The United States Data Federation is a project focused on making it easier to collect, combine, and exchange data from disparate sources.

The problem

Federal agencies often need to collect data from state and local governments, other federal agencies, and other data providers, which are used to support policy or budget decisions, operational efficiencies, or published in aggregate form for other data users. Making it easy and convenient for data providers to submit their data while maintaining quality, accuracy, and completeness can often be a challenge.

About the U.S. Data Federation Project

The U.S. Data Federation Project 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.

The goal of the project is to **document repeatable processes, develop reusable tooling, and curate resources** to support federated data projects, making it easier for data providers to submit interoperable, high quality data to federal agencies.

We define a federated data project as **an effort in which a common type of data is collected or exchanged across complex, disparate organizational boundaries**. This type of data management effort is common in our distributed style of government, but currently most such efforts are undertaken in isolation and lack infrastructural support or repeatable best practices.

Our work

In **Phase 1**, we interviewed a variety of distributed data management projects and synthesized findings in a [Data Federation Framework](). Our findings suggested that reusable tooling and processes would benefit future federated data efforts.

In **Phase 2**, we partnered with the USDA Food & Nutrition Service and built a functional prototype of a reusable data validation tool that allows users to submit data via a web interface or API to be validated against a set of customizable rules in real time.

In **Phase 3**, we will develop a software and service offering for data federation by
1. curating and communicating best practices and resources;
2. refining and expanding the prototype tool developed in Phase 2; and
3. validating the tool and other potential reusable tools with additional use cases.

The best practices and resources are intended to include guides and repeatable processes around data governance, organizational coordination, and standards development in federated environments. The reusable tools are intended to include capabilities around data validation, automated aggregation, and the development and documentation of data specifications.

Partner with us!

As we develop reusable tools and resources, **we are seeking partners who are improving existing federated data management efforts or launching new ones**. We want to help you navigate the challenges you encounter as a way of testing our tools and developing new ones.

We're looking for partners who:

- Collect data from external stakeholders, e.g., states or local governments, non-profits, other agencies, etc.
- Have specific data validation criteria that must be met by submitting entities
- Are charged with aggregating and acting as stewards for collected data
- Want to improve data federation across government in order to better serve the public