

SYM-AM-19-059



**PROCEEDINGS
OF THE
SIXTEENTH ANNUAL
ACQUISITION RESEARCH
SYMPOSIUM**

**WEDNESDAY SESSIONS
VOLUME I**

**Acquisition Research:
Creating Synergy for Informed Change**

May 8–9, 2019

Published: April 30, 2019

Approved for public release; distribution is unlimited.

Prepared for the Naval Postgraduate School, Monterey, CA 93943.



ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL

Evaluating the Use of Public Data Sources to Improve Acquisition Processes: A Market Research Use Case

Dorcas L. Lasalle—Principal Investigator and Lead Acquisition and Program Management Analyst at The MITRE Corporation, supports a variety of sponsors with full acquisition life cycle support services. She has more than 15 years of experience in acquisition and contracting, both federal government and industry. She is a former Veterans Affairs and Government Accountability Office contract specialist and former Acquisition Specialist with the Department of Defense Office of Inspector General. She holds a BS in Hospitality Management from the University of Central Florida. She was FAC-C Level certified in Contracting while in civil service. [dlasalle@mitre.org]

Kristin Fitzgerald—is a Senior Data Scientist in the Model-Based Analytics Department within MITRE. She is passionate about finding ways to apply machine learning and data analytics to a variety of public service applications, including veterans' benefits, child welfare, the opioid epidemic, and the federal acquisitions process. She earned a BSE in Operations Research and Financial Engineering from Princeton University and is currently completing an MS in Analytics through the Georgia Institute of Technology. [kafitzgerald@mitre.org]

Abstract

This research describes how a FAR Part 10 market research report can be generated by integrating acquisition data from multiple government-maintained public databases into a single portal. This effort builds upon the wealth of federal acquisition data made public through several initiatives to increase government transparency and data sharing, including the DATA act of 2014. The overall study aims to demonstrate how inefficiencies in the acquisition process can be addressed through tailored design of data-driven decision support tools.

Introduction and Motivation

Acquisition processes are time and resource intensive and rely heavily on staff experience and expertise. Unfortunately, this expertise can be hard to find. In particular, the process for creation of market research reports is not standardized; they are typically compiled manually by relatively inexperienced staff who may not have the time or knowledge to integrate relevant data from different sources. As a result, given the overall acquisition strategy and execution rely heavily on elements of this market research, there is potential for significant downstream decision-making inconsistencies and delays.

Over the past few years, however, the federal government has made a large amount of acquisition data publicly available. The data is predominantly historical in nature and presents an opportunity to help develop automated decision support tools, in this instance supporting market research. Unfortunately, these datasets are of varying size and scope and are typically siloed. Even when valuable and relevant datasets are identified, they may be difficult to access or the relationships between them are not immediately clear. Greater research efforts are required to better understand the overall landscape of this data and its potential for practical use.

This research addresses a preliminary first use case application of public data sought to aid the development of a FAR Part 10 market research report by integrating acquisition data from multiple government-maintained public databases into a single portal.



Background and Past Research

A literature review to understand the current applications of data analytics and use cases for data-driven acquisition decision making determined that

- a. The use of predictive analytics is being studied within government agencies as well as academia and private industry:
 - Dai and Li (2016) discussed the development of applications for non-government “armchair auditors” to analyze acquisitions data; as data is not in a consistent format, analysis is difficult without a standardized application. Applications included data reliability, suspicious supplier detection, abnormal pricing, and abnormal bidding.
- b. Research has focused on using analytics to improve the following acquisitions functions:
 - *Cost estimations/budget overruns:* Adoko, Mazzuchi, and Sarkani (2015) proposed a predictive model that analyzes the impact of system performance, Technology Readiness Level (TRL), schedule, risk, and reliability on the Nunn-McCurdy significant cost overrun guidelines. Tracy and White (2011) generated models to estimate the cost of completion of contracts at varying stages of completion. Morgan (2013) uses data analytics to evaluate the use of performance-based contracts as a cost-saving measure. Reed, Keller, and Fallon (2016) review cost per dollar obligated measurement in defense contracts and show the use of analytics to show trends beyond dollars spent.
 - *Requirements development:* Dargan et al. (2014) developed a statistical model of the relationship between requirements quality and operational results, using data from the DoD and DHS.
 - *Performance:* Knudsen and Blackburn (2016) propose a predictive model to examine project schedule performance. Apte, Rendon, and Dixon (2016) explored how the DoD can leverage acquisition data, specifically contractor performance information, in identifying drivers of success in services acquisition using big data techniques. Guillaume-Joseph and Wasek (2015) used historical aspects of software project failure to develop a predictive model that can be used in acquisitions.
 - *Regulation:* Patrignani (2014) evaluated the impact of changes to the FAR in 2009 on contractor misconduct focusing on the impact of penalties for misconduct using statistical analysis. Tkach (2017) examined USSOCOM's acquisition and procurement processes, policies, and challenges and provides insight into nontraditional DoD contracting by studying historical data.
- c. A variety of analytical methods are being used:
 - *Natural Language Processing (NLP)/text mining:* Gao, Singh, and Mehra (2012) developed a tool (Contract Miner) to extract data from service contracts using NLP. Yang et al. (2013) developed a prototype NLP tool to analyze contract service agreements (CSAs). Chalkidis, Androutsopoulos, and Michos (2017) studied how legal contract element extraction can be automated using NLP and machine learning.



- *Predictive regression*: Miller (2012) examines the use of text mining in acquisitions management and combines that with predictive regression models to determine cost estimate changes.
- *Bayesian/statistical modeling*: Knudsen and Blackburn (2016) used a Bayesian model to predict schedule performance.
- *Agent-based simulation*: Schwenn et al. (2015) introduced a research methodology for examining the U.S. weapon procurement system as a complex adaptive system (CAS) and using agent-based modeling (ABM) to identify significant causal factors that contribute to the performance of the procurement system.

While the research indicated that federal acquisitions data can be analyzed to improve performance, there was limited evidence of its use at an agency- or government-wide level.

Data Sources and Integration

Phase one of the project involved collecting data from a variety of public sources, identifying relationships between the sources, and completing any required data cleaning and integration. The team collected over 20 GB of data dating back to 2013 from the following public acquisition data sources:

- a. USA Spending—Contract spending records for the U.S. government, 2013–2017;
- b. System for Awards Management (SAM)—List of all contractors eligible to contract with the federal government and an exclusions list of contractors excluded from contracting with the government;
- c. Federal Awardee Performance and Integrity Information System (FAPIIS)—Contractors' performance and integrity records;
- d. Federal Procurement Data System—Next Generation (FPDS-NG)—Government-wide procurement and spending database;
- e. FBO.gov—General portal of entry for competitive acquisitions and corresponding documentation to include requirements; and
- f. Interagency Contract Directory—Procurement and spending between U.S. government agencies.

As shown in Figure 1, we identified existing relationships between the sources as well as primary keys between the datasets that allowed them to be linked. In most cases, the linking key was the Dun & Bradstreet (DUNS) contractor number and/or the contract number. The Interagency Contract Directory was not used in the final Market Research prototype, so it is excluded in the figure.



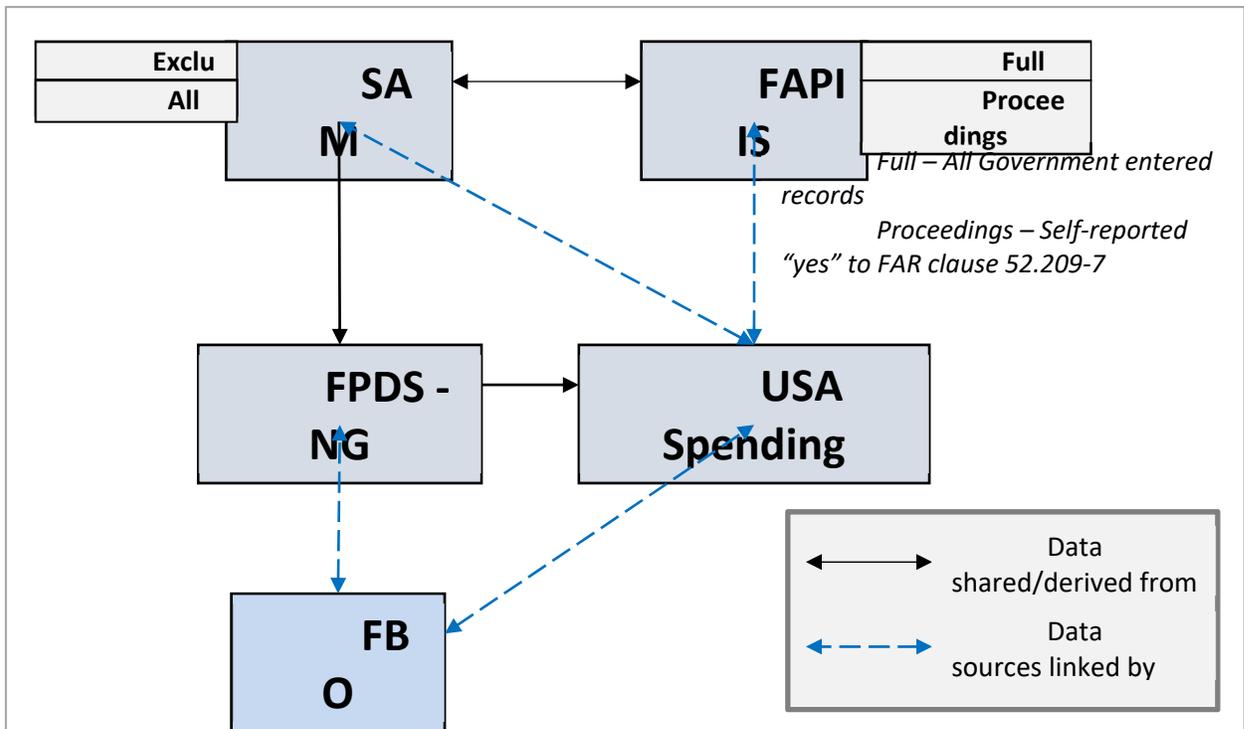


Figure 1. Data Source Map

Significant data cleaning was required, including the removal of duplicate records, type coercion, correction of data quality issues, and removal of nonsense characters. Data was stored using a combination of PostgreSQL and MongoDB databases.

Market Research Portal Overview

Phase two of the project consisted of developing the Market Research Portal prototype, a use case that demonstrated the data integration capabilities of the data warehouse. The team used standard agile development practices and leveraged both subject matter experts and user interactions to perform usability studies and evaluate features.

The Market Research Portal provides a market research report template as well as a system that supports an approvals process, review of older reports, and tailored instruction or template modifications by a supervisor. The key improvement of this portal over existing technology is that it automatically brings in data from the sources described previously. Based on the desired acquisition category (as defined by NAICS and/or PSC code), the portal filters to relevant vendors registered to do business with the government and provides contract history. The portal assists the user in identifying potential contractors who may or are already providing supplies and services to the government for similar requirements with additional refinement parameters such as dates and dollar amount. The portal also allows the user to search for similar requirements from other agencies and develop their own initial requirements validated by research. Lastly, the portal allows email correspondence with industry. The team applied agile development practices to update the application using feedback received from sponsors and user groups. All features were designed to meet minimum report requirements based on widely used agency templates and user feedback sessions.



A diagram of the project technical architecture is included in the appendix.

Walkthrough of Prototype Features

The set of features described next was created based on a series of user stories developed by subject matter experts, which were then adjusted or supplemented by user testing and feedback.

1. **Login and Report Overview:** The system is protected by a user login and password; this can be tailored to a specific government agency or to use a CAC card. Once logged in, a user encounters a list of generated reports. The user has an option of creating a new report, editing an existing report, sharing a report (e.g., for approval), exporting a report, or deleting a report.
2. **Report Creation:** When a user creates a new report, they give the report a title and limit the scope of the report using a NAICS and/or PSC code. These codes are later used for database search and aggregation. The user supplies a narrative description of services and other relevant background information. They also provide specifications of supplies/services as well as a list of requirements. Additionally, they specify the period of performance.
3. **Vendor:** The user is then able to explore the set of potential vendors. This vendor list is already filtered based on the NAICS/PSC information initially provided. The user may search the vendors for specific keywords and may also filter the vendors by whether they have been awarded contracts previously, by various small business categories, by the dollar value of past contracts, and by whether they have performance or integrity issues recorded (based on FAPIIS records and SAM exclusions).

For each vendor, the user has the opportunity to examine information about prior contracts and select vendors they are interested in researching further.

4. **Research:** For each vendor selected, the user may enter notes, assign requirements to the vendor, and view vendor contact information. The portal also gives the user the option to send a standardized email to each vendor asking for answers to questions and to get quotes. The research phase also allows the user to search FBO for similar requirements by NAICS or description and also provides a list of Government-Wide Acquisition Contracts (GWACs) that meet the requirements.
5. **Analysis:** In analysis, the information from all the previous phases are merged together into a single summary that the user can annotate and view. The analysis section has the greatest room for growth based on individual organizations' needs (discussed below).
6. **Recommendations:** Finally, the user can summarize their findings into a recommendation, which may include details such as contract type, contract vehicle type, solicitation strategy, key differentiators for source selection, and the identification of strong candidate contractors. The report can then be exported to a pdf, saved for later, or sent for approval to others.

Additional potential features are discussed in our conclusions.



User Testing and Review

Initial user testing leveraged MITRE subject matter experts who had extensive experience in creating market research reports as previous federal employees or contracted support staff. We also received feedback from several sponsors during the design and build process. Consistently, we were told that a tool like this would reduce the time required to complete the highly manual portion of the market research process (finding, accessing, extracting the data from many sources, then aggregate and execute reviews of information) and would enable users to dedicate more time into their thought-driven analysis and benchmarking of strategy and approach to the acquisition. Many of the features discussed previously were derived through user testing and interviews, including the ability to contact potential vendors through the tool. Test users also helped to identify which contract information was most relevant for their analysis, and which vendor characteristics they would like to be able to use as a filter. One feature that was highlighted frequently as important and increasingly relevant was the ability to determine a contractor's small business status and to compare that vendor's contract history with other vendors.

In the future, we will quantify the value added by this tool by running a controlled experiment. In the experiment, several users will be asked to create a market research report using either a simple template (that is similarly formatted to the portal report) or using the designed prototype leveraging all the data sources. Users will be asked to (a) report time required to complete the report and (b) describe how well they believe they were able to complete each report and their experience building the report.

Conclusions and Recommendations

It is evident that leveraging historical acquisition data can play a part in expediting administrative acquisition functions, in this case as part of the fully standardized market research report, which can then inform decisions. The realm of decisions aligned to the data sources selected in this research addresses estimates, contractor responsibility, integrity and performance, alternatives for contract type, contract vehicle type, and meeting small business goals. The generated report is another stepping stone towards standardizing the onset of educated consumerism (market expertise) in the public sector. The prototype is designed to demonstrate the viability of such a tool when it leverages these types of data sets. An assessment of technologies and sources should be made when making a business case for such a tool on a case by case basis for each agency. The MITRE Corporation–developed prototype demonstrates the value added by integration of public acquisition data sources into a single, easy-to-use system. Since MITRE is a manager of the Federally-Funded Research and Development Center, this prototype design and utilization is available to the public sector for its application completely free of charge. Transfer of the prototype will identify frameworks, templates, and processes needed to adopt the construct for federal agency use.

Ongoing research will include going beyond the data synthesization and into how artificial intelligence and predictive analytics may be applied to better customize acquisition decision-making.

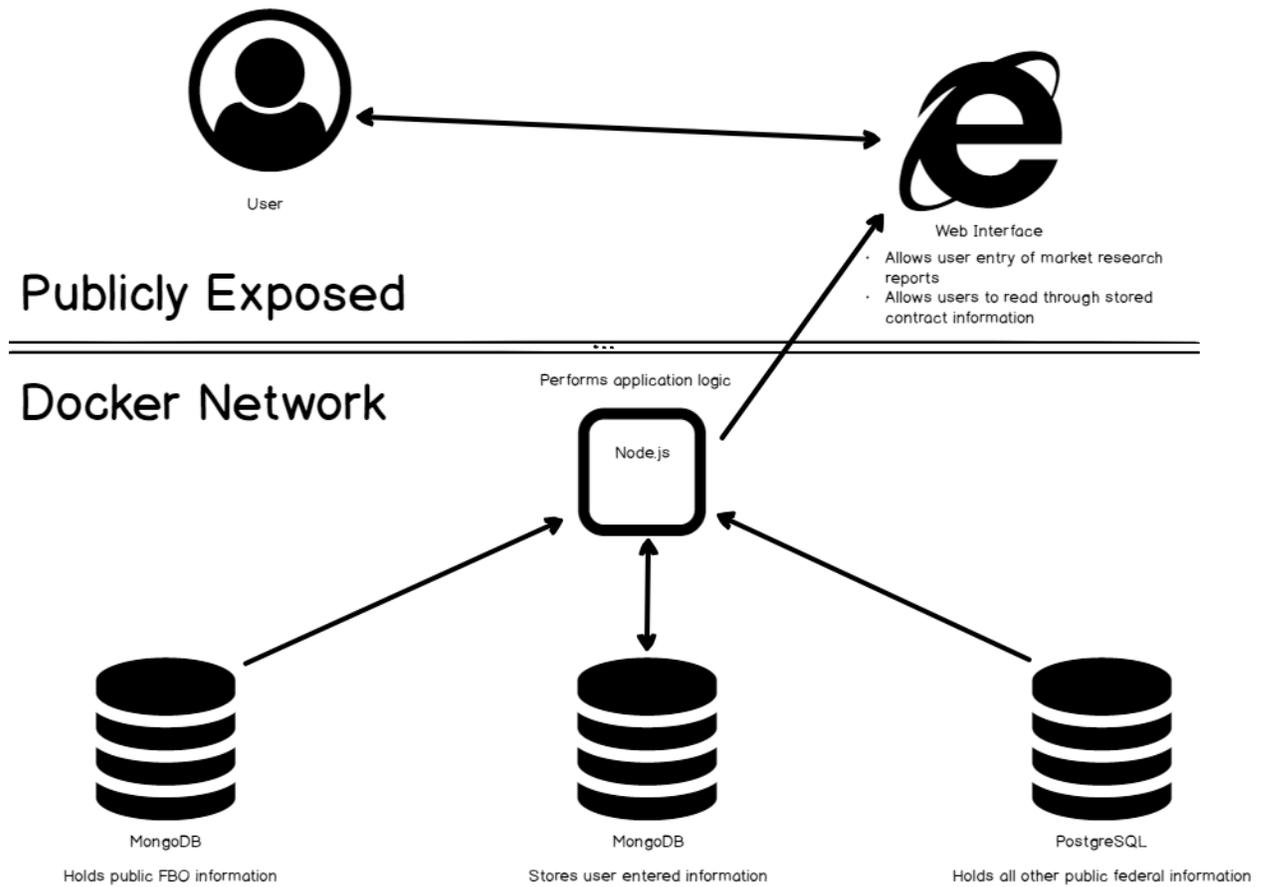


References

- Adoko, M. T., Mazzuchi, T. A., & Sarkani, S. (2015). Developing a cost overrun predictive model for complex systems development projects. *Project Management Journal*, 46(6), 111–125.
- Apte, U., Rendon, R., & Dixon, M. (2016). *Big data analysis of contractor performance information for services acquisition in DoD: A proof of concept*. Monterey, CA: Naval Postgraduate School.
- Chalkidis, I., Androutsopoulos, I., & Michos, A. (2017). Extracting contract elements. In *Proceedings of the 16th International Conference on Artificial Intelligence and Law* (pp. 19–28).
- Dai, J., & Li, Q. (2016). Designing audit apps for armchair auditors to analyze government procurement contracts. *Journal of Emerging Technologies in Accounting*, 13(2), 71–88.
- Dargan, J. L., Campos-Nanez, E., Fomin, P., & Wasek, J. (2014). Predicting systems performance through requirements quality attributes model. *Procedia Computer Science*, 28, 347–353.
- Gao, X., Singh, M. P., & Mehra, P. (2012). Mining business contracts for service exceptions. *IEEE Transactions on Services Computing*, 5(3), 333–344.
- Guillaume-Joseph, G., & Wasek, J. S. (2015). Improving software project outcomes through predictive analytics, part 2. *IEEE Engineering Management Review*, 43(3), 39–49.
- Knudsen, K. T., & Blackburn, M. (2016). A knowledge and analytics-based framework and model for forecasting program schedule performance. *Procedia Computer Science*, 95, 319–326.
- Miller, T. (2012). *Acquisition program problem detection using text mining methods*. Wright-Patterson Air Force Base, OH: Air Force Institute of Technology.
- Morgan, C. (2013, Summer). Best practices of developing performance-based acquisitions to save the U.S. federal government money. *Journal of Contract Management*, 77–86.
- Patrignani, J. O. (2014). *Impact of federal acquisition regulation change on contractor misconduct*. Minneapolis, MN: Walden University.
- Reed, T., Keller, J., & Fallon, J. (2016). *Organization analytics: Taking cost-per-dollar-obligated (CPDO) measures to the next level in defense contracting* (No. SYM-AM-16-068). McLean, VA: Beyond Optimal Strategic Solutions.
- Schwenn, K., Colombi, J., Wu, T., Oyama, K., & Johnson, A. (2015). Toward agent-based modeling of the U.S. Department of Defense acquisition system. *Procedia Computer Science*, 44, 383–392.
- Tkach, B. (2017). *Special operations contracting: 21st century approaches for service and technology acquisition* (Joint Special Operations University Report 17-5).
- Tracy, S. P., & White, E. D. (2011). Estimating the final cost of a DoD acquisition contract. *Journal of Public Procurement*, 11(2), 190–205.
- Yang, D. et al. (2013). A natural language processing and semantic-based system for contract analysis. In *Proceedings of the 2013 IEEE 25th International Conference on Tools with Artificial Intelligence* (pp. 707–712).



Appendix. Market Research Portal Architecture





ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CA 93943

www.acquisitionresearch.net