use cases. This work will be focused on further validating and building the initial product offering using real use case examples.
Week 12 - 20	Enter formal sales and service agreements with two additional partners. Stand up supportive service offerings.	Service and tool consulting and development of two demonstrable examples of the product working in the field. These formal, deployed examples become the core product offerings.
Week 20 - 30	Formal launch of services to support ingest tool deployment and / or acquisition.	Formal launch of enhanced data ingest tool and hosting services for federal partners.

Potential use case: Data.gov / Office of Management and Budget

OMB needs better ways to validate data when submitted through web interfaces, and provide other agencies more transparency across initiatives. Current systems are outdated and fragmented. If agencies could coordinate together across initiatives data would become more immediately actionable.

Thanks!