This repository    Search       Explore   Gist   Blog   Help                    dazzaji



**Watch** 18    **Star** 6    **Fork** 249

fedspendingtransparency / **fedspendingtransparency.github.io**
forked from project-open-data/project-open-data.github.io

# What are some of the considerations that Treasury and OMB should take into account when establishing data standards? #8

**New issue**

**⊙ Open**   **rmaziarz** opened this issue on Dec 1, 2014 · 4 comments

**rmaziarz** commented on Dec 1, 2014    Owner  +1

*No description provided.*

**rmaziarz** added the **data elements question** label on Dec 1, 2014

**bsweger** commented 12 days ago    +1

Though not directly related to the data standards, it would be helpful to know if there are plans to include any derived/calculated elements in the next iteration of federal spending information.

For example:

- providing the corresponding fiscal year for dates (where it makes sense)
- breaking down addresses into atomic geographic components, ideally accompanied by standard codes (*e.g.*, FIPS) that allow linking to other data
- geocoding addresses

Obviously you can't be in the business of handling every indiviual request for a calculated field, but doing a few that are widely beneficial would make it easier and more efficient to use the data.

Either way, as you finalize the standards, it would be useful to know the plans for derived fields and, if they're in play, what the guidelines are for determining their inclusion.

**HerschelC** commented 12 days ago    +1

One of the most important considerations with any standard is how to audit compliance with the standard. The standard must be comprehensive to leave little gray area or wiggle room. More specifically in the case of the DATA Act - an important consideration is ensuring the quality of the standardized data. Quality likely being defined as fitness for purpose.

The process of establishing a standard is of little value without rigorous quality control measures. How you will measure data quality should be part of data standard definitions. How data defects are adjudicated should be a process established alongside the data standard since the people setting the standard are likely the people to deal with violations of the standard.

**HerschelC** commented 12 days ago    +1

When dealing with a migration of existing data standards to a new standard, the idea of "global" and "local" standards could be implemented. A global standard applies across all stakeholder communities whereas a local standard applies to a subset that for whatever reason is not able to meet the global definition. In theory we prefer global standards for everything; in practice there are sometimes nuances because either the data does not exist, is not easily accessible, or the local organization cannot use the

**Labels**

**Milestone**
No milestone

**Assignee**
No one assigned

**Notifications**
⊙ Subscribe
You're not receiving notifications from this thread.

**4 participants**

**Export Issue**

global definition for operational reasons. That's ok - so long as there is a clear local definition. Don't stop progress waiting on perfect data alignment.

The significant downside to local definitions is that it prevents cross-organization analysis which requires apples-to-apples comparisons. I often use the "local" option as a stepping stone for organizations working towards the global definition but aren't quite there yet. Local definitions should be the rare exception. Using a github analogy, the local definitions could be "forks" from the master definition. Work on migrating the local definition to the global definition progresses until the local definitions are eventually pulled back into the master. The group using the local definition continues to take a hit on quality until they've migrated to the global definition.

An option often considered in data standards is to simply have both a global definition AND a local definition for cases where the local organization prefers to view the world in their own way. This is a different scenario than the above where the organization can't get to the global standard. In this case it's a different view that is desired, generally for operational reasons. This is fine and even encouraged since the local view likely enables the local organization to better manage their business while at the same time aligning with the global view of the world – it's just a data crosswalk.

**klguentert** commented 4 days ago                                            +1

As someone that will be asked to provide data, my concern is with reducing administrative burden while providing accurate, consistent, reliable data. Much of the data is already in federal systems and some of that data is data the recipients originally provided when asking for the funding.

For research institutions, we provide much of the data through agency systems (like SAM or grants.gov). It's critical that the data standards consistently utilize what has already been built (grants.gov data standards or RPPR data standards).

Agencies should be required to consistently use these data standards and there should be oversight to ensure that agencies are using them as intended.

Not related to data standards, please do not use FFATA reporting as the model for how the reporting will work. It is a clunky system that does not utilize existing federal data very efficiently.

In addition, something to consider is the zip plus 4 requirement. That does not always get you the "location of the work" information that you want. What happens if the work is being done in the Gulf of Mexico by an institution in Connecticut. Zip plus 4 will give you the institution in Connecticut that's receiving the funding but, the work is conducted somewhere else (without a zip plus 4) and impacting a completely different economy.

| Write | Preview | | MD Markdown supported | Edit in fullscreen |

Leave a comment

Click or drag files here
POWERED BY ZENHUB

Attach images by dragging & dropping, selecting them, or pasting from the clipboard.

**Comment**