Data Analytics within E-procurement can Enhance Procurement Integrity

By: Sheryl Goodman & Thomas Caulfield

The Misconduct:

Two Special Agents from the United States (U.S.) are conducting an investigation regarding an employee (hereafter referred to as Mark Swindle) of a federal contractor taking bribes from a construction company that is performing work on behalf of the federal contractor. The Special Agents have jurisdiction to investigate as the construction company is doing maintenance work on buildings of the federal contractor that are dedicated to only federal government contracts. The suspicion by the Special Agents was that if bribes were being paid, there would be a secondary scheme were Mark Swindle could have allowed for inflated or fake cost (i.e. mischarging) to be submitted by the construction company which in turn would be passed onto the U.S. government as in-direct cost[1].

In the early stages of the investigation, with the assistance of the federal contractors, the special agents conducted a basic data analytical search of the federal contractor's e-procurement system for all award decisions of Mark Swindle to identify if a pattern of irregularities existed. The analytics quickly revealed over a three-year period, Mr. Swindle awarded 116 out of 124 maintenance work-orders to the bribe paying construction company. These work-orders were supposed to be openly competed between seven different companies in the local area. Seeing the results, the decision was to perform

---

[1] Indirect costs represent the expenses of doing business that are not readily identified with a particular contract, project function or activity, but are necessary for the general operation of the organization and the conduct of activities it performs.

the same data analytical searches of the other 33 maintenance staff with the same authority as Mr. Swindle. This time six more employees were identified as also selecting the same construction company during the same three-year period over all other companies in the local area.

Federal Prison:

With the results of data analytics, the special agents questioned each of the individuals to which they each separately admitted to their level of involvement in the misconduct. In the end, the construction company owner – the mastermind of the scheme, and seven employees of the federal contractor were convicted for various federal crimes, including the Anti-Kickback Statute. The court sentenced the construction company owner to 13 months confinement in a federal prison and a $1.3 million fine. The employees of the federal contractor were fined a total of $800,000. One employee was sentenced to six months confinement and the others received supervised probations.   The investigation didn't just end with criminal convictions, the U.S. Department of Justice charged the federal contractor, under the Civil False Claims Act, with failing to ensure sufficient oversight on payments to the federal government. The company pleaded guilty and agreed to a $2.2 million fine.

Tip of the Iceberg of Data Analytics:

With the enormous amount of digital data being collected and stored by e-procurement systems, data analytics allows us to go well beyond being reactive and simply supporting an investigation and now allows us to be more proactive in identifying potential misconduct, reducing potential reputational damages, identifying financial losses faster, performing more effective contract oversight, improving internal processes, and achieving the ultimate goal of enhancing procurement integrity[2].

Data analytics can also be used to identify improper payments, which are not always a result of misconduct. Improper payments are ones that should not have been made or that were made in an incorrect amount (including overpayments and underpayments). For fiscal year 2017, the U.S. Government Accountability Office reported that federal entities estimated they improperly paid about $141 billion across 90 programs in 21 agencies. This total was down from $144 billion for fiscal year 2016, but up from $137 billion for fiscal year 2015[3]. These reported lost billions of improper payments in the federal government do not represent those only in procurements, but includes multiple social programs. However, with an estimated $10 trillion on average spent each year

---

[2] The term *procurement integrity* refers to honest, fair, impartial, and legal contracting, free from corruption, fraud, and/or improper payments.
[3] These estimates were identified on GAO's website.

globally on just government contracts, the risks of improper payments remains a significant concern[4].

Data Analytics:

Data analytics typically refers to the use of analytical techniques to identify trends, patterns, anomalies, outliers, and exceptions in data. For several years, organizations have used various data analytical techniques and approaches in a full range of ways from marketing, customer surveys, and influencing financial investment. Data analytics eventually became a technique to help ensure payment integrity like that used by credit card and healthcare providers. It is also being used as a tool in identifying potential misconduct or non-compliance to internal policy.

Depending on the data analytics used, it may involve a single or multiple techniques. More sophisticated techniques are also being designed where the data analytical results improves its own capabilities (i.e. machine learning or artificial intelligence). Below is a list of four common data analytics techniques that can be deployed against procurement data. Some may identify indicators of a single misconduct scheme; some will identify potential indicators of multiple schemes; and others will identify indicators of both a misconduct scheme and potential improper payments.

> ❖ Parameter Analytics
> ❖ Distribution Analytics
> ❖ Social Networking Analytics
> ❖ Predictive Analytics

How to Start:

The first step to any data analytics approach is "knowing the data." Knowing the data is not just understanding what data your e-procurement system is collecting, but also its availability, the reliability of the data, its format (i.e. structured with predetermined data sets or free flowing textual or both), and how the data can be analyzed to identify indicators of specific schemes of misconduct or improper payments. For example, if the e-procurement system is limited to only collecting vendor information (i.e. vendor management system) and contracting data up to the awarding of the contract, your data analytics may be restricted to misconduct like shell companies, bid rotation, market sharing, leaking of bidding information and others. If your e-procurement system does not include post award transactional data (i.e. payments to the vendor), then you are missing a critical component of identifying risk for improper activity.

If your e-procurement system has both pre-award and post-award data, your analytics can be much more complex throughout the enterprise. For example, if your e-

---

[4] This estimate was reported by a Lead Procurement Specialist for Public Integrity and Openness, World Bank Group, at the 2019 United Nations Conference of State Parties.

procurement data collects the date of contract award, the date of the first contract payment, and each payment amount, you will be able to identify any vendor that received payment(s) prior to contract award and as well as any vendors receiving payments inconsistent with similar contracts. With regard to improper payments, if your e-procurement system can identify the number of invoices received and the amounts, compared to the number of payments made and the amounts, then with data analytics any differences could be identified as questionable.

Along with knowing the data, you also need to know the "corruption and fraud schemes" that are traditionally being deployed in today's contracts. Schemes like bribes, gratuities, extortion, false invoicing, multiple invoicing, and conflict of interest are all examples of known schemes. However, with properly structured data analytics, even the lesser known schemes like unbalanced bidding or anti-competitive schemes can also be identified. For example, to identify an unbalanced bidding scheme your e-procurement system would need to collect cost proposal data consistently from each bidder down to each contract deliverable, then with data analytics any outliners from one bidder compared to all the other bidders would stand-out (i.e. unbalanced bidding).

Along with knowing the schemes, you have to know how to design "an analytical approach" that will indicate the specific scheme which may be or has occurred. Knowing how to develop a scheme specific analytical approach is one of the most difficult aspects of data analytics of e-procurement data as there are just over 40 different corruption and fraud schemes that can be deployed in today's contracts. Each of those schemes can further be done a thousand different ways. Designing an effective analytical approach requires a strong understanding of the indicators (i.e. red flags) of each of the corruption and fraud schemes.

Techniques:

Parameter Analytics

Parameter Analytics is a technique that uses a limited number of data sets combined in a simple Boolean condition[5] (true or false – yes or no) followed by an action. A very simple example of such a condition would be – "is the name of the requester of the goods and that of the procuring official the same – if yes identify?" Parameter Analytics and building "Boolean" conditions will normally start from a known statement – "if the contract is structured for quarterly payments the maximum number of payments is four (4)." The "Boolean" condition would be "is the number of payments greater than four?"

This type of analytics could be used effectively in establishing both potential indicators of misconduct, for example "identify instances when a contract has an invoice that has

---

[5] Boolean conditions are queries that compare two values with each other (e.g. with == or >=) and then returns the value true or false or yes or no. Which values are interpreted as true or false depends on the data type.

multiple dates," "identify instances when a contract has an invoice that has two different checks associated with it," or "identify when the vendor is located at a non-business address or not listed on the internet." Along with identify potential misconduct, these say parameter analytical searches could also identify improper payments that were caused by human error.

Other examples of Boolean conditions could be – "was a vendor entered into the e-procurement system and did their first payment come within 15 days or less" or "does any invoice have multiple tax identification numbers," or "does any contract have a change-order within 20 days after contract award?" Each of these can be filtered as a "true or false" – "yes or no" (Boolean condition) and whenever positive identify. It is important to remember that data analytics does NOT identify misconduct, it can only identify indicators or questionable events and therefore each may require additional follow-up to determine the true cause.

Distribution Analytics

Distribution Analytics relies on accurate historical data, known patterns and thresholds. This type of analytics is looking for outliers throughout the e-procurement data, like payments by percentage higher than prior purchases for the same items or a greater number of contract modifications than that of a similar contract. These outliers are more commonly referred to as "anomalies" and can be identified when looking at multiple prior procurement occurrences and is identified whenever the same type of procurement is not consistent. For example, identify when the time between invoice payments is abnormally low, or there is a higher number of payments of a similar contract, or a higher number of purchases under a multiple-award-contract for a product compared to a judgmental sample. Another example could be identifying the percentage of sole-sourcing being performed by a procurement official compared to others. All of these are based on a deviation from what is normal.

Distribution analytics can also be used across an entire organization or even a Country like what Instituto Mexicano para la Competitividad (IMCO), a Mexican non-governmental organization, had conducted to spot suspicious patterns in the Mexican government's procurement awards[6]. They identified suspicious patterns related to which vendors had historically won what percentage of public contracts, who the awarding government organizations were, who were being paid what and when. Using the Mexican federal government's procurement data, IMCO found that at the time of the 2012 presidential election, 73 companies had won 29 percent of all the public contracts awarded that year. But four years later, in 2016, the same group of companies had won barely one percent of all contracts awarded. During the same period in 2012, a second group of companies had won only 2% of the total amount, but in 2016 they ended up winning over 32%. These are concerning statistical changes given the stability normally found in public procurement

---

[6] IMCO, Índice de Riesgos de Corrupción: El sistema mexicano de contrataciones públicas, https://imco.org.mx/articulo_es/indice-riesgos-corrupcion-sistema-mexicano-contrataciones-publicas/

markets. The questionable pattern IMCO unearthed demonstrates the value in enterprise-wide data analytics and its capabilities. These finds were provided to the Anti-Corruption Authority of the country for further investigation.

Social Network Analytics

Social Network Analytics, also called "link analysis", is a technique used to determine the probability of a relationship among different entities and/or individuals. This analytical approach will normally generate a web-like graphic display that visualizes and links the type and strength of connections between organizations or people. These connections might include shared addresses, phone numbers, company affiliations, family relationship, or other business connections. Social network analytics could enable contracting officers and other procurement officials to react immediately when a connection is discovered and attempt to prevent or limit the impact.

Social Network analytics can be designed with a strong reliance on unstructured data such as blogs, video services, social networks, discussion forums, and others. An example of this type of social network analytics could be using public media networks like Twitter, Facebook and/or LinkedIn to track potentially inappropriate relationships. Social networking analytics is designed with the underling reality that most misconduct schemes occurs between people familiar with each other and have previously established relationships.

Predictive Analytics

Predictive analytics is the gold star of techniques as it can forecast a potential future occurrence of improper conduct or non-compliance to policy or improper payments. Predictive analytics has the potential of being one of the more effective contract management techniques. However, predictive analytics is not easy as it requires assigning a numerical value of risk (i.e. risk factor) to individual procurement actions, combining the risk factors together, compare the results to historical procurement data of the organization, and establish a range of low to high risk. What makes this analytical approach so difficult is the identification of risk factors and properly assigning a numeric value to them sufficiently to reduce any false positives.

To give an example, first let's start with the following risk factor statement - "using a range of 1 to 10, with 10 being the higher risk, assign a score based on the number of standard deviations away from the number of contract modification (mods) for a similar contract. The higher the percentage of deviation the higher the risk score." The predictive analytical technique would then automatically combine this risk factor with others like the following:

- Number of mouths for the period of performance – the greater the number the higher the risk;

- Contract value – the greater the amount the higher the risk;
- Procurement method used – sole-sourcing and emergency purchasing assigned higher risk,
- Number of follow-on contracts – the higher the number the greater the risk;
- Degree of complexity of the effort – low complexity low risk;
- Prior incidents with a vendor – five or more prior incidents of poor performance a higher risk; and
- Number of days from contract award to the first contract mod compared to similar contracts – over 40 days low risk, between 20 and 39 days medium risk, under 20 days high risk.

When combined, it will provide a statistical probability of the contract's degree of risk, which can be displayed to visually give a procurement professional an "early warning" and an opportunity to determine what may need to occur. The statistical probability can be visually displayed to identify each of the numerical values for each risk factor, giving the procurement professional a better understanding of the specific risk condition. The visual display can be configured to perform this type of analytics across a division of the organization, a region of the country, or to a specific type of contract. This type of predictive analytics is currently being performed within a small number of entities.

Conclusion:

Honest, fair, impartial and legal contracting, free from misconduct like corruption, fraud or failure to comply with policy is achieving true Procurement Integrity. It also includes ensuring all payments are consistent with the conditions in the contract and are ones that should be made in their correct amount. It is no longer a question of how can data analytics enhance procurement integrity and identify improper payments, rather it's how can we maximize the various techniques. Parameter, Distribution, Social Networking, Predictive analytics and other not mentioned now give us the ability to be more fiscally responsible, provide greater insight to any high-risk contracts, and identify at a faster rate potential misconduct.

Data analytics is not the end all solution, but when combined with well-qualified procurement professionals, in a setting committed to procurement integrity, its potential is limitless. If you would like more information on e-procurement data analytics, or ensuring you are operating in a procurement integrity environment, please contact the authors at www.procurement-integrity.net.