

White Paper: Analysis of Small Business Utilization within the Information Technology Category

ATARC Data Analytics and AI Lab
May 2022

Contributors:

Ellery Taylor, GSA
Richard Schreiber, GSA
John Polishuk, Acuity Systems - nAbleCloud LABS
Raj Srivatsan, SymphonyAI
Travis Barton, SymphonyAI
Eric Murray, SymphonyAI
Nicole Mandes, ATARC

Table of Contents

I.	Lab Charter	1
II.	Problem Statement.....	1
III.	Objectives	1
IV.	Results	1
	Spend Analysis Prototype	2
	1. GSA Business Goal	2
	2. Scope of Spend Analysis Prototype.....	2
	3. Major Findings based on Spend Analysis	2
	3.1. GSA Findings	2
	3.2. Cross-Department Findings.....	3
	3.3. DOJ Findings.....	4
	3.4. DOD Findings	5
	3.5. HHS Findings.....	5
V.	Next Step: Building Spend Analysis Applications.....	5

I. Lab Charter

The Advanced Technology Academic Research Center (ATARC) Data Analytics and Artificial Intelligence Lab (ADAAL) will develop, research, communicate, and manage a set of use cases provided by Agencies. The ADAAL will provide an easy to use, secure lab environment provided by nAbleLabs where the Government can leverage state of the art tools to quickly understand the potential use of Analytics and/or AI to better provide actionable information to aid decision making across the Federal Government. The Lab also provides a single environment where industry and Federal agencies can collaborate more effectively.

II. Problem Statement

Small Business utilization by Federal agencies needs to improve (based on the Biden Administration Statement¹). There is a lack of connection and networking between government and Small Businesses. To better understand this issue, we need to increase the usage of existing data by leveraging data analytics and AI/ML. A better understanding of the data will result in targeted strategies to maximize small businesses participation in the Information Technology category/sector.

III. Objectives

The ATARC Data Analytics and AI Lab intended to accomplish the following objectives:

- Analyze the fas_procurement data (and more data sets used to enrich) to better understand the small business utilization marketplace - deliverable would be a market report
- Determine the value of leveraging AI/ML software to create products that would help customer X increase metric Y
 - Hypothetical example: User-level software/program that helps Small Business connect with (or find resources for) partnering with the government.
- Present findings to the ATARC Analytics and AI Lab and the Federal Agency community for further input

IV. Results

The ATARC Data Analytics and AI Lab presents the following summary of its findings:

¹ <https://www.whitehouse.gov/briefing-room/statements-releases/2021/12/02/fact-sheet-biden-harris-administration-announces-reforms-to-increase-equity-and-level-the-playing-field-for-underserved-small-business-owners/>

Spend Analysis Prototype

1. GSA Business Goal

The US General Services Administration (GSA) wants to improve the utilization of small businesses on federal contracts. In support of this business goal, they plan to use data analytics and machine learning approaches to analyze procurement data obtained from the Federal Acquisition Service (FAS).

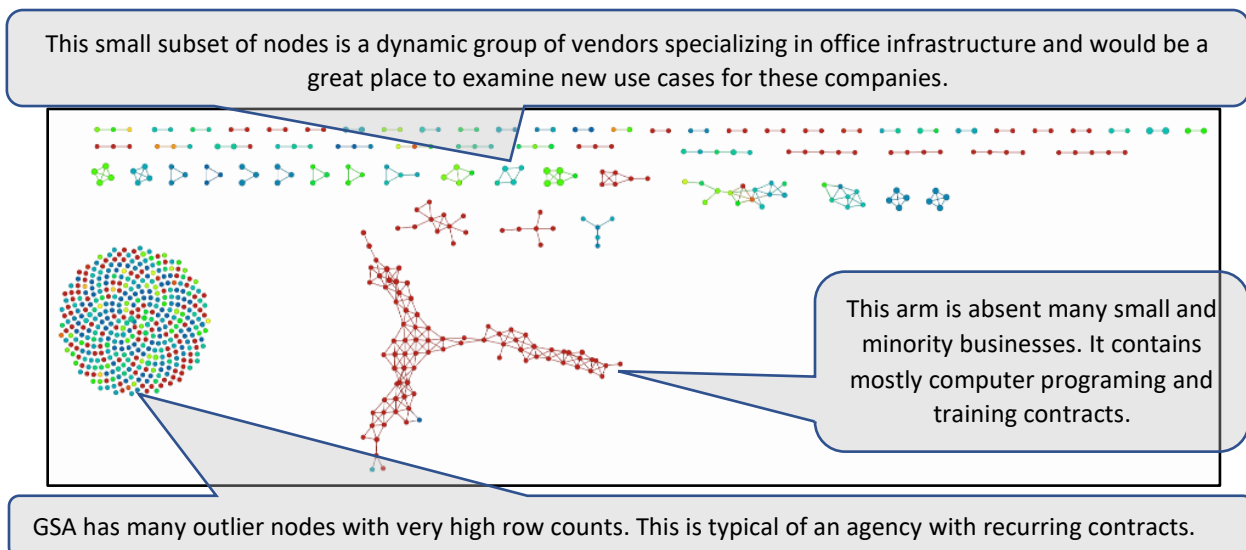
2. Scope of Spend Analysis Prototype

GSA asked the **ATARC ADAAL** to build a Spend Analysis prototype focused on analyzing contracts awarded by various agencies over the past few years. The prototype was focused on analysis of data provided for 376,000 contracts representing awards made over the 4-year period from 2017 through 2021, by GSA, DOJ, HHS, and DOD. Data for each contract included flags indicating whether the awardee belonged to certain business categories – American Indian owned business, African American owned business, Small Business, Small Disadvantaged Business – as well as information on the obligated amount, vendor DUNS number, etc. Note that the data reviewed was for a collection of ITC contracts and a few other vehicles. As such it may not be fully representative of the state of IT procurement.

This prototype was built using SymphonyAI’s award-winning, patented **EurekaAI Platform** which brings together the only commercially available implementation of Topological Data Analysis (TDA) and a broad range of machine learning, statistical, and geometric algorithms to create a visual representation of all the data points in a large data set. This visual representation helps rapidly uncover critical patterns and relationships. By identifying the geometric relationships between data points, TDA offers an easy and efficient way of partitioning data to understand underlying properties that characterize the segments and sub-segments within the data. This helps quickly identify previously unknown segments or cohorts, anomalies, and unknown-unknowns, and detect hot spots that cannot be identified using standard analytical techniques.

3. Major Findings based on Spend Analysis

3.1. GSA Findings



High-Level Summary

- GSA has a much better ratio of small, disadvantaged businesses to the total population than other agencies reviewed.
- It has the lowest average contract award amount of the agencies that were reviewed
- It tends to have a lot more recurring contracts than other agencies that were examined

Category Analysis

- GSA IT outsourcing contracts are more often associated with large businesses and high dollars obligated. This is true for other agencies also
- There are several industry categories that can improve their small and disadvantaged businesses utilization, such as Repair, Training, and Coding Services

Contract Vehicles

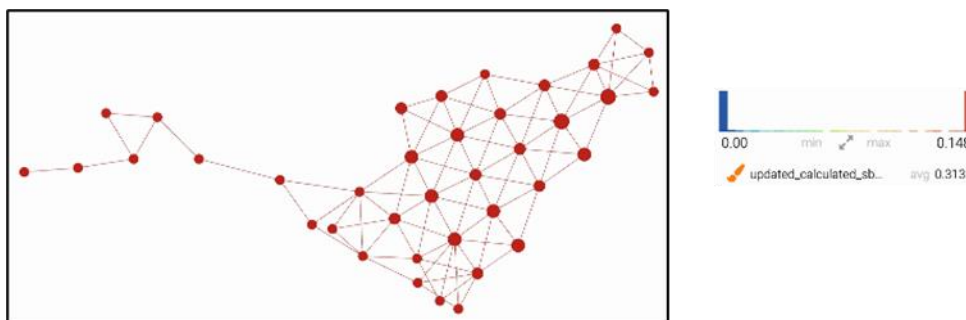
GSA has a large number of contracts funded through other agencies. This would make it difficult to implement any recommended changes, especially since the largest contracts use GSA purely for its contracting vehicles.

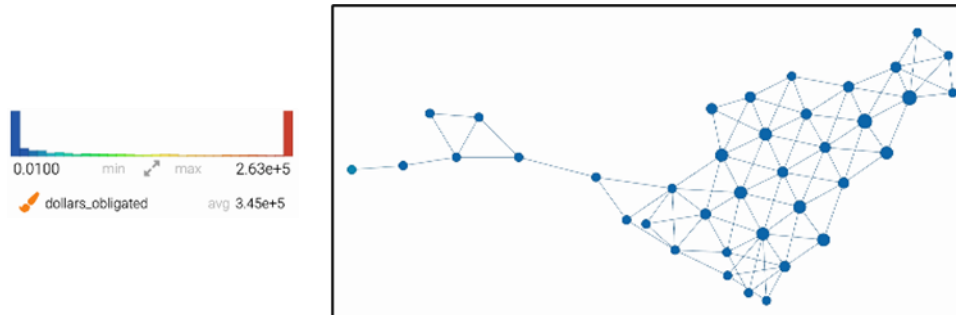
Key Takeaways

GSA could improve small business utilization by offloading general computer-related services to small and disadvantaged businesses.

3.2. Cross-Department Findings

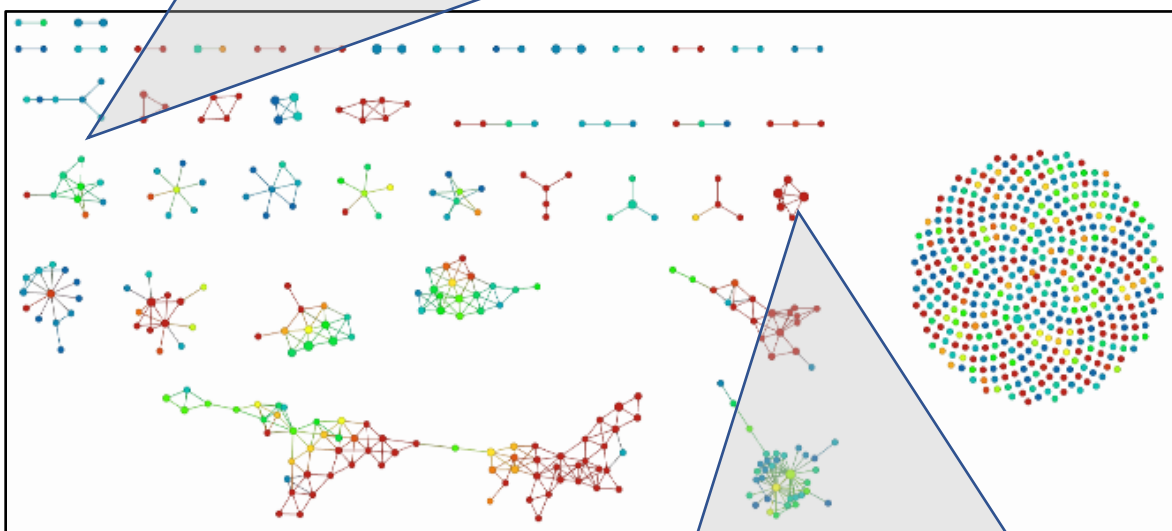
Our model picked up subtle patterns within the data that grouped vendors who claim to be small, disadvantaged businesses but are without any socio-economic information.





3.3. DOJ Findings

DOJ has a large string of contracts in the IT hardware space where small businesses are present in large numbers but have low-dollar obligations.



DOJ spend is imbalanced with a small number of areas being a huge part of their spend. Such high dollar spending could mean small shifts towards small business would have enormous impacts on small business

High-Level Summary:

- DOJ has twice the average contract award amount as compared to other agencies
- DOJ's percentage allocation to small businesses is less than half that of other departments
- DOJ's small business contracts generally have high contract award amounts

Category Analysis:

- IT Outsourcing contracts are largely associated with high-dollar contracts
- There are many Telecom and IT Hardware contracts in low-dollar obligated regions

- IT Software is equally represented across low-dollar and high-dollar regions

Other Findings:

- EPA designation “NOT REQUIRED” is the norm, with only 4% having any EPA designation
- Vendors in the high-dollar obligated region have little in common with the rest of the data

Key Takeaways:

- DOJ could improve dollar allocations by diversifying their allocations
- Within DOJ, small businesses are a niche, rather than the norm

3.4. DOD Findings

- DOD is behind other agencies by percent for small, disadvantaged business contracts
- Where small businesses are present, they have very high dollars obligated compared to the rest of the data

3.5. HHS Findings

- HHS is above average in most key categories though it suffers from the same global issue of IT Outsourcing and Hardware
- HHS has a high number of unique contracts that might make it difficult to redistribute

V. Next Step: Building Spend Analysis Applications

As demonstrated by the prototype, SymphonyAI’s Eureka platform provides great insights into small business allocations by federal agencies and makes it easy to identify areas for improvement. Agencies could license this software tool for use by in-house data scientists that can be trained on how to use our software. However, to make the insights from Eureka actionable by end users that are not data scientists, we recommend creation of two distinct applications: an application for government agencies and an industry-facing application.

We recommend creation of an application for small business advocates and procurement analysts to visually explore identify areas for improvement. Each user would begin with a dashboard which can be configured to highlight their specific areas of interest. From the dashboard, the application would allow users to quickly identify hot spots and perform analysis across a variety of dimensions including category, department, type of small business, geography and even item and contracting office. With our unsupervised machine learning and visualization capabilities, this application can be configured and deployed without the need for a data scientist or any special training for end users.

We also recommend creation of an application which small businesses can use to identify areas where

the government is looking to increase awards to small businesses. This industry-facing application would allow small businesses to enter their small business status, their geography and procurement categories of interest as well as other relevant information. From there, the system would provide guidance to small businesses, directing them towards areas where the government is looking to increase small business utilization. Like the first application, our powerful unsupervised machine learning would not require government data scientists to configure the application and small businesses could use the application without special training.