

Big-Data Analytics for Federal Procurement

On May 1, 2014, the Executive Office of the President completed a thorough evaluation of Big-Data benefits and risks that may transform the way we live, work and its effects on the relationships between government, businesses and citizens. Big-Data already impacts our lives as many businesses utilize Big-Data tools to better assess the demands and needs of their customers.

A few statements contained within the Big-Data evaluation can better define the potential benefits; *'In particular, Big-Data Analytics present an important opportunity to increase value and performance for the American people in the delivery of government services. Big-Data also holds enormous power to detect and address waste, fraud and abuse, thereby saving taxpayer money and improving public trust. Big-Data can further help identify high performers across government whose practices can be replicated by similar agencies and programs; and may deliver new insights into effective public sector management.'* [1]

How could Big-Data Analytics tools improve federal procurement in delivering insights for effective public sector management? The purpose of this document is to present our Big-Data Analytics for the Federal Procurement use case and briefly demonstrate how a government agency could use our Big-Data Analytics tools approach to answer this question. We hope our approach and strategies will help in the evaluation of other federal government programs and beyond.

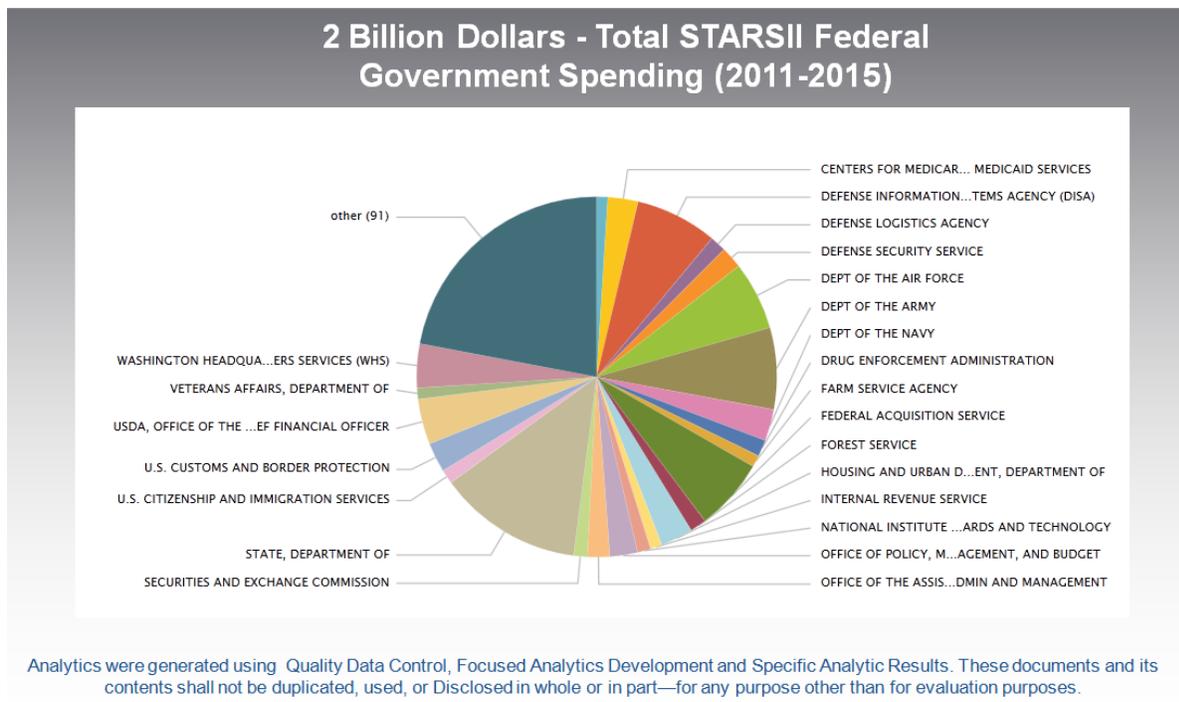


Figure A illustrates the total amount awarded by government agencies for IT Services over the last 4 years. Each slice of the pie can be drilled down to more granular analysis of the data.

Our research team cast our fishing net into the open sea of federal data sets containing contract information in order to capture the skeletal structural program domain of the GSA GWAC 8a STARSII program use case. Our approach in using Big-Data Analytics tools for this use case consisted of three phases:

- A. Quality Data Control
- B. Focused Analytics Development
- C. Specific Analytic Results

A. Quality Data Control – The federal procurement data collected and processed in the assorted traditional government systems is stored in legacy data warehouses in a variety of formats. This data does not need to be refined or normalized to fit into a Big-Data central repository. This data can be extracted from the government legacy data warehouse and migrated into a Big-Data repository very easily using an intuitive tool with a drag-and-drop interface. There are Big-Data analytical toolsets available that quickly process all types of data sets.

‘Government CFOs say their agencies’ services are at risk due to rising requirements, diminishing resources and the lack of quality data’ [2]

The challenge is in the process of extracting quality data and utilizing it within the Big-Data Analytics platform. Data must be verified for accuracy and completeness to ensure that it is error free. Critical decisions are made based upon the belief that data is accurate. Ensuring accurate data is required to create the trust of each individual customer.

Thoroughness for completeness and accuracy was pervasive throughout our systematic data cleansing approach when defining and extracting our data domain.

Our team was able to verify the data for accuracy and completeness using different mining strategies. Furthermore, the data was reconciled and validated with other government published procurement data and reports.

B. Focused Analytics Development – Our team considered three elements in the development of our analytics. First, our target audience was identified. Second, the information needs of our target audience were identified. Third, the Big-Data Analytics framework required by our target audience was defined for optimal performance.

1. A target audience should be defined early in order to avoid wasting time and resources needed in later phases of present and future analytics development. Our target audiences identified for our purposes were federal agency management, federal agency program staff, federal contractors (present and future) and the general public.

The target audience’s information needs will appear similar at a macro-level-view but become more specific as needs are further defined. By defining each target audience’s detailed informational needs the proper analytics can be developed and shared between relevant data users. This development process creates an open information environment using a single data source based on the level of detail shared between the different audiences.

2. Accurate analytics must be driven by performance improvements’ goals and objectives with measurable outcomes. Consideration of performance based analytics early will effectively support

improved program and project management.

3. Our procurement data was migrated into our Big-Data platform. We were then presented with an arsenal of available advanced data analytics tools in the same Big-Data Analytics platform. These user-friendly exploratory data analytics tools gave our team ample freedom to categorize user defined groupings that could be drilled down to granular levels for further analysis.

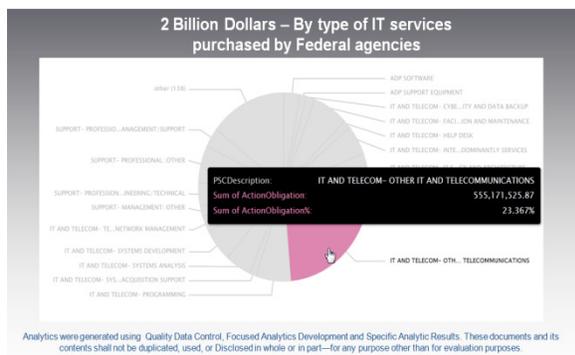


Figure B illustrates how Big-Data dashboards allow you to hover over a slice of the pie chart to see the amount spent by each type of IT service for a federal agency.

C. Specific Analytic Results – Finally our analytic results were customized to meet the target audience’s needs. We used graphs, dashboards, including data-driven spreadsheets and database formats. Our analytic results can be used to support various information needs and may drive decisions at various levels for projects, programs and agencies.

Once we defined our data modeling we were able to display the historical data results to predict new trends for its program. Trending analysis in a Big-Data Analytics platform

permits government entities the ability to perform comparisons and sensitivity statistical data analysis.

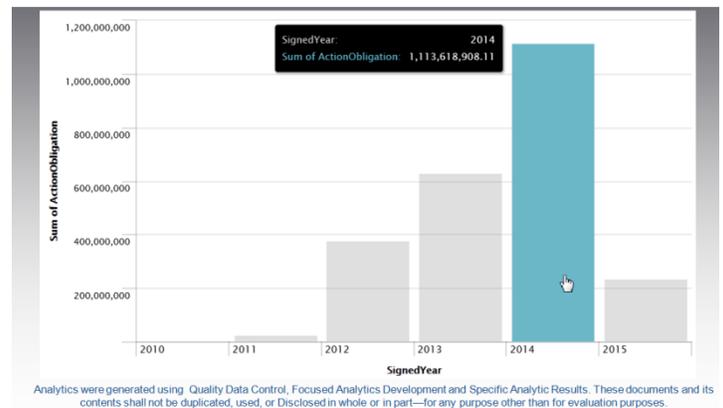


Figure C illustrates This Big-Data dashboard shows a bar chart to see the trend of the total amount spent per year.

We used rich visual tools to report our conclusions, create customized dashboards, and display trending analysis. Data can easily be organized and visualized to discover and share new insights with the appropriate users and audiences.

Conclusion

The White House Big-Data initiative has made its federal government data more easily available. Big-Data Analytics tools have tremendous power to consume this federal open data and scale up data analysis capabilities at a fraction of the cost and time required by traditional solutions.

The Big-Data Analytics approach presented in this white paper may be used for other government programs. Agencies may begin to cast their fishing net into the open sea of federal Big-Data central repository and capture their skeletal structural program domain.

‘Developing systems that can analyze Big-Data does not have to stretch agency budgets’ [3].

The sky is the limit with the myriad of opportunities offered by Big-Data Analytics tools. Big-Data Analytics empowers departments and agencies in the federal contracting market place with program compliance data not previously available.

Big-Data Analytics tools offer astonishing and powerful opportunities to unlock previously inaccessible insights from new and existing data sets in a cost effective manner.

REFERENCES

1. Executive office of the President, 2014. BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES
[online] Available at: https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf
[Accessed July 22, 2015]
2. Elena Malykhina, 2013. “Government CFOs: Bad Data Quality Hobbles Agencies”
[online] Available at:
<http://www.informationweek.com/government/leadership/government-cfos-bad-data-quality-hobbles-agencies/d/d-id/1111886?>
[Accessed July 22, 2015]
3. Colby Hochmuth, 2014. “Could big-data analytics improve federal procurement?”
[online] Available at:
<http://fcw.com/articles/2014/10/10/big-data-procurement.aspx>
[Accessed July 22, 2015]

About the Authors

Juan Kuanfung has over 30 years of information technology consulting experience and has provided services and support for Fortune 500 corporations and government agencies. Juan is the founder and CEO of Quanteq, Inc. Quanteq focuses on delivering IT solutions to address government regulatory requirements, customer phasing requirements and internal business process efficiencies. He holds dual Master's degrees in Economics and in EDP Management Systems from the University of Illinois. Juan serves as one of the advisors for Voice of Big Data; <http://www.voiceofbigdata.com/advisoryteam.html> and is the Programs Chair for NCMA Gateway Chapter serving the Greater St. Louis area; <http://ncma-stl.org/>. Juan is in the board of directors of Pronet International Gifts & Scholarships that helps veterans in need of furthering their education.

Tom Ryan is a professional engineer with over 37 years of project and program management in both the public (28 years) and private (9 years) sectors. As a professional engineer, Tom has lead major transportation projects and programs in various phases of planning, design, construction and operations. Tom’s experience in delivering transportation technologies projects like freeway and arterial management systems made for a nice transition to being the contract manager for Quanteq. Tom holds a Bachelor’s degree in Civil Engineering from Missouri Science and Technology. Tom has served on several regional and local transportation boards for ITS and ITE Chapters and is currently very active in volunteering at a local food bank.