Corporate Partner
Advisory Group
Research Series

Leveraging Data Analytics
in Federal Organizations

AGA®

Acknowledgements

About the Authors

Helena G. Sims is a co-author of this report. Most of her 35 years of professional experience have been devoted to forging partnerships among public and private organizations. Since 2007, she has served as Director of Intergovernmental Relations for AGA. Prior to her work with AGA, she was Managing Director of Public-Private Partnerships for a banking trade association that played a key role in fostering a nationwide system for the electronic delivery of government benefits. She served as Washington Office Director for the National Association of State Auditors, Comptrollers and Treasurers and has also worked for the Council of State Governments, the U.S. Senate and the Florida Legislature. Helena earned bachelor's and master's degrees in speech communication from Florida State University.

Steven E. Sossei, CPA, is a co-author and the primary researcher for this report. He is a retired Director of State Audits from the New York State Comptroller's Office, where he worked for 35 years. Steve has extensive experience in auditing all types of government and not-for-profit operations. In his role as Director of State Audits, he was responsible for directing the audits of New York's Medicaid program, public employee's health insurance plan and New York City's school system. He also worked extensively with the National Association of State Auditors Peer Review program, working on more than 20 peer reviews during his career. Steve also served as AGA's Director of Education from 2010 through 2011. He earned his bachelor's and master's degrees in accounting from the State University of New York at Albany.

Sponsor: AGA is proud to recognize Accenture for supporting this effort.

AGA's Corporate Partner Advisory Group Research Program

One of the roles of professional associations like AGA is to develop new thinking on issues affecting those we represent. This new thinking is developed out of research and draws on the considerable resources and experiences of our members and counterparts in the private sector – our Corporate Partners. These organizations all have long-term commitments to supporting the financial management community and choose to partner with and help AGA in its mission of advancing government accountability.

AGA has been instrumental in the development of accounting and auditing standards and in generating new concepts for the effective organization and administration of government financial management functions. The Association conducts independent research and analysis of all aspects of government financial management. These studies make AGA a leading advocate for improving the quality and effectiveness of government fiscal administration and program performance and accountability.

Our Thought Leadership Library includes more than thirty completed studies. These in-depth studies are made possible with the support of our Corporate Partners. Download complimentary reports at www.agacgfm.org/researchpublications.

Table of Contents

Executive Summary

Data analytics is a powerful tool that can help government agencies reduce fraud, waste and abuse. The commercial sector has used data analytics for years to improve decision making, achieve better financial outcomes and improve customer service. The use of data analytics is growing at a rapid rate. The International Data Corporation, a provider of market intelligence in the information technology field, estimates that the business analytics market for software, hardware and consulting services is expected to grow at an 8 percent rate worldwide, reaching nearly $33 billion in 2012.

AGA set out to determine how the federal government is using data analytics and what it is doing with the resulting information. We interviewed eight agencies and surveyed a broad spectrum of federal financial officials. From this, we learned that some federal agencies have embraced data analytics and have demonstrated the benefits of integrating analytics tools into their operations. As a result, the federal government is in a position to build on the analytic advances it has already made. Some organizations are poised to share their capabilities with other federal organizations and possibly with other levels of government that implement federally funded programs. However, there is no clear plan to leverage the government's investment in data analytics.

From our agency interviews, we found:

■ The Food and Nutrition Service (FNS) within the U.S. Department of Agriculture reduced the rate of food stamp trafficking from about 2.5 percent of food stamp benefits to about 1 percent by using data analytics. (Trafficking is the buying or selling of food stamp benefits for cash.) Based on benefit levels of $75.6 billion in FY 2011, we estimate that this would translate into $1.1 billion of benefits that were not trafficked last year. Its data analytics system has allowed FNS to quickly identify merchants who traffic in food stamps and remove them from the program. The reduction in food stamp benefit trafficking helps ensure that funds are spent on their intended purpose.

■ In a matter of months, the Recovery Operation Center (ROC) at the Recovery Accountability and Transparency Board implemented a powerful data analytics system to screen those who received much of the $800 billion appropriated in stimulus funds. Using its high-powered system, the ROC identified recipients who had previous brushes with the law or who were receiving multiple awards.

■ The Center for Program Integrity (CPI) at the Centers for Medicare & Medicaid Services developed a predictive analytics system for Medicare fee-for-service payments that screens $450 million in Medicare claims each day. The system is just beginning to bear fruit but should have a major impact on identifying and reducing Medicare fraud.

We also surveyed federal financial officials on the development and use of data analytics in their operations. Two-thirds of the respondents reported the use of data analytics in operations, with nearly all systems focused on financial performance, improper payments and identifying high-risk investigative targets. This is not surprising, as the federal government has launched a major initiative to reduce the annual rate of improper payments, currently estimated at $125 billion annually. The remaining third of respondents, who said they had not implemented a data analytics system, identified multiple reasons for the delay. Sixty-seven percent of respondents to this question cited a lack of budget, 53 percent cited a lack of staff and 33 percent said they were unsure of how to start developing a data analytics system. (Respondents were able to provide more than one response.)

Our research identified the following common characteristics in successful data analytics adoption:

Leadership—Not surprisingly, the most important factor for success was strong leadership at the top of the organization. Successful projects began with a strong commitment from leadership and a clear articulation of the project's goals. The U.S. Postal Service's Inspector General, David C. Williams, led data analytics efforts in his organization and remained actively involved in the system's development. He served as a champion for the project, monitoring the work at a high level and providing a clear vision for the system. It is critical for leaders to help an organization clearly define what it wants to accomplish through data analytics. Once a system's goal is defined, leadership is critical to execution. Engaged leaders help break down silos within organizations, reach out for technical assistance and propel the organization toward its goal.

Data Availability and Resources—All of the organizations we interviewed were able to primarily rely on in-house data to draw effective conclusions. Some organizations supplemented their internal data with information from outside sources, such as Dun & Bradstreet – a global source of commercial information on business. Depending on the data analytics project's size and objectives, an organization can create the computer systems that analyze the data in different ways. The incremental approach, which uses existing software and computer system capacity, is likely to be the least costly approach to system design, development and implementation. The U.S. Postal Service's Office of Inspector General (USPS-OIG) took this approach, while the Recovery Accountability and Transparency Board was able to build an entire new system that included outside data.

Staff Knowledge and Skills—All of the organizations used a combination of

> Two-thirds of the respondents reported the use of data analytics in operations, with nearly all systems focused on financial performance, improper payments and identifying high-risk investigative targets.

internal and contracted staff to complete the project. Staff within the organization typically provided business expertise or program knowledge. They understood the problems that needed to be solved, the laws and regulations governing the program and the data and resources available within their organization. Outside expertise or knowledge about statistics, modeling, and software design or computer code was typically needed for analytical work. The interpretation of the results was generally accomplished by a combination of internal and contracted staff.

Integration of Data Analytics into the Work—To be successful, data analytics systems must prompt action. They can support management decisions, help spot irregularities and direct the work of audit and investigative staff. Successful data analytics systems are seamlessly integrated into the organization's operations. In turn, information revealed by the system prompts action and is used to update and improve the system's models and algorithms. An effective data analytics system is not static, but continually enhanced.

Through our research, we found that various agencies and financial organizations have used data analytics at varying levels, from simple to advanced, with the purpose of solving a business need. Progress in adopting data analytics has been inconsistent across the federal government for a number of reasons. Lack of leadership, scarce budget resources

or the need to acquire the appropriate expertise can slow a project.

The Government Accountability Transparency Board (GATB) recently identified a number of factors that have inhibited the development of data analytics. For instance, the GATB reported that data is often collected in different ways by different systems across federal agencies, thus making analysis difficult. Although the GATB made several recommendations for strengthening data collection methods, many of these recommendations take time to implement. But some short-term actions might bear results. For example, the Recovery and Transparency Board's ROC has already begun providing services to other organizations. In another instance, the USPS-OIG's Risk Assessment Data Repository System could be modified by other agencies and used in screening for contract and worker compensation fraud.

Based on our research, we are making five recommendations regarding data analytics integration within the federal government:

Transfer Analytics Success among Agencies—The federal government should build upon demonstrated data analytics successes. It should strive to leverage the experience gained by those who have implemented data analytics systems to help other organizations overcome deficiencies in budget, staffing and experience.

Advance Education and Enhance Capabilities—While it appears most government leaders have a general understanding of data analytics, they are not yet champions of the process. Federal leaders must continue to be educated on the benefits and uses of data analytics and the actions needed to implement data analytics in their agency. One possible key to building on success, advancing education and enhancing federal capabilities is to develop a "How-to" or "Best Practices" guide. Since respondents identified overcoming budgetary and staffing obstacles, the guide could assist the 33 percent who indicated they were unsure of how to start developing a data analytics system. The guide could also reduce agencies' risk in implementing data analytics systems.

Focus on Performance and Outcomes—Organizations should expand the use of analytics to include performance and outcomes. Only one of the eight organizations interviewed was using data analytics to measure performance, yet the 2010 amendments modernizing the Government Performance and Results Act (GPRA) establish performance indicators for measuring or assessing progress toward required agency performance goals. The other seven organizations were using data analytics to prevent and detect improper payments, obtain information on financial performance and identify instances of fraud and abuse for investigation and audit.

Continually Update Analytics Systems—Data analytics systems must be continually monitored and updated to reflect lessons learned. Algorithms, routines and models should be updated based on experience to ensure that the results will continue to improve and drive organizational action. Successful data analytics systems are not static.

Simplify Procurement—A guide to procuring data analytics systems or consulting services should be developed (ideally by a neutral third party) based on the experiences of federal officials that have procured systems and the vendors who provide these services. This guide could include information on the services that can be procured, ways to determine the correct data analytics method, provisions that significantly increase or decrease the cost of a contract and lessons learned from previous procurements.

Explore Intergovernmental or Shared Services—Federally funded programs implemented at the state or local government levels should be reviewed to determine whether a collaborative arrangement can be developed for data analytics. A central system administered at the federal level might be more effective and cost-efficient than stand-alone systems throughout the country.

Objectives, Scope and Methodology for the Study

The objective of this study is to provide the government accountability community with information on the successful development of data analytics systems within a sample of federal agencies. We sought to identify common factors and characteristics of agencies and programs that were successful in establishing and implementing a data analytics project. Our study focused on management and organizational issues, including agency or program leadership, funding, human capital for acquiring and sustaining the project, and the need for program and process reengineering. We also asked about the procurement methods used, if any, for both software and consulting contractors. Other areas of inquiry included sources of data, change management issues, the basic data analytics techniques employed by the systems, how system results were used and the success of each program.

As part of our research, we interviewed officials from eight offices within six federal agencies. We questioned officials on a variety of topics including the factors that led them to conclude they should develop a system, the basic process they employed to develop the system, the contributors to their success and lessons learned. We also interviewed several firms that had conducted data analytics work for federal agencies.

We conducted interviews with the following federal organizations:

- Recovery Accountability and Transparency Board, Recovery Operations Center (ROC)
- U.S. Department of Agriculture, Food and Nutrition Service (FNS)
- U.S. Department of Defense, Defense Finance and Accounting Service (DFAS)
- U.S. Department of Defense, Defense Logistics Agency, Office of Operations Research and Resource Analysis (DORRA)
- U.S. Department of Defense, United States Navy, Naval Sea Systems Command, Office of Fraud Deterrence and Detection (NAVSEA-OFDD)
- U.S. Department of Education, Office of Inspector General (ED-OIG)
- U.S. Department of Health and Human Services, Centers for Medicare and Medicaid, Center for Program Integrity (CPI)
- U.S. Postal Service, Office of Inspector General (USPS-OIG)

We also surveyed federal financial officials including chief financial officers, inspectors general and deputy chief financial officers to gain information on their activities related to integrating data analytics into their operations.

Our work did not concentrate on the technical aspects of the systems that were developed. We did not ask whether the systems were using a particular analytical software package, were web accessible and/or used predictive data modeling of a certain type. Rather, we dealt with the overall processes used to develop the system and attempted to identify the common factors that appeared to contribute to successful development.

Introduction

Businesses and governments alike have always faced economic pressures. Businesses are constantly faced with demands to increase market share, improve profitability, secure resources and develop the latest product. Competitive pressures also stem from factors such as the globalization of markets, the rise of Internet commerce and improved automation. These pressures have contributed to the rise of data analytics in businesses, whose leaders have found that they must quickly transform data into information that can be acted on by managers and executives.

Governments are also under economic pressure. Revenues are dropping, and citizens' needs in areas such as improved transportation infrastructure and unemployment compensation are increasing. Financial strains are prompting citizens to scrutinize spending and question where money is going and what it is buying. News outlets are asking for more information and a variety of specific reports. And citizens are asking the government and its agencies to reduce waste, identify and eliminate fraud, and redesign programs to improve efficiency and effectiveness.

At the federal level, this demand for a more efficient and effective government has also been reflected in two important pieces of legislation, the Government Performance and Results Act (GPRA) and the Improper Payments Elimination and Recovery Act (IPERA). GPRA, enacted in 1993 and amended in 2010 with the GPRA Modernization Act, is designed to improve government operations by requiring agencies to set relevant and meaningful goals for their programs, measure their actual results against those goals, and report on their progress. To implement GPRA, federal agencies produce strategic and performance plans, measure their results and conduct gap analysis.

IPERA, also enacted in 2010, is designed to reduce improper payments. An improper payment is one paid in the wrong amount, paid to the wrong person and/or paid for the wrong reason. In a November 2011 press conference, the director of the Office of Management and Budget reported that improper payments in FY 2011 were estimated to be $115 billion, down from an estimated $125 billion in FY 2010. IPERA set a goal to reduce improper payments by $50 billion in 2012.

To reduce improper payments, government officials need information. The government collects tremendous amounts of data, but information in its raw form is not very useful. To be effective, data must be collected in a timely manner, analyzed quickly and presented in an understandable format. Some government agencies are already using this process of data analytics. For instance, the IRS has been employing data analysis for years in questioning the accuracy of income tax returns. It has matched wages reported by employers and interest earnings reported by banks to individual income tax returns since the mid-1970s. Other government agencies, such as the Census Bureau, have analyzed census and income tax return data to publish trend reports. The difficulty with these systems, however, has been the lag between when data has been collected and when it is analyzed and provided to management to make decisions or take some form of action. In the current environment, managers need information to make decisions quickly. This is where data analytics delivers value.

Part One: Work Has Started in Federal Agencies

Two research reports released by AGA last year identified opportunities to deploy data analytics in government operations.[1] AGA's report "Improper Payments: Not Just the Purview of the CFO Anymore?" noted that one of the root causes for improper payments is the tension that exists between a program's desire to make speedy payments and the delays that are caused by ensuring a payment is fully supported. The report concluded that "(r)educing improper payments costs money, and it is not clear how much. If the mission is to redistribute wealth or help the poor and aged, many federal officials and members of Congress feel that erring on the side of paying out the benefits quickly is better than cold-hearted efficiency. Until the prevalent culture that creates this attitude changes, the best attitude may be that a certain level of improper payments is a 'cost of doing business.' The use of business analytics through large databases and data mining holds much promise of reducing improper payments, but to some extent its widespread use depends on people's willingness to accept some loss of privacy."

A key finding in the second report, "Using Performance Information to Drive Performance Improvement," indicated that opportunities to deploy data analytics more effectively exist in operations related to GPRA.[2] This report found that "(a) few agencies reported having automated performance information systems, with one reporting that its system has built-in business intelligence and data-mining capabilities. In most instances, however, the data are collected with data calls and entered into Excel spreadsheets. The agencies would prefer to have an automated performance information system, but lack of resources appears to be the primary obstacle."

The need for improved data analytics systems exists, and our interviews at federal agencies indicate that some agencies have and are continuing to bring data analytics into their operations. Overall, trends in the use of data analytics are positive.

Our interviews revealed that much of the work in data analytics is focused on financial resource monitoring and control, with an emphasis in two specific areas: preventing and detecting improper payments and identifying instances of fraud and abuse for investigation and audit. These advances have produced important results and are to be applauded. However, only one of our case study agencies, the Defense Logistics Agency, developed a system focused on measuring the agency's performance in fulfilling its mission of supplying its customers. Additional examples of data analytics being deployed to improve program effectiveness and efficiencies can be found in a report issued in November 2011 by the Partnership for Public Service entitled "FROM DATA TO DECISIONS – The Power of Analytics," which highlighted federal agencies that are implementing data analytics to improve program effectiveness and efficiency.

FEDLINK's Centralized Procurement Contracts

Other federal agency actions also reveal increased interest in data analytics. FEDLINK, which operates under the Federal Library and Information Center Committee (FLICC) within the Library of Congress, provides centralized services to the FLICC, as well as to other federal operations. FEDLINK executes centralized procurement contracts that can be used by most federal agencies. FEDLINK issued a Request for Information (RFI) on December 1, 2011, to gather information on the extent of data analytics services and the kind of software available in today's marketplace. The FEDLINK official said that its customers are most interested in contracts for data analytics services and software. The information gathered through the RFI was used to prepare a series of Requests for Proposals that were issued in 2012 with the aim of awarding a series of centralized contracts for data analytic services and software for use by federal agencies.

Survey of Federal Financial Officials

We also surveyed federal agency officials in February and March 2012 to ascertain the extent and focus of data analytics activities in their agencies. Our survey was sent to the offices of inspectors general, chief financial officers and other federal financial officials. Results from the 39 responses we received have been incorporated throughout this report. Our survey asked about the extent to which data analytics had been implemented in their operations, views on leadership support for the use of data analytics and the levels of expertise in developing and using data analytics in the agency. We also asked questions about any barriers that might have slowed the development of data analytics in agencies.

The first survey question asked if an agency had implemented data analytics to improve decision making. As shown in Figure 1, two-thirds of the federal officials responded that their agencies had.

This response can be viewed as a positive sign for data analytics in government. We also asked respondents, who could select more than one answer, to identify their areas of focus. The answers indicate that data analytics have been implemented most frequently in four key areas:

1. The detection and prevention of improper payments for IPERA
2. Information on financial performance and/or budgeting decisions
3. The identification of vendors or contractors for further audit and investigation
4. Staff deployment

However, it is equally important to note that one third of the respondents reported that their agency had not implemented data analytics to improve decision making. We asked these respondents what factors have prevented them from proceeding with data analytics. (Respondents were able to provide more than one answer to our question.) The top three answers provided were:

1. A lack of budget resources **(67%)**
2. A lack of appropriate staff **(53%)**
3. An uncertainty as to how to develop a data analytics system **(33%)**

With budget limitations, a lack of staff resources and an uncertainty as the top three factors preventing progress on data analytics, it makes sense to leverage the experience gained by those who have implemented data analytics systems to help other organizations overcome these obstacles.

The continued integration of data analytics into all facets and levels of government has the potential to make a marked difference in governmental accountability and performance. With the effective use of data analytics, governments will be better able to demonstrate that they are monitoring funds in a prudent manner and are achieving effective results. Increasingly, taxpayers want to know that governments are spending funds wisely and that small problems are being identified before they become big problems. Today, improvements in data collection, increases in Internet speed and decreases in data storage costs have allowed information to be collected and stored closer to the time of its generation. These advancements will allow managers to have more timely information and data on events and operations to make better decisions.

FIGURE 1: HAS YOUR FEDERAL AGENCY IMPLEMENTED DATA ANALYTICS TO IMPROVE DECISION MAKING?

NO **33%**

YES **67%**
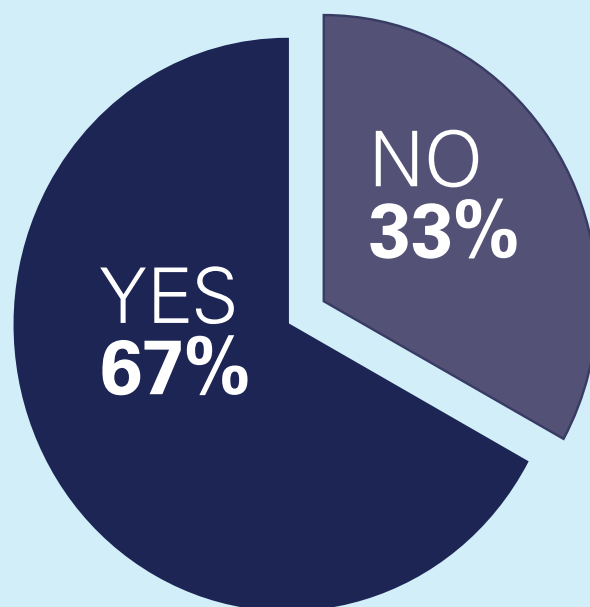
Part Two: What Is Data Analytics?

The goal of the data analysis process (or data analytics) is to support decision makers by identifying patterns and trends and by highlighting useful information. Data analytics can help to identify and/or develop information that would otherwise not be discernible by simply examining raw data.

A variety of data analytics methods are available; no one standard method applies to all situations. Rather, the correct data analytics method must be selected from among a variety of possibilities. Identifying the best option can be a complex task, and it may be useful to use a team approach in making a determination.

In some cases, off-the-shelf software, like spreadsheet programs, might be sufficient. For example, a government might determine which vendors to audit by arraying the total dollar value of payments to each vendor from high to low. The auditors could then review the list and select the top 10 vendors for an audit of payment accuracy and service quality.

Data matching could also be considered another form of data analysis. Data matching involves finding similar entities across disparate databases and matching the information to show exceptions. For example, a government entity that pays for medical services for government employees might match the claims it receives for dental services to a file of licensed dental providers before it authorizes payment. Unlicensed dentists or mismatches would be subject to audit follow-up or rejection. Government agencies have successfully used data matching techniques for many years.

FIGURE 2: BASIC STEPS IN DATA ANALYSIS

Data Collection → Data Cleanup → Initial Data Analysis → Revision of Analysis Based on Initial Results → Output from Data Analysis

Data analytics is a buzz word or catch phrase that has come to be associated with the field of data analysis. It generally refers to a variety of processes and techniques all focused on improving the value of information for decision makers. The basic process of analyzing data is depicted in the graph above.

However, these techniques can also have a downside by producing more false positives than the agency can effectively handle. (A false positive would be an exception that appears to be an error based on an initial analysis, but is found not to be an error when investigated.) If there are too many false positives, an agency's staff will expend valuable resources reviewing proper and correct transactions.

False positives can be reduced by using more advanced forms of data analytics. This requires better data, better systems and a more highly trained staff. Shown here is a list of data analytics techniques Elder Research, Inc., has used in providing training for

AGA. (Elder Research officials said that the list was based on the Eight Levels of Analytics identified by the SAS Institute Inc.) The list classifies the data analytics techniques as either descriptive or predictive. Descriptive techniques are those that describe the population being examined. Predictive techniques use patterns discovered by examining historical data to identify previously unidentified risks and opportunities.

The list arranges data analytics techniques from the most elementary level, number 1, to the more advanced level, number 9. Standard Reporting, listed as number 1, includes any reports generated by a system intended for management use in monitoring operations. For example, a standard report from an accounts receivable system would be a list of overdue accounts receivable. Management would use this type of standard report to determine whether overdue accounts receivable is growing or remaining stable and to identify which accounts should be sent to collection. This would constitute data analytics because it provides management with information for making decisions based on insights that are not discernible by simply looking at individual transactions. The standard report is just the first step in data analytics. As the analysis becomes more sophisticated, management may develop special queries or drill-downs to produce more useful information, such as the geographic location that produces the highest level of overdue accounts receivable.

Predictive techniques such as predictive modeling, link analysis and text mining are more advanced techniques that are designed to predict outcomes. Predictive data analytics methods are commonly used in the computation of credit scores. A credit score assigns a numerical rating to an individual that will predict the risk of default associated with loaning the person money.

Data mining, another term frequently associated with the field of data analytics, is both a descriptive and predictive data analytics technique. Data mining is defined as the examination of large sets of data, using a variety of techniques to discover patterns in data that were unknown to users. For instance, a clustering data mining technique would be used to discover transactions that appear to fall outside of the population's norms. Clustering techniques assign data into clusters so that the objects are grouped similarly by a number of characteristics. This same analysis would identify objects that fall outside of the clusters. Researchers then use the information learned from investigating the results to further refine the data analysis. Figure 4 illustrates the typical data mining cycle.

FIGURE 3: ANALYTIC TECHNIQUES

Descriptive

1. Standard Reporting

2. Custom Reporting or "Slicing and Dicing" the Data (Excel)

3. Queries/Drilldowns (SQL, OLAP)

4. Dashboards/Alerts (Business Intelligence)

5. Statistical Analysis

6. Clustering (Unsupervised Learning)

Predictive

7. Predictive Modeling

8. Optimization and Simulation

9. Next Generation Analytics: Text Mining and Link Analysis[3]

FIGURE 4: DATA MINING CYCLE
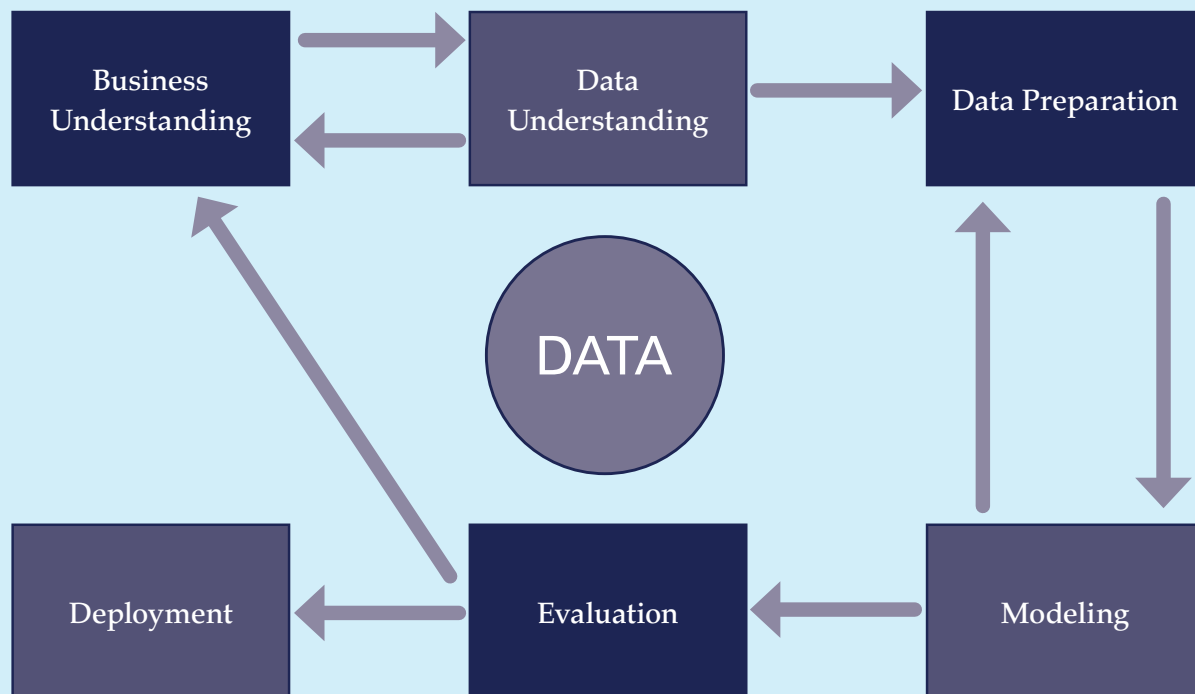
As depicted in the graphic above, data mining work starts with the organization gaining an understanding of its business and the associated risks. The next phase involves understanding the data and its qualities, quantities and sources. Data preparation for analysis includes cleaning to identify and correct errors. Data analysis follows, and exceptions (or outliers) are identified. Exceptions are evaluated, frequently in some form of investigation or audit, to confirm that they are true exceptions. The results of the evaluation phase are then factored back into the process to further refine the results and the models.

Part Three: Key Questions for Implementers

This part examines key questions that organizations should consider when designing, developing and implementing a data analytics system. The questions, which are based on lessons learned from our interviews and through our survey, are:

- What is the goal of the data analytics project, and is leadership committed to the project and its goals?
- What data will be needed for the data analytics work, and is this data available? In addition, what challenges must be overcome before the data can be effectively analyzed?
- Does our staff have the knowledge and skill sets needed to design, develop and implement our analytics processes and systems?
- Has the agency considered the need for cultural and organizational skills?
- Does the organization have systems and software to implement a solution, or will they need to be acquired?

We begin this part by providing detail on the oldest system among those we examined — the ALERT system developed by the Food and Nutrition Service (FNS) of the U.S. Department of Agriculture. We chose ALERT as an illustrative example because it is a highly mature system. Based on the ratings below, the ALERT system operates at a high level with respect to both its level of maturity and the degree to which FNS relies on its results.

Level of System Maturity

High level — Analytics are well established, and the results are trusted by decision makers and integrated into operations.

Medium level — Analytics are still being refined, results are still subject to caution and integration into operations is still in process.

Low level — Analytics have been developed, testing is in process and integration is at a low level.

The ALERT System at FNS: An Early Data Analytics Project

The Supplemental Nutrition Assistance Program (SNAP), formerly known as the Food Stamp Program, has existed in various forms since the early 1960s.[4] To understand the issues that organizations typically weigh when developing data analytics processes and systems, we began by examining how FNS developed its Anti-Fraud Locator using the EBT Retailer Transactions (ALERT) system. While data analytics may be used for a variety of purposes, most agencies' first systems are generally geared toward detecting fraud or reducing erroneous payments — payments made in the wrong amount, to the wrong person or at the wrong time.

In the 1980s, FNS began testing of an electronic benefit transfer system (EBT) to replace the antiquated paper coupons for distributing benefits. EBT was developed to use the infrastructure already being developed by the commercial sector for processing debit and credit card transactions at various retail outlets. The EBT system was deployed on a state-by-state basis across the United States. Maryland was the first state to fully implement an EBT card system, going operational in 1993. Deployment continued across the country with full deployment by 2004, when California came online. EBT was designed in part to eliminate the stigma attached to receiving SNAP benefits and to reduce administrative costs associated with printing, storing, distributing and accounting for the paper coupons. It was estimated that SNAP was handling 1 billion pieces of paper annually; consequently, backroom operations were costly to operate and provided very little information on where benefits were redeemed.

When states began issuing EBT cards, recipients no longer had to stand in line each month to obtain their benefits. Like a debit card, an EBT card is secured by a PIN; benefits are automatically available via the EBT card each

month. EBT is a savings for government because coupons — which were like a second form of currency that was used once and then destroyed — did not have to be delivered to distribution sites via armed guard and secured in vaults. Merchants like the system because they do not have to count and deposit the coupons.

Along with the development of the EBT system, FNS modernized the system for enrolling merchants in SNAP. This new enrollment system improved the agency's ability to screen prospective SNAP merchants, facilitating the detection of merchants who had prior difficulties with the program. The new enrollment system also provided better demographic information on each merchant, the merchant's location, number of outlets, volume of sales, etc. This merchant enrollment system, known as the Strategic Tracking and Redemption System (STARS), combined with the EBT system to provide FNS with transaction-specific data including, but not limited to, a merchant's identification number, transaction date and time, and amount of the transaction. FNS obtains transaction information daily from the three contractors that operate the states' EBT systems.

During development of the EBT system, FNS personnel at the regional and central offices recognized that this new system would provide information that could be valuable in monitoring the program. FNS monitors a number of risks associated with SNAP benefits. The most basic challenge for the program is ensuring that recipients are eligible. This task is handled by the states that enroll recipients and monitor their continuing eligibility. However, other risks involve eligible recipients using benefits to purchase illegal items such as tobacco, and benefit trafficking whereby paper coupons are sold by benefit recipients to individuals or merchants for cash and not exchanged for eligible food items.

> FNS has been successful with implementing its ALERT system in large part because it had the support of mid-level managers who understood the power inherent in transaction-level data.

Before the advent of EBT, it was no secret that trafficking was a problem. Investigators and the press reported that it was not unusual for coupons to be sold by some recipients for 50 cents on the dollar. Trafficking was more difficult to detect in the paper coupon system because information was not available for investigators or auditors to analyze. With the advent of EBT benefits, the methods for trafficking changed because benefits were provided electronically. What FNS needed was a system to analyze the data and provide this information for investigation by FNS staff in various regional offices. This resulted in a project that would develop the ALERT system.

ALERT took more than two years to develop and the process was evolutionary. FNS used a continuous feedback loop in developing the system, whereby models were updated based on the results of investigations. This helped them improve the system overall. For the first time, people were able to see where the money was spent, when it was spent, and how much was spent on a comprehensive basis. In short, ALERT facilitated "market research" by FNS personnel.

The system was developed through a team of in-house business experts from FNS regional offices, investigative staff and outside contractors. FNS staff from regional offices and the central office combined to provide the business talent for the ALERT development

team. FNS staff had knowledge of the types of problems they were seeking to identify, what patterns and trends in the data could indicate that a vendor might be trafficking in benefits, and which business rules, if violated, would be problematic. This knowledge was translated by the FNS staff and contractors into automated systems that could screen and analyze the EBT data. The system looked for various patterns, computed results for all EBT merchants, and provided a list of exceptions for follow-up. Initially, the contractor provided a wide range of services for FNS, from developing the computer programs to performing the analyses to hosting ALERT on its own computer system. As it has evolved, FNS now operates ALERT on its own systems. The contractor continues to provide system maintenance and enhancements.

Our discussion with FNS personnel revealed several factors that appear critical to the initial and sustained success of ALERT. FNS has been successful with implementing its ALERT system in large part because it had the support of mid-level managers who understood the power inherent in transaction-level data. With the advent of EBT, the agency had access to a large amount of data, and because management understood its value, they were supportive of building a system to analyze the data and reengineering processes to better work with this new information. FNS

also obtained funding for the system because it aligned closely with the development of EBT.

FNS also had a clear goal in mind for the system. It wanted to improve its approach to monitoring possible fraud and abuse by merchants. The ALERT system was intended to transform the primary investigative focus from the use of food stamps for buying ineligible items into one focused on wholesale, larger frauds involving benefit trafficking. According to an FNS official, there was no easy way of identifying trafficking targets under the earlier paper system. The ALERT system dramatically changed the methods for the selection of investigative targets.

Expert outside assistance was also sought to develop the routines for the new system. This was critical, as in-house resources were not available to provide the technical assistance needed to develop the system and determine which techniques should be employed. When it came to designing the system, creating ways to analyze the data, developing the software, and building and operating the ALERT System, FNS reached out to private-sector contractors. FNS recognized that it would need outside skills to develop the software routines and functionality that it would need to analyze the EBT transactions. FNS staff also benchmarked their analytical approach against systems developed by credit and debit card processors for identifying fraud and abuse. This process led them to carefully consider which factors used by these

firms were common to their analysis. FNS officials identified differences in their program's operations versus credit and debit card operations, which allowed them to further strengthen the system. FNS also constructed a system that could rely solely on information available in its own systems. This avoided issues of having to negotiate with other agencies for information, and allowed development to proceed unencumbered by the wait for data.

Based on our work with FNS and other organizations, below are several key questions that we believe agencies should consider before implementing data analytics. These questions are explained in detail below.

Question: What is the goal of the data analytics project, and is leadership committed to the project and its goals?

U.S. Department of Agriculture, Food and Nutrition Service

Establishing goals and securing leadership for any new initiative is often the most critical aspect of a project's success. This was certainly true of the data analytics projects researched for this report. Clear goals establish realistic bounds and reasonable expectations for projects. Leaders establish the vision for the project and inspire employees' support.

In our interviews with FNS officials, it was clear that the leadership not only

stood behind the project but that they had clearly articulated the goals and expected benefits of ALERT before starting work on the system. FNS sought to improve the detection of SNAP benefit trafficking by identifying better targets for investigation. The effect upon SNAP would be to reduce the incidence of fraud in the program.

FNS leadership further demonstrated commitment to the ALERT system by securing funding and outside expertise to assist in the system's development. It also supported the redeployment of some staff from regional offices to the ALERT project.

Our research at the other federal agencies confirmed the critical importance of supportive leadership and the clear articulation of project goals and benefits. Not surprisingly, all of the officials we interviewed were able to articulate the goals of their data analytics projects and provide examples of demonstrable leadership support.

U.S. Postal Service, Office of the Inspector General

The second agency we interviewed was the United States Postal Service Office of the Inspector General (USPS-OIG). USPS-OIG started a project to develop the Risk Assessment Data Repository (RADR) to aid in the selection of areas for investigation and audit. RADR was developed to assist in four distinct areas: contract fraud, healthcare fraud, mail theft and financial fraud. According to USPS-OIG official we interviewed, the goal of the data analytic work is to identify possible fraudulent activities in a number of critical areas for investigative staff. This should result in more productive investigations and better use of staff.

Leadership for the work at USPS-OIG started at the top of the organization with the inspector general, David C. Williams. According to Bryan Jones, USPS-OIG Director of Countermeasures and Performance Evaluation at the Data

> Clear goals establish realistic bounds and reasonable expectations for projects. Leaders establish the vision for the project and inspire employees' support.

Mining Group, the inspector general led the efforts at their office and remains actively involved in the system's development. He served as a champion for the project — monitoring the work at a high level and providing a clear vision for the system. He also reinforced the need for RADR throughout his organization, and assisted in overcoming problems and promoting collaboration with other USPS-OIG departments.

U.S. Department of Education, Office of Inspector General

The Department of Education's Office of the Inspector General (ED-OIG) also had a clear goal and champion for its project at the top of the organization. The project's goal is to provide support for the audit and investigative operations by using predictive models to identify audit and investigative targets. One part of the system is used to identify higher education student aid fraud. It analyzes a wide variety of information and ranks transactions for investigative follow-up. The second part of the system analyzes information related to federal funds given to local education agencies. The analysis results in rankings that are used to select local agencies for audit. ED-OIG officials said champions of the data analytics project include Inspector General Kathleen S. Tighe and Assistant Inspector General for Information Technology Audits and Computer Crime Investigations Charles Coe. Mr. Coe, recognizing that analytical projects require a multiyear effort, secured support from the top levels of the OIG, to ensure that these projects continue to move forward.

Centers for Medicare and Medicaid Services, Center for Program Integrity

At the Centers for Medicare & Medicaid Services (CMS), a clear goal for the project was set by the passage of the Small Business Jobs Act of 2010 (SBJA), and strong support for the work was provided by CMS leaders. The SBJA, signed into law in September 2010, contained a

---

FIGURE 5: EXCERPT— SMALL BUSINESS JOBS ACT OF 2010

(a) USE IN THE MEDICARE FEE-FOR-SERVICE PROGRAM.—

The Secretary shall use predictive modeling and other analytics technologies (In this section referred to as ''predictive analytics technologies'') To identify improper claims for reimbursement and to Prevent the payment of such claims under the Medicare fee-for service Program.

(b) PREDICTIVE ANALYTICS TECHNOLOGIES REQUIREMENTS.—

The Predictive analytics technologies used by the Secretary shall—

(1) capture Medicare provider and Medicare beneficiary activities across the Medicare fee-for-service program to provide a comprehensive view across all providers, beneficiaries, and geographies within such program in order to—

    (A) identify and analyze Medicare provider networks, provider billing patterns, and beneficiary utilization patterns; and

    (B) identify and detect any such patterns and networks that represent a high risk of fraudulent activity;

(2) be integrated into the existing Medicare fee-for-service program claims flow with minimal effort and maximum efficiency;

(3) be able to—

    (A) analyze large data sets for unusual or suspicious patterns or anomalies or contain other factors that are linked to the occurrence of waste, fraud, or abuse;

    (B) undertake such analysis before payment is made; and

    (C) prioritize such identified transactions for additional review before payment is made in terms of the likelihood of potential waste, fraud, and abuse to more efficiently utilize investigative resources;

(4) capture outcome information on adjudicated claims for reimbursement to allow for refinement and enhancement of the predictive analytics technologies on the basis of such outcome information, including post-payment information about the eventual status of a claim; and

(5) prevent the payment of claims for reimbursement that have been identified as potentially wasteful, fraudulent, or abusive until such time as the claims have been verified as valid.

---

specific provision, Part II, Section 4241, titled "Use of Predictive Modeling and Other Analytics Technologies to Identify and Prevent Waste, Fraud and Abuse in the Medicare Fee-for-service Program." This provision provided CMS with a clear direction to develop a predictive analytic system that would prevent the payment of potentially wasteful, fraudulent or abusive claims. The law clarified that the

system was to implement real-time, pre-payment claims analysis that would allow the agency to identify fraudulent claims prior to payment. The SBJA formally directed CMS's transition from a system that pays Medicare claims and then seeks to recover improper payments, often referred to as "pay and chase," to a system that enables CMS to prevent payment of bad claims by screening and

analyzing all Medicare Fee-for-service claims prior to releasing payment. The law further mandated that CMS implement this system to screen Medicare claims for 10 states by July 1, 2011.

CPI officials said HHS and CMS leadership provided strong support for the project. Dr. Peter Budetti, Deputy Administrator and Director for Program Integrity, ensured that resources were dedicated to this critical objective. He was personally and visibly involved while providing strategic direction throughout the implementation. He encouraged all components within CPI to work together and participated in business process re-engineering sessions to develop appropriate responses for working with the results of this new transforming technology. His lead-by-example approach supported CMS's strategy to realign the internal organizational structure the previous year, consolidating the Medicare and Medicaid program integrity groups under a unified CPI. In a series of Congressional hearings in the spring of 2011, Dr. Budetti continually stressed the importance of these new data analytics programs, stating that "[g]iven the changing landscape of health-care fraud, any successful technology will need to be nimble and flexible, identifying and adjusting to new schemes as they appear." Dr. Budetti also testified shortly after the implementation of the new data analytics tools on July 12, 2011, stating "[t]he new authorities given to us by Congress and the experience of private sector industries in combating fraud have greatly enhanced our capacity to carry out this task."

The implementation of new data analytics technology has received significant support from the highest levels of the Department of Health and Human Services. HHS Secretary Kathleen Sebelius announced the launch of the new technology at a joint fraud summit with the Department of Justice. The Secretary announced that "[a]s long

as we continue to aggressively put these tools to work preventing and prosecuting fraud, we can continue to protect and strengthen Medicare's future." Then-CMS Administrator Donald Berwick, MD, also noted that "[t]his new technology will help us better identify and prevent fraud and abuse before it happens and helps to ensure the solvency of the Medicare Trust Fund."

Recovery Accountability and Transparency Board

The American Recovery and Reinvestment Act of 2009 (ARRA) created the Recovery Accountability and Transparency Board (RATB) with two goals in mind: to provide transparency of recovery-related funds, and to detect and prevent fraud, waste and mismanagement. The board originally included 13 members, with a chairperson appointed by the president and 12 inspectors general. ARRA contained ambitious transparency and reporting requirements. Reporting was done online and published by the RATB on its publicly accessible website, Recovery. gov. The RATB, under the leadership of the chairperson, formed the Recovery Operations Center (ROC) and tasked it with quickly developing an advanced data analytics system to aid in the prevention and detection of ARRA fund fraud, waste and mismanagement. The project required the RATB to develop in a very short time frame a data analytics system that could screen the recipients of ARRA funds for past problems as well as identify new, potential problems with recipients of ARRA funds. ROC officials said the chairperson's leadership was invaluable to the project's success. The chairperson remained actively involved in the ROC's development and laid out the vision for the ROC's work.

Naval Sea Systems Command

The U.S. Navy's Naval Sea Systems Command (NAVSEA) is beginning an ambitious project to develop a data analytics system to detect fraud and abuse.

High-level leaders in NAVSEA have made the data analytics project a high priority and clearly communicated this to the entire command. The high priority of the system has been demonstrated through the creation of the Office of Fraud Deterrence and Detection (OFDD) within NAVSEA. One of OFDD's principal tasks is to develop a data analytics system that will help NAVSEA identify potential procurement fraud for further investigation.

Our interview revealed that leadership support has been a significant factor in the project's progress. NAVSEA's leadership has made the implementation of the data analytics program one of the top three priorities in its plans for FY 2012 and 2013. It received this high ranking because OFDD's data analytics project is in alignment with NAVSEA's CORE goals of safeguarding taxpayer funds and maintaining good stewardship of its resources. NAVSEA's leadership has also supported the project by implementing a mandatory contracts review process that has reshaped NAVSEA's procurement practices for the better.

This reshaping allows OFDD to obtain critical procurement insight and access to the data needed to perform data analytics. Because OFDD, Naval Criminal Investigative Services and Navy Audit all investigate and audit NAVSEA operations, they have developed a cooperative agreement to avoid conflicts in their respective roles and the duplication of effort. The agreement has fostered a strong, cooperative working relationship among the three units.

Survey Results on Leadership

All of the federal officials we interviewed emphasized the value of setting goals and the need for leadership to set a positive tone at the top. They noted that leadership, whether coming from a vision at the top, a strategic planning process or mid-management, was essential. Discussions with firms that

have worked on data analytics projects in both the private and the public sectors confirmed the need for strong leadership that will champion the project and articulate its vision and goals. However, less than one-fourth of our survey respondents indicated that leadership support for using data analytics was high. While it appears most government leaders have a general understanding of data analytics, they are not yet champions of the process.

Over two-thirds of our respondents rated leadership's support for using data analytics at the medium level, meaning that they exhibit a general understanding of what analytics mean and how they can assist an agency. It would seem that federal leaders must continue to be educated on the benefits, the uses of, and the actions needed to implement data analytics work in their agency.

HOW WOULD YOU RATE YOUR AGENCY IN THE AREA OF LEADERSHIP SUPPORT FOR USING DATA ANALYTICS?

High level — Champions of analytic processes or systems, with deep analytical experience and background.	**23%**
Medium level — General understanding of what analytics means and how they can assist the agency.	**68%**
Low level — No clear understanding of what analytics mean and how they can be used.	**9%**

> While it appears most government leaders have a general understanding of data analytics, they are not yet champions of the process.

Question: What data will be needed for the data analytics work, and is this data available? In addition, what challenges must be overcome before the data can be effectively analyzed?

The development of a data analytics process or system begins with a clear vision of the desired goals, and progresses to an assessment of what data is available for analysis and whether this data will be sufficient to meet the agency's goals. If it is determined that sufficient data is not available, agencies must then determine where they will generate or obtain it (another agency or a commercial vendor might have the needed information). In addition, data must be cleansed to ensure it is useful. (See Figure 2 on page 11.)

U.S. Department of Agriculture, Food and Nutrition Service

The ultimate design of ALERT was influenced by the availability of data already, or soon to be, in the possession of FNS. FNS was able to balance the system's goals with the information readily available. Data collected from the new EBT system and the new vendor enrollment system, STARS, provided the data for analysis in ALERT. This allowed FNS to proceed in development with the certainty that sufficient data would be available. For example, hypothetically FNS might have sought information from the Internal Revenue Service that could assist in its analysis; yet access to this information is highly restricted and may have been impossible to obtain, which could have delayed the project.

FNS is not alone in data self-sufficiency. Our interviews with officials in other agencies disclosed that almost all of them were able to develop their systems solely using internal data. Few had to obtain information from another government agency or a commercial vendor. The chief exception to data self-sufficiency was the need to obtain corporate information, like the nine-digit Dun & Bradstreet D-U-N-S number that uniquely identifies an individual business. Data self-sufficiency saves time by allowing development work to move ahead without having to negotiate with other government agencies, seek legislative authorization to access government information or purchase the information from a commercial vendor.

U.S. Department of Education, Office of Inspector General

The ED-OIG collects a great deal of information and is an excellent example of an organization that can primarily rely on data from within its agency. The ED-OIG collects information from eight different Department of Education data sources including systems for processing student aid applications and student loan information, as well as payment information from the accounting systems. The ED-OIG has an experienced analytic team and understands the value of information available internally. This awareness allows them to focus on the

mechanics of collecting and preparing the data for analysis.

The complexity of ED-OIG's data sources are identified in a chart provided by Edward Slevin, director of the Computer Assisted Assessment Techniques Division at the ED-OIG. Combining information from a large number of data sources into a single database for analysis can present some unique challenges in the data cleanup phase of a project. Some of the information in ED-OIG's system is

non-numeric, such as street addresses. If one part of the system's analysis is to match addresses across records, this can be problematic. For example, an address could be identified on records ranging from such representations as 1st Street, First Street or Furst Street in the database. The ED-OIG solved this problem by employing a solution from the U.S. Postal Service to clean up and standardize addresses. Another useful aspect that was built into this system is investigator email alerts. For example,

if an investigator in San Francisco conducted a series of analytical queries with regard to a particular lead, and months later another investigator — say in Miami — conducted a series of analytical queries surrounding the same set of data, the ED-OIG's system would generate an email alert to the original investigator advising of this mutual area of interest. This has proven to be extremely valuable in terms of increasing the effective use of limited investigation resources.

FIGURE 6: DATA SOURCES FOR ED-OIG'S SYSTEM
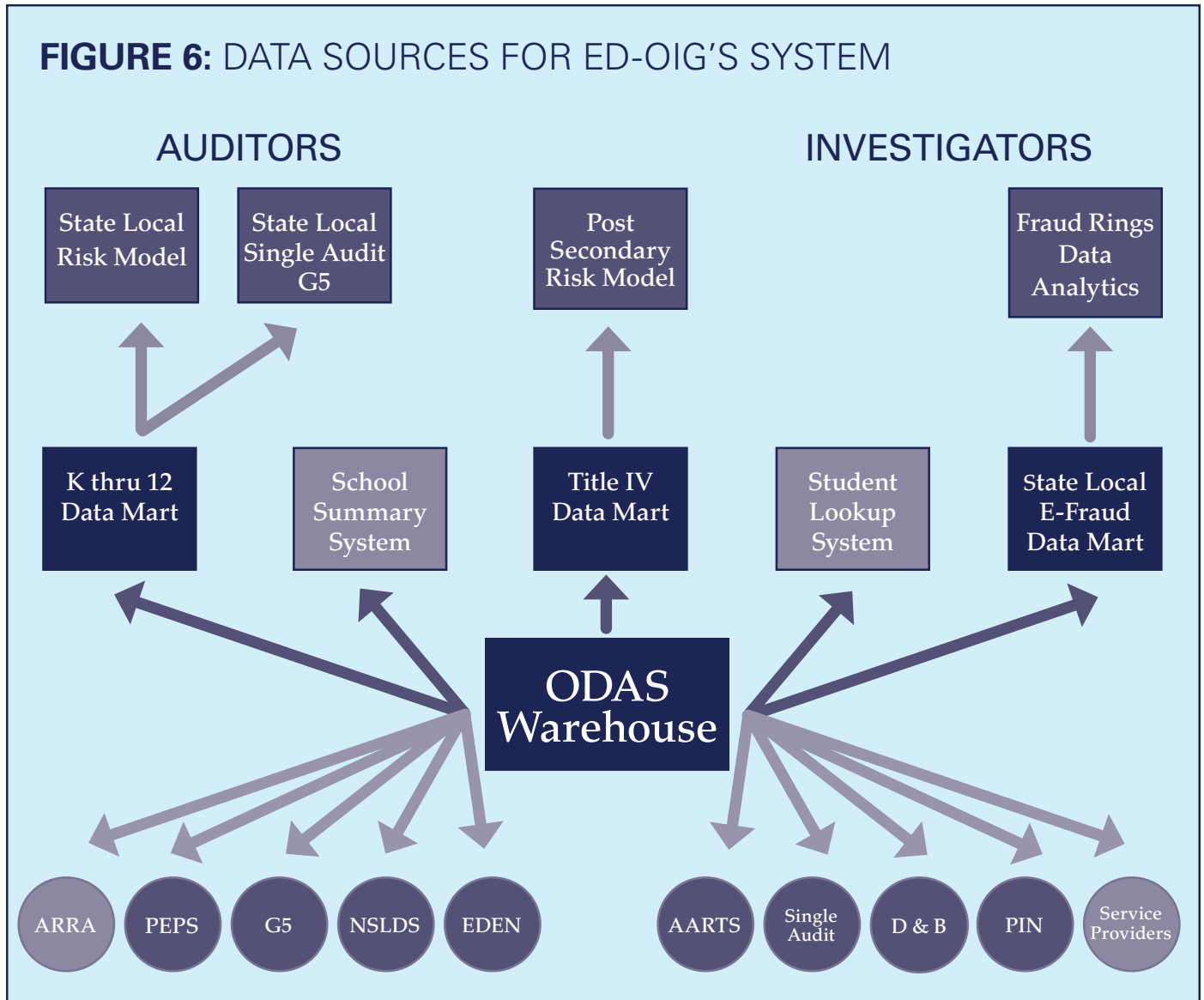
U.S. Postal Service, Office of Inspector General

USPS-OIG began developing the RADR system in 2009 with a similar approach. The RADR system is designed to examine transactions related to health care, contract, mail and financial transactions with the aim of identifying those that have a high probability of fraud or abuse. USPS-OIG developed the RADR system to analyze information already available to USPS, including information that was made available by the U.S. Department of Labor.

Defense Logistics Agency

In a similar manner, the Defense Logistics Agency (DLA) Office of Operations Research and Resource Analysis (DORRA) established a data analytics system to assist DLA leadership in managing operations using information that was available from its existing systems such as the Enterprise Resource System. DLA provides a full spectrum of logistics, acquisition and technical services to the military, providing nearly all of the consumable items used by military forces. This includes such things as food, fuel, uniforms, medical supplies, construction equipment and spare parts. DLA established a data analytics system to aid in managing its operations. The system provides performance metrics on a routine basis to agency management and leadership. The number of metrics and the detail provided to management or leadership varies depending upon their responsibility. DORRA officials said some managers or leaders receive information on as many as 50 metrics to use in managing operations.

Centers for Medicare and Medicaid Services, Center for Program Integrity

CPI's systems to review Medicare fee-for-service payments and provider enrollment applications also had sufficient data for the data analytics project. The Fraud Prevention System that analyzes fee-for-service payment claims, deployed nationwide on June 30, 2011, uses information from the Medicare payment processing system and other CMS systems for performing the analysis. CMS officials said obtaining data for analysis was not a problem, but that handling the volume of the data presented challenges because they receive about 4.5 million Medicare fee-for-service claims each day.

Defense Finance and Accounting Service

Similarly, the Defense Finance and Accounting Service (DFAS) had information readily available for data analysis but lacked the means to draw the information together before it implemented the Business Activity Monitoring (BAM) system in 2008. This system was designed to focus on reducing improper payments through data matching and analysis. BAM is a commercial off-the-shelf (COTS) system that interacts with existing payment systems to integrate information, analyze it, and create exceptions for follow-up and correction before payments are made. BAM also allows DFAS personnel to modify and customize the various analysis routines.

Question: Does our staff have the knowledge and skill sets needed to design, develop and implement our analytics processes and systems?

It is critical for agencies to assess their in-house talent. Our interviews revealed that a variety of talent is needed to develop, design and implement a data analytics system. Agencies should anticipate that they will need both those with business expertise and those with analytical expertise.

Business expertise was generally defined as expertise in the programs and areas upon which the system will focus. These individuals need to have an expert knowledge of the business (program) being analyzed, the rules that govern the program, past problems noted in the program, any risks to the program and knowledge of the data available for the project.

Analytical expertise was generally defined as expertise in the field of data analysis. This included individuals with expert knowledge in fields such as statistics, data analysis and data modeling. This also included expertise in using software or developing computer code to analyze data.

Needed expertise loosely breaks down as follows:

Business expertise means experience, knowledge or training in:

- The agency and the program being analyzed
- Regulations governing the program
- Past problems and risks associated with the program
- The type of data being collected by the agency or is available from other sources
- The quality of the data being collected

Analytical expertise means experience, knowledge or training in:

- Statistics, mathematics and data analysis
- Software design
- Analyzing data and developing predictive models
- Developing computer code

In nearly all cases, the federal agencies interviewed for this research provided the business expertise for the project and hired a contractor to provide most, if not all, of the analytical expertise. Generally, expert staff from units within the agency were assigned to work on the projects with contractor personnel. The contractor provided staff that were trained and had experience in statistics, data mining, data analytical techniques and the various software packages that are used to analyze and

display data. The result is a collaborative approach with each group bringing in the talent that is needed for the work.

Our survey of federal officials also confirmed that contractors were used to assist in the development of data analytics, with 50 percent of the respondents stating that contractors assisted them in the development of data analytics.

U.S. Department of Agriculture, Food and Nutrition Service

For example, at FNS, expert staff from the regional offices and the headquarters were assigned to the ALERT project to work with a vendor hired to develop the system. FNS staff provided insights into what types of transactions or patterns of transactions could be indicative of trafficking in SNAP benefits. The contractor staff then developed computer routines that searched for more than five different patterns in the EBT data. The resulting data was accumulated by merchant and ranked using a risk weighting system

developed by FNS staff. The results were provided to FNS field staff for investigation and validation, with the actual investigation results used to further refine the ALERT system.

U.S. Department of Education, Office of Inspector General

Similar methods were used in the ED-OIG. Business talent was provided by the ED-OIG staff with experience in investigations and audit, which provided insights into what types of issues and problems they were seeking to detect. They also assisted in reviewing and refining the data analytics system to provide better results.
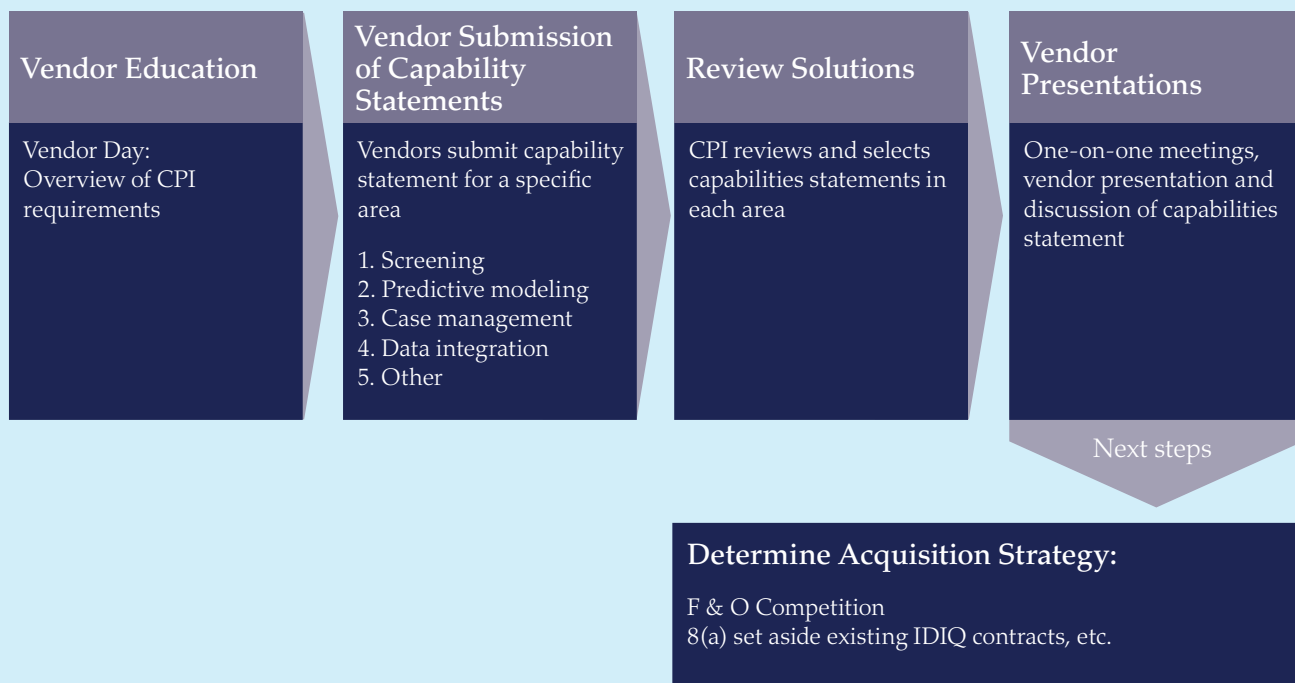
Centers for Medicare and Medicaid Services, Center for Program Integrity

We found that the CMS Center for Program Integrity used a somewhat different approach to procuring the talent it needed to complete its project. CMS created the CPI to serve as the

focal point for all programs that deal with the prevention and detection of fraud and abuse in the Medicare and Medicaid programs. CMS had developed a strategic plan to reorganize and begin development on data analytics, work that was switched into high gear in September 2010, when President Obama signed into law the Small Business Jobs Act. The law required that CMS issue a request for proposals and contract with private companies to conduct predictive modeling and other analytics technologies to identify and prevent payment of improper claims submitted under Medicare Parts A and B. The deployment was to encompass the claims from at least 10 states and be online by July 1, 2011, a 10-month time frame.

CPI began development of two data analytics systems, one to review Medicare fee-for-service claims called the Fraud Prevention System (FPS) and another to automate the process for screening Medicare provider enrollment

FIGURE 7: CPI ROADMAP

Vendor Education	Vendor Submission of Capability Statements	Review Solutions	Vendor Presentations
Vendor Day: Overview of CPI requirements	Vendors submit capability statement for a specific area 1. Screening 2. Predictive modeling 3. Case management 4. Data integration 5. Other	CPI reviews and selects capabilities statements in each area	One-on-one meetings, vendor presentation and discussion of capabilities statement

Next steps

Determine Acquisition Strategy:

F & O Competition
8(a) set aside existing IDIQ contracts, etc.

applications. The fee-for-service system for Medicare claims would represent a major change for CMS processing of Medicare claims. In 2010, many of the advanced controls over Medicare claims were oriented toward a "pay and chase" system. The law mandated that CMS would implement predictive analytics and place more emphasis on the prevention of problems while still providing information for the pay-and-chase system controls. The CPI data analytics system was developed in a short period of time by CPI using an approach that combined outside business and data analytics talent with in-house talent in the same areas.

With a tight time frame, CPI was forced to think creatively in acquiring talent and skills for the work. Through a two-phase "Industry Day" initiative, CPI engaged contractors to assist in the development of the data analytics systems. Figure 7 from the CPI Industry Day presentation illustrates the approach to securing contractors to assist in system development.

As shown in Figure 7, CPI provided information to vendors on CPI's requirements and asked interested vendors to submit a statement through a Request for Information process as to their firm's ability to meet CPI's needs in any one of five areas. CPI staff reviewed the proposed solutions and subsequently used these capabilities to inform the requirements. CPI acquired both business talent and analytical talent with the contracts. CPI officials said that the contractors brought in not only experts in analyzing data and developing predictive models but also experts in the field of health care and health-care fraud to work on the project. CPI also included CMS expert staff in the project, including experts in the areas of health care, Medicare and Medicare fraud.

CMS deployed the Medicare fee-for-service system nationwide on June 30, 2011, three years ahead of schedule. This exceeded the requirements of

the Small Business Jobs Act, which mandated deployment in only 10 states by July 1, 2011. CPI officials emphasized the importance of the Industry Day approach to their success. They felt that Industry Day allowed them to secure well-qualified contractors that understood their priorities and vision for the future, which in turn facilitated contractor interaction in the "Command Center" with inside experts to produce a system quickly and effectively.

Survey Results on Knowledge and Skill Sets

Our survey of federal officials confirmed that assistance is needed in data analytics, as the skills of the agency's current work force were not rated at a high level in most instances. In response to the question, "How would you rate the data analytics skills of the organization?", only 18 percent of the respondents rated the analytical skills of their organization at a high level, with most rating the skills at a medium or low level in their organization:

High level — Highly skilled work force, data analytics techniques are understood across the agency or department.	**18%**
Medium level — Pockets of skilled analytical professionals, focused on individual business units or programs	**50%**
Low level — Some skilled analytical professionals across the organization.	**32%**

Based on our survey results, most federal agencies will need to acquire some amount of analytical expertise from contractors to implement data analytics solutions in their operations.

Question: Has the agency considered the need for cultural and organizational skills?

Firms that work in the area of data analytics and analysis also discussed the need for organizations to consider cultural and organizational issues when integrating data analytics into an organization. One firm in the area of data analytics discusses the need for achieving "cultural alignment" to drive high performance and sustained results.[5] "Cultural alignment" for this firm constitutes a number of considerations. This firm speaks about creating a culture in the organization that values analytics and respects the data, and combines this with a pervasive curiosity for information. If this is not done, the pockets of existing analytical talent quickly grow disillusioned and, because they are not integrated into the business as a whole, fail to deliver much strategic value. A culture of analytics can be built by focusing on five elements:

- Respect for data: Organizations with analytical cultures demonstrate a profound respect for data and fact-based decision-making.
- Pragmatic decision-making: On the other hand, organizations with analytical cultures know the limits of data and do not get stuck in "analysis paralysis."
- Drive to optimize: An analytical organization is fundamentally curious — about what others are doing in the market, about performance patterns and root causes, and about new and better ways to do things.
- Collaboration and transparency: An analytical culture is marked by collaboration and information sharing across organizational boundaries.
- Rewards for analytics: Individuals are recognized and rewarded for their analytical capability, including not only the quality of analyses and

insights, but also the breakthrough business results achieved by putting them into action.

Other experts make similar points. One not only spoke of the need for a champion, but also stressed the importance of making data analytics an integral part of the organization's work.

U.S. Department of Agriculture, Food and Nutrition Service

Evidence of a cultural realignment within FNS is the ALERT system's long-term acceptance and integration into FNS operations. An example of the cultural alignment is in the procedural change for removal of merchants deemed to be trafficking in SNAP benefits. All retailers who participate in SNAP must be licensed by FNS and may be banned from the program for offenses including trafficking or failure to comply with program rules. At the start of ALERT, retailers could not be disbarred from the program, no matter how compelling the analytical information, without an actual field investigation. Today, retailers can be barred from the SNAP program based solely upon information provided by the ALERT system. The effectiveness of ALERT has prompted confidence in the system and has resulted in a cultural realignment in support of data analytics within FNS.

Centers for Medicare and Medicaid Services, Center for Program Integrity

CPI officials said they are addressing cultural alignment in their work by using the "Command Center" to integrate CPI and CMS staff into the work. The Command Center was launched after the FPS system went live in June 2011. The Command Center is used to share technical details and expectations with investigators and analysts, work collaboratively on the results from the system, refine and improve the analytic models and associated business processes, and create consistent investigative approaches and appropriate actions.

CPI has established a system to rotate key staff from other CMS units through the Command Center to work on the new systems. Officials say they plan to maintain this approach to develop new systems and improve and expand the current system.

U.S. Postal Service, Office of the Inspector General

USPS-OIG addressed the issue of cultural alignment by placing the unit developing the data analytics system into the investigative unit that would use the results of the analysis. During the initial phase of RADR development, the system was housed in a support unit outside of the investigative unit that would use the system's output. During the course of system development, the RADR system and its development team were moved to the investigative unit. USPS-OIG officials felt that this instilled ownership by the investigative staff and better aligned the investigative unit's work with the new system. As the system has matured and shown results, the RADR team has expanded the system's role to work with the audit side of USPS-OIG in addition to the investigative unit. To facilitate this change, the RADR team's place in the organization has been moved to a support unit that reports to the chief technology officer at USPS-OIG. It was felt that this change would facilitate RADR development as the development team will be supporting operations in two units rather than one.

Survey Results on Cultural and Organizational Skills

In our survey of federal officials we asked, "How would you rate the integration of data analytics processes and systems into the agency's operations?"

High level — Data analytics techniques are fully integrated into management, budgeting and planning functions.	**8%**
Medium level — Data analytics techniques are used inconsistently; there is some linkage to management, budgeting, and planning functions.	**46%**
Low level — Data analytics processes are conducted in silos, with little consistency or standardization.	**46%**

As shown, only 8 percent of respondents think that data analytics are fully integrated into key functions. This indicates that data analytics has not yet achieved the acceptance and integration that will be needed to achieve lasting organizational success.

Question: Does the organization have systems and software to implement a solution, or will they need to be acquired?

Starting data analytics processes and systems may present issues related to data storage and the selection of analytic tools. These issues were addressed in different ways by the organizations we interviewed. Data storage was commonly addressed either by establishing a data warehouse or by storing data in files on servers. In determining how to handle the information, agencies need to consider the volume of data, the diversity of the data and the kinds of analysis they plan to perform. Larger solutions may require extensive investments in data warehouses and software, whereas smaller solutions may not require much up-front investment. Two different approaches are highlighted below.

U.S. Department of Education, Office of Inspector General

The ED-OIG used a planned approach in developing its data analytics system that called for the development of a data warehouse and the acquisition of a

commercial software suite designed for data analytics work.[6] The ED-OIG determined that this approach best met its needs and would provide the most effective platform for future development. As presented in the report, the ED-OIG uses information from eight different systems in its various analysis components.

U.S. Postal Service, Office of the Inspector General and U.S. Department of Agriculture, Food and Nutrition Service

In developing the RADR system, USPS-OIG decided not to establish a data warehouse in the first phase of the development process. Instead, staff focused on collecting data and constructing models for the health-care fraud part of the RADR system. USPS-OIG felt that, given its level of resources, this approach would allow construction of data analysis models to begin sooner than if a data warehouse was created during the first phase. However, USPS-OIG recognized the need for a data warehouse and the efficiencies it could provide, and is currently constructing one for the RADR system. USPS-OIG also opted not to procure a software suite for its systems. Instead, a decision was made to construct models and interfaces using software products that were already licensed to the larger USPS enterprise. FNS also used this approach in the ALERT system, using existing software licensed to FNS to construct its routines. Both approaches were successful and have merit.

Summary of Organizations Reviewed and Stage of Deployment

The following table captures highlights of the eight organizations interviewed for this report. We have summarized each data analytics system's purpose, the stage of development and the extent to which it has been deployed within the organization.

FIGURE 8: SUMMARY OF SYSTEMS BY INTERVIEWEE

ORGANIZATIONS INTERVIEWED	PURPOSE OF SYSTEM	STAGE OF DEVELOPMENT	EXTENT OF DEPLOYMENT
Recovery Accountability and Transparency Board **Recovery Operations Center (ROC)**	ROC – Identifies potential fraudulent transactions by analyzing ARRA funds information using 22 different data sets and analytics	Development Complete	Deployed across all ARRA funding, pilot testing against remaining federal funds
U.S. Department of Agriculture, Food and Nutrition Service **Anti-Fraud Locator using EBT Retailer Transactions (ALERT)**	Identifies food stamp trafficking by analyzing EBT transactions using pattern identification and risk rankings	Development Complete	Deployed and fully integrated into FNS operations
U.S. Department of Defense, Defense Finance and Accounting Service (DFAS) **Business Activity Monitoring System (BAM)**	BAM – Identifies improper payments before they are made by matching data across systems, looking for patterns of payments indicative of fraud from past transactions, etc.	Development Complete	Deployed across nearly all DFAS payment systems. Results are integrated into DFAS operations
U.S. Department of Defense, Defense Logistics Agency, Office of Operations Research and Resource Analysis (DORRA) **Enterprise Business System (EBS)**	EBS – Provides the DLA leadership with information on performance metrics based upon information from the agency's systems	Development Complete	Deployed across DLA
U.S. Department of Defense, United States Navy, Naval Sea Systems Command, Office of Fraud Deterrence and Detection (NAVSEA-OFDD)	Planned system will be used to identify cases of potential fraud and abuse related to contract procurement	In Planning Stages	Not applicable
U.S. Department of Education, Office of Inspector General (ED-OIG) **E-Fraud Analytical Model (EFAM) & State and Local Education Agencies Risk Model (SLRM) Systems**	EFAM – Identifies fraud in higher education assistance programs by using data mining techniques SLRM – Analyzes information on federal funds provided to state and local education agencies to develop a risk model prototype	EFAM – System in Testing SLRM – Development Under Way	Deployment started.
U.S. Department of Health and Human Services, Centers for Medicare and Medicaid, Center for Program Integrity (CPI) **Fraud Prevention System (FPS)**	FPS – Analyzes Medicare payments, prior to payment, to identify potential fraud and abuse. FPS uses predictive modeling techniques.	Development Complete	Deployment across all Medicare fee-for-service payments; system integration has been ongoing for one year.
U.S. Postal Service, Office of Inspector General (USPS-OIG) **Risk Assessment Data Repository (RADR)**	RADR – Analyzes activity in four areas to identify potential fraud and abuse by using data mining and predictive analytics. Results referred to investigators.	Development Complete	Deployed and fully integrated into USPS-OIG operations

Part Four: How is Success of the Systems Measured?

No standard approach exists for measuring the success of data analytics solutions. Most agencies measure the outputs from their systems and the increase or decrease in the inputs that are required to achieve results. For example, a data analytics system implemented by an OIG investigative unit may measure the change in the number of fraud cases investigated and prosecuted before and after system implementation as a way of measuring success. It might combine this number with the cost of staff, or combine staff resources with the number of fraud cases identified in investigations, to compute a return on investment. CPI officials indicated that more than 510 new investigations have been started as a result of these results. In addition, 336 existing investigations are now being supported by real-time FPS data. FPS data has also led to more than 400 direct interviews with providers who may be participating in potentially fraudulent activity. The bottom line is that output measures are necessary and needed, but long-term results must also be measured.

Often an agency may measure the effectiveness of its system by looking at such things as the decline in the amount of waste, fraud and abuse uncovered by the system. While this may be an effective measure, it might also indicate that the system has lost effectiveness.

Fraud and abuse schemes change rapidly, and the methods that are used to detect them must change as well. Other factors might also be in play, such as declines in the quality of data or the effectiveness of the investigative staff. ED-OIG officials said the issues might be mitigated by a well-established and reliable line of communication between the respective system owner and the analytical team to ensure that all significant system enhancements are shared. Otherwise it is quite possible that critical aspects of previously successful analytical engines rely on now-"stale" segments of data.

One method to measure the effects is to periodically examine the program on a wide scale to see if a certain element is declining, remaining stable or increasing. This is done at USDA for the SNAP program. USDA implemented the ALERT system to detect trafficking in SNAP benefits. (Trafficking is defined as the illegal sale of SNAP benefits for cash or other non-allowable items.) The first assessment in 1993 determined that about $811 million in program benefits, or 3.8 percent of the benefits redeemed, were lost due to trafficking. Since then, FNS has completed four additional periodic studies and is currently completing a new study to estimate the rate at which benefits are lost due to trafficking. The most recently completed study for 2006 – 2008

concluded that the trafficking rate was 1.0 percent, or $330 million. Based on benefit levels of $75.6 billion in FY 2011, we estimate that this would translate into $1.1 billion of benefits that were not trafficked last year. This method does not measure system outputs but rather outcomes of FNS work on the food stamp program.

Malcolm Sparrow, PhD, Professor of the Practice of Public Management at Harvard's Kennedy School of Government, has studied and written extensively on health-care fraud. In his publications, Dr. Sparrow writes about the lack of solid measurements to back up the estimates on the extent of fraud in the health-care industry. He advocates that accurate information on the extent of fraud can only be achieved by a series of random, rigorous audits of health-care providers.[7] In this way, studies can estimate the extent of fraud in order to measure future progress.

Absent a system such as that advocated by Dr. Sparrow, it will be challenging for federal agencies to measure the effectiveness of data analytics systems that are designed to detect and prevent fraud. Measuring direct outcomes will only provide some of the answers regarding data analytics system effectiveness. The authors of an AGA research report, "Using Performance Information to Drive Performance Improvement," identified

a similar concern when reviewing the use of performance measures in federal agencies. The report found that most measures used by agencies relate to program activities and outputs, and it concluded that measures used to manage a program day-to-day can cause difficulty when no outcome measures linked to these programs exist. People want to know program effectiveness as well as how much money has been saved. People want to know if fraud detection programs are also reducing fraud.

Part Five: The Case for Government-wide Leadership

Our survey asked federal officials, "If you have not begun using data analytics in your operations, what are the principal reasons for this?" A frequent response was that they were awaiting government-wide solutions in some areas. Our research noted that data analytics solutions have been developed, which many agencies could use for their operations. Deployment of such systems would save time and reduce costs while allowing development to proceed on additional systems. Two prime examples of this were found in the work by USPS-OIG and the RTAB Recovery Operations Center.

An official at ED-OIG with extensive experience in data analytics systems development said he believed the future for many aspects of data analytics would be in the cloud environment, managed by experienced individuals who would be able to meet the needs of organizations in a cost-effective manner. Officials at USPS-OIG also expressed similar sentiments. One individual said that many of the inspector general operations throughout the federal government are not very large, and developing data analytics systems can be beyond their budgets. This official advocated sharing existing systems as a means to reduce costs and provide immediate results.

Possibilities for Government-wide Solutions

Two solutions follow that could be quickly deployed in organizations across government with concerted action by federal officials.

U.S. Postal Service, Office of Inspector General

RADR is used by USPS-OIG to identify instances of potential fraud for further investigation, resulting in better utilization, a greater return on investment and the reduction of waste, fraud and abuse of USPS resources. The

FIGURE 9: HIGH-LEVEL MODEL ARCHITECTURE
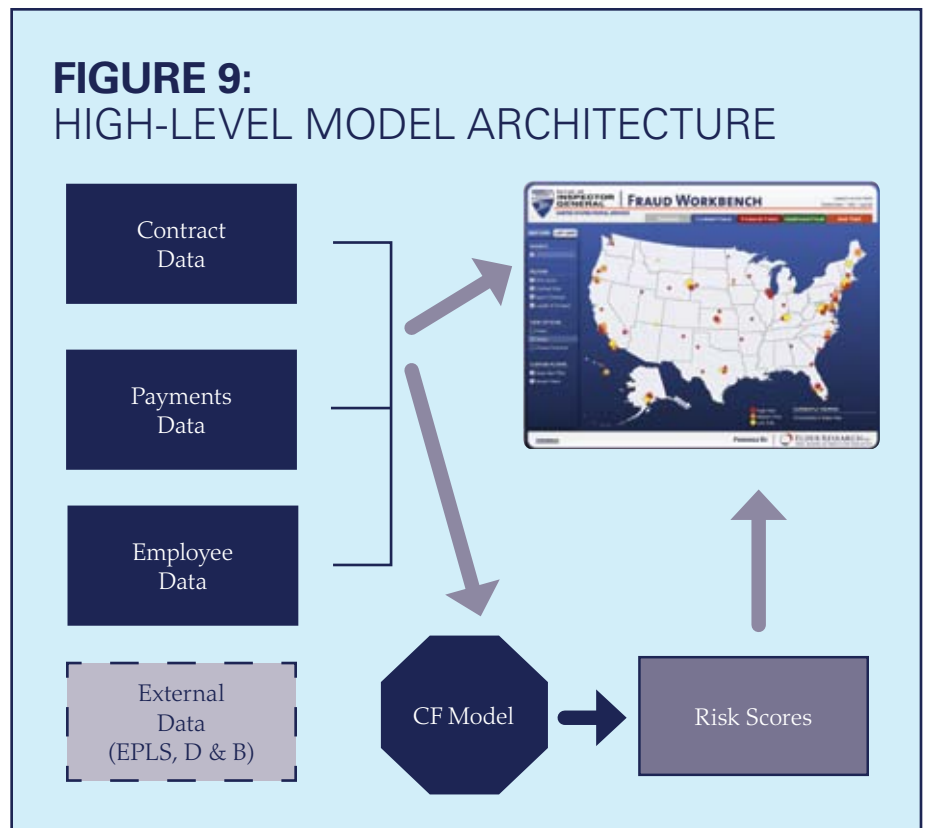
USPS-OIG RADR system has four components, two of which address issues unique to USPS operations. However, the remaining two RADR components are designed to address issues that are common across many federal agencies, namely worker compensation fraud and contract fraud. It would seem that these components of RADR could be deployed across other governmental agencies in a cost-effective manner.

The RADR system for contract fraud works in the general manner illustrated in Figure 9. Data from a number of in-house USPS information systems are collected and passed through a contract-fraud model, which merges the data with information from commercially available sources to rank the contracts in terms of risk for fraud. RADR uses more than 30 risk ranking criteria to score and rank each contract. Users can access the system online, and data is displayed through a visual interface.

USPS-OIG officials believe that the visual display makes information more accessible and improves how effectively the results of the system's analysis can be communicated. The RADR system displays the results of its work in a visual map of the Unites States, using red dots to indicate "hot spots" and yellow dots to indicate "warm spots" for contract fraud. Red spots indicate a high number of contracts that have tripped many of the system's triggers and are thus more likely to indicate fraud. The size of the dot also indicates the dollar value of the risk in the area. Users can further refine the results by filtering information through the visual display. The user can also click on a spot to access the details of the RADR analysis and examine the contract details associated with the risk rating.

The RADR system uses a similar approach to reviewing worker compensation payments. Investigators are able to display a map of potential worker compensation fraud hot spots and can rank and review all associated cases to determine if further investigation is warranted. The risk rating system will consider such factors as the length of payments, the frequency of treatment and the severity of the injury in deciding which payments to investigate.

USPS-OIG is using this rating system to change its organization from a reactive to a proactive organization. For contract fraud, the RADR system has started to transform the work of USPS-OIG investigators in a number of ways. USPS-OIG is now analyzing every contract, instead of taking a random sample or waiting for a lead to initiate an investigation. USPS-OIG is also moving its investigative unit from a subjective assessment of risk to a more unified approach.

USPS-OIG officials have demonstrated the system for and shared their knowledge with other federal agencies. They believe the RADR system could be

FIGURE 10: SYSTEM INTERFACE
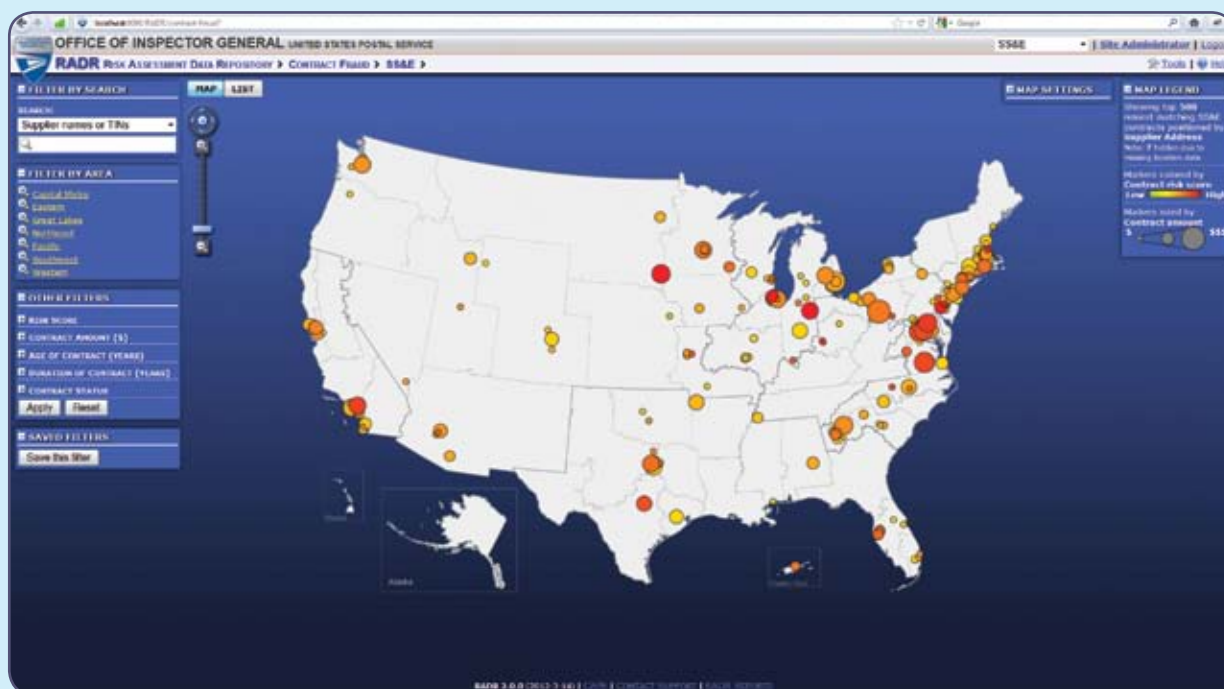
modified and deployed in a cost-effective manner across federal inspectors general and chief financial officers' communities to assist in the identification of contract and worker compensation fraud and abuse.

The work of the Recovery Operations Center shows another example of cost-effective deployment solutions.

Recovery Accountability and Transparency Board

The Recovery Accountability and Transparency Board (RATB) was created by the American Recovery and Reinvestment Act of 2009 (ARRA) with two goals: to facilitate transparency on the use of recovery-related funds and to detect and prevent fraud, waste and mismanagement. ARRA was enacted with provisions that required transparency and quarterly reporting of fund use and activities. Recipients were required to report information on their spending and other activities via the federalreporting.gov website. The RATB published this information on Recovery.gov.

To assist in the prevention and detection of fraud, waste and mismanagement in ARRA funds, the RATB formed the Recovery Operations Center (ROC). The ROC developed a data analytics system that is used to screen recipients of ARRA funds. In its screening process, the system uses 22 separate outside data sets from other government agencies and commercial vendors. ARRA recipients are screened to detect anomalies, which are used to identify potential high risk entities and unlikely circumstances. Risky organizations may range from those with a bad record for accountability in past government programs to entities that have re-established themselves with a new identity so that they can continue nefarious operations. One example of unlikely circumstance is where several companies have the same fax number and address, but are not in the same type of business.

Information that is obtained by the ROC through new or updated data sets are processed against the historical data to determine if transactions that seemed benign are now potentially problematic.

Potential problems are screened by data analysts using information at the ROC's disposal. After screening, the analysts provide the potentially problematic transactions to the appropriate inspector general, U.S. Attorney, or other law enforcement or audit office for follow-up and review. The ROC has concluded initiatives that include reviews of provider enrollment data, veteran's disability payments and Service Disabled Veteran Owned Small Businesses assistance. The ROC has begun a pilot program to allow inspectors general and program officials access to the ROC data sets and tools through a web portal called FederalAccountability. gov. Until recently, the ROC's systems were limited to working only on ARRA funds; however, changes to the RATB's appropriations law now allows for the testing and development of the ROC systems outside of ARRA.

The Current Government-wide Approach: Government Accountability and Transparency Board

Similar questions regarding government-wide solutions were addressed by the Government Accountability and Transparency Board (GATB). The GATB, formed in June 2011 by Executive Order (E.O.) 13576, Delivering an Efficient, Effective and Accountable Government, has as its mission to identify implementation guidelines for integrating systems that support the collection and display of government spending data, ensuring the reliability of those data and broadening the deployment of fraud detection technologies, including those proven successful during the implementation of

ARRA. Specifically, E.O. 13576 directed the GATB to work with the RATB to extend its successes and lessons learned to all federal spending.

The GATB issued an interim report in December 2011 with recommendations in three broad areas that are foundational to cross-government work in the field of data analytics. The first recommendation dealt with the adoption of a cohesive, centralized accountability framework to track and oversee spending. The report concluded that such a universal framework would have many benefits in the efforts to fight fraud, including increased collaboration, rapid innovation, less costly development efforts and data sharing.

The second recommendation advised that government must rationalize the way it collects and displays spending data by consolidating and streamlining technology platforms. The report states the existing universe of federal reporting systems and applications, developed over a long period by the federal government, is large, complex and costly. According to the GATB, federal reporting systems and applications must be integrated in a manner similar to that employed by the RATB. The RATB uses a limited set of data elements and established data standards and has developed highly scalable systems to accommodate changing reporting and display requirements. The RATB has also implemented an extremely aggressive reporting schedule for recipients and has migrated its solution to a cloud computing environment, which significantly increases efficiency while reducing operation and maintenance costs. In addition, the RATB has developed close working relationships with stakeholders.

Third, the GATB recommends that the government migrate to a universal identification system for all federal awards. The GATB found that no requirement exists for the

standardization of award IDs across the federal government and that some agencies do not have universal IDs. Performing analytics on inaccurate and non-standardized data often leads to erroneous conclusions and costs the government untold resources (in wasted money and human capital) that could be used more efficiently and effectively.

The GATB has a number of recommendations and observations specific to adopting and implementing improvements. For instance, the report recommends that the federal government identify and revise legal authorities that inhibit data matching and data analysis. The report also cites the work of the RATB and the ROC as a basic model for transparency and accountability. This includes FederalAccountability.gov, a web-based, secure portal through which both program and enforcement authorities gain access to its forensic and analytical capabilities. The GATB believes that FederalAccountability.gov has applicability beyond the Recovery Act for identifying and mitigating fraud, waste and abuse related to federal funds.

Many of the GATB's recommendations deal with transparency and accountability. However, other short-term solutions have been developed that already apply to the work of many federal agencies and departments. One example discussed earlier is the RADR system, which has components that deal with contract and workers compensation fraud. Work has also started to expand the reach of the RATB's ROC, which has developed a large number of valuable data analytics routines that can be used across the federal government.

One OIG official observed that, because many OIG operations are not very large, developing data analytics systems might not be possible. He suggested some type of shared service approach across OIG communities. Another OIG official suggested that services offered in a cloud environment

offering broad accessibility at a reduced cost. The question: How should the efforts be organized and, if offered in some consolidated manner, where they would be housed?

If fully implemented, recommendations in the GATB's December 2011 report related to transparency and accountability across the government and the tracking of federal funds could facilitate data matching between agencies.

Part Six: Recommendations

- **Build on Success:** The federal government should strive to leverage the experience gained by those who have implemented data analytics systems to help other organizations overcome deficiencies in budget, staffing and experience. Rather than assuming that each agency should develop its own data analytics solution from cradle to grave, the federal government should consider shared service arrangements enabling agencies with successful systems to build upon the work of others. Prime examples of this are the contract fraud systems developed by USPS-OIG, the data analytics routines developed by the ROC and the ED-OIG system for identifying fraud among federal fund award recipients. Powered by technology, these systems could be used across the spectrum of federal agencies, including the inspector general and CFO communities.

- **Continue Education:** While it appears most government leaders have a general understanding of data analytics, they are not yet champions of the process. Federal leaders must continue to be educated on the benefits and uses of data analytics and the actions needed to implement data analytics in their agency.

- **Focus on Performance and Outcomes:** Organizations should expand the use of data analytics to include performance and outcomes. Only one of the eight organizations interviewed was using data analytics to measure program performance; the others were primarily using it to prevent and detect improper payments, obtain information on financial performance and identify instances of fraud and abuse for investigation and audit. Given the current emphasis on reducing government spending, data analytics could play a vital role in helping decision makers determine where scarce dollars could be most effectively used.

- **Procurement:** A guide to procuring data analytics systems or consulting services should be developed (ideally by a neutral third party) based on the experiences of federal officials that have procured systems and the vendors who provide these services. This guide could include information on the services that can be procured, ways to determine the correct data analytics method, provisions that significantly increase or decrease the cost of a contract and lessons learned from previous procurements.

- **Explore Intergovernmental Services:** Federally funded programs implemented at the state or local government levels should be reviewed to determine whether a collaborative arrangement can be developed for data analytics. A central system administered at the federal level might be more effective and cost efficient than stand-alone systems throughout the country.

Endnotes

1. "Improper Payments: Not Just the Purview of the CFO Anymore?" AGA. September 2011.

2. "Using Performance Information to Drive Performance Improvement." AGA. December 2011.

3. Antonia de Medinaceli, Director of Business Analytics and Fraud Detection, Elder Research, Inc. Presentation during AGA audio conference, *Data Mining to Prioritize Investigations of Contract Fraud*. March 7, 2012.

4. SNAP provided nutrition assistance to more than 21 million households and disbursed more than $71.8 billion in benefits in FY 2011. SNAP is administered with the assistance of the state government agencies. State officials determine recipient eligibility, and FNS enrolls retailers in the payment system and remits payments to the retailers through an electronic benefit transfer system.

5. *Why Leadership and Culture Matter: Helping Turn Analytic Insights into Outcomes and Drive High Performance*. Accenture. May 20, 2011.

6. A data warehouse is a database created for reporting and analysis purposes only. Data is stored in either a dimensional or a normalized approach, and multiple users can access the data. A data mart is a subset of a data warehouse.

7. Sparrow, Malcolm K. *License to Steal How Fraud Bleeds America's Health Care System*. Westview Press, 2000.

Project Participants:
Government Agencies

U.S. Department of Agriculture

Jeffrey N. Cohen
Deputy Associate Administrator,
Supplemental Nutrition
Assistance Program
Food and Nutrition Service

Thomas (Tim) O'Connor
Associate Administrator, Special
Nutrition Programs
Food and Nutrition Service

Shelly Pierce
Program Analyst,
Benefit Redemption Division
Food and Nutrition Service

U.S. Department of Defense, Defense Finance and Accounting Service

Rebecca S. Beck
Director, Accounting, Columbus

Terri Schulze
Director, Site Support Office, Columbus

William McGee
Improper Payments Functional Lead
Business Activity Monitoring (BAM)

U.S. Department of Defense, Defense Logistics Agency

Lawrence Vadala
Office of Operations Research
and Resource Analysis

U.S. Department of Defense, Department of the Navy

Willie White
Chief, Anti-Fraud Officer
NAVSEA Office of Fraud Deterrence
and Detection, SEA 00F

Laura Crawford
Anti-Fraud Project Manager
NAVSEA Office of Fraud Deterrence
and Detection, SEA 00F

Steve Derry
Business Operations Manager
NAVSEA Office of Fraud Deterrence
and Detection, SEA 00F
U.S. Department of Education,
Office of the Inspector General

Edward Slevin
Director
Computer Assisted Assessment
Techniques Division

U.S. Department of Health and Human Services, Center for Medicare & Medicaid Services

David Nelson
Director, Data Analytics and
Control Group
Center for Program Integrity

Kelly Gent
Deputy Director, Data Analytics and
Control Group
Center for Program Integrity

Devin Williams
Director of the Command Center Division
Data Analytics and Control Group
Center for Program Integrity

U.S. Postal Service, Office of the Inspector General

Bryan Jones
Director, Countermeasures and
Performance Evaluation
Data Mining Group
Recovery Accountability and
Transparency Board

Alan F. Boehm
Assistant Director
Recovery Operations Center

Ed Martin
Deputy Assistant Director
Recovery Operations Center

AGA CPAG RESEARCH REPORTS

No. 1, March 2005:   *Audit Federal Financial Controls: Sooner Rather than Later?*

No. 2, July 2005:   *Financial Management Shared Services: A Guide for Federal Users*

No. 3, November 2005:   *Trends in Technology*

No. 4, April 2006:   *The Federal Purchase Card: Use, Policy and Practice*

No. 5, June 2006:   *Challenges in Performance Auditing: How a State Auditor with
Intriguing New Performance Authority is Meeting Them*

No. 6, June 2006:   *PAR—The Report We Hate to Love*

No. 7, February 2007:   *The State Purchase Card: Uses, Policies and Best Practices*

No. 8, March 2007:   *Federal Real Property Asset Management*

No. 9, May 2007:   *Should State and Local Governments Strengthen Financial Controls
by Applying SOX-Like Requirements?*

No. 10, April 2007:   *Process-Based Financial Reporting*

No. 11, May 2007:   *The State Travel Card—Uses, Policies and Best Practices*

No. 12, June 2007:   *Trends in Technology—2007 Review*

No. 13, June 2007:   *The Federal Travel Card—Uses, Policies and Best Practices*

No. 14, January 2008:   *21st Century Financial Managers—A New Mix of Skills and Educational Levels?*

No. 15, July 2008:   *SAS 70 Reports: Are they Useful and Can They Be Improved?*

No. 16, September 2008:   *XBRL and Public Sector Financial Reporting: Standardized Business Reporting:
The Oregon CAFR Project*

No. 17, November 2008:   *Characteristics of Effective Audit Committees in Federal, State and Local Governments*

No. 18, January 2009:   *Grants Management: How XBRL Can Help*

No. 19, February 2009:   *Procuring Audit Services in Government: A Practical Guide to Making the Right Decision*

No. 20, March 2009:   *Performance-Based Management*

No. 21, June 2009:   *Trends in Technology—2009 Review*

No. 22, September 2009:   *Managerial Cost Accounting in the Federal Government:
Providing Useful Information for Decision Making*

No. 23, November 2009:   *State and Local Governments' Use of Performance Measures to Improve Service Delivery*

No. 24, June 2010:   *Creating an Interactive Single Audit Database*

No. 25, July 2010:   *Redefining Accountability: Recovery Act Practices and Opportunities*

No. 26, September 2010:   *The Maturity of GRC in the Public Sector: Where Are We Today? Where Are We Going?*

No. 27, June 2011:   *Trends in Technology: 2011 Report, The Information Explosion*

No. 28, September 2011:   *Improper Payments: Not Just the Purview of the CFO Anymore?*

No. 29, December 2011:   *Using Performance Information to Drive Performance Improvement*

Download complimentary reports at www.agacgfm.org/researchpublications.

*Advancing
Government
Accountability*

Association
of Government
Accountants

2208 Mount Vernon Avenue
Alexandria, VA 22301

703.684.6931
800.AGA.7211
703.548.9367 (fax)

www.agacgfm.org
agamembers@agacgfm.org