

Comparative Analysis of USAspending.gov and FPDS

General Services Administration | Office of Information Technology Category (ITC)

Tung Lin, Data Engineer Fellow | *UCLA, Cognitive Science*

Keywords:

discovery, process automations, data cleaning, data integration

Summary:

While Federal Procurement Data System (FPDS) is ITC's current standard for procurement data, Tung conducted a discovery report on USAspending.gov, a Congress-mandated tool that tracks government spend. Leveraging **Python** and **SQL**, Tung revealed data relationships and quirks that helped the data team better understand contract data. Following up on findings and recommendations of her own report, Tung built a keyword analysis tool using **Google Apps Script** and the USAspending API, allowing users to pull, summarize, and visualize data **in less than 3 minutes** – a process that typically takes over an hour using FPDS data.

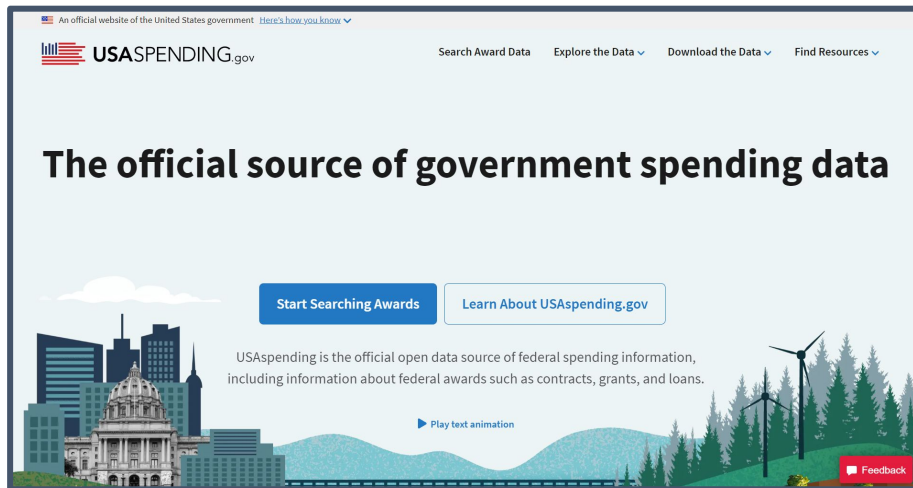


U.S. General Services Administration
OFFICE OF INFORMATION TECHNOLOGY CATEGORY

USAspending or FPDS

Tung Lin - QT1D

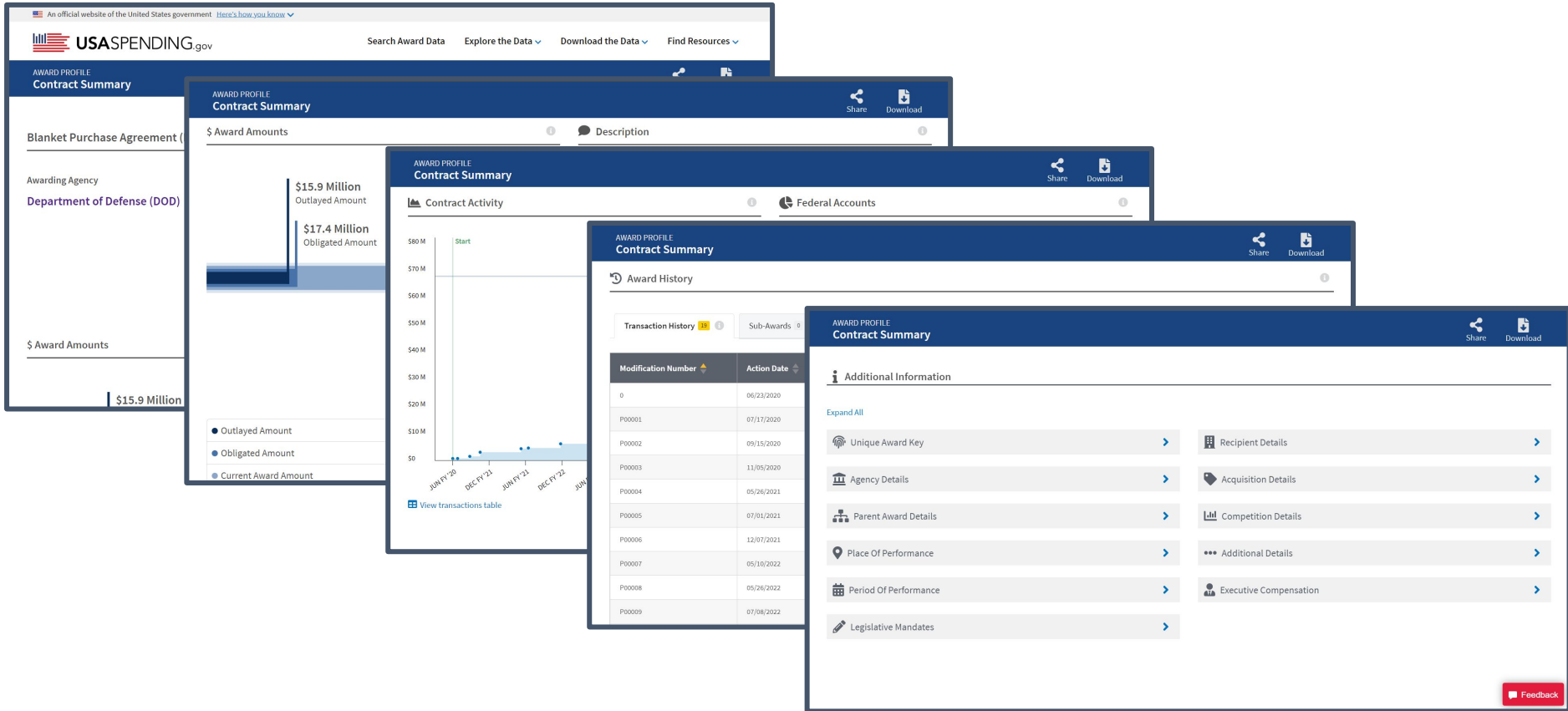
USAspending is a Congress-mandated tool to track government spend



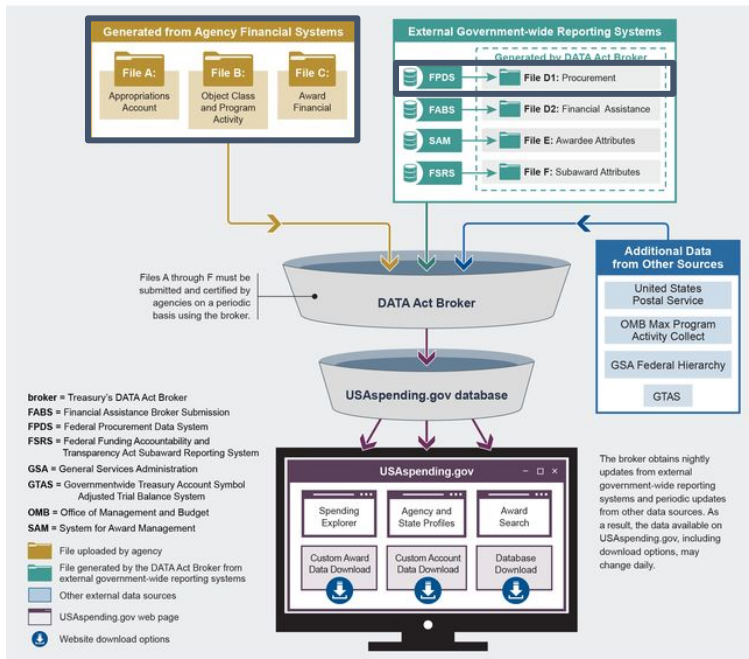
One stop shop:

- Contracts
- Contract IDVs
- Grants
- Financial Assistance
- Subawards
- Loans

USAspending is a Congress-mandated tool to track government spend



The Why: USAspending may have data that FPDS does not

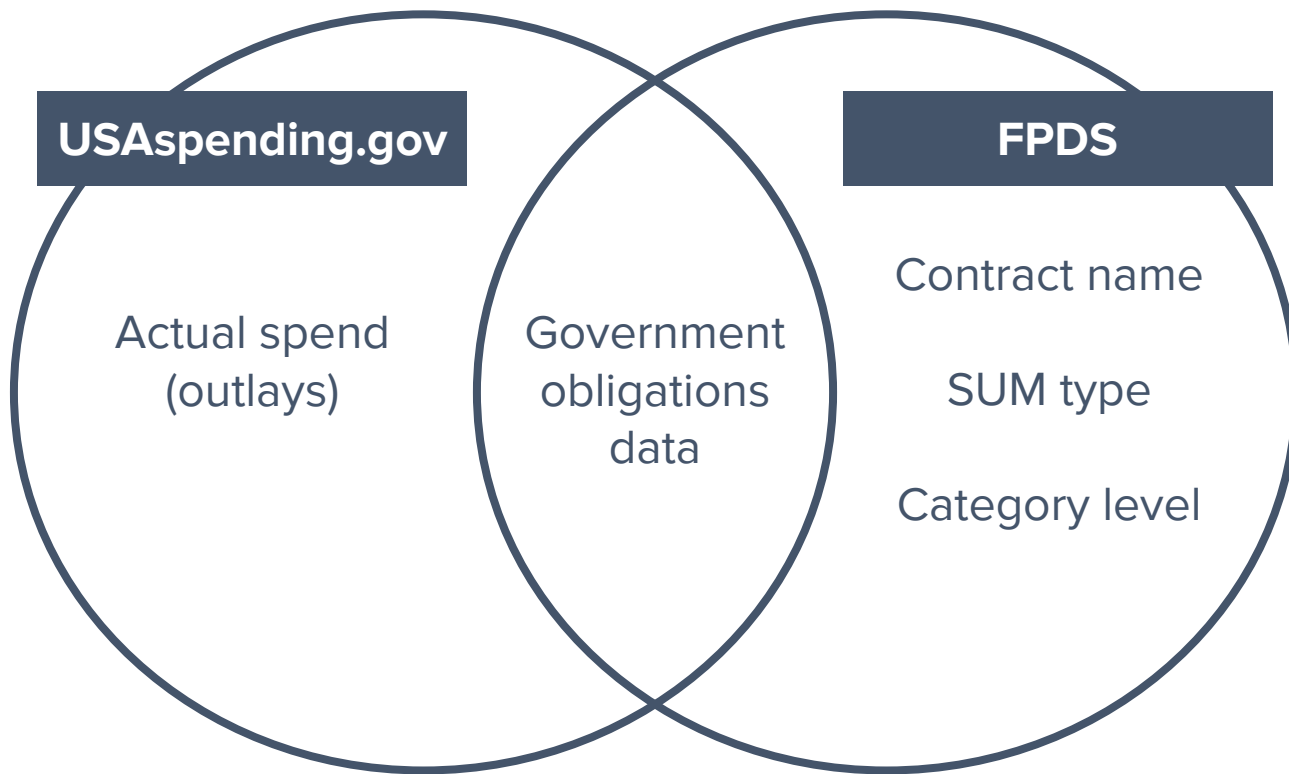


USAspending caters to a wide audience:

- A consolidation of many sources
 - FPDS
 - Agency Financial Files
- Potential to improve reporting precision

Source: GAO analysis of USAspending.gov. | GAO-22-104702

USAspending has additional outlay data, but FPDS is designed for analysis



How to use USAspending:

1. Not replacing FPDS
2. Enhancing our current processes:
 - a. API
 - b. Website
 - c. Policies and infrastructure

USAspending and FPDS Comparative Analysis (AH256) Discovery

Part I - Executive Summary

Background

The ITC Data team (QT1D) uses procurement data to perform various recurring or single request analyses. One of the main data sources is FAS Procurement Data, an internal database. The data is curated by Enterprise Data Architecture (EDA) using the Federal Procurement Data System (FPDS). On the other hand, [USAspending.gov](https://www.usaspending.gov) is the official source for government spend data, maintained by the Department of Treasury, but QT1D haven't had the time to explore this resource.

Problem Statement

QT1D doesn't have in-depth knowledge about how and when to use USAspending, if at all. Additionally, there is a lack of understanding of some FPDS columns due to lack of documentation and time to investigate.

Summary of Findings

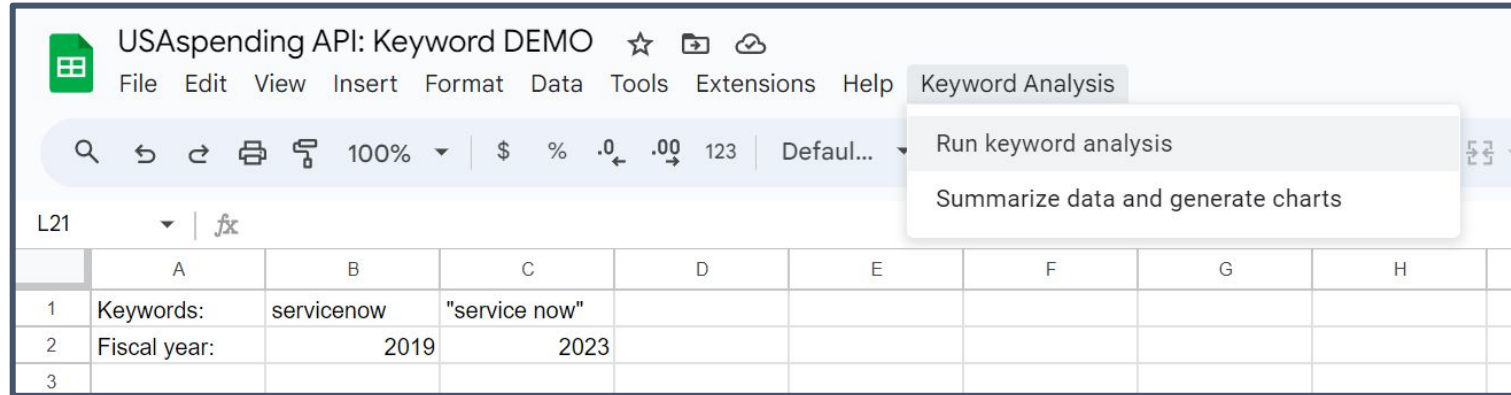
1. **Prime Award Data Integrity:** Both sources extract procurement data from FPDS Atom. USAspending adds description tags while FAS Procurement Data adds contract vehicle details. USA updates daily while FAS updates every 2 weeks.
2. **USAspending Unique Features:** The API is a fast method of calculating counts with one request. USAspending is enriched with other data sources that include financial awards, Disaster Emergency Funding Code (DEFC), and congressional districts data.
3. **Data Accessibility:** USAspending limits API requests to 50,000 data rows and downloads to 500,000 records. There is a USAspending database but getting access is projected to be difficult. There are no limits for query pulls from FAS Procurement Data.
4. **Data Management:** USAspending is the official tool from Congress, which comes with legislature to improve data standards. The broader user profile of USAspending also translates to better support and documentation of the system. FAS Procurement Data occasionally updated by EDA with limited data documentation.

Prioritized Recommendations

1. Build a comprehensive data and terminology dictionary for FAS Procurement Data.
2. Expose USAspending.gov website to QT1D as a tool to quickly share and evaluate individual awards.
3. USAspending is not viable as a data source for current project needs; however, we can document a few unique features for future requests and reference.

Using Google Sheets:

- Start a keyword analysis



The screenshot shows a Google Sheet titled "USAspending API: Keyword DEMO". The menu bar includes File, Edit, View, Insert, Format, Data, Tools, Extensions, and Help. The "Extensions" menu is open, showing the "Keyword Analysis" extension. Below the menu bar is a toolbar with various icons for search, undo, redo, print, and zoom. The spreadsheet grid shows columns A through H and rows 1 through 3. The data in the grid is as follows:

	A	B	C	D	E	F	G	H
1	Keywords:	servicenow	"service now"					
2	Fiscal year:	2019	2023					
3								

Using Google Sheets:

- Start a keyword analysis

The screenshot shows a Google Sheet titled "USAspending API: Keyword DEMO". The menu bar includes File, Edit, View, Insert, Format, Data, Tools, Extensions, Help, and Keyword Analysis. A toolbar shows search, undo, redo, print, and zoom (100%) icons. A dark overlay box in the center reads "Running script" with "Cancel" and "Dismiss" buttons. The sheet has columns A through H. The first three rows contain data for a keyword analysis:

	A	B	C	D	E	F	G	H
1	Keywords:	servicenow	"service now"					
2	Fiscal year:	2019	2023					
3								

Using Google Sheets:

- Start a keyword analysis

A1	fx	Award ID							
	A	B	C	D	E	F	G	H	I
1	Award ID	Mod	Recipient Name	Action Date	Transaction Amc	Awarding Agenc	Awarding Sub A	Award Type	generated_interr
2	86615423C0000	0	BOWHEAD ENT	2023-09-30	3334661.01	Department of H	Department of H	DEFINITIVE CO	CONT_AWD_86
3	19AQMM18F42	P00010	CARAHSOFT TE	2023-09-30	-0.04	Department of S	Department of S	DELIVERY ORD	CONT_AWD_19
4	75R60223F8017	P00001	NEW TECH SOL	2023-09-29	-965852.66	Department of H	Health Resource	DELIVERY ORD	CONT_AWD_75
5	75R60223F8017	0	NEW TECH SOL	2023-09-29	965852.66	Department of H	Health Resource	DELIVERY ORD	CONT_AWD_75
6	19AQMM23F75	0	THUNDERCAT	2023-09-29	19583.32	Department of S	Department of S	BPA CALL	CONT_AWD_19
7	19AQMM23F76	0	THUNDERCAT	2023-09-29	213189.83	Department of S	Department of S	BPA CALL	CONT_AWD_19
8	75P00123F8023	0	DAS FEDERAL I	2023-09-29	796560	Department of H	Office of the Ass	DELIVERY ORD	CONT_AWD_75
9	W9124J23F018	0	CARAHSOFT TE	2023-09-29	1005630.51	Department of D	Department of th	DELIVERY ORD	CONT_AWD_W

Using Google Sheets:

- Start a keyword analysis

USAspending API: Keyword DEMO

File Edit View Insert Format Data Tools Extensions Help

Keyword Analysis

Run keyword analysis

Summarize data and generate charts

	A	B	C	D	E	F	G	H	
1	Award ID	Mod	Recipient Name	Action Date	Transaction Amount	Awarding Agency	Awarding Sub Agency	Award Type	general
2	86615423C0000	0	BOWHEAD ENT	2023-09-30	3334661.01	Department of H	Department of H	DEFINITIVE CO	COMMODITY
3	19AQMM18F42	P00010	CARAHSOFT TE	2023-09-30	-0.04	Department of S	Department of S	DELIVERY ORD	COMMODITY

Using Google Sheets:





- Start a keyword analysis



- 1 hour to less than 3 minutes

The bigger picture of USAspending.gov

- FFATA and DATA Act
- High visibility and accessibility
- GAO audits

Recommendations for Executive Action		
<u>Agency Affected</u>	<u>Recommendation</u>	<u>Status</u>
Department of the Treasury	<div> Priority Rec. </div> <p>The Secretary of the Treasury should develop and implement approaches to make known data limitation disclosures more prominent and easily accessible to USAspending.gov users. (Recommendation 4)</p>	<div> Closed – Implemented </div>

Current Impact and Future Work

- Increasing awareness of available FPDS-related tools
- Discovering key nuances to better understand contract data
- Demonstrated potential tooling opportunities

What I Learned

- Intricacies of government procurement
- Real world data and problems
 - Python, SQL, JavaScript
- Federal work environment