



## ABOUT THE ACADEMY

The National Academy of Public Administration is a non-profit, independent organization of top public management and organizational leaders who tackle the nation's most critical and complex public management challenges. With a network of nearly 700 distinguished Fellows and an experienced professional staff, the Academy is uniquely qualified and trusted across government to provide objective advice and practical solutions based on systematic research and expert analysis. Established in 1967 and chartered by Congress in 1984, the Academy continues to make a positive impact by helping federal, state and local governments respond effectively to current circumstances and changing conditions. Learn more about the Academy and its work at [www.NAPAwash.org](http://www.NAPAwash.org).

Word cloud on cover generated from Dialogue feedback using Wordle (<http://www.wordle.net>).

*A Report by a Panel of the*

NATIONAL ACADEMY OF  
PUBLIC ADMINISTRATION

*For the Recovery Accountability and Transparency Board*

**National Dialogue on Innovative Tools to Prevent  
and Detect Fraud, Waste, and Abuse**

**PANEL**

**Alan Shark,\* Chair**

**Frank Chellino\***

**Franklin Reeder\***



*\* Academy Fellow*

December 2011

## **Officers of the Academy**

**Diane M. Disney**, *Chair of the Board*  
**Robert J. Shea**, *Vice Chair*  
**David F. Garrison**, *Secretary*  
**Sallyanne Harper**, *Treasurer*  
**Dan G. Blair**, *President and Chief Executive Officer*

## **Study Team**

**Joseph P. Mitchell, III**, *Director of Project Development*  
**Laurie J. May**, *Project Director*  
**Faith Gibson**, *Research Associate*  
**Matthew Thomas**, *Research Associate*

---

The views expressed in this report are those of the Panel. They do not necessarily reflect the views of the Academy as an institution.

National Academy of Public Administration  
900 7<sup>th</sup> Street, N.W.  
Set 600  
Washington, DC 20001-3888  
[www.napawash.org](http://www.napawash.org)

Published December 2011

Printed in the United States of America

Academy Project Number: 2167-001

## TABLE OF CONTENTS

<b>FOREWORD</b> .....	<b>v</b>
<b>ACRONYMS</b> .....	<b>vii</b>
<b>EXECUTIVE SUMMARY</b> .....	<b>ix</b>
<b>Consolidated List of Recommendations</b> .....	<b>x</b>
<b>INTRODUCTION</b> .....	<b>1</b>
<b>Putting the Dialogue in a Larger Context</b> .....	<b>2</b>
<b>SECTION I: ANALYSIS OF DIALOGUE RESULTS, INTERVIEWS, AND RESEARCH</b> .....	<b>5</b>
<b>Distribution of Ideas and Content</b> .....	<b>5</b>
<b>Issues and Themes</b> .....	<b>9</b>
<b>The Prevention/Detection Continuum: An Analytic Framework</b> .....	<b>14</b>
<b>SECTION II: RECOVERY BOARD DATA CHALLENGES</b> .....	<b>17</b>
<b>A Uniform System for Identifying Federal Contracts and Grants</b> .....	<b>17</b>
<b>Federal Tax Delinquency History</b> .....	<b>17</b>
<b>Personally Identifiable Information</b> .....	<b>18</b>
<b>SECTION III: RECOMMENDATIONS AND CONCLUDING REMARKS</b> .....	<b>21</b>
<b>Recommendations for the Recovery Board</b> .....	<b>21</b>
<b>Recommendations for Government-wide Consideration</b> .....	<b>26</b>
<b>Concluding Remarks</b> .....	<b>29</b>
<b>APPENDICES</b>	
<b>APPENDIX A: PANEL AND STAFF BIOS</b> .....	<b>31</b>
<b>APPENDIX B: INDIVIDUALS INTERVIEWED</b> .....	<b>33</b>
<b>APPENDIX C: DIALOGUE METHODOLOGY AND METRICS</b> .....	<b>35</b>
<b>Planning and Execution of the Dialogue</b> .....	<b>35</b>
<b>Measuring Traffic and Participation</b> .....	<b>38</b>
<b>APPENDIX D: DIALOGUE STATEMENT OF PURPOSE AND POLICIES</b> .....	<b>45</b>
<b>APPENDIX E: IDEAS/COMMENTS SUBMITTED IN DIALOGUE</b> .....	<b>49</b>

**This page intentionally left blank.**

## FOREWORD

*"The American people must be able to trust that their Government is doing everything in its power to stop wasteful practices and earn a high return on every tax dollar that is spent."  
June 2011 Executive Order*

To manage effectively, today's public administrators must utilize data-driven decision-making, engage in continuous improvement, and conduct broad stakeholder engagement. These themes—long championed by independent Panels of the National Academy of Public Administration (the Academy)—came together in this report for the Recovery Accountability and Transparency Board. Although the Board uses a wide range of tools to prevent and detect fraud in its oversight of stimulus funds, it engaged the Academy to conduct a national online dialogue, expert interviews, and other research to determine if other promising tools, technologies, and strategies were available.

Given the nation's long-term fiscal challenges, the federal government should utilize the most advanced tools and techniques to collect and analyze a wide array of data in order to ensure that tax dollars are used as intended. The Academy convened a three-member expert Panel to oversee this work. This report not only summarizes key insights gained from the October 2011 dialogue and other research, but also includes the Panel's recommendations—some of which are directed specifically to the Board, while others are for the Legislative and Executive Branch to consider adopting. The Panel believes its recommendations will help the Board and the rest of the federal government expand the suite of prevention/detection tools, improve federal spending data, and increase the public's trust in government.

I extend my appreciation to Chairman Earl Devaney and the Recovery Board for providing the Academy with an opportunity to assist in this ground-breaking effort to take federal data management and analysis to the next level. Thanks as well to the Academy Panel members for their excellent work and important insights. I would also like to acknowledge the contributions of dialogue participants, private sector technologists, and government officials who participated in the study and generously shared their time, expertise, and perspectives. Finally, I want to acknowledge our study team for their significant contributions to this thorough and insightful report.



Dan G. Blair

President and Chief Executive Officer

**This page intentionally left blank.**



## ACRONYMS

<b>Academy, The</b>	National Academy of Public Administration
<b>CBO</b>	Congressional Budget Office
<b>DATA Act</b>	The Digital Accountability and Transparency Act of 2011
<b>DOJ</b>	Department of Justice
<b>Dialogue, The</b>	The National Dialogue on Innovative Tools to Prevent and Detect Fraud, Waste and Abuse
<b>FAR</b>	Federal Acquisition Regulations
<b>GAO</b>	Government Accountability Office
<b>GATB</b>	Government Accountability and Transparency Board
<b>GIS</b>	Geographic Information Service
<b>IRS</b>	Internal Revenue Service
<b>MPP</b>	Massively Parallel Processing
<b>NAICS</b>	North American Industry Classification System
<b>NLP</b>	Natural Language Processing
<b>OLAP</b>	Online Analytical Processing
<b>OMB</b>	Office of Management and Budget
<b>Panel</b>	A Panel of National Academy of Public Administration Fellows
<b>Partnership Fund</b>	Partnership Fund for Program Integrity Innovation
<b>PII</b>	Personally Identifiable Information
<b>Recovery Board, The</b>	The Recovery Accountability and Transparency Board
<b>Recovery Act</b>	American Recovery and Reinvestment Act of 2009
<b>ROC</b>	Recovery Operations Center
<b>SSN</b>	Social Security Number
<b>SQL</b>	Structured Query Language

**This page intentionally left blank.**

## EXECUTIVE SUMMARY

On February 17, 2009, Congress enacted the American Recovery and Reinvestment Act of 2009 (Recovery Act) to help stimulate the United States economy by investing in projects expected to provide long-term economic benefits.<sup>1</sup> The Recovery Act established the Recovery Accountability and Transparency Board (the Recovery Board) to coordinate and conduct oversight of covered funds and to prevent fraud, waste, and abuse. The Recovery Board, a non-political agency, has two goals:

- To provide transparency of Recovery-related funds; and
- To detect and prevent fraud, waste, and mismanagement.

It is staffed primarily by personnel on loan from other federal organizations and a small number of individuals serving on term appointments. The Recovery Board began meeting regularly in May 2009 and will sunset in September 2013.

In September 2011, the Recovery Board, recognizing rapid technological advancements in the commercial sector, partnered with the National Academy of Public Administration (the Academy) to host a National Dialogue on Innovative Tools to Prevent and Detect Fraud, Waste, and Abuse (the Dialogue). This Dialogue is part of the Recovery Board's continuing commitment to identifying and using state-of-the-art tools as it enhances accountability, and improves oversight of Recovery Act funding. The Dialogue reached out to technologists, thought leaders, and other interested parties and sought to engage them in a discussion about the most promising advances in the fight against fraud, waste, and mismanagement.

The Academy hosted the public Dialogue from 8 AM October 17<sup>th</sup> through 8 AM on October 24<sup>th</sup>, during which time there were 953 visits from 591 unique individuals; 53 users registered to participate. Dialogue participants used the opportunity to share their expertise and knowledge and submitted 36 ideas via the public dialogue and 20 comments. Nine participants elected to use an email link to submit ideas directly to the Academy. A Panel of Academy Fellows directed the analysis of the Dialogue, supplemental interviews, and research, and identified two sets of recommendations: the first warranting further exploration by the Recovery Board and the second for potential government-wide follow-up. (See Appendix A for brief biographies of the Academy Panel and project staff.)

Table 1 summarizes the Panel's recommendations.

---

<sup>1</sup> Pub. L. No 111-5, 123 Stat.115

**TABLE 1 – RECOMMENDATIONS**

<b>Part A- FOR RECOVERY BOARD CONSIDERATION</b>
<b>1. Increase emphasis on predictive analysis, particularly to prevent and detect contract/grant fraud.</b>
<b>2. Increase use of sophisticated textual analysis tools to mine the abundance of narrative information that is unstructured.</b>
<b>3. Increase data sources, particularly state and local governmental data and proprietary business data, to improve data validation.</b>
<b>4. Work across government to establish and publicize more consistent performance metrics for fund recipients and increase transparency of outcomes for tax dollars spent.</b>
<b>Part B – FOR GOVERNMENT-WIDE CONSIDERATION</b>
<b>5. Consider establishing a permanent, centralized portal for data to enhance federal data management and analysis.</b>
<b>6. Evaluate ways to expedite the sharing of aggregated federal data to enhance federal predictive modeling.</b>
<b>7. Consider establishing a uniform system for identifying federal contracts and grants to improve tracking of federal payments to recipients.</b>
<b>8. Explore regulatory changes to require applicants for federal funding to sign a waiver allowing access to their tax records.</b>

## INTRODUCTION

In 2009, the Recovery Board partnered with the Academy to host an online dialogue focused on soliciting broad public input on how the Recovery.gov website should display information about Recovery Act funding in the most transparent and accessible manner possible. This Recovery Act-mandated website broke new ground by relying on recipient, rather than agency-provided data collected via FederalReporting.gov. Using technology to aggregate the collected data, the public-facing website, Recovery.gov, is able to provide user-friendly, digestible, and up-to-date information on how and where taxpayers' money is spent. The results of that initial study can be found at <http://www.napawash.org/wp-content/uploads/2009/09-07.pdf> .

In 2011, the Recovery Board requested that the Academy provide similar insights into how to enhance its already aggressive efforts to prevent and detect fraud, waste, and abuse in Recovery Act spending. This interactive, online Dialogue focused on gathering innovative tools and strategies from those with relevant knowledge and expertise and asked five key questions:

- What management practices, policies and programs and incentives would improve financial stewardship and help prevent and detect fraud, waste and abuse?
- What specific governmental, public or proprietary data sources could help the Recovery Board prevent and detect fraud, waste and abuse?
- What technologies or systems do you think would be effective in integrating and aggregating diverse types of data?
- What types of risk models would identify entities receiving Recovery Act funds as being most vulnerable to fraud, waste and abuse?
- What types of performance metrics could be applied to Recovery Act-funded programs and recipients to increase oversight and accountability?

When the week-long Dialogue concluded, the Academy analyzed the input and supplemented the insights gained from it with expert interviews and research to ensure that future enhancements to the Recovery Board's oversight efforts reflect the best ideas of America's leading thinkers. (See Appendix B for a list of individuals interviewed for this effort.)

Prior to the Dialogue's launch, the Recovery Board and the Academy invited participants to the Dialogue site by conducting targeted outreach via email and telephone contact. The Academy reached out to 906 individuals and organizations, including representatives of the media/press, industry, academia, associations and foundations, and state, local, and federal officials. Using the guidance provided by the Recovery Board, the Academy project team initially identified 670 individual email addresses for those in the targeted audience; these included Academy Fellows with experience in such occupations as accounting, auditing, forensic accounting, governmental management, law enforcement, relevant academic fields, IT, cloud computing, risk models, performance metrics, and intelligence analysis. The Academy also sent an email invitation to the more than 600 Fellows and made personal contact with about two dozen individuals with relevant experience to invite them to participate and to encourage their colleagues as well. In addition, the Academy project staff made more than 100 outreach phone calls. (See Appendix C, Dialogue Methodology and Metrics for additional information and Appendix D for Dialogue statement of purpose and policies.)

Between October 17<sup>th</sup> and October 24<sup>th</sup>, 2011, the Dialogue received 953 visits from 591 unique visitors, which equates to an average of 136 visits from 84 unique visitors each day. Direct traffic<sup>2</sup> to the site accounted for almost 60 percent of all visits made, indicating that most visitors accessed the dialogue directly, as opposed to finding it through a search engine.

Many of the traffic metrics from the Dialogue were lower than originally anticipated. In the view of the Panel, these lower metrics are likely attributable to: (1) the compressed time period for outreach, and (2) the specialized nature of the Dialogue subject matter. Approximately half of all visitors came to the homepage and left without exploring the individual Dialogue forums. However, when the other traffic metrics are considered in relation to this relatively high “bounce rate”<sup>3</sup> of 48.27 percent, site activity among those who were interested in the subject matter appears comparable to other Academy dialogue initiatives. When those visitors who only visited the homepage are removed from the calculation of average page views and length of time on site, the metrics show that the remaining 52 percent who went beyond the homepage visited an average of six pages and spent significantly more time on the site than the four minutes recorded for all participants.

The Dialogue platform was open for participation twenty-four hours a day during this period. It allowed users to submit ideas in five different discussion forums and to rate, tag, and comment on other users’ public submissions.

This report contains the Academy’s analysis of the Dialogue results, as well as recommendations for Recovery Board and broader government follow-up.

## **PUTTING THE DIALOGUE IN A LARGER CONTEXT**

Both the Recovery Board’s efforts over the last several years and this Dialogue are most appropriately viewed in the context of federal efforts to stem improper payments and curtail fraud, waste, and abuse. By strengthening financial management controls and facilitating the improved prevention and detection of improper payments, the government can better ensure that taxpayer dollars are spent wisely and efficiently. Taxpayers are able to report suspected incidents of fraud, waste, and abuse via FraudNet—established in 1979 and expanded in recent years to receive allegations via the internet, fax, or letter.

In 2010, the federal government’s improper federal payments—including funds to the wrong recipient, incorrect amounts, lack of documentation, and so on—were estimated to be \$125 billion dollars – or about 5.5 percent of all payments. This represented an increase of \$15 billion from 2009. In evaluating the programs with highest known error rates, the most susceptible programs appear—unlike the Recovery Board—to make payments primarily to individuals.<sup>4</sup>

---

<sup>2</sup> Direct traffic measures the percentage of visits to the site that came from users clicking an email link or directly typing the URL into their web browser.

<sup>3</sup> The percentage of visits that entail only visiting the first page of a site is termed the “bounce rate.”

<sup>4</sup> Recognizing that contracts and grants are also vulnerable, on June 18, 2010, the President issued a memorandum directing the establishment of a “Do Not Pay List.”

The top five programs in terms of improper payment amounts for 2010 were Medicare, Medicaid, Unemployment Insurance, Medicare Advantage (Part C), and Supplemental Social Security Income (SSI).

In July of 2011, as part of the “Campaign to Cut Waste,” the President announced the launch of a new board, the Government Accountability and Transparency Board (GATB). Its mission—broader, yet clearly parallel to that of the Recovery Board — is to root out misspent tax dollars and make government spending more accessible and transparent for the American people. The GATB is currently chaired on an interim basis by the Chair of the Recovery Board.<sup>5</sup>

There is also continuing interest on the legislative front to strengthen payment integrity and provide a statutory authority for the GATB. The Digital Accountability and Transparency Act of 2011—known as the DATA Act (H.R. 2146)—is currently pending in Congress. Introduced in June of 2011, the legislation would establish a new independent board within the Executive Branch that would replace the Recovery Board. The bill calls for the new board to track all federal spending on a single website and requires the use of consistent government-wide data standards. In a September 16, 2011 report, the Congressional Budget Office (CBO) estimated that implementing the bill would cost \$575 million over the 2012-2016 period. Of that amount, CBO estimated that about \$325 million would be required to improve the government’s current efforts to collect and report on financial data. CBO went on to say that enacting the DATA Act “could increase revenues from the collection of civil and criminal penalties and direct spending of those amounts,” but estimated that “the net budgetary impact of any additional collections would be negligible for each year.”<sup>6</sup> CBO did not include an estimate of federal dollars saved in prevention of fraud, waste, and abuse.

Consistent with all of these efforts and ongoing funding to reduce improper payments and improve federal information systems is the Recovery Board’s focus on finding the most innovative strategies, tools, and technologies to prevent and detect fraud, waste, and abuse among recipients of its federal contracts and grants. While focused on Recovery Act funds, the Recovery’s Board’s efforts are likely to have a far broader impact. Whereas the improper payment program has, by design, focused on the high risk payments to individuals, the Board is now the primary laboratory for testing data sources, fraud patterns, risk models, and predictive technologies with applicability to contracts and grants.

---

<sup>5</sup> See July 28, 2011 White House press release at <http://www.whitehouse.gov/the-press-office/2011/07/28/white-house-launches-governme...>

<sup>6</sup> Congressional Budget Office, “H.R. 2146, DATA Act: As ordered reported by the House Committee on Oversight and Government Reform on June 22, 2011,” September 16, 2011, p. 1.

**This page intentionally left blank.**



## SECTION I. ANALYSIS OF DIALOGUE RESULTS, INTERVIEWS, AND RESEARCH

The Dialogue provided participants from a broad spectrum of disciplines and perspectives an opportunity to share their knowledge and expertise and help the Recovery Board to identify new tools and strategies by which it might prevent and detect fraud, waste, and abuse of Recovery Act funds and improve the oversight of Recovery Act spending. The Dialogue also encouraged participants to share ideas related to oversight of all federal spending, particularly if those ideas had applicability to Recovery Act spending. Participants who preferred to submit an idea without sharing it in the public Dialogue had the option to send solution narratives directly to the Academy using a link in the public Dialogue.

The Dialogue consisted of the following five discussion forums, in which participants could submit ideas and rate and comment on the ideas of others:

- Management;
- Data Sources;
- Technologies;
- Risk Models; and
- Performance Metrics

In addition to insights gained from the Dialogue and concomitant submissions, the Academy gained insight from interviews with 15 industry thought leaders and governmental officials and related research. At the conclusion of the process, a Panel of Academy Fellows directed the analysis of these combined efforts. This section analyzes those results.

### DISTRIBUTION OF IDEAS AND CONTENT

In all, participants submitted 36 ideas via the public Dialogue and 20 comments. Nine participants provided solution narratives through the direct submission mechanism. Table 2 depicts the distribution of ideas, comments, and direct submissions among the five topic areas and shows that the highest level of interest by far was in the management arena, with technologies ranked second.

**TABLE 2 - DISTRIBUTION OF DIALOGUE IDEAS, COMMENTS, AND  
DIRECT SUBMISSIONS**

	<b>Mgt.</b>	<b>Data Sources</b>	<b>Technologies</b>	<b>Risk Models</b>	<b>Performance Metrics</b>	<b>Other</b>	<b>Total</b>
Ideas	15	5	6	4	6		36
Comments	9	3	4	2	2		20
Direct Submissions	2	1	2	1	1	2	9
<b>Total</b>	<b>26</b>	<b>9</b>	<b>12</b>	<b>7</b>	<b>9</b>	<b>2</b>	<b>65</b>

Table 3 provides a brief summary of each idea discussed in the public Dialogue and a tally of votes cast by Dialogue participants. Note that while there were five established categories, Dialogue participants often made suggestions across those boundaries or had different perspectives as to what, for example, might be construed as a management strategy. Seven of the 36 ideas received six or more positive votes. The idea receiving the highest number of votes (11) was Idea 23, which advocated the use of multiple technological approaches to address the increasing complexity of fraud networks and schemes. Appendix E provides a verbatim recounting of the Dialogue discussion.

**TABLE 3 - DIALOGUE IDEA SUMMARY**

<b>MANAGEMENT</b>	<b>Positive Votes</b>	<b>Negative Votes</b>
<b>Idea 1</b> - Separate qualification and allocation functions.	2	0
<b>Idea 2</b> - Design funds distribution phase carefully pre-announcement.	2	0
<b>Idea 3</b> - Provide some portion of funds only after demonstration of performance.	2	0
<b>Idea 4</b> - Reimburse for actual expenses. Do not use fixed price contracts for universities.	0	5
<b>Idea 5</b> – Hold universities accountable.	1	2
<b>Idea 6</b> - Strengthen systems and audit processes to assure timely de-obligation of unneeded funds.	4	0
<b>Idea 7</b> – Do more to collect accounts and taxes receivable.	3	0
<b>Idea 8</b> - Focus prevention and detection on business and individual recipients applying for government benefits, grants, and loans because their motivation is self- interest, and they pose a higher risk.	7	5
<b>Idea 9</b> – Do not keep secrets. Hold quarterly public forums for civil servants to answer questions from the public.	2	4
<b>Idea 10</b> – Agencies should be compelled to fully audit their books, not take a sample.	0	4
<b>Idea 11</b> – Pull as much of the data on to a single integrated platform to enable data mining. Build statistical models to analyze patterns.	4	0
<b>Idea 12</b> – The government needs to provide full organizational access to detailed financial data throughout the accounting cycle. Eliminate information black holes. People need to be able to trust the data.	3	0
<b>Idea 13</b> – The government needs internal controls for the full range of program activities, not just accounting and finance and to identify and control risks.	6	0

<b>Idea 14</b> – Integrity means complete forthrightness in all dealings.	2	1
<b>Idea 15</b> – How does an organization really measure return on investment in fraud detection?	1	1
<b>DATA SOURCES</b>	<b>Positive Votes</b>	<b>Negative Votes</b>
<b>Idea 16</b> - Before declaring any person or organization eligible to receive funds, the government should pre-certify, run a credit check, a check of criminal files, and relevant licenses. Follow-up with relevant databases on theory that negative tendencies repeat.	1	0
<b>Idea 17</b> - Identify types of files containing names of those applying for benefits and determine if computer matches with those files should be mandatory before payments are made.	5	0
<b>Idea 18</b> – Can we use Whitelisting or crowd sourcing to aid in reduction of fraud?	3	1
<b>Idea 19</b> - What state and local data bases could provide information on negative actors?	7	0
<b>Idea 20</b> - What about industry negative databases (blacklists), business license data, crime data, social media data, claim forms, notes, historic data (sales, transaction, purchase, etc.), payment records, location information, etc.?	6	0
<b>TECHNOLOGIES</b>	<b>Positive Votes</b>	<b>Negative Votes</b>
<b>Idea 21</b> – Are there any affordable, robust solutions that automate fraud detection other than those offered by Thomson Reuters and SAS EBI?	2	1
<b>Idea 22</b> - What analytical software tools are most useful?	3	0
<b>Idea 23</b> – Use multiple approaches that are flexible to support new requirements and that can address increasing complexity of fraud networks and schemes. Requirements: advanced analytic techniques, such as geospatial mapping and analysis, social network analysis, in-database data mining, and text analysis.	11	0
<b>Idea 24</b> – Use a framework, such as MapReduce, to explore large data sources that have not been fully cleansed and structured.	8	0
<b>Idea 25</b> - Integrate and aggregate data at the most detailed level possible. Use a scalable, massively parallel processing (MPP) relational database system to grow the system without sacrificing performance.	6	0
<b>Idea 26</b> – Use artificial intelligence?	4	0

<b>RISK MODELS</b>	<b>Positive Votes</b>	<b>Negative Votes</b>
<b>Idea 27</b> – Use content analysis based on key phrases associated with fraud to review new proposals and try to identify vulnerabilities.	2	0
<b>Idea 28</b> – On what risk models/ risk management organizations do you rely?	3	1
<b>Idea 29</b> - Identify good and bad examples of fraud and look for differences and sources. Are there guidelines/rules available to assess compliance, guide exploratory analysis, and identify initial fraud, waste, or abuse cases?	4	0
<b>Idea 30</b> – Use unstructured data as input to add power to analytics. Note that text data from documents, e.g. emails, funding applications, and contracts, requires some pre-processing. Do not focus solely on structured data.	5	0
<b>PERFORMANCE METRICS</b>	<b>Positive Votes</b>	<b>Negative Votes</b>
<b>Idea 31</b> – Are online dashboards used for organizing and assessing data?	2	0
<b>Idea 32</b> - Have a rubric to evaluate the effectiveness of consultants and contractors; minimum requirement is to discern relative age/experience relative to task; screen out those organizations formed just to get at funds.	2	0
<b>Idea 33</b> - Provide card access vehicle for employees to give anonymous feedback on grants and contracts and consider adding incentive if information proves useful in identifying fraud.	4	1
<b>Idea 34</b> - Measure outcomes at stakeholder level where the stimulus money actually goes. This will help identify such frauds as wage theft.	2	0
<b>Idea 35</b> - Make annual recovery audits mandatory at each cabinet level and quasi-governmental agency to reduce \$125 billion in annual improper payments. Upload software in every agency’s accounts payable system.	2	1
<b>Idea 36</b> - Each program should have explicit, measurable input and outcome goals from the outset, with logical relationships between outputs and outcomes. Require periodic reports of negative trends/failures and determine and address reasons.	7	1

## ISSUES AND THEMES

The most relevant feedback focused on the topics of data, risk models, technologies, and performance metrics. To collect additional tools and strategies, the study team interviewed 15 individuals and researched topics of interest to the Panel and the Recovery Board. After analyzing the results of the Dialogue, interviews, and supporting research, the Academy organized the contributions into several themes aligning with the last four dialogue categories, and to a category the Panel has captioned “Organizational Awareness and Strategy.” The ideas set forth in this section are not always mutually consistent nor do they necessarily reflect the views of the Panel.

### **Data**

Issues related to data were a common topic of discussion in both the Dialogue and the interviews conducted by the Academy study team. Ideas tended to focus around three themes: Data Sources, Data Integration, and Data Cleansing and Validation.

#### *Data Sources*

- Identify the data sources available for mining and their relationship to the problem at hand.
- For grants and sub-contracts, the government must have data on expenditures at the lowest level.
- Increase the types of databases that are used: e.g. state and local databases, industry negative databases (blacklists), Whitelists (positive registers of members or those with access/recognition); business license data, crime data, social media data (mass online collaborations or “crowd sourcing,” community bulletin boards, online forums); claim forms, notes, historic data (sales, transaction, purchase), payment records, location information.
- Use proprietary databases as needed to get revenue, employment, location, and product/line business information for publicly and privately held companies to match that information with service/product offered to the government. If the business does not have sufficient expertise, staff, or history of offering that product or service, it may be a good indicator of possible inability to perform or possibly of fraud.
- Use large data sources even if they have not been fully cleansed and structured.
- Do not focus solely on structured data; unstructured data adds power to analytics. Text data from documents, emails, funding applications and contracts may require some pre-processing.

### *Data Integration*

- Create a complete data dictionary — essential for common understanding of terminology — as the first step prior to integration.
- Integrate and aggregate the data at the most detailed level possible. Data silos do not facilitate the best decision-making.
- Pull as much of the data on to a single integrated platform to enable data mining and real-time fraud detection.
- The key to detecting fraud is discovering the relationship among seemingly unrelated data elements.
- In centralizing data, an organization gains a single version of the truth. Storing data once and accessing it many times through a variety of methods to answer even unanticipated queries reduces redundancies and costs and improves synergy.

### *Data Cleaning and Validation*

- People need to be able to trust the data; data preparation and cleansing are critical. Resolve missing values, poor/incorrect data entry, and inconsistencies from multiple sources. Remove data ambiguities and limit fraud. “Fraudsters” thrive on ambiguity.
- Assign unique, but consistent, identifiers to anonymize data containing Personally Identifiable Information (PII) so analysts can see patterns without violating PII constraints.
- Validate data used in decision-making against other government or third-party sources to determine accuracy before making payment.

### **Risk Models**

For this project, the Recovery Board was especially interested in collecting ideas on possible risk models that could be adopted. Most of the feedback received focused on the iterative nature of the risk modeling process. The ideas listed below are arranged in order of logical implementation to illustrate the process of risk model development. This is not meant to be an exhaustive list of all the steps involved.

- Use compliance guidelines and rules to guide analysis and identify initial fraud cases.
- Build an accurate model by basing it on historical training data from known fraud cases, including both good and bad examples.

- Consider decision trees, neural networks, Bayesian networks, support vector machines, random forests, and regression models.
- Evaluate and validate a model prior to deployment to make sure it will perform as expected.
- Once deployed, monitor the model's performance on new models. Automate as much of the monitoring as possible to free up staff time.

## **Technologies**

Most of the ideas related to technologies fell into three sub-categories: Alignment of Technology with Need, Specific Technologies, and Technology Strategy. These ideas came from the Dialogue, interviews and supplemental Academy research and are not intended to represent an exhaustive list of specific technologies that the Recovery Board could use to prevent or detect fraud, waste, and abuse.

### *Alignment of Technology with Need*

- First articulate the problem that the organization is trying to solve; then pick a suite of tools that addresses those needs within practical constraints.
- Do not “follow the shiny pebble.” The newest tool being used by others does not necessarily work for your organization's challenges.
- Do not put the focus on proprietary software just because your organization already owns it and paid a lot of money for it. Find the tool that will work most effectively for what you are trying to achieve. Do not force feed data through your existing technology just because you have done that historically.

### *Specific Technologies*

- Consider Open Source software; it can often be more effective than proprietary software and at a lower cost.
- Data matching, which focuses primarily on inconsistencies in individual data elements across two or more databases, is not the same thing as data mining. The more sophisticated data mining technologies allow an organization to create a model using statistical and mathematical algorithms to predict future fraud patterns based on a wide array of data elements and patterns.
- Use a scalable, massively parallel processing (MPP) relational database system for the integration, aggregation, and preparation of the data. This will help the organization to explore large data sources and better exploit social relationships within the predictive modeling framework without sacrificing performance.

- Grant applications provide a rich source of unstructured textual data. Use content/text analysis to identify untrustworthy patterns and indicators of fraud.
- Do not overly rely on alert systems based on statistical outliers; they produce too many false alarms.

### *Technology Strategy*

- Use multiple technology approaches so you can be flexible to changing requirements.
- Investing in fraud prevention/detection technology is not expensive and is highly likely to have a good return on investment.

### **Performance Metrics**

Feedback collected on performance metrics generally focused on three main topics: measuring outcomes, measuring outputs, and monitoring performance. In addition to the ideas listed below, Dialogue participants and interviewees emphasized the importance of establishing explicit, measurable outcome and output goals in advance of funding projects.

#### *Measuring Outcomes*

- When applying performance metrics to spending, measure the value added by the project, not just the outputs produced.
- Measure outcomes at the stakeholder level to help identify fraud among the actual stimulus recipients, e.g. wage theft.

#### *Measuring Outputs*

- Include measures of timeliness, cost overruns, and sufficiency of delivered product.
- For infrastructure projects, have clear regulatory compliance and pollution prevention metrics.

#### *Monitoring Performance*

- Make annual Recovery Board audits mandatory to reduce improper payments.
- Withhold a portion of allocated funds until after a demonstration of performance.
- Upload software in every agency's accounts payable system.



- Require periodic reports of negative trends/failures.

## **Organizational Awareness and Strategy**

The Dialogue generated few ideas directly related to management practices that could help mitigate fraud, waste and abuse. However, between the Dialogue and the interviews conducted, observations and ideas emerged related to organizational awareness and strategy that could enhance accountability. Some of the ideas offered are for government-wide implementation and not within the domain of the Recovery Board.

### *Organizational Awareness*

- Information silos and “turf wars” discourage good decision-making and make it easier for fraudsters to succeed. Foster a culture of collaboration.
- To engage people in the fight against fraud, waste and abuse, they need to understand the associated costs and “buy in” to its prevention. The risks must be presented in a factual, credible way that is related to an individual’s domain and clearly identifies the risk for exposure.
- Because fraudsters modify their behaviors quickly, organizations need to be able to respond rapidly to prevent and detect fraud.

### *Organizational Strategies*

- Fraud typically occurs with a systemic or management error that is then exploited by fraudsters. Management needs to be agile and employ timely “closed loops” processes to minimize the opportunities for fraud. This means that when fraud is detected, organizations must change vulnerable business processes, add required controls, make the fraud visible, and systematize continuous improvements.
- While there are systems in place by which federal agencies collect monies owed, they are often cumbersome, slow, and of lower priority. Consider making information public about the amount of money uncollected by each agency to increase accountability and timeliness.
- Mirror the private sector and evaluate every transaction for validity and risk; focusing on small statistical samples is insufficient.
- Approach the use of PII and related legal constraints from a different perspective; focus not on the individual information but on the valuable aggregate patterns that can be derived.

**THE PREVENTION/DETECTION CONTINUUM:  
AN ANALYTIC FRAMEWORK**

In addition to the above insights, the Panel believes that looking at the prevention/detection continuum in an analytic framework is most helpful in identifying next steps, possible gaps, and opportunities for future Recovery Board investment.

According to experts in the field, analytics fall into two basic categories: descriptive and predictive.

Table 4<sup>7</sup> depicts the framework, with the six descriptive and three predictive levels. It also provides selected examples (in Levels 1-5) of current Recovery Board activities that might fall within each level; highlighted areas are new areas of opportunity (in Levels 6-9) for the Recovery Board to explore as it enhances its efforts.

**TABLE 4 – ANALYTIC FRAMEWORK LEVELS AND EXAMPLES**

<b>Recovery Board Activity Examples/Potential Opportunities<sup>8</sup></b>		
<b>Descriptive Techniques, Level 1-6</b>		
1	Standard Reporting	Manual review of daily postings of Contracts in FedBizOpps.gov; review of Central Contractor Registration/DUNS number; Excluded Parties List System; data matching.
2	Custom Reporting (Excel)	DOJ database review of Fraud Hotline complaints and recurring issues; data slicing and dicing.
3	Queries/drilldowns (SQL/OLAP)	Search of non-fixed price and non-competitive contracts by agency, state, and contract type; identification of recurring issues.
4	Dashboards/alerts	Procurement data system monitoring; business intelligence; link analysis; visualization of data; fraud mapping; GIS.
5	Statistical Analysis	Recovery Operations Center identifies non-obvious relationships, risk factors, and leads for investigations/audits.
6	Clustering (unsupervised learning)	Advanced analytic techniques including “mining” networks of connected individuals to score them for potential fraud, e.g. to determine how many fundamentally different types of behaviors are in the data and what they generally look like.
<b>Predictive Techniques, Levels 7-9</b>		
7	Predictive Modeling	Advanced data mining technique using statistical and mathematical algorithms, such as neural networks, decision trees, nearest neighbors, linear and logistic regression, etc. This kind of modeling could help determine which contracts are most likely to be fraudulent.

<sup>7</sup> Elder Research, Inc., “Using Data Mining to Detect and Prevent Fraud, Waste, and Abuse,” November 1, 2011 presentation to the National Academy of Public Administration, and inspired by “Eight Level of Analytics,” by SAS Institute, Inc.

<sup>8</sup> Light gray shading indicates potential opportunities for Recovery Board analysis.

8	Optimization & Simulation	Extend modeling functions through creation of “what-if” scenarios; examine and test decisions prior to making them; include elements of uncertainty and variability in forecasts of process performance. This kind of analysis could, for example, help determine what number of investigators to put on each case to maximize expected return.
9	Next Generation Analytics: Text Mining & Link Analysis; Ensemble Modeling	Plans for Textual Analysis of PDF files (e.g. contract/grant applications/audits); advanced text mining, often called Natural Language Processing (NLP) or Statistical NLP, which uses more robust software, a knowledge of the language syntax/structure/grammar, as well as statistical information derived from the text. This level of analysis could help reveal a coordinated set of people likely to be a fraud ring.

**This page intentionally left blank.**

## **SECTION II. RECOVERY BOARD DATA CHALLENGES**

During the course of the Recovery Board Dialogue effort, the Academy became aware of three major challenges related to data collection, analysis, and utilization:

1. Lack of a single alpha numeric numbering system for federal contracts and grants;
2. Legal barriers to using federal tax delinquency history to limit eligibility for receipt of federal contracts and grants; and
3. Challenges associated with Personally Identifiable Information (PII).

The Panel believes each of these warrants further discussion.

### **A UNIFORM NUMBERING SYSTEM FOR FEDERAL CONTRACTS AND GRANTS**

The Recovery Board effort to increase transparency and accountability in federal spending brought to light a challenging federal contracts management organizational scheme. For example, each of the 29 federal agencies dispersing Recovery money has its own contracts/grants numbering system. Some have letters; some a mix of letters and numbers. Some use semi-colons; some do not. Each has its own syntax, rationale, and historical records. There is thus no uniform schema for identifying recipient organizations. These disparate systems and organizational frameworks make it far more time-consuming and confusing than it needs to be and stymie the analysis of the relationship between contracts, contractors, performance, and delinquency. It took the Recovery Board a year to identify clearly the organizations receiving Recovery Act money. In its effort to untangle the data, the Recovery Board found 16,000 data mismatches in the varying contract tracking systems. There is currently a “harmonization” feasibility system ongoing.<sup>9</sup> There is likely to be resistance to such an effort, which no doubt will be costly and time-consuming. Absent a universal federal contract tracking mechanism, however, efforts to prevent and detect fraud, waste, and abuse in federal contracts/grants arena will be less timely, less effective, and more costly.

### **FEDERAL TAX DELINQUENCY HISTORY**

In April 2011, the U.S. Government Accountability Office (GAO)<sup>10</sup> found that at least 3,700 Recovery Act contract and grant recipients were estimated to owe more than \$750 million in known unpaid federal taxes as of September 2009. These tax delinquent entities nevertheless received over \$24 billion in Recovery Act funds. Current law does not prohibit the awarding of contracts or grants to entities because they owe federal taxes. Section 6103, the confidentiality provision of the Internal Revenue Service (IRS) Code,

---

<sup>9</sup> The Department of Defense currently has a single system for numbering its contracts, which account for roughly two-thirds of federal contracting dollars. It is possible that a government-wide system may be able to build upon this existing system.

<sup>10</sup> GAO-11-485, “Recovery Act: Thousands of Recovery Act Contract and Grant Recipients Owe Hundreds of Millions in Federal Taxes,” p.1., April 2011.

protects tax information from disclosure to other parties except under very limited, defined circumstances. These exceptions do not allow the IRS to share delinquency information with other federal agencies unless the taxpayer consents. Some federal agencies, such as the Small Business Administration and the Veterans Administration, have required that applicants seeking eligibility for specific programs waive their Section 6103 rights and provide access to their federal tax history. The Recovery Board did not elect this option. This lack of access to taxpayer data hinders the Recovery Board's ability to oversee Recovery Act spending.

In an April 4, 2011 letter to GAO, John Higgins, the Recovery Board's Director of Accountability, explained further: "We could also utilize the unpaid tax data and a host of other risk-relevant data to create a risk-based model upon which government agencies could rely in making their own expenditure determinations. Through such a proactive approach, the Recovery Board could further engage its tools to prevent fraud and waste of government funds, rather than merely detect problems after they occur."

## **PERSONALLY IDENTIFIABLE INFORMATION**

OMB defines Personally Identifiable Information (PII) as "information which can be used to distinguish or trace an individual's identity, such as their name, Social Security Number, biometric records, etc. alone or when combined with other personal or identifying information which is linked or linkable to a specific individual date and place of birth or mother's maiden name, etc."<sup>11</sup> Because federal agencies maintain significant amounts of potentially personal information, the government has a special duty to protect that information from loss and misuse. The Privacy Act requires that each agency establish: (1) rules of conduct for those maintaining any system of records, particularly those containing PII; and (2) appropriate administrative, technical, and physical safeguards to ensure their security and confidentiality.<sup>12</sup>

The Recovery Board, like all federal agencies, maintains voluminous systems of records and is very much aware of the legal and regulatory requirements surrounding these systems of records and the penalties associated with violation. As the Recovery Board seeks to use a broader array of data from a variety of sources and to link and aggregate that data, it faces multiple PII challenges. Among those challenges is the perennial need to balance personal information safeguards and the government's legitimate need to access that data to perform its oversight and law enforcement functions. In striking this balance, government officials are less likely to collect the data in the first place and more cautious about sharing the data with federal partners. Federal database managers and risk modelers face the additional challenge of conveying to agency counterparts the important distinctions between anonymous, aggregated data, with its powerful predictive value, and protected PII. Successfully overcoming this hurdle and coming to a shared understanding are critical to improvements in the management and analysis of federal data.

---

<sup>11</sup> See M-07-16, Letter from Clay Johnson, III, Deputy Director for Management, Office of Management and Budget, "Safeguarding Against and Responding to the Breach of Personally Identifiable Information," June 22, 2007.

<sup>12</sup> See Privacy Act of 1974, 5USC§552 as amended.

**The Panel will offer recommendations with regard to each of these data challenges in the following section.**

**This page intentionally left blank.**



## **SECTION III. RECOMMENDATIONS AND CONCLUDING REMARKS**

This section offers two sets of Panel recommendations: (1) on issues warranting further exploration by the Recovery Board and (2) on issues of government-wide significance, for consideration by the Congress or the Administration, as appropriate for broader adoption.

### **RECOMMENDATIONS FOR THE RECOVERY BOARD**

Based on feedback received in the Dialogue, solutions submitted by participants, Academy research, and insights from supplementary interviews and discussions with experts and governmental officials, the Academy Panel recommends that the Recovery Board pursue the following ideas which surfaced during the course of this effort.

---

#### **RECOMMENDATION 1**

---

#### **RECOMMENDATION 1 – Increase emphasis on predictive analysis, particularly to prevent and detect contract/grant fraud.**

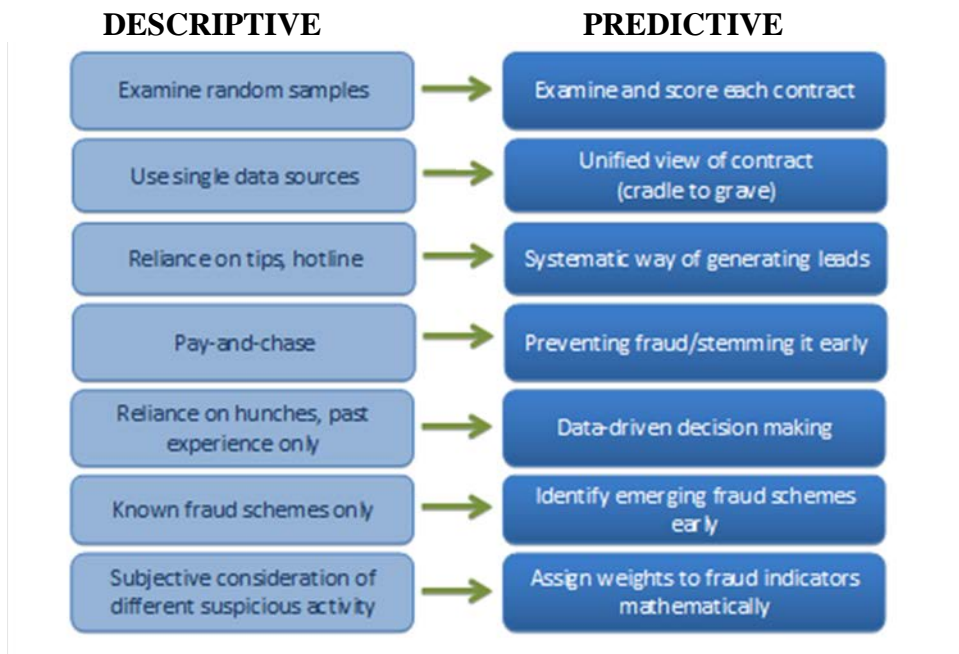
The Recovery Board’s efforts in the field of prevention and detection of fraud, waste, and abuse have served as a laboratory for government efforts to increase transparency and accountability in government spending—particularly funding for contracts and grants. The Recovery Board’s Recovery Operations Center (ROC), designed with the best technology available in 2009, when it became operational, uses a risk analysis model and a number of predictive analysis tools on a wide range of variables to identify risk areas that might be susceptible to fraud or waste and to identify non-obvious relationships. The Panel believes that the Recovery Board can enhance its already aggressive efforts by investing further in predictive analysis and focusing on systems, tools, and technologies that will easily translate to other federal procurements — an area ripe for additional oversight because of its complexity, volume, and enormous economic impact.

Figure 1<sup>13</sup> describes the benefits of using predictive versus descriptive analytics for fraud investigations and differentiates between the use of random samples, single data sources, reliance on tips and hotlines, hunches and past experience, known fraud schemes, and more subjective consideration of suspicious activity versus a more thorough, predictive approach based on risk models, model monitoring, and recalibration, data mining, and link/network analysis.

---

<sup>13</sup> Source: Elder Research, Inc., “Using Data Mining to Detect and Prevent Fraud, Waste, Abuse,” Presentation to the National Academy of Public Administration, November 1, 2011.

**FIGURE 1 – BENEFITS OF USING PREDICTIVE VS. DESCRIPTIVE ANALYTICS FOR CONTRACT FRAUD INVESTIGATIONS**



While federal agencies focused primarily on payments to individuals have been making headway in preventing and detecting high-dollar value fraud, waste, and abuse, the Recovery Board has focused on preventing and detecting fraud, waste, and abuse in the less frequently quantified area of funding of contracts and grants. The experience of these past few years has put the Recovery Board in a position to transfer the lessons it has learned to the rest of the federal sector and facilitate enhanced financial stewardship.

Based on private and public sector parallels, the Panel believes that this investment<sup>14</sup> will likely have a high return. For example, an organization that used basic analytic tools typically found one case of actual fraud for every red flag alert; in elevating the organizational level of analytics and using predictive techniques, such as non-linear algorithms and decision trees, it found 25 cases of actual fraud for every 100 alerts. Another major corporation recovered \$20 million in the first nine months of an anti-fraud data analysis. In another example, an insurance company used data mining techniques to look for patterns in data that identified previously unseen fraud schemes. In one year, the company uncovered more than \$9 million, for a roughly 400 percent increase in the amount of recovered fraud dollars.<sup>15</sup>

<sup>14</sup> The Panel did not attempt to determine the actual cost of purchasing analytic services or technologies. However, in the course of interviews, various providers and users roughly estimated the cost for a six to twelve month period as being no more than \$200,000 to \$300,000 for sophisticated technologies with support services.

<sup>15</sup> Teradata Corporation, “Building a Fraud Analytics Ecosystem: Critical Success Factors,” November 2011, p.7.

According to service providers, public sector investment has had similar results. During the 2009-2010 tax year the Australian Tax Office combined a range of analytical and expert business models to identify a variety of frauds; they reviewed and assessed \$2.36 million individual returns, stopped 26,000 for review, and protected \$74 million in refunds.<sup>16</sup> The State of Texas tax department increased its productivity by 37 percent and recovered \$800 million in revenue after implementing a system to identify those with underpaid taxes. Missouri recovered \$100 million in a similar initiative. The Internal Revenue Service (IRS) has recovered hundreds of millions over the course of a three-year effort to root out non-compliant taxpayers.

In addition to preventing fraud, predictive analysis has been shown to be a valuable tool in facilitating the rapid approval/processing of requests by those entities deemed to be of low risk. For example, the Social Security Administration attributes the elimination of a two-year back-log in applications for disability insurance claims to streamlined analysis of potentially fraudulent and meritorious applications. The Recovery Board, therefore, has an opportunity to set a federal standard for identifying low as well as high risk applicants and speeding delivery of government services to meritorious applicants. Cutting improper payments/awards will increase federal productivity and response times for meritorious, low- risk applicants.

---

## RECOMMENDATION 2

---

### **RECOMMENDATION 2 – Increase use of sophisticated textual analysis tools to mine the abundance of narrative information that is unstructured.**

As part of this increased emphasis on predictive analysis, the Recovery Board should focus particularly on automated textual analysis and mining of unstructured data. These Natural Language Processing (NLP) or Statistical NLP tools have the potential to streamline bureaucratic processes as well as prevent and detect fraud. Forensic analytic experts estimate that 80 percent of data is unstructured. New tools are available that will automate the scanning of lengthy government documents replete with this unstructured, semi-structured, as well as more standard structured data in rows and columns. The tools can convert free-form text into relational tables and fuse this data with structured data. For example, one Dialogue participant suggested that using such textual analysis on written materials perfected by Bernie Madoff in his now-known Ponzi scheme might help modelers identify future frauds. These tools even have the capability to find grammatical patterns and risk indicators in languages other than English. The Panel recommends that the Recovery Board build upon its current efforts by investing in such tools and

---

<sup>16</sup> Teradata Corporation, “Building a Fraud Analytics Ecosystem: Critical Success Factors,” November 2011, p.7.

technologies to identify red flags in contract/grant applications and build models to predict those applications with highest risk and those with least risk.<sup>17</sup>

---

### RECOMMENDATION 3

---

#### **RECOMMENDATION 3 – Increase data sources, particularly state and local governmental data and proprietary business data, to improve data validation.**

Dialogue participants identified a range of additional data sources that warrant serious exploration by the Recovery Board. Specifically, the Panel believes that state and local databases hold great promise; web-scraping tools are available to pull quality state and local data into the Recovery Board net. Given that some \$280 billion of Recovery Act funds are administered through state and local governments<sup>18</sup> and that the states are the largest reporters under the Recovery Act,<sup>19</sup> it makes sense to partner with them in this accountability effort.

Likewise, proprietary databases with company-specific information on business locations, industry sub-categories (North American Industry Classification System or NAICS codes), numbers and types of employees, salaries, revenue, and technological investment can add great texture and meaning in risk modeling. This additional data has the potential to enhance the predictive modeling capability and help the Recovery Board determine if an applicant is likely to be able to perform promised contractual obligations and deliver a product. Furthermore, the Panel believes that there is promise in private industry volunteered data. For example, the banking industry of its own accord agreed to provide the federal government with information on payroll deposits to help track illegitimate unemployment insurance claims. The Panel believes that the Recovery Board should actively seek out other similar partnerships. According to those in the banking/financial community, two areas that typically provide huge opportunities for fraud detection are: (1) detailed transactional financial histories and (2) data sources that identify individuals who have fallen off the grid, who may have relocated, died, or gone underground to avoid payment of debts. As a cautionary note, governmental use of proprietary databases will likely require the establishment of a “Chinese data wall” to ensure that the government is not inappropriately in possession of proprietary data and

---

<sup>17</sup> Dialogue participants noted that necessary insights often require a combination of tools. While the Panel is not endorsing any specific tools, participants cited some tools as worthy of consideration for unstructured analysis, iConnex and Endeca. In addition, participants mentioned IDEA and ACL for forensic accounting; Clementine, Darwin, and S-Plus for data mining; and MapReduce as a framework for exploring large data sources that have not been fully cleansed/structured. In addition, there are new tools emerging that enable imagery and video data mining, such as Enterprise Speech Intelligence (ESI), Insightful Miner, and PolyAnalyst.

<sup>18</sup> Government Accountability Office, Letter from Gene Dodaro to The Honorable Mitch McConnell, November 30, 2009. GAO-10-216R, Recovery Act.

<sup>19</sup> See testimony of the Honorable Earl E. Devaney before the Committee on Oversight and Government Reform, United States House of Representatives, June 14, 2011; [http://www.recovery.gov/About/board/Documents/Devaney\\_Testimony\\_06142011.pdf](http://www.recovery.gov/About/board/Documents/Devaney_Testimony_06142011.pdf).

that use of such data is consistent with federal privacy laws. This may argue for third-party analysis of such non-public databases.

---

## RECOMMENDATION 4

---

### **RECOMMENDATION 4 – Work across government to establish and publicize more consistent performance metrics for fund recipients and increase transparency of outcomes for tax dollars spent.**

Dialogue participant discussion on the need for increased performance metrics resonated with the Panel. While federal contracts typically spell out specific deliverables and deadlines, the Panel believes that contracts do not consistently capture basic performance metrics. Even contract deliverables get buried in a sea of paperwork and are not routinely shared with the public. The Panel believes that the Recovery Board should work across government to capture and publicize the basic information (e.g. product timeliness, product delivery, regulatory compliance, and cost overruns) in a consistent format and require the establishment of outcome goals and tangibles for each funding recipient.

Consistent with increasing transparency and accountability, the establishment of more consistent, quantifiable measures for each fund recipient will also facilitate prevention and detection of fraud, waste, and abuse. Recent media investigations into improper payments in federal housing programs highlight the need for improved data verification and management.<sup>20</sup> While Recovery.gov currently pinpoints the name and geographic location of funding recipients, provides project descriptions, and amount of funding, the Panel believes that requiring consistent output performance metrics and relevant project documentation, such as permits, would enable the Recovery Board to more easily communicate these results. This transparency would enhance the public's understanding of the added value of each project and facilitate public reporting of inconsistencies.

---

<sup>20</sup> See "Finding more flaws in HUD's accounting: 'Unreliable' data, larger-scale problems for housing program," *The Washington Post*, November 7, 2011, p.1.

## **RECOMMENDATIONS FOR GOVERNMENT-WIDE CONSIDERATION**

The Academy's primary charge was to host the Dialogue, conduct related interviews and research, and make recommendations to the Recovery Board. While performing these tasks, however, the Panel identified a number of issues of government-wide significance and here offers its views and the following recommendations for consideration of the Executive/Legislative branches. The Panel recognizes that these recommendations raise implementation issues which will require further study.

---

### **RECOMMENDATION 5**

---

#### **RECOMMENDATION 5 – Consider establishing a permanent, centralized portal for data to enhance federal data management and analysis.**

While federal agencies have legitimate legal and mission concerns they nonetheless also tend to be territorial in the ownership and use of their data systems. The Panel believes that if government is to enhance its fraud prevention and detection efforts and maximize the use of federal resources, the time has come to centralize the data gathering and analysis of such data. Leading-edge practitioners endorse such efforts that minimize data movement and therefore decrease time and save costs.<sup>21</sup> The financial community has long used such data networks to authenticate the identity of purchasers and authorize transactions and as the foundational element of their fraud prevention and detection program.

The Panel believes that the establishment of such a federal clearinghouse should be grounded in three basic tenets of good government: consistency, coordination, and correctness.

With a single data portal, the opportunity for the collection of consistent data would be greatly enhanced and the data and predictive models would therefore be transferable to a wider range of federal entities. Creating a government standard for consistent data would minimize the time, effort, and financial resources associated with data preparation.

Increased coordination also makes practical sense. Several existing governmental entities might be able to take this function on and might seek to do so. The Panel, however, believes that the optimal home for this function is in an independent entity with permanent professional staff with expertise in analysis of learned behaviors. While this consolidation will likely create efficiencies of scale not possible under the current construct, the Panel is not advocating the establishment of another major administrative institution, but rather a small clearinghouse or coordinating unit that recognizes and respects agency specific needs for customized reports and retention of agency-specific

---

<sup>21</sup> Teradata Corporation, "Building a Fraud Analytics Ecosystem: Critical Success Factors," p. 1. November 2011.

sensitive data. Each data set requires agency expertise for full context. The Panel believes that the current state, with a multiplicity of essentially uncoordinated agency data systems, stymies efforts to predict fraud patterns; and that, conversely, the integration of data systems and establishment of a central portal will facilitate the development of accurate and cost effective predictive models that will thwart fraud and otherwise repetitive patterns.<sup>22</sup>

Data quality is critical to building efficient predictive models. The Recovery Board's scrubbing and validation of recipient-provided data is a prime example of the value of such an investment. Cleaning and preparing data often take up to 70 to 80 percent of a typical data mining project.<sup>23</sup> Experts focus on identifying and correcting missing values, poor/incorrect data entry, and reconciling disparate data sources. The Panel believes that having a centralized and dedicated resource will facilitate cleaner, more usable federal data and be more cost-efficient than having each agency examine its own data independently. It will also facilitate the pre-payment validation of data used in federal decision-making, including pre-screening for program eligibility.<sup>24</sup>

---

## RECOMMENDATION 6

---

### **RECOMMENDATION 6 – Evaluate ways to expedite the sharing of aggregated federal data to enhance federal predictive modeling.**

The Panel believes that as federal entities come to a fuller understanding of the power of predictive modeling, the government should encourage the sharing of aggregated data from the widest array of federal sources. The Recovery Board has made great strides in this area by sharing its fraud prevention and detection insights with its federal partners. The information has not always flowed as freely in the reverse.

In building a better predictive model to prevent and detect fraud, waste, and abuse, federal officials do not need individually-specific data or PII; rather, modelers make the data anonymous—by assigning unique, consistent identifiers to individual data and scrambling it to see the aggregated patterns. The Panel recognizes and appreciates the value of agency-owned and operated databases and the need for database owners to

---

<sup>22</sup> In February 2010, the General Services Administration announced a continuing effort to make the federal acquisition process more efficient and transparent and awarded a \$74.4 million contract to consolidate nine acquisition databases that track pre-and post-award contract data across the entire federal civilian and Department of Defense acquisition communities. Lessons learned from this attempt should inform future data portal efforts. See <http://www.gsa.gov/portal/content/104150>

<sup>23</sup> Elder Research, Inc. October 24, 2011, "National Dialogue on Innovative Tools to Prevent and Detect Fraud, Waste, and Abuse," p.2.

<sup>24</sup> For example, GAO found that \$17 million in Recovery Act money went to ineligible 8(a) firms who misrepresented their eligibility for the program. See *8 (a) Program: The Importance of Effective Fraud Prevention Controls*; GAO-11-440T, March 3, 2011.

customize data and protect its integrity, but believes that the sharing of aggregated federal data is critical to enhanced financial stewardship and accountability.

---

## RECOMMENDATION 7

---

### **RECOMMENDATION 7 – Consider establishing a uniform system for identifying federal contracts and grants to improve tracking of federal payments to recipients.**

The Panel believes that the federal government needs a single organizational scheme for tracking its awards and procurements if it is to improve its ability to track federal payments, funding recipients, and reduce improper payments. With much history here to undo and bureaucratic hurdles and resistance more than likely, the Panel acknowledges that implementation of a change of this magnitude will require judicious evaluation of the current pilot effort and years beyond to effect.

Nevertheless, the Recovery Board’s experience with rampant data mismatches makes a compelling argument for change. Under the current construct, with each agency having its own disparate system, it is extremely difficult to unearth fraud, waste, and abuse in contracts and grants. Organizations that strive to rely on data-driven decision-making cannot operate effectively with mismatched data.

---

## RECOMMENDATION 8

---

### **RECOMMENDATION 8 – Explore regulatory changes to require applicants for federal funding to sign a waiver allowing access to their tax records.**

Acknowledging the Panel’s limited scope and analysis of this difficult issue, the Panel nevertheless believes that federal accountability would be well served if applicants for federal funding were consistently required to sign a waiver allowing access to their tax records. While individual agency programs can invoke such a requirement by giving legal notice to applicants, and a few have done so,<sup>25</sup> creating such a requirement for all federal contracts, grants, and cooperative agreements would require changes to the Federal Acquisition Regulations (FAR) for contracts and to OMB Circulars for grants and cooperative agreements. While this is not a new topic, but rather one hotly debated and reported upon, the Panel believes that requiring this waiver as part of the application process is a common sense and appropriate solution; it is forthright and puts applicants seeking federal funds on notice that relevant tax information will be taken into account in

---

<sup>25</sup> Among agencies requiring applicants for selected programs to allow access to their IRS records are the Small Business Administration (8a program) and the Veterans’ Administration (Vet Business Verification).



evaluating risk. It has the added benefit of enabling the government to consider the resultant tax delinquency data in its risk modeling.

## **CONCLUDING REMARKS**

In initiating this Dialogue, the Recovery Board has evidenced its commitment to continuous improvement of its data-driven decision-making. The Panel believes that by implementing these recommendations, expanding its data sources, and investing in some of the newest analytic tools and strategies used by the private sector, the Recovery Board will enhance its already effective efforts in the prevention and detection of fraud, waste, and abuse and continue to raise the standard for federal accountability and transparency. Moreover, the Panel believes that governmental exploration of the broader, second set of recommendations is important in continuing to improve federal financial stewardship and management practices. In combination, these actions have the potential to further reduce improper federal payments and put more taxpayer dollars to their intended use.

**This page intentionally left blank.**

## APPENDIX A PANEL AND STAFF BIOS

### PANEL

**Alan R. Shark, D.P.A.\*** *Chair* – Executive Director, Public Technology Institute. Former President and CEO, American Mobile Telecommunications Association; Associate Executive Director, Marketing & Communications, Water Environment Federation; Director of Marketing, North American Telecommunications Association; Vice President for Marketing and Communications, American Resort Development Association; Vice President for Marketing, Voice Computer Technologies Corporation; Director of Research and Information Services, National School Boards Association; Director of Programs, Association of Governing Boards of Universities and Colleges; Coordinator, State and Organizational Relations, American Association of State Colleges and Universities; Seabees, U.S. Navy, Vietnam service.

**Frank J. Chellino\*** – President and Chief Executive Officer, Langley-Hunt International, Inc. Former positions with U.S. Drug Enforcement Administration: Special Agent in Charge, Miami Division Office; Special Agent in Charge, Washington Division Office; Deputy Assistant Administrator, Office of Inspections; Assistant Special Agent in Charge, Washington Division Office; Supervisory Senior Inspector, Office of Professional Responsibility; Unit Chief, Office of Security Programs.

**Franklin S. Reeder.\*** – President, The Reeder Group. Former Director, Office of Administration, The White House. Former positions with U.S. Office of Management and Budget: Deputy Associate Director for Veterans Affairs and Personnel; Assistant Director for General Management and Deputy Assistant Director; Chief, Deputy Chief, Information Policy Branch; Policy Analyst; Chief, Systems Development Branch. Former Deputy Director, House Information Systems, Committee Staff, Committee on House Administration, U.S. House of Representatives. Former positions with U.S. Department of the Treasury and U.S. Department of Defense focusing on information technology and systems.

*\*Academy Fellow*

### STAFF

**Joseph P. Mitchell, Ph.D.,** *Director of Project Development* – Joe Mitchell is responsible for managing the Academy's project development process to ensure that all studies are designed appropriately from a methodological, budgetary, and personnel standpoint; providing highly skilled technical advice and expertise on proposals, concept papers, and project deliverables; identifying opportunities for the Academy to improve the functioning of governments at all levels and advancing the organization's strategic agenda; and assessing internal competency/skill needs and developing strategies to address these issues. Dr. Mitchell also serves as a trusted advisor to the Academy President and CEO, Fellows, and clients. Dr. Mitchell received a Ph.D. in Public

Administration from the Virginia Polytechnic Institute and State University, a Master of Public Administration from the University of North Carolina at Charlotte, and a Bachelor of Arts from the University of North Carolina at Wilmington.

**Laurie J. May, *Project Director*** – Laurie May has directed complex studies—many mandated by Congress—for the National Academy of Public Administration as well as other organizations since 2004. The Academy studies have focused on various management and other issues at organizations such as the U.S. Patent and Trademark Office, the Centers for Disease Control and Prevention, the National Institutes of Health, the National Aeronautics and Space Administration, and the White House’s Office of National Drug Control Policy. From 2009 through 2011, Ms. May directed a market sizing study, commissioned by the U.S. Department of Transportation Research and Innovative Technology Administration (RITA), to develop accurate and comprehensive estimates of the breadth and size of the U.S. and North American Intelligent Transportation Systems (ITS) markets. Prior to her time at the Academy, Ms. May served as the Director of the Organizational Management and Integrity Staff at the U.S. Environmental Protection Agency (EPA). As a senior EPA program management official, Ms. May directed a staff and provided organizational leadership and policy direction for the full range of management issues. She has served as a confidential management advisor to numerous Presidential appointees across EPA and is the recipient of EPA's Excellence in Management Award. Ms. May is a Phi Beta Kappa graduate of Duke University.

**Faith Gibson, *Research Associate*** – Faith Gibson has been with the Academy since June 2011. Prior to joining the Academy staff, Faith worked in the public education and non-profit sectors. She has research experience and interests in the areas of policy formation and implementation, civic engagement and participation, program evaluation and education policy. Ms. Gibson received her Master of Public Administration degree from the College of Business and Public Administration at Old Dominion University in Norfolk, Virginia, and a Bachelor of Arts degree in Communications/Public Relations from Georgia State University in Atlanta, Georgia. Currently, Ms. Gibson is working on her Ph.D. in Public Administration and Policy at Virginia Polytechnic Institute and State University’s Center for Public Administration and Policy at the National Capital Region Campus in Alexandria, Virginia.

**Matthew Thomas, *Research Associate*** – Matt Thomas has served as a Research Associate on past Academy studies for the Department of Homeland Security, the General Services Administration, the Department of Energy, the Coalition to End Childhood Lead Poisoning, and Amtrak. Prior to joining the Academy, he served as an administrative staff assistant for LogiCom Project Management and the American Association of Naturopathic Physicians. Mr. Thomas holds a BA in Political Science from Tulane University.

## **APPENDIX B INDIVIDUALS INTERVIEWED**

Gregor Bailar, former CIO Capital One Financial Group and former CIO NASDAQ

Stanley J. Czerwinski, Director, Intergovernmental Relations, U. S. Government Accountability Office

Rick Eng, Solutions Manager, Eastport Analytics

Bill Franks, Chief Analytics Officer, Global SAS Program Teradata Government Systems

Isaiah Goodall, Director, Washington DC Office, Elder Research Inc.

Gregory D. Kutz, Director of Audit Services, Forensic Audits and Investigative Services, U. S. Government Accountability Office

John Larson, Vice-President, Public Sector Consulting Unit, IHS Global Insight

Antonia de Medinaceli, Director, Business Analytics and Fraud Detection, Elder Research, Inc.

Christopher Mihm, Managing Director of Strategic Issues, U. S. Government Accountability Office

Dermot O'Sullivan, Federal Civilian Account Representative, Teradata Government Systems

Carol Patey, Assistant Director, U. S. Government Accountability Office

Joseph Pika, Policy Analyst, Office of Federal Financial Management, U. S. Office of Management and Budget

Mark Williams, Founder/Director, Eastport Analytics

William Woods, Director, Acquisition and Sourcing Management, U. S. Government Accountability Office

Paul Wormeli, Executive Director Emeritus, IJIS

**This page intentionally left blank.**

## **APPENDIX C**

### **DIALOGUE METHODOLOGY AND METRICS**

In addition to the substantive findings discussed earlier in this report, the Dialogue generated valuable lessons about the process of online stakeholder consultation. Based on key metrics, this appendix provides important lessons learned from the planning and execution of the Dialogue, as well as analysis of web traffic and participation.

#### **Planning and Execution of the Dialogue**

In the weeks before the Dialogue's launch, the Academy worked with the Recovery Board to understand its goals and objectives for the Dialogue and translate them into meaningful content for the website. Synteractive, the provider of the online platform, also participated in this process to ensure that site content was aligned with the structure of the platform and that Dialogue site functionality met the needs of the engagement. Simultaneously, the Academy worked with the Recovery Board to identify stakeholder communities that could provide valuable feedback and developed a strategy to communicate to them the value of participating.

#### ***Content Development***

The first step in the Dialogue planning process was to develop a clear statement of purpose that would guide content development and outreach messaging for the initiative. At the beginning of the project, the Recovery Board indicated its desire to use the Dialogue platform to collect ideas on specific tools and strategies to prevent and detect fraud, waste and abuse. In initial discussions, the Academy proposed focusing the Dialogue on tools and strategies to prevent and detect fraud, waste, and abuse government-wide; however, the Recovery Board believed this approach would go beyond its mandate and preferred to focus the initiative on tools and strategies to enhance accountability of Recovery Act funding exclusively. The text from this statement of purpose is available in Appendix D.

Concurrent with the development of the Dialogue statement of purpose, the Academy worked with the Recovery Board to brainstorm the key topics on which it wanted to solicit input from participants. After some discussion, the Academy and the Recovery Board identified five types of tools and strategies for preventing and detecting fraud, waste and abuse around which it wanted to solicit substantive ideas.

- Management
- Data Sources
- Technologies
- Risk Models
- Performance Metrics

To encourage participant submission of feedback on all five topics, each topic had its own assigned discussion forum within the Dialogue. In addition, the Academy created a

tailored prompt question for each forum to help participants understand the type of idea they were being asked to contribute. These prompt questions included:

- What management practices, policies, programs, and incentives would improve financial stewardship and help prevent and detect fraud, waste, and abuse?
- What specific governmental, public, or proprietary data sources could help the Recovery Board prevent and detect fraud, waste, and abuse?
- What technologies or systems do you think would be effective in integrating and aggregating diverse types of data?
- What types of risk models would identify entities receiving Recovery Act funds as most vulnerable to fraud, waste and abuse?
- What types of performance metrics could be applied to Recovery Act funded programs and recipients to increase oversight and accountability?

During the initial planning phase of the Dialogue, the Academy also collaborated with the Recovery Board to generate additional content that would encourage and inform Dialogue participants. Prior to the Dialogue going live, the Academy and the Recovery Board worked together to develop homepage text that would communicate the value of the Dialogue and encourage the submission of substantive ideas. The final version of the homepage text that appeared on the Dialogue is excerpted as follows:

*In its continuing efforts to develop innovative fraud prevention tools, the Recovery Accountability and Transparency Board is hosting an online dialogue. While the scope of the dialogue is oversight of Recovery Act funds, your comments and ideas related to oversight of all federal spending are critical to identifying the best management practices, data sources, technologies and systems, risk models, and performance metrics to improve oversight of Recovery Act spending.*

In addition, the Academy worked with the Recovery Board to develop Terms of Use and privacy and moderation policies to ensure participants were fully informed of how their feedback would be used. These three documents can be found in Appendix D.

### ***Platform Customization***

To host the Dialogue, the Academy used Social Rally, a cloud-based, software-as-a-service, crowd sourcing application created by Synteractive. The Dialogue platform, which was hosted at [www.FedAccountabilityDialogue.org](http://www.FedAccountabilityDialogue.org), allowed participants to submit and tag “Ideas,” comment on the ideas of others, and vote the best submissions to the top. Visitors to the site who wished to participate were required to register an account, which involved creating a custom username and password, and providing an email address.



While the Social Rally platform was especially well-suited to hosting this type of online engagement, it required some platform customization. Most of these customizations were aesthetic in nature; however, one significant modification requested by the Recovery Board was the addition of a mechanism for Dialogue participants to submit ideas directly. In partnering with the Academy to conduct this Dialogue, the goal of the Recovery Board had been to collect specific tools and strategies that it could use to prevent and detect fraud, waste and abuse. One of the Board's initial concerns was that the public nature of this type of engagement would discourage participants from sharing potentially sensitive or proprietary information that could prove most useful. To accommodate this need, the Academy worked with Synteractive to integrate a direct email submission feature into the Dialogue platform. This allowed participants to submit ideas via email directly to the Academy without having them appear on the public Dialogue site. Although the majority of the ideas that were collected through the Dialogue were submitted on the public site, the direct email submission feature did generate valuable feedback that informed the findings of this report.

### ***Conducting Outreach***

Prior to the launch of the Dialogue, the Academy designed and executed a comprehensive outreach strategy that included a diverse community of stakeholders who could provide substantive feedback on potential tools and solutions to prevent and detect fraud, waste, and abuse. Based on the objectives of the Dialogue, the Academy and the Recovery Board identified a number of key groups that they wanted to engage early in the planning process. These groups included:

- Private-sector companies with expertise in fields related to the prevention and detection of fraud, waste and abuse, including: forensic analytics, banking, insurance, cloud computing, geographic information systems (GIS), and database management;
- Academia, particularly those in the fields of business and public administration;
- Federal, state, and local government leaders;
- Academy Fellows with experience in accounting, auditing, governmental management, law enforcement, information technology, cloud computing, risk modeling, performance metrics, and intelligence analysis;
- Private and public sector professional associations;
- Foundations and think tanks; and
- Members of the press and authors of relevant blogs.

At the heart of the Academy's outreach effort was an email campaign designed to reach each constituency with knowledge of potential tools and strategies for preventing and detecting fraud, waste, and abuse. In the week before the Dialogue went live, the Academy began to send outreach emails to 670 relevant stakeholders from across industry, academia, government, the media, and the non-profit sector. This email campaign continued throughout the week that the Dialogue was live. As outreach efforts continued, the Academy contacted numerous stakeholders by phone to solicit their participation personally. The Academy also reached out to its network of more than 600 Fellows to solicit their ideas and gain their assistance in reaching out to their professional networks.

In total, between the beginning of the outreach campaign on October 12 and the conclusion of the Dialogue on October 24, 2011, the Academy sent almost 4,800 emails and made more than 100 phone calls to approximately 1,350 relevant stakeholders, registered users and Academy Fellows. In addition, several other organizations and members of the media pushed information about the Dialogue out to their members and audiences, including the American Society of Public Administration, Government Executive Magazine, and Federal News Radio, which aired an October 18<sup>th</sup> interview about the Dialogue and replayed it during the week the Dialogue was open.

In conjunction with this extensive email and telephone campaign, the Academy used social networking tools, such as Facebook and Twitter, to reach constituencies interested in the prevention and detection of fraud, waste and abuse, but outside Academy and Recovery Board networks.

As discussed in the Traffic and Metrics section of this appendix, outreach efforts did not yield the level of traffic or participation initially anticipated. Generally, when preparing to execute an online dialogue, the Academy spends two to four weeks contacting potential participants to generate interest in the engagement. Unfortunately, due to a compressed project timeline and unexpected delays in the final approvals of outreach materials and platform functionality, initial communication about the Dialogue to external stakeholders was postponed until five days before the dialogue went live. This abbreviated outreach timeframe was a likely contributor to the relatively low level of traffic and participation. The experience illustrates the importance of ensuring sufficient time to conduct an effective outreach campaign.

### **Measuring Traffic and Participation**

In addition to the results submitted by participants, the Dialogue captured two broad categories of data: traffic and participation.

- **Traffic metrics** generally measure the amount of overall traffic to an activity on the site, including metrics such as Unique Visitors,<sup>26</sup> Total Visits, and Page

---

<sup>26</sup> Unique visitors (or absolute unique visitors) represent the number of unduplicated (counted only once) visitors to the website over the course of a specified time period. Although each visitor is identified as unique, it constitutes a unique visit from an IP address. Thus, an individual could have visited the dialogue

Views. The Academy used Google Analytics to capture this information. All traffic information was collected and reported in the aggregate. Also captured were measures of visitor engagement with the site, including “bounce rate”—a measure indicating the “percentage of single-page visits or visits in which the person left [the] site from the entrance (landing) page.”<sup>27</sup>

- **Participation metrics** measure active involvement in the Dialogue. Participation metrics collected for this Dialogue include registered users,<sup>28</sup> ideas, comments, and votes.

### **Site Traffic**

The Dialogue opened to the public on October 17, 2011 and closed on October 24, 2011. During the seven days it was live, the Dialogue received 953 visits from 591 unique visitors—an average of 136 visits from 84 unique visitors each day. The average participant spent approximately four and a half minutes on the site and viewed three and a half pages per visit. The bounce rate for the Dialogue was 48.27 percent, meaning that slightly less than half of all visitors left the Dialogue without moving beyond the homepage. Direct traffic<sup>29</sup> to the site accounted for almost 60 percent of all visits made, indicating that most visitors accessed the Dialogue directly, as opposed to finding it through a search engine. Table C- 1 summarizes Dialogue traffic metrics.

**TABLE C-1. DIALOGUE TRAFFIC METRICS**

<b>Metric</b>	<b>Data</b>
<b>Visits (visits/day)</b>	<b>953 (136)</b>
<b>Unique Visitors (new visitors/day)</b>	<b>591 (84)</b>
<b>Avg. Page Views</b>	<b>3.43</b>
<b>Avg. Time on Site in Minutes</b>	<b>4:22</b>
<b>Bounce Rate (%)</b>	<b>48.27%</b>
<b>Direct Traffic (%)</b>	<b>58.13%</b>

Many of the traffic metrics from the Dialogue were lower than originally anticipated. The compressed time period for outreach, as well as the specialized nature of the Dialogue subject matter, may have played a role in limiting the number of visitors who came to the site and explored the individual forums. However, when the other traffic metrics are

---

site from three separate computers or IP addresses. In this case, Google would count each visit as a unique visitor.

<sup>27</sup>The bounce rate is the percentage of visits that entailed only visiting the first page of the site. This metric provides an indication of how much users felt enticed to view other pages and engage with the site.

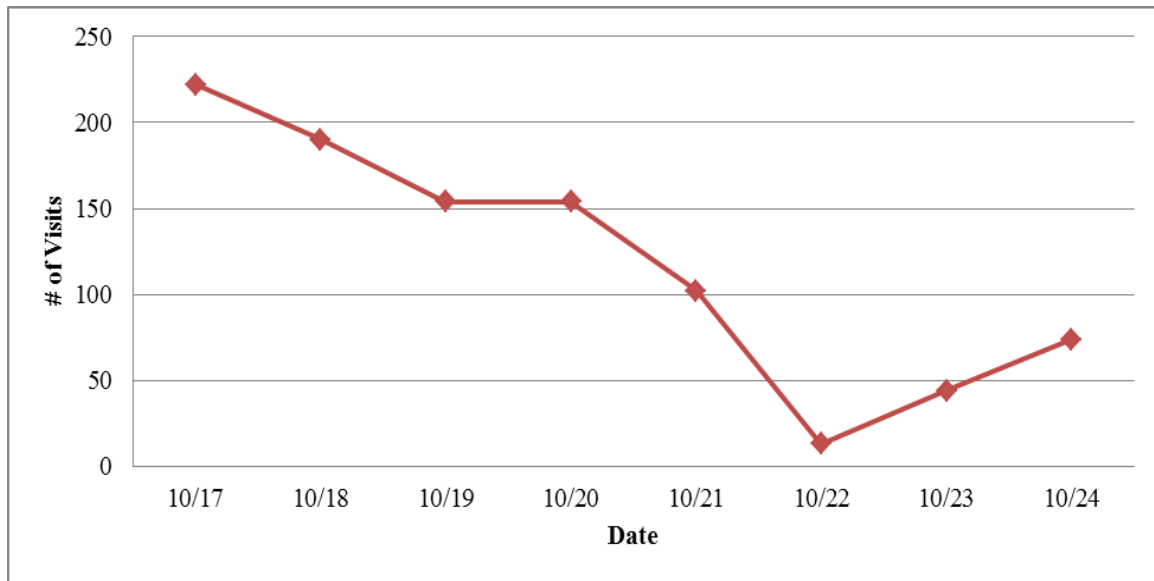
<sup>28</sup>A registered user is any individual who creates an account on the dialogue site; registration is necessary in order to submit, rate, or comment on any ideas on the site.

<sup>29</sup> Direct traffic measures the percentage of visits to the site that came from users clicking an email link or directly typing the URL into their web browser.

considered in relation to the bounce rate of almost 50 percent, site activity among those who were interested in the subject matter appears comparable to other Academy dialogue initiatives. When the 48.27 percent of visitors who only visited the homepage are removed from the calculation of average page views and length of time on site, the metrics show that the remaining 52 percent who went beyond the homepage visited an average of six pages and spent significantly more time on the site than the four minutes recorded for all participants.

Figure C- 1 shows the number of site visits each day that the Dialogue was live.

**FIGURE C-1. SITE VISITS OVER TIME**



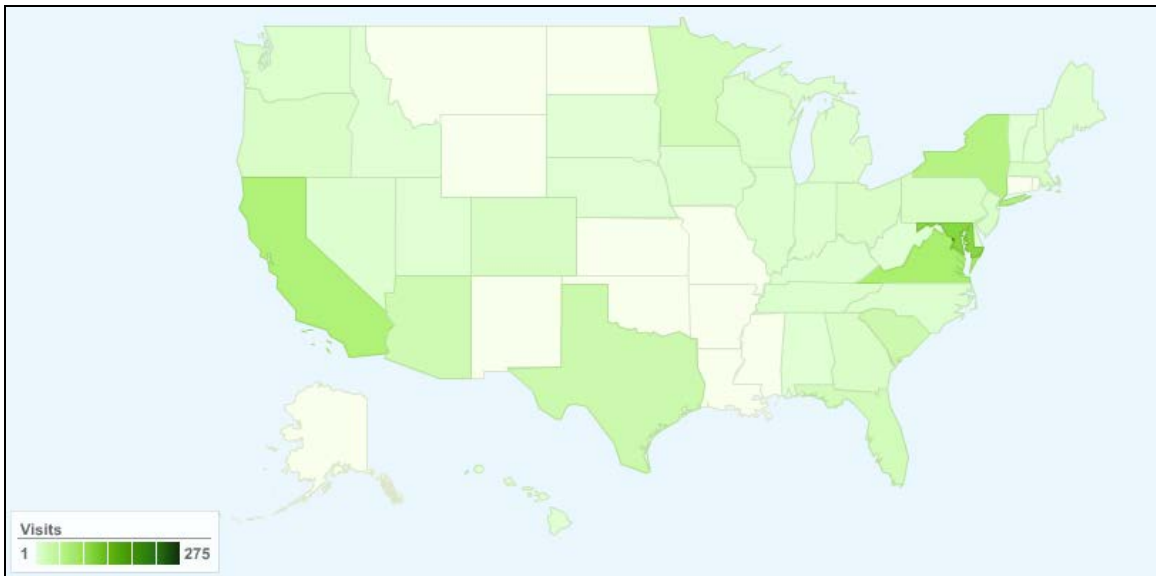
As Figure C- 1 illustrates, Dialogue site traffic reached its peak on the day the Dialogue went live and dropped gradually as the week continued. This aligns with site traffic data from past Academy dialogues, which followed similar patterns. Site traffic dropped precipitously on Saturday, October 22<sup>nd</sup>, which is also not unusual. The figure does show a notable increase in traffic on Sunday, October 23<sup>rd</sup>; this was likely due to the “last chance” email that was sent that morning.

In the one week that it was live, participants from the United States and 22 other countries visited the National Dialogue. Within the U.S., Google Analytics captured site visits from 37 states and the District of Columbia. Table C-2 provides a list of the states with the highest number of recorded site visits. Figure C- 2 maps this information.

**TABLE C-2. HIGHEST NUMBER OF SITE VISITS BY STATE**

Rank	State	Visits
1	District of Columbia	275
2	Maryland	123
3	Virginia	83
4	California	75
5	New York	65
6	Texas	31
7	South Carolina	30
8	Arizona	27
9	Florida	22
10	Minnesota	21

**FIGURE C- 2. DIALOGUE TRAFFIC BY STATE**



*\*Darker shades of green indicate a higher number of visits from that state.*

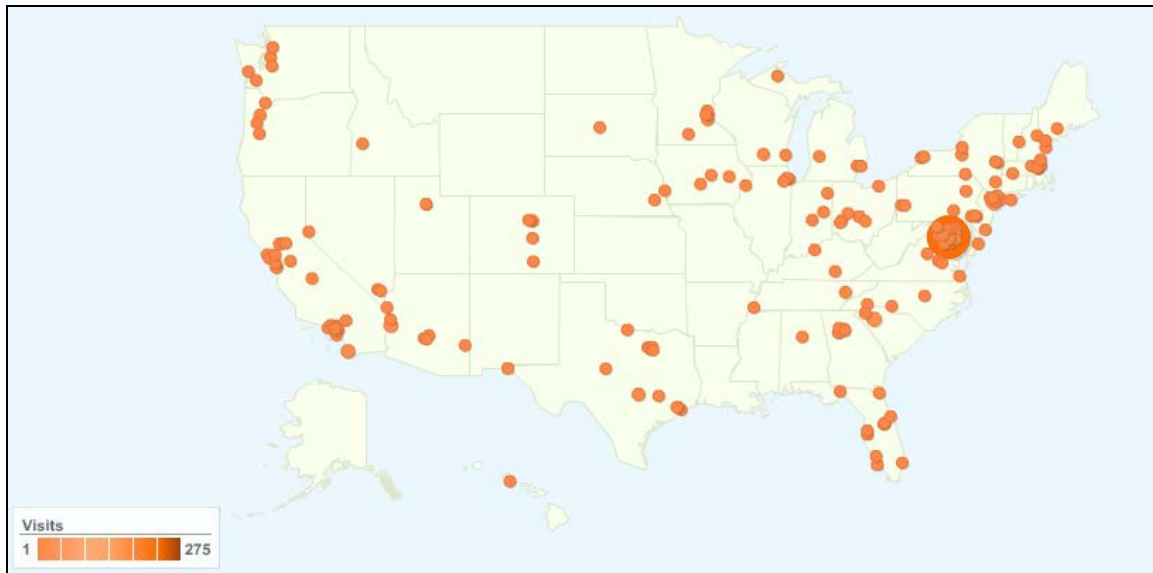
As illustrated in the preceding table and figure, the largest contributors to Dialogue traffic were from the Washington, DC multi-state area, California, and New York. Collectively, these states provided 65 percent of all site visits, with 29 percent coming from the District of Columbia alone. Receiving such a high concentration of traffic from the DC-area and the nation’s two most populous states was certainly not intended; however, it may be explained by the specialized nature of the feedback solicited and the narrow appeal of the Dialogue subject matter to those with backgrounds and expertise in government and finance.

In addition to state-level data, Google Analytics also captured traffic data for individual cities and towns. During the one-week period that the Dialogue was live, the site received visits from 208 cities and towns across the country. Table C- 3 provides a list of the cities with the highest number of recorded site visits. Figure C- 3 is a map displaying the city-level data.

**TABLE C-3. HIGHEST NUMBER OF SITE VISITS BY CITY**

Rank	State	Visits
1	Washington, DC	275
2	Bethesda, MD	73
3	New York, NY	45
4	Arlington, VA	29
5	Laurens, SC	29
6	San Diego, CA	20
7	Parker, AZ	19
8	Austin, TX	12
9	Rosemount, MN	10
10	Dunn Loring, VA	10

**FIGURE C-3. DIALOGUE TRAFFIC BY CITY**



*\*Larger, darker dots indicate a higher number of visits from that city.*

Once again, the preceding table and figure show a preponderance of visits from the D.C. area. However, it should be noted that, when viewing traffic data at the city-level, it is clear that the Dialogue drew visitors from a larger area than when looking at the state-level data alone. This may suggest that high-level interest in the Dialogue’s subject matter was broader than other indicators show.

***Dialogue Participation***

In the seven days that the Dialogue was live, 53 participants submitted 28 ideas, 20 comments, and 153 votes through the online forum. In addition, during the “Soft-launch” phase that took place in the five days before the Dialogue went live, Dialogue Catalysts submitted an additional eight ideas and one vote, for a total of 36 ideas and 154 votes. The Dialogue also allowed participants to submit ideas directly through an Academy

email address. Eight participants submitted nine ideas using this option. Table C- 4 provides details on participation metrics.

**TABLE C- 4. DIALOGUE PARTICIPATION METRICS**

<b>Metric</b>	<b>Data</b>
<b>Registered Users</b>	<b>53</b>
<b>Registered Users as % of Unique Visitors</b>	<b>9%</b>
<b>Total Ideas</b>	<b>36</b>
<b>(per day)*</b>	<b>(4.00)</b>
<b>(per registered user)</b>	<b>(0.68)</b>
<b>Total Comments</b>	<b>20</b>
<b>(per day)</b>	<b>(2.86)</b>
<b>(per registered user)</b>	<b>(0.38)</b>
<b>Total Votes</b>	<b>154</b>
<b>(per day)*</b>	<b>(21.86)</b>
<b>(per registered user)</b>	<b>(2.91)</b>

*\*Number includes only those ideas and votes that were submitted during the seven days that the Dialogue was live.*

Feedback across the individual Dialogue forums was not distributed equally. As Table C- 5 illustrates, the “Management” forum had the most activity by far with 15 ideas, 9 comments, and 60 votes; approximately 40 percent of all three types of feedback received in the Dialogue. By contrast, the “Risk Models” forum received only 4 ideas, 2 comments, and 15 votes, with only slightly higher levels of participation in the other three forums. .

**TABLE C-5. PARTICIPATION METRICS BY FORUM**

<b>Forum</b>	<b>Ideas</b>	<b>Comments</b>	<b>Votes Submitted</b>
<b>Management</b>	<b>15</b>	<b>9</b>	<b>60</b>
<b>Data Sources</b>	<b>5</b>	<b>3</b>	<b>23</b>
<b>Technologies</b>	<b>6</b>	<b>4</b>	<b>35</b>
<b>Risk Models</b>	<b>4</b>	<b>2</b>	<b>15</b>
<b>Performance Metrics</b>	<b>6</b>	<b>2</b>	<b>21</b>
<b>TOTAL</b>	<b>36</b>	<b>20</b>	<b>154</b>

The disparity in the distribution of feedback may have resulted from the more general types of ideas that participants could submit in the “Management” forum. Rather than proposing specific technological and analytical solutions, some may have felt more comfortable submitting higher-level conceptual ideas. Another possibility, which has arisen in other multi-forum Academy Dialogues, is that participants were drawn to the “Management” forum simply because it was the first one that appeared as they reviewed the forums from left to right.

The National Dialogue on Innovative Tools to Prevent and Detect Fraud, Waste, and Abuse offered a unique opportunity to engage a broad community of stakeholders around

specific ways to enhance accountability of Recovery Act funding. Although participation in the Dialogue was not as high as was originally anticipated, this engagement has provided important lessons on engaging the public around a specialized topic over an extremely short timeframe and generated some valuable results. Ultimately, the feedback received through both the Dialogue platform and direct Academy email address helped inform the development of the Academy's report on potential tools and solutions to prevent and detect fraud, waste and abuse.



## **APPENDIX D**

### **DIALOGUE STATEMENT OF PURPOSE AND POLICIES**

#### **Statement of Purpose**

The Recovery Accountability and Transparency Board (Recovery Board) is collaborating with the National Academy of Public Administration (the Academy) in an effort to solicit input on how to prevent and detect fraud, waste, and abuse of Recovery Act funds. Participation from key stakeholder groups will be conducted through a week-long online forum from October 17-24. The objective of this dialogue is two-fold:

- To identify best management practices, data sources, technologies and systems, risk models, and performance metrics associated with Recovery funds; and
- To improve oversight of Recovery Act spending.

While the scope of the dialogue is oversight of Recovery Act funds, we are interested in ideas related to oversight of all federal spending to see if they are applicable to Recovery spending. Upon conclusion of the dialogue, the Academy will conduct an in-depth analysis for the Recovery Board. Ultimately, the most promising technologies and practices will support the Recovery Board's mission of increasing the transparency and accountability of federal funds.

#### **Moderation Policy**

This online dialogue allows you the opportunity to fill out custom text fields, which are publicly visible. While we invite open participation and diverse viewpoints, the main goal of this dialogue is to answer the overarching question about possible technologies and techniques to prevent and detect fraud, waste, and abuse. The site, therefore, operates a moderation policy to ensure that your comments are on topic and not harmful to others. Moderators reserve the right to delete comments that are not relevant to the topic and/or contain the following:

- Threats or incitements to violence
- Obscenity
- Duplicate posts
- Posts revealing your own or others' sensitive/personal information (e.g., Social Security numbers)
- Information posted in violation of federal law
- Our desire is to remove as few posts as possible while ensuring that a focused, constructive discussion takes place.

If you have a complaint about an item of user-generated content on this site, please contact us.

## **Privacy Policy**

We are committed to ensuring the privacy of our users. In order to post ideas, comment on ideas, or vote on ideas, this dialogue requires you to create a personal profile. Although you must submit an e-mail address when creating a profile, we will not display this e-mail address on the site or share it with anyone else. The only information from your profile that is visible to other users is your chosen username, your chosen geographic zone, and your avatar. Your username and avatar will be visible next to every idea and comment that you add to the dialogue. The information gathered when creating your profile is not validated. Text entered as ideas or comments in the forum will immediately be visible to the public, but may be subject to post-moderation if the Academy believes that they do not abide by the terms of use (see the terms of use and the moderation policy).

Additional data, such as what pages on the site you've visited and the length of time you visited them, is collected anonymously for the purposes of analyzing visitor traffic; none of this data can be associated with your individual user identity within the dialogue. This data may be made available in the aggregate to visitors of the website, developers, and dialogue sponsor organizations including but not limited to those within the federal government.

## **Terms of Use**

### ***Acceptance of Terms***

By accessing this website (the "Site"), including all information, documents, communications, files, text, graphics, software, and products available through the Site (collectively, the "Materials"), you are agreeing in good faith to the following terms. If you do not wish to be bound by these terms and conditions, you may not access or use the site. If you breach any of the terms of this Agreement, your authorization to use this Site automatically terminates, and any Materials downloaded or printed from the Site in violation of the Terms of Use must be immediately destroyed.

The National Academy of Public Administration (the Academy) may modify this agreement at any time, and such modifications shall be effective immediately upon posting of the modified agreement on the site. You agree to review the agreement periodically to be aware of such modifications and your continued access or use of the site shall be deemed your conclusive acceptance of the modified agreement.

### ***Nature of Materials***

The Materials are samples that have been provided in order to advance the public interest for educational and guidance purposes only. They are not standard or model forms and should not be used as such. Because these documents have been drafted to meet the unique program needs of the Academy and to satisfy specific requirements applicable to this dialogue, the actual content of a particular document may not be appropriate for your use. Therefore, while these documents may serve as an excellent starting point for drafting or revising similar documents, independent judgment and, where appropriate, the advice of competent legal counsel is strongly recommended.

### ***Use of Materials***

This Site is for informational purposes only and is not intended to provide specific commercial, financial, investment, accounting, tax, or legal advice. It is provided to you solely for your own personal, non-commercial use and not for purposes of resale, distribution, public display or performance or any other uses by you in any form or manner whatsoever. Unless otherwise indicated on this website, you may display, download, archive, reformat and print a single copy of any information on this website for such personal use. Some of the Materials on this Site are subject to specific restrictions with respect to use. You expressly agree that you will comply with any and all such restrictions.

### ***Intellectual Property***

The Materials and Services on this Site are protected by copyright, trademark, patent, and/or other intellectual property laws, and any unauthorized use of the Materials or may violate such laws. You agree to use the Materials in compliance with all applicable laws, governmental regulations and guidelines, including without limit, all copyright, trademark and other intellectual property laws. You understand that the ownership of the materials is retained by the originators of those materials and that this agreement does not constitute the originators waiver of any rights. The originators are the beneficiaries of, and may independently enforce, this agreement. Except as expressly provided herein, neither the Academy nor the originators grant any express or implied rights to use the Materials.

### ***Warranties and Disclaimers***

You assume total responsibility and risk for your use of the site. The Academy is providing access to these materials and related information "as is" and does not make any express or implied warranties, representations or endorsements whatsoever with regard to the materials or related information. This site could include technical or other mistakes, inaccuracies or typographical errors. It is solely your responsibility to evaluate the accuracy, completeness and usefulness of all opinions, advice, services, merchandise and other information provided through the site generally. You expressly understand and agree that the Academy disclaims any and all responsibility or liability for the accuracy,

content, completeness, legality, reliability, or operability or availability of information or material in the site. The Academy further disclaims any responsibility or liability for any harm resulting from downloading or accessing any information or material through the products and services, including, without limitation, for harm caused by viruses or similar contamination or destructive features and the Academy shall not be liable for any cost or damage arising either directly or indirectly from your use of the site.

If you elect to send private information to the Academy through the Solution Summary email address, rather than posting that information on the online forum, the Academy cannot guarantee the confidentiality of content submitted. We recommend that you do not submit trade secrets, or privileged, confidential commercial or financial information.

### ***Indemnification***

You agree to indemnify and hold the Academy, and its officers, agents and employees, harmless from any claim or demand, including reasonable attorneys' fees, made by any third party due to or arising out of your use of the Site, your connection to the Site, your violation of the Terms of Use, or your violation of any rights of another person or entity.

### ***General***

The Terms of Use and the other rules, guidelines, licenses and disclaimers posted on the Site constitute the entire agreement between the Academy and you with respect to your use of the Site. If for any reason a court of competent jurisdiction finds any provision of the Terms of Use, or portion thereof, to be unenforceable, that provision shall be enforced to the maximum extent permissible so as to effect the intent of the parties as reflected by that provision, and the remainder of the Terms of Use shall continue in full force and effect. Any failure by the Academy to enforce or exercise any provision of the Terms of Use or related right shall not constitute a waiver of that right or provision. The section titles used in the Terms of Use are purely for convenience and carry with them no legal or contractual effect.

## APPENDIX E IDEAS/COMMENTS SUBMITTED<sup>30</sup> IN DIALOGUE

### Management

#### ***Idea 1 - Separate qualification and allocation functions.***

Because enormous political pressure will be brought to bear on the people distributing the funds, it would be advisable to separate the qualification and allocation functions.

Applicants would be vetted by a qualifications staff for adherence to guidelines and for suitability as revealed through a background check. Then the allocators would deal only with those who had been determined to be qualified. Such separation of tasks will not guarantee success but will help minimize the opportunity for political interference and manipulation. It will also permit the background checking to be done by specialists in that area, while the allocation is done by those with broader policy perspectives.

2 Positive Votes

Authored by [DrD1](#)

10/23/2011 8:27:53 AM

0 [Responses](#)

Tags [distribution](#), [accountability](#)

#### ***Idea 2 - Design funds distribution phase carefully pre-announcement.***

The process of distributing funds sets the tone for what happens later. If political considerations force rapid distribution, then the ability to maintain control can be lost. The distributors are rewarded for giving large sums to individual recipient organizations, as that gets the money out, though safeguards are often ignored in such transactions. If the large sums are then redistributed, money is wasted in covering unnecessary administrative costs. From a management perspective, then, the wiser approach is to design the distributing phase carefully before announcements are made.

2 Positive Votes

Authored by [DrD1](#)

10/23/2011 8:22:23 AM

0 [Responses](#)

Tags [accountability](#), [distribution](#)

---

<sup>30</sup> Material in this Appendix is verbatim from the Dialogue. The Academy has not made any grammatical, spelling, or other edits.

**Idea 3 - Provide some portion of funds only after demonstration of performance.**

Under CETA governmental agencies and nonprofits used governmental funds to put people on their own payrolls, calling that job creation, but provided little or no training. Under the successor, JTPA, fund recipients had to find private-sector jobs for people, and received funds only AFTER the placement. This was much harder, but the jobs tended to last and funds were not wasted. The Recovery system needs a similar mechanism to provide at least some funds only after the demonstration of performance. To do otherwise is to ensure quick, shallow fixes with little staying power and real chances for self-enrichment.

2 Positive Votes

Authored by [DrDI](#)

10/22/2011 5:49:22 PM

0 [Responses](#)

Tags [accountability](#)

**Idea 4 - Reimburse for actual expenses. Do not use fixed price contracts for universities.**

The feds need to take a long, hard look at why they engage in fixed price contracts with universities, especially no bid fixed price contracts. What's wrong with reimbursing for actual expenditures?

5 Negative Votes

Authored by [Liz](#)

10/21/2011 10:58:41 AM

0 [Responses](#)

**Idea 5 - Hold universities accountable.**

Unless and until universities hold faculty accountable for their compliance (or lack thereof), nothing will change, administrators will expend time, money and effort to find fraud and the muckety mucks will come up with the bucks to fix it, sweep it under the rug and the PI will engage in the behavior again (and again).

1 Positive Vote, 2 Negative Votes

Authored by [Liz](#)

10/21/2011 10:57:28 AM

0 [Responses](#)

**Idea 6 - Strengthen systems and audit processes to assure timely de-obligation of unneeded funds.**

When goods or services are finally received, agencies are supposed to deobligate any remaining obligations. How often is that not done. Unliquidated obligations is an opportunity for chicanery. Systems and the audit process need to be strengthened to assure the timely deobligation of unneeded funds.

4 Positive Votes

Authored by [halsteinberg](#)

10/20/2011 1:51:39 PM

1 [Responses](#)

*Comment 6.1 - Reallocate unused funds semi-annually during mi-session review or at end of contract/grant.*

Do we need a de-obligation process anymore? De-obligations were a result of paper-book keeping systems of the pre-ERP era. Why shouldn't unused funds be re-allocated semi-annually in line with the mid-session reviews or rescinded when unused at the end of the grant/contract. is it more a desire to spend the budget someone fought for or a need to spend as much as possible to get as much result as possible?

[maforman](#)

**Idea 7 - Do more to collect accounts and taxes receivable.**

The Federal government is owed close to \$100 billion in accounts and taxes receivable; \$500 billion in loans receivable after deducting the subsidy allowances and not counting the GSE's mortgage backed securities; and \$1,6 trillion in loan guarantees. Most agencies have a vigorous program for collecting these amounts, but more can always been done. A smaller, but still significant amount is due to the government from the placement of IPAs and other types of reimbursable agreements with organizations outside the government. The nature of these agreements often precludes agencies from billing and attempting to collect the amounts due. This is another source of revenues.

3 Positive Votes

Authored by [halsteinberg](#)

10/20/2011 1:48:25 PM

0 [Responses](#)

**Idea 8 - Focus prevention and detection on business and individual recipients applying for government benefits, grants, and loans because their motivation is self-interest and they pose a higher risk.**

Why is it that every government funded “safety net” or attempt at “stimulus” is accompanied by an outrageous amount of fraud and abuse? Indeed, there are no known exceptions in history. The answer, of course, is quite simple: it’s all about which participants have skin in the game. Those that apply for government benefits, grants and loans have it; those that administer these things on behalf of taxpayers don’t. Human nature is what it is. Motivated self-interest will always be way out in front rule implementation and compliance monitoring. Only stakeholders and free market participants are capable of policing fraud and abuse. Once this simple truth is understood we no longer need the best and brightest of the most educated to discover new monitoring metrics. The fixes become obvious, even to the dullest among us.

7 Positive Votes, 5 Negative Votes

Authored by [rockvfaz](#)

10/19/2011 9:42:04 AM

5 [Responses](#)

**Comment 8.1 - The bulk of fraud is committed by business. Use effective monitoring and audit policies to stop it.**

I have a rational self interest to see the job created by stimulus and to see it properly monitored as both a citizen and an administrator of a private sector organization. I agree, there will be those who attempt to game the system, but the bulk of the fraud that I have witnessed was committed by business, and could have been easily stopped by more effective monitoring and audit policies.

[jond](#)

**Comment 8.2 - The need for new rules and larger and larger monitoring and compliance increases the bureaucracy.**

Should I have stated: "Those, including businesses, that apply for government...." "Those" was meant to be generic, i.e., any one or any entity that applies. In a round-about way you make my case: We need "...more effective monitoring and audit policies" until we discover we need even more effective monitoring and audit policies...as we implement new rules and build a larger and larger monitoring and compliance bureaucracy.

[rockvfaz](#)

**Comment 8.3 - Unless there is something to gain or lose, you will never develop meaningful accountability.**

He is right! Skin in the game changes everything. Unless there is something to gain or lose, you will never vbe able to develop any meaningful accountability.

[Kingboone](#)



*Comment 8.4 – I don't accept the premise that every government-funded safety net is accompanied by abuse.*

The statement is so fraught with question begging assumptions that it really doesn't even warrant response -- and yet here we are.... To respond is to accept the idea that "every government funded "safety net" or attempt at "stimulus" is accompanied by an outrageous amount of fraud and abuse" -- followed with the contention that there are "no known exceptions in history"! And yet the post declares the "simple truth" that only "stakeholders and free market participants are capable of policing fraud and abuse".... These contentions seem like the kind of overstated rhetoric of a political campaign rather than the basis for discussion and analysis. There is no use arguing the point, for any response helps reinforce the rhetorical form of the statement -- ("In a round-about way you make my case"). All programs -- safety net, stimulus, or otherwise -- are designed with inherent flaws, and they can be gamed because they are perceivable as games. Some programs are even designed to be gamed -- which is why many policies create those strange things called "incentives". Under the logic of the post, the entire US tax code, every public works program, all social programs (corporate welfare included) are bundle of fraud and abuse waiting to happen (well, of course they are!). We have long known that policies in the form of laws, rules, regulations are imperfect. So what is the news here? (More)

[mdubnick](#)

*Comment 8.5 – De-regulation didn't minimize abuse.*

The news seems to be that it would be best to let stakeholders and free markets handle such programs -- or perhaps it is that we should eliminate the programs because they cannot be effectively policed by those charged with implementing the programs. The "simple truth", however, does not seem to hold water -- one has merely to look at the consequences of de-regulation, especially in the financial markets, to see how effective stakeholders and free market types deal with fraud and abuse. It seems that rather than police, these folks feed on games they create by fostering fraud and abuse that generate the bubbles that eventually burst. There are interesting issues to address in this forum about problems in the administration of governance and how we might improve their effectiveness and efficiency. It is too bad we have to be distracted by rhetorical statements such as the initiating post in this thread.

[mdubnick](#)

**Idea 9 - Don't keep secrets. Hold quarterly public forums for civil servants to answer questions from the public.**

Quarterly public forums for all civil servants will balance the budget. When folks have to answer questions from the public (about their public job) without having prepared remarks from lawyers will root out the truth. Agencies keep secrets that are not secret.

[www.militec.blogspot.com](http://www.militec.blogspot.com)

2 Positive Votes, 4 Negative Votes

Authored by [spankyhazel](#)

10/18/2011 7:45:45 AM

0 [Responses](#)

Tags [quarterly-public-forums-for-civil-servants-will-balance-the-budget](#)

**Idea 10 – Agencies should be compelled to fully audit their books, not take a sample.**

One of the keys to successful recovery audits is to fully audit an agency's books. Unfortunately the current standard is "sampling" which generally equates to shooting a rifle in the air and hoping a bird flies into your bullet. Thankfully, Senator Carper (D-DE) will be offering an improvement to his already successful IPERA law (PL 111-204) during this fall legislative session. This bill will compel agencies to contract with professional, external service providers like APEX Analytix, for example, that have software that identifies, prevents and improves the improper and overpayment situation that currently hamstring many agencies. Now the fox won't guard the hen-house anymore (a la Sarbanes-Oxley) and each invoice and payment will be thoroughly inspected for the potential for fraud via an efficient software solution.

<http://carper.senate.gov/public/index.cfm/pressreleases?ID=457fc5ce-06b7-4d3e-8dfc-4350a5c1d8e4>

4 Negative Votes

Authored by [mceoxley](#)

10/17/2011 6:53:57 PM

0 [Responses](#)

Tags [-ipera](#), [-senatorcarper](#)

**Idea 11 – Pull as much of the data on to a single integrated platform to enable data mining. Build statistical models to analyze patterns.**

To reduce waste, make the most of the limited personnel resource dedicated to catching fraud --- too often the snarl of overwhelming data causes them to spend all their time gathering data, and too little time analyzing it. So, first, pull as much of the data as possible onto a single integrated platform which allows special fraud investigations units to use data mining tools to quickly cut through the mountains of data. Reduce fraudulent payouts by building statistical models which can analyze historical patterns of potentially bogus requests, related to, for example suspicious addresses.

4 Positive Votes

Authored by [monica smith](#)

10/17/2011 4:15:51 PM

0 [Responses](#)

**Idea 12 – The government needs to provide full organizational access to detailed financial data throughout the accounting cycle. Eliminate information black holes. People need to be able to trust the data.**

- People have to be able to trust the financial data they are seeing, and understand its underlying components before they can be made accountable and act to change the results. Having access to detailed financial data throughout the accounting cycle (not just at close of month) and throughout the organization with fiscal responsibility enables accountability by eliminating information black holes. True oversight requires confidence in decision making supported by an accurate representation of what is currently happening.

3 Positive Votes

Authored by [monica smith](#)

10/17/2011 1:06:17 PM

0 [Responses](#)

[financial-management](#), [finance-amp-performance-management](#), [financial-insight](#), [management-reporting](#)

**Idea 13 – The government needs internal controls for the full range of program activities, not just accounting and finance and to identify and control risks.**

The application of internal controls to more than accounting and financial transactions. In other words, for various program activities, such as determining eligibilities, providing services, making grants, developing model internal control plans encompassing the specification of objectives, identification of possible risks, and the identifications of appropriate controls to address the risks/

6 Positive Votes

Authored by [halsteinberg](#)

10/17/2011 11:12:23 AM

0 [Responses](#)

Tags [his](#)

***Idea 14 – Integrity means complete forthrightness in all dealings.***

“Having integrity means more than simply the absence of deception. It means we are completely forthright in all our dealings. We say what needs to be said, not simply what people want to hear.” Scott Cook, Founder, Intuit

2 Positive Votes, 1 Negative Vote

Authored by [fhc](#)

10/15/2011 5:34:08 AM

0 [Responses](#)

***Idea 15 – How does an organization really measure return on investment in fraud detection?***

ROI analysis on the implementation of a Fraud Detection Management solution is problematic. What would be beneficial to an organization that needs to measure ROI accurately?

1 Positive Vote, 1 Negative Vote

Authored by [dc1191](#)

10/12/2011 7:43:27 PM

3 [Responses](#)

Tags [roi-analysis](#)

***Comment 15.1 – Measure ROI by improvement in number of frauds detected compared to random or other techniques.***

It is a very interesting question. I guess one way of measuring ROI might be to see the improvement in number of frauds detected compared to random (or any other past used) technique. Of course, one have to pay attention to the FP (False Positive) and FN (False Negative) ratios.

[clusty](#)

Comment 15.2 – ROI has to be based on reduction or elimination of current fraud. Measuring current fraud loss is critical.

This is a very broad question, but should be approached as follows: 1. Define the existing fraud/waste/abuse problem within the agency or process. (i.e. you need to know very clearly what actions the solution will address) 2. Estimate/quantify the current \$ loss resulting from the problem. (as refined as possible, so it can be tied directly to the problems identified in 1 above) 3. Define how the solution will address 1 above as directly as possible. (i.e. the solution will reduce the number of wrong amounts issued, the number of payments made for warranted products, the number of payments that do not comply with contract terms, etc., which cuts down on improper payments resulting from internal errors.) 4. Estimate/quantify the reduction in current \$ loss resulting from application of the proposed solution to the problem. (the more tightly you can tie items 1-3 to 4 the more solid the argument) While this is a pretty generic response, in the end the ROI has to be based on reduction or elimination of the current fraud/waste/abuse. It can be hard to quantify, but the steps above provide the conceptual basis for understanding the value of the solution in terms of ROI. Obviously it's critical that the current \$ loss is measured (step 2). If you don't know how much they're losing today, it's impossible to evaluate a solution based on ROI. I hope this helps. It's never an easy question to answer.

[monica smith](#)

Comment 15.3 – To make a good business decision, you need historical data and a clear picture of the magnitude and nature of the losses as well as an understanding and costing of the operational process. It is more difficult to identify savings benefits. The problems can be overcome using correct evaluation methodologies.

Making an investment decision in any area of business is a challenge and fraud management is no different in this respect; it is often more difficult to identify the benefits of savings than it is to forecast the positives involved in increasing revenue streams. This does not mean that it is a problem that cannot be overcome with the correct evaluation methodologies and data to support it. Firstly it is critical that any organization has a clear picture of both the magnitude and nature of its fraud losses, which in turn requires effective identification and labeling of fraudulent transactions over a period. How long should the period be? Well it depends on a number of factors but typically 12 months would be considered an ideal, to account for seasonality, but the more data the better. Where speed of action is required, then shorter intervals may well be satisfactory. It is also vital that there is a clear understanding of how the actual detection of fraud and the operational process to prevent loss will interact and at what stage of the fraud lifecycle financial losses can be controlled. In simple terms, merely identifying suspicious behavior does not necessarily deliver any financial benefit. Most organisations will also experience some intangible benefit from enhanced fraud detection, whether by virtue of improved customer service or enhanced reputation, and these should also be considered as providing a return on investment. Once armed with a clear picture of current fraud losses and an understanding of the operational process (and its concomitant costs) it is a matter of simulating the proposed fraud detection system on the historical data to provide a number of scenarios for the likely outcomes. Mostly this will be a question of identifying at what stage these past frauds would have been identified and how much of the eventual loss could have been prevented or recovered by the operational process. In essence, the methodologies and mechanics for evaluating the benefits of fraud detection are not difficult. The challenge mostly resides in having the historical data available to undertake the process. In fact it is a truism that data and data quality are key considerations in the world of fraud detection and management, irrespective of the context being debated.

[riberry](#)

## **Data Sources**

***Idea 16 - Before declaring any person or organization eligible to receive funds, the government should pre-certify, run a credit check, a check of criminal files, and relevant licenses. Follow-up with relevant databases on theory that negative tendencies repeat.***

Before a semi-finalist is ruled eligible for a high position, any competent headhunter or search committee will run a credit check, a check of criminal files, and a check of relevant licenses. The government should do the same before declaring anyone or any organization eligible to receive funds. Once eligible recipients have been determined, then additional checks of relevant databases could be run, on the theory that negative tendencies are likely to be repeated. Honesty, after all, is not a trait that can be proven; but dishonesty can be identified.

1 Positive Vote

Authored by [DrD1](#)

10/22/2011 5:58:32 PM

0 [Responses](#)

Tags [precertification](#)

***Idea 17 - Identify types of files containing names of those applying for benefits and determine if computer matches with those files should be mandatory before payments are made.***

The government computer matches many data sources before making benefit and other payments, e. g., recent death files, lists of Incarcerated persons, recently bankrupt and other defunct companies. Attempts should be made to identify the types of files containing names of persons applying for benefits, grants, loans, contracts, etc. and determinations made as to whether computer matches with those files should be made mandatory before payments are made.

5 Positive Votes

Authored by [halsteinberg](#)

10/20/2011 4:08:40 PM

0 [Responses](#)

***Idea 18 - Can we use WhiteListing or crowd sourcing to aid in reduction of fraud?***

WhiteListing or crowd sourcing are terms that we have become familiar with over the recent years. Can we use these capabilities to aid in the reduction of fraud?

3 Positive Votes, 1 Negative Vote

Authored by [dc1191](#)

10/17/2011 2:10:22 PM

1 [Responses](#)

Comment 18.1 – Analyze blog content for complaints about funds recipients.

I have blogs may be good to find out if there are complaints against a company or vendor but need to verify.

[joserf](#)

**Idea 19 - What state and local data bases could provide information on negative actors?**

In addition to common federal databases, what state and local databases could provide information on negative actors, fraud history, etc?

7 Positive Votes

Authored by [James Duginske](#)

10/17/2011 10:46:33 AM

1 [Responses](#)

Comment 19.1 – State licensing, contracting, audit and AG/IG data bases will provide useful information.

State Licensing Databases; State Contracting Vendor Debarment, Single Audit Data, Criminal History, Attorney General, Inspector General

[joserf](#)

**Idea 20 - What about industry negative databases (blacklists), business license data, crime data, social media data, claim forms, notes, historic data (sales, transaction, purchase, etc.), payment records, location information, etc?**

Industry commonly uses data provided by the government or governmental organizations to support fraud detection including secret service, FBI, state attorneys general, local law enforcement. Depending on the agency and fraud being detected, data may include industry negative databases (blacklists), business license data, crime data, social media data, claim forms, notes, historic data (sales, transaction, purchase, etc.), payment records, location information, etc.

6 Positive Votes

Authored by [Bill Franks](#)

10/14/2011 8:37:04 PM

1 [Responses](#)

Comment 20.1 – Is personal credit history germane? Can it help identify those who commit fraud?

Is a person's credit history maintained by the three credit agencies usually included in these investigations? Is this information germane to individuals that commit fraud?

[dc1191](#)



## **Technologies**

### ***Idea 21 - Are there any affordable, robust solutions that automate fraud detection other than those offered by Thomson Reuters and SAS EBI?***

There are many analytical software tools available to prevent and detect fraud. What programs do you find most useful?

2 Positive Votes, 1 Negative Vote

Authored by [scarver](#)

10/18/2011 9:11:47 AM

[2 Responses](#)

#### ***Comment 21.1 – Is there a central data base or organization that rate tools and software?***

What about any sites or organizations that rate analytical tools and software? If there was a central database listing and exploring the tools, then not only would we know what programs are out there, but how they work, what their capabilities are and their rating...

[ebuckley](#)

#### ***Comment 21.2 – There are many tools, but they need to be considered in the context of available data, type of fraud, and skill sets needed to use tools.***

There are many tools in the market place that can help prevent and detect fraud. Tools need to be considered in context with the type of available data (e.g. structured vs. unstructured), the type of fraud under investigation (e.g. contract fraud, financial transaction fraud), the skillsets necessary to operate these tools (e.g. business analyst, statistician, engineer) and the amount of data that needs to be analyzed. Very often, a combination of tools is required to obtain the necessary insights. Tools worthy of consideration include: Forensic accounting - IDEA, ACL Unstructured analysis – iConnex, Endeca Data mining - Clementine, Darwin, S-Plus

[rickyeng](#)

### ***Idea 22 - What analytical software tools are most useful?***

There are a few business intelligence solutions that support automating the detection of fraud schemes. Thomson Reuters and SAS EBI provide such a capability in what is known as the fraud framework. Are there any other affordable, robust solutions that offer similar capabilities?

3 Positive Votes

Authored by [pnauroth](#)

10/17/2011 3:20:40 PM

[0 Responses](#)

**Idea 23 - Use multiple approaches that are flexible to support new requirements and that can address increasing complexity of fraud networks and schemes.**

**Requirements: advanced analytic techniques, such as geospatial mapping and analysis, social network analysis, in-database data mining, and text analysis.**

Fraud is constantly evolving and typically flows to the path of least resistance, therefore, the technologies employed for data aggregation and integration should be flexible to support new requirements as they emerge. In addition to being flexible, the technologies should also address the increasing complexity of fraud networks and schemes, which requires advanced analytic techniques such as geospatial mapping and analysis, social network analysis, in-database data mining, and text analysis. Key to this is that just one or two approaches for detecting fraud is not enough, multiple approaches must be employed.

11 Positive Votes

Authored by [Bill Franks](#)

10/14/2011 8:31:57 PM

1 [Responses](#)

Tags [data-aggregation-and-integration](#)

*Comment 23.1 – IGs and investigators made redundant buys. Many needed tools are routinely used for counter-terrorism.*

such tools are routinely used for counter-terrorism to find people as part of a terrorist network. Shouldn't the Intel community make its sw tools immediately available for use by Inspectors General to identify networks of fraud? It seems the biggest problem here is that IGs and investigators have to make redundant buys, wasting time and money on tools that the government already bought.

[maforman](#)

**Idea 24 - Use a framework, such as MapReduce, to explore large data sources that have not been fully cleansed and structured.**

Frameworks that assist with exploring large data sources that haven't been fully cleansed and structured should be utilized. MapReduce is one such framework which is fairly new and growing in adoption.

8 Positive Votes

Authored by [Bill Franks](#)

10/14/2011 8:30:00 PM

0 [Responses](#)

**Idea 25 - Integrate and aggregate data at the most detailed level possible. Use a scalable, massively parallel processing (MPP) relational database system to grow the system without sacrificing performance.**

Data should be integrated and aggregated at the most detailed level possible to support differing levels of management reporting, analytics and insight. Utilizing a scalable, massively parallel processing (MPP) relational database system for the integration, aggregation, and preparation of the data is an important consideration given that there will be a lot of data and the ability to grow the system without sacrificing performance will be an issue. Full disclosure: I work for such a company. However, I intend the concept and architecture be the focus rather than any specific vendor for this forum.

6 Positive Votes

Authored by [Bill Franks](#)

10/14/2011 8:27:04 PM

1 [Responses](#)

*Comment 25.1 – Data must be timely, integrated, with analysis making sense of unanticipated linkages. Detailed information loses meaning in aggregation.*

I'd like to that the importance of creating a flexible data environment that allows for constant improvement in the data environment along the following lines: • Timely: Data must be as current as possible. • Integrated: Data must be integrated to the extent possible. Any meaningful analysis must be able to make sense of unanticipated linkages between transactions. • Detailed: Information loses much of its meaning in the aggregation process. While aggregates are necessary to summarize conclusion, detailed data is essential to achieving true understanding.

[monica smith](#)

**Idea 26 - Use artificial intelligence?**

Can artificial intelligence be used in fraud detection?

4 Positive Votes

Authored by [dc1191](#)

10/12/2011 7:48:05 PM

0 [Responses](#)

Tags [fraud-detection](#)

## **Risk Models**

### ***Idea 27 - Use content analysis based on key phrases associated with fraud to review new proposals and try to identify vulnerabilities.***

Content analysis can provide powerful insights. It might be worthwhile, for example, to submit prospectuses and proposals by such people as Bernie Madoff to be submitted to a computerized analysis to identify commonalities. Those key phrases associated with fraud can then be used in a review of new proposals to try to identify untrustworthy tendencies.

2 Positive Votes

Authored by [DrDI](#)

10/22/2011 6:03:03 PM

0 [Responses](#)

Tags [content-analysis](#)

### ***Idea 28 - On what risk models/ risk management organizations do you rely?***

What risk models are currently available? What risk management organizations or agencies do you rely on?

3 Positive Votes, 1 Negative Vote

Authored by [ebuckley](#)

10/17/2011 1:47:35 PM

1 [Responses](#)

*Comment 28.1 – Risk factors must include inherent risk, history of fraud, magnitude or reputation risk.*

I prefer a voting one to get various factors must include inherent risk, history o fraud, and magnitude or reputation risk

[joserf](#)

### ***Idea 29 - Identify good and bad examples of fraud and look for differences and sources. Are there guidelines/rules available to assess compliance, guide exploratory analysis, and identify initial fraud, waste, or abuse cases?***

How many known cases of fraud have been identified already? Many analytics begin by having known “good” and “bad” examples and looking for differences. A source of known examples falling into each group is important. Alternatively, what guidelines or rules are available that can be used to assess compliance, guide exploratory analysis, and identify initial fraud, waste, or abuse cases?

4 Positive Votes

Authored by [Bill Franks](#)

10/14/2011 8:35:49 PM

0 [Responses](#)

**Idea 30 - Use unstructured data as input to add power to analytics. Note that text data from documents, e.g. emails, funding applications, and contracts, requires some pre-processing. Don't focus solely on structured data.**

Text data from documents such as emails, funding applications, contracts, and other sources should be used as inputs to fraud analysis. Organizations often focus purely on structured data, but unstructured data like text can add a lot of power to the analytics. It requires some pre-processing to extract important information from the text to feed the analysis.

5 Positive Votes

Authored by [Bill Franks](#)

10/14/2011 8:33:52 PM

1 [Responses](#)

*Comment 30.1 – Terms such as variances, shifting funds, change orders, multipliers, and overhead, sole source pre-negotiated rates are often disguised terms for fraud.*

Fraud terminology is often disguised as variances, shifting funds, change orders, multipliers and overhead, sole source pre-negotiated rates.

[joserf](#)

## **Performance Metrics**

**Idea 31 - Are online dashboards be used for organizing and assessing data?**

Installing a new process or developing a new product often entails the hiring of consultants or contractors. Having a rubric to evaluate the effectiveness of such entities would help curb fraud, waste, and abuse. At a minimum, the vetting process for awards should be able to discern the relative age and experience of such entities to help determine which have the requisite experience and which might have been formed just to gain access to funds.

2 Positive Votes

Authored by [DrDI](#)

10/23/2011 8:40:15 AM

0 [Responses](#)

Tags [accountability](#), [consultants](#)

**Idea 32 - Have a rubric to evaluate the effectiveness of consultants and contractors; minimum requirement is to discern relative age/experience relative to task; screen out those formed just to get at funds.**

Ideally, each grant or contract would include a mechanism whereby every employee could provide feedback anonymously. Providing each with a card or a unique number that could be used to access a reporting terminal could give voice to people who can see what is happening but have no mechanism to report it. Adding an incentive if the information proves out could be a useful addition.

2 Positive Votes

Authored by [DrD1](#)

10/22/2011 6:10:46 PM

0 [Responses](#)

Tags [employee-access](#)

**Idea 33 - Provide card access vehicle for employees to give anonymous feedback on grants and contracts and consider adding incentive if information proves useful in identifying fraud.**

Outcomes need to be measured at the stakeholder level where the targeted stimulus was meant to go. For a construction project, this means the construction worker (one could argue the contractor as well, but they make their money off a percentage of the total bid, so the business stimuli exist when the system works properly). A typical project might have a contract specialist who acts as a compliance officer, but who's main duty is to complete the project, not root out wage theft. The enforcement agency, DOL, will only investigate a complaint by a worker. The workers are either job scared, or the process is unknown to them so they make no complaint. Therefore, the targeted funds never make it to the hands of the workers. The outcomes are measured internally, at the DOL or the contracting agency level. For a construction project, this means that the true outcome (jobs) is never truly measured. While not all funds are construction related, they do represent a large percentage. As a third party, representing workers, I have witnessed wage theft on multiple projects in the Tampa area, across multiple agencies. Many of these have gone unresolved. And, it is happening as I write this.

4 Positive Votes, 1 Negative Vote

Authored by [jond](#)

10/19/2011 9:39:52 AM

0 [Responses](#)

**Idea 34 - Measure outcomes at stakeholder level where the stimulus money actually goes. This will help identify such frauds as wage theft.**

How can online dashboards be utilized for organizing and disseminating data? Would these be an effective means of tracking fraud, waste, and abuse of Recovery funds?

<http://www.federalnewsradio.com/?nid=242&sid=2596204>

2 Positive Votes

Authored by [ladyinred](#)

10/18/2011 10:07:19 AM

1 [Responses](#)

Tags [online](#), [dashboards](#), [performance](#), [metrics](#)

*Comment 34.1 – Would “open government” and “crowd oversight” help identify fraud?*

Open Government and "crowd oversight"... Would opening "the books" for the public to see all aide in identifying fraud or fraud patterns? Would it reduce fraud? Are there examples of this practice?

[dc1191](#)

**Idea 35 - Make annual recovery audits mandatory at each cabinet level and quasi-governmental agency to reduce \$125 billion in annual improper payments. Upload software in every agency’s accounts payable system.**

Federal agencies cumulatively admitted to \$125 Billion in improper payments in FY'10. While the "supercommittee" meets to discuss ways to cut inefficient programs, we are strongly encouraging them to stop this hemorrhaging source of fraud against taxpayer dollars by implementing fraud prevention technology and mandatory annual recovery audits at each cabinet-level and quasi-governmental agency. That's \$125B in taxpayer dollars we can save annually by uploading software in every agency's accounts payable system.

<http://paymentaccuracy.usaspending.gov/high-priority-programs>

<http://www.apexanalytix.com/industry/government/default.aspx>

2 Positive Votes, 1 Negative Vote

Authored by [mceoxley](#)

10/17/2011 7:00:40 PM

0 [Responses](#)

**Idea 36 - Each program should have explicit, measurable input and outcome goals from the outset, with logical relationships between outputs and outcomes. Require periodic reports of negative trends/failures and determine and address reasons.**

Each program should have explicit, measurable output and outcome goals defined at the outset. The outcome goals can be in the future, i. e., beyond the term of the program, and they can be affected by programs other than the one for which measures are being established. However, there should be a logical relationship described between the outputs and the outcomes. The program should then report the actual outputs (and outcomes if possible) periodically. The reasons for any negative trends or failure to reach the performance goals should be determined, articulated, and addressed.

7 Positive Votes, 1 Negative Vote

Authored by [halsteinberg](#)

10/17/2011 11:21:55 AM

1 [Responses](#)

Tags [hal](#)

*Comment 36.1 – The most difficult part is to establish metrics that validly measure outcome goals.*

The most difficult part is to establish metrics that validly measure outcome goals. Can anyone contribute thoughts for a non-for profit organization.

[dwho](#)







**900 7<sup>th</sup> Street, N.W.  
Suite 600  
Washington, DC 20001-3888  
Phone: (202) 347-3190  
Fax: (202) 393-0993  
[www.napawash.org](http://www.napawash.org)**