

THE UNITED STATES
DATA FEDERATION



What it is

THE GOAL

We want to make managing
federated data projects
in government **easier.**

federated data projects

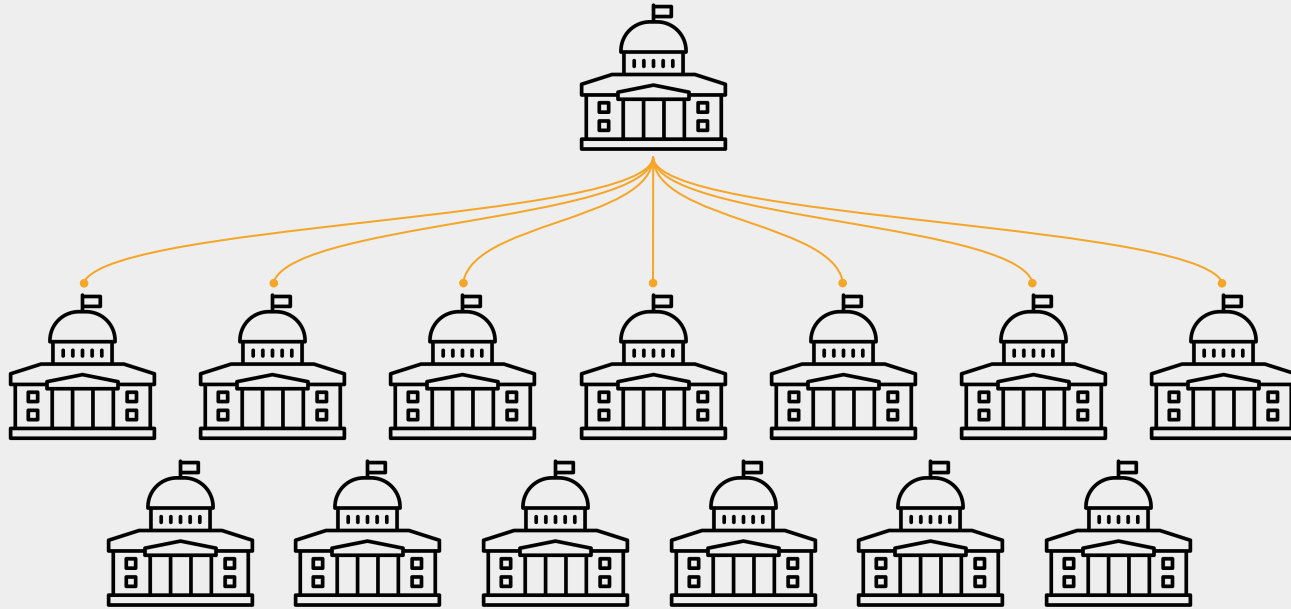
=

efforts in which a **common type of data**
is collected or exchanged across
disparate organizational boundaries

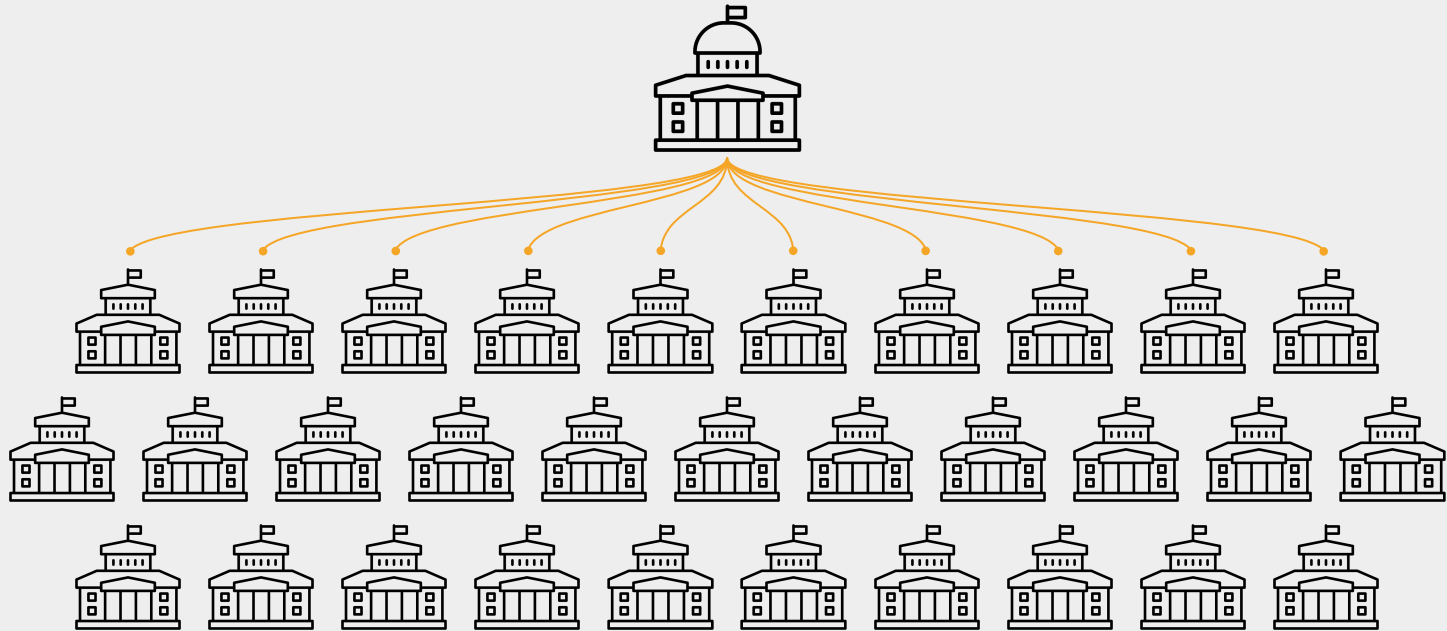


In our distributed style of
government, this happens
all the time.

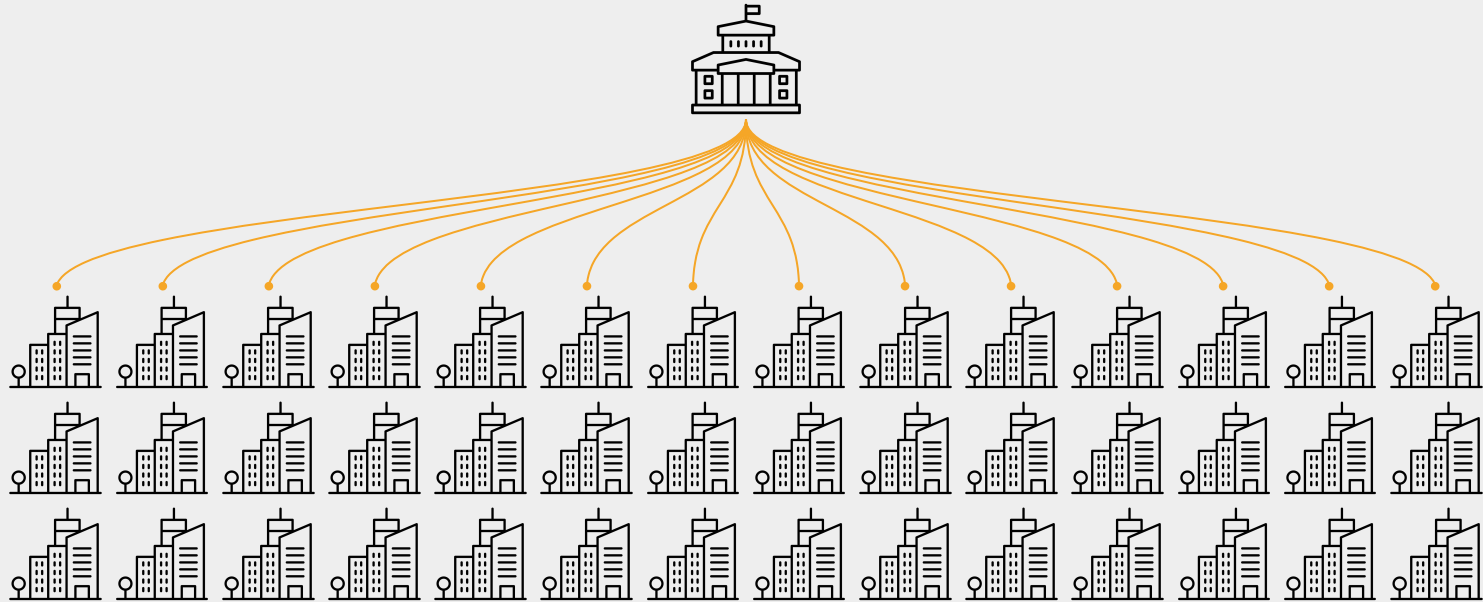
It happens when a federal agency
collects data from other federal agencies



It happens when a federal agency
collects data from states



It happens when states
collect data from municipalities



These data may

- inform policy and budget decisions,
- support operational efficiencies, or
- be published in aggregate form for other data users.

THE PROBLEM

Making it easy and convenient for data providers to submit their data while maintaining quality, accuracy, and completeness can often be a challenge

ANOTHER PROBLEM

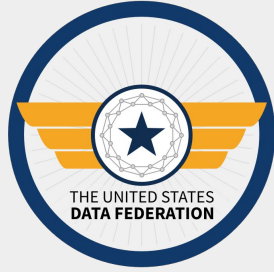
Most federated data projects are done in isolation and lack infrastructural support or repeatable best practices

THE IDEA BEHIND THE U.S. DATA FEDERATION

Develop **reusable** tooling,
processes, and resources to
support federated data projects



Our work



The United States Data Federation is an initiative of the GSA Technology Transformation Services (TTS) 10x program, which funds technology-focused ideas from federal employees with an aim to improve the experience all people have with our government.

PHASE ONE

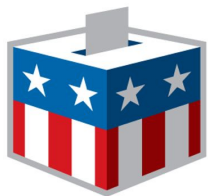
Understand the problem space

We interviewed examples of successful federated data projects



DATA.GOV

code_



Open311

We developed a maturity model

Dimensions		Beginnings	Growth	Mature
Impetus	<i>The business or legal need that sets the project in motion</i>	Grassroots	Policy	Law
Community	<i>The people who provide and consume the data</i>	First Adopters	Trending	Maintenance
Specification	<i>The definition of what data needs to be provided and what format it should be in</i>	Dictionary	Machine Readable	Standardized
Application	<i>A software application optimized for demonstrating and providing the value of the data to end users</i>	Alpha	Beta	Production

Details available on [GitHub](#).

We assembled a playbook

1. Policy should be focused on processes and outcomes, not implementation
2. Identify use cases with demonstrated demand
3. Develop a killer app
4. Allocate proper resources
5. Implementation should be driven by a single empowered team
6. When possible, deliver value directly to data owners
7. Start with simple technologies
8. Nurture early adopters
9. Support compliance tooling

Details available on [GitHub](#).

PHASE TWO

**Pick a use case &
build a prototype**

We partnered with the USDA Food & Nutrition Service



Build out a data service for form FNS-742,
which collects data from about 20,000
school districts across the country on the
school lunch and breakfast programs

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 persons are required to respond to a collection of information unless it contains a valid OMB control number. The valid OMB number for this collection is 0584-0026. 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.

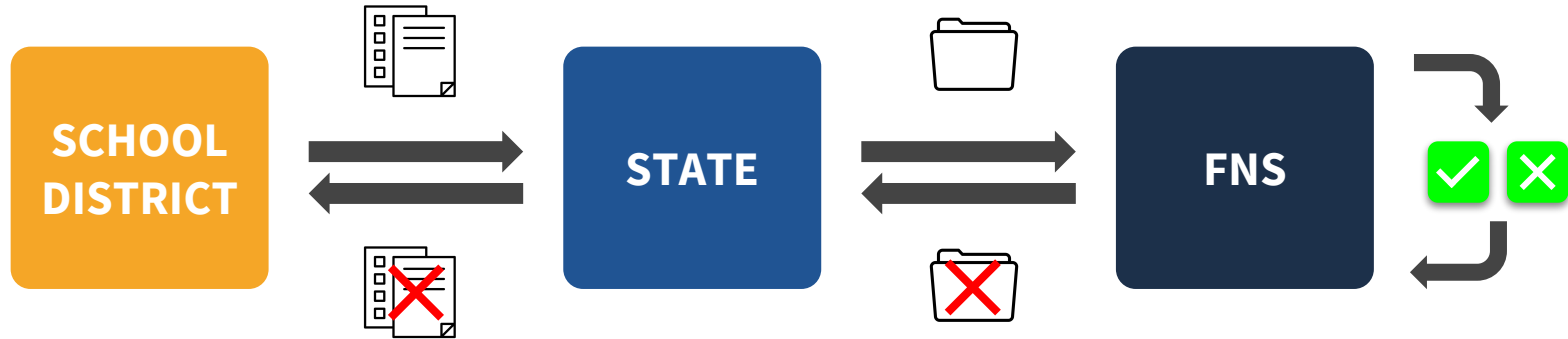
State Agency Name:	SFA ID#:	Type of SFA: <input type="checkbox"/> Public <input type="checkbox"/> Nonprofit/Private	School Year: From: 20 To: 20
SFA Name:		SFA City:	SFA Zip code: <table border="1" style="display: inline-table; width: 100px; height: 20px; vertical-align: middle;"></table>

Section 1	Total Schools, Residential Child Care Institutions (RCCIs) and Enrolled Students	**All SFAs must report Section 1**	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
		1-2a: RCCIs with day students (Report ONLY day students in 1-2aB):	1-2aA	1-2aB
		1-2b: RCCIs with NO day students:	1-2bA	1-2bB

Section 2	SFAs with schools operating alternate provisions	**ONLY SFAs with alternate provisions must report Section 2**	A. Number of Schools AND Institutions	B. Number of Students
		2-1: Operating Provision 2/3 in a BASE year for NSLP and SBP:	2-1A	2-1B
		2-2: Operating Provision 2/3 in a NON BASE year for NSLP and SBP:	2-2A	2-2B
		2-2a: Provision 2/3 students reported as FREE in a NON BASE year:		2-2aB
		2-2b: Provision 2/3 students reported as REDUCED PRICE in a NON BASE year:		2-2bB
	2-3: Operating the Community Eligibility Option:	2-3A	2-3B	

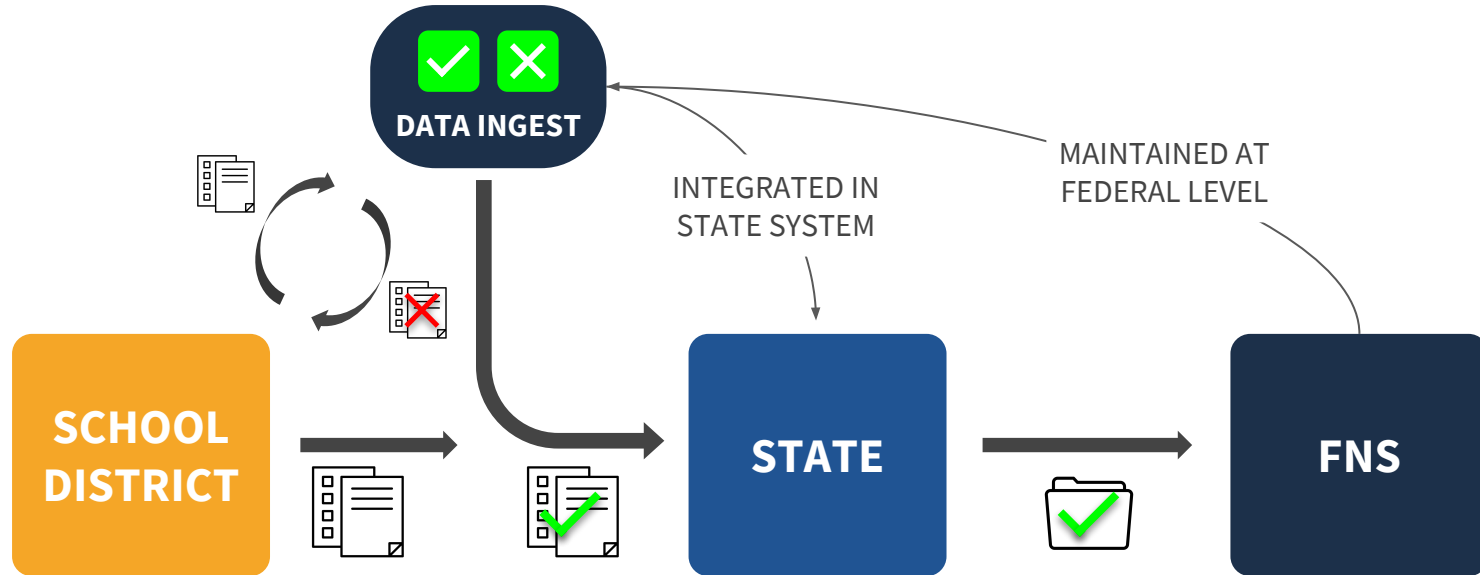
CURRENT STATE

Data validation occurs at the federal level; making corrections is a difficult and time-consuming process



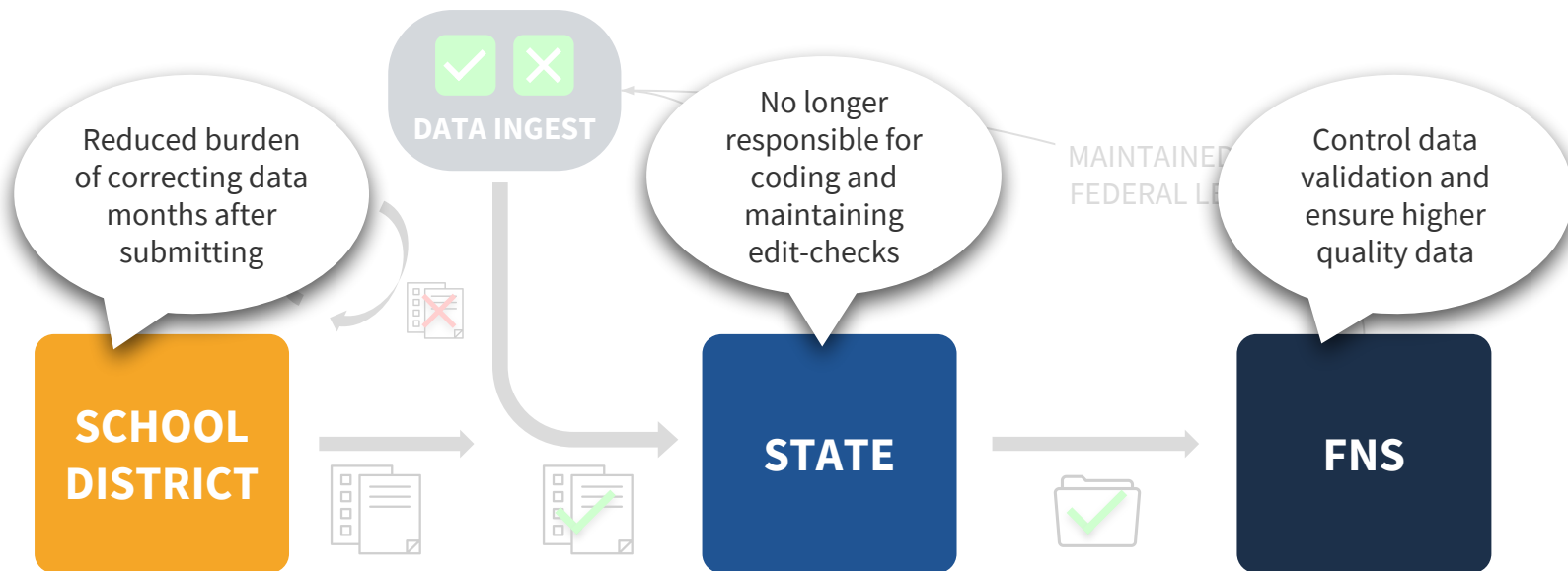
FUTURE STATE

Integrate federally-maintained
data validation in the state system



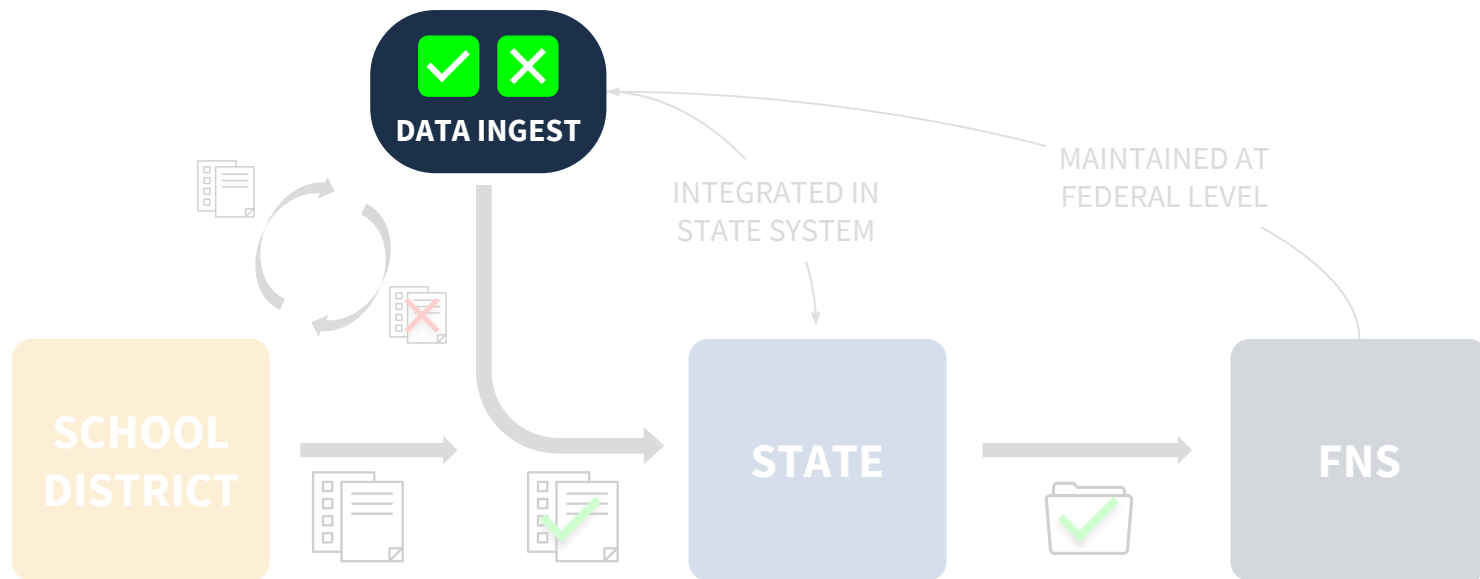
BENEFITS AT ALL LEVELS

These tools can benefit data providers, aggregators, and users



OUR PROTOTYPE

A customizable data aggregation and validation tool accessed via API



See the code on [GitHub](#).

PHASE THREE

**Make it real &
make it last**

In Phase 3, we will go both deep and broad



Build out the prototype
developed in Phase 2
and implement it in
two state systems



Find other use cases,
curate resources and best
practices, and develop a
plan to sustain this project

We are seeking partners who
are improving existing or
launching new federated data
management efforts.

Does this sound like you?

If so, we want to hear from you!

- **We collect data from external stakeholders**, e.g., states or local governments, non-profits, other agencies, etc.
- **We have specific data validation criteria** that must be met by submitting entities
- **We are charged with aggregating** and acting as stewards for **collected data**
- **We want to better serve the American public** by improving data federation across government

CONTACT THE TEAM

data-federation@gsa.gov



FOLLOW OUR WORK ON GITHUB

<https://github.com/18F/data-federation-project>



VISIT OUR (WIP) WEBSITE

<https://federation.data.gov/>



Thank you!

Noun Project icons by Made x Made, ATOM & DewDrops