Idea Papers for Strengthening Federal Spending Transparency: A Working Conference to Develop a Plan of Action

May 20, 2010
Washington, D.C.

Contents

The papers in this document are the views of their authors and have not been edited for content by OMB Watch.

CREATING A 21ST CENTURY AUTHENTICATED ELECTRONIC FEDERAL FINANCIAL FRAMEWORK

WHITE PAPER

FINANCIAL SPENDING TRANSPARENCY

May 2010

Prepared by:
Henry N. Dreifus
Dreifus Associates Limited

Vision: Establish a 21st century, standards-based, secure electronic financial framework across government that provides accurate, interoperable, fully auditable electronic accounting with openness and transparency.

STATEMENT OF THE PROBLEM

The Federal Budget is one of the most complex and essential instruments needed to run the country. The steps and efforts it takes to prepare, justify, approve, appropriate, distribute and report results of the federal budget is a hugely complex and time-consuming task for both the Legislative and Executive Branches.

Significant gaps exist today in the financial framework that supports this effort across the organizations and the different components. These gaps impede accuracy, accountability, visibility, precision and timeliness, consequently reducing the effectiveness of decision-making by stakeholders at all levels and incurring wasteful and unnecessary expenditures. The Federal Budget begins at the activity / agency level which is rolled-up into the "Presidents Budget" which is annually submitted to Congress as a "starting point" to determine how tax payer monies should be allocated to meet the needs of the nation. Although there is a tremendous amount of good work being done across the Executive branch – including developing electronic budget creating tools and ever improving accounting systems, there is no unifying vision to build a connected model to harmonize these initiatives and thus create an even greater accretive return for tax payer investment.

Background—A Modern "Golden Spike"

Congress enacted the Pacific Railway Act of 1862 *(U.S. Statutes at large, vol. XII, p. 489)* creating the vital transcontinental link that is considered one of the greatest American technological feats of the 19th century. At Promontory Summit in Utah on May 10, 1869 the final link was completed with a ceremonial "golden spike" connecting the eastern and western portions of the rail system to create a fully connected, single gauge standard rail network which historians credit as the leading economic reason the United States was able to capitalize on the Industrial Revolution. Today our national rail network remains critical to commerce, national security and our continued national success.

The current federal financial framework functions as a series of largely unconnected networks, analogous to the rail networks prior to the Pacific Railway Act, with neither consistent standards nor interoperable electronic "rails". Creating a more precise, efficient and modern framework

that standardizes and connects the 21st century electronic rails, as was done with the physical rails in 1869, will establish a more accountable and transparent Government, enabling our Executive and Legislative leadership to succeed as world class managers of public resources. Bridging the electronic rails between Congress and the Executive Branch with an electronic "golden spike" can become one of the greatest 21st century Government modernization accomplishments, in keeping with the both the Federal Funding Accountability and Transparency Act of 2006 and the current administration's commitment to an open and transparent Government[1].

Current Challenges

As in many large enterprises, the Federal financial system evolved as a series of incompatible computer systems developed and built at different times for each entity rather than holistically across the Government. Because individual systems are not interoperable, actual funds executed cannot be easily linked with a corresponding activity or appropriation with the required precision and timeliness. Because the individual components are not electronically connected, this current financial infrastructure makes it difficult to achieve transparency and accountability which results in ineffectiveness that creates significant, unnecessary cost burdens.

Developing a way to better describe outcomes not just outputs is a challenge for any enterprise, but especially for the Federal Government due to its large size, complexity and interdependencies – both within and between agencies and as well with the State and Local Governments. Another challenge is the ability to align the program elements to across the entire life cycle – from budget development/submission to Congress to final appropriations – through the passage and signing into law as well as performing oversight and measuring outcomes. Achieving an information age approach that provides for a more transparent and accountable outcome can be accomplished only when the Legislative and Executive branches have a better set of tools through which information can be contextually processed. This "life cycle" is still tremendously a manual and partially semi-automated process, one that has many disconnected parts.

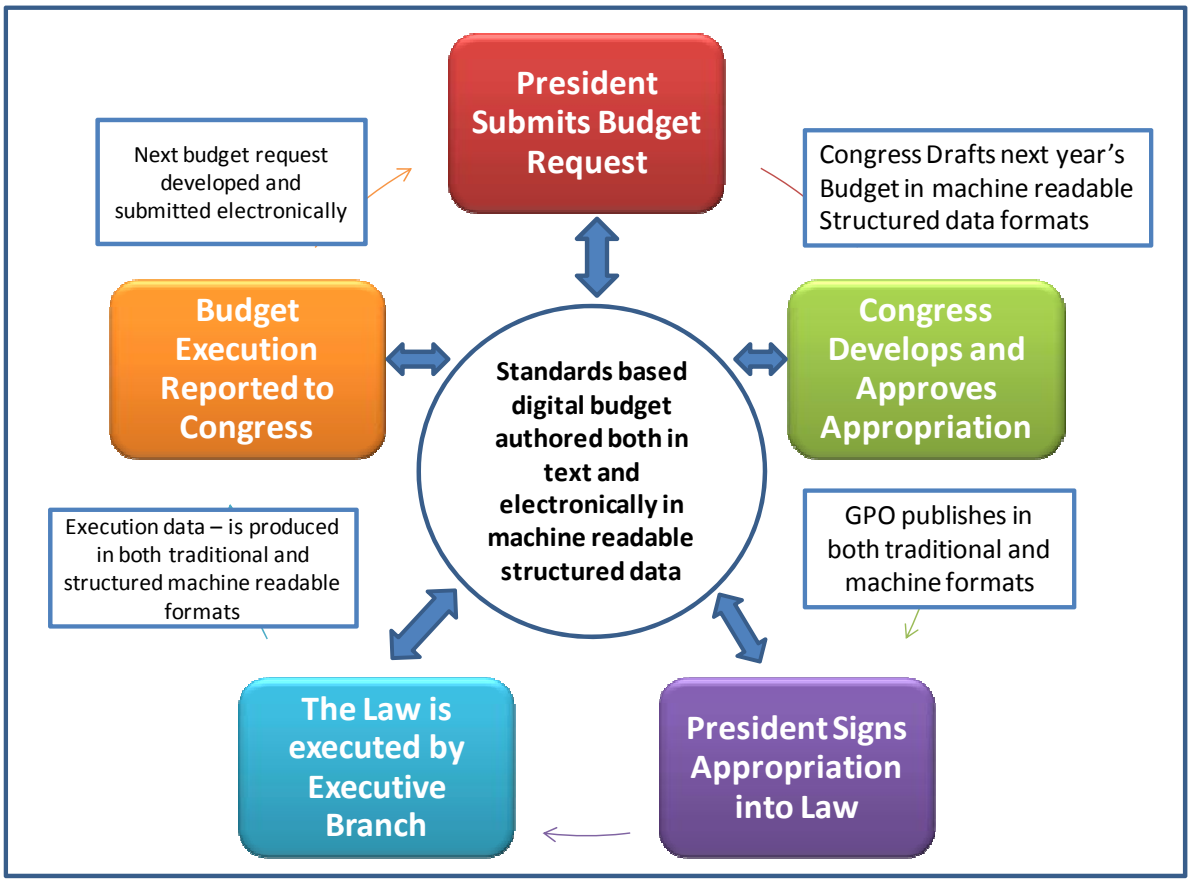
Today an urgent need exists to improve the many shortcomings in the federal budget process. The Peterson-Pew Commission on Budget Reform notes, amongst other things, there is a need to make the accounting standards more modern and consistent to improve transparency and sustainability. According to the Peterson-Pew Commission, "significant emphasis on budget reform over the next few years is imperative given the current economic conditions, the tremendous amount of borrowing taking place to help stabilize the economy, and the large and growing promises for Social Security, Medicare and Medicaid benefits".[2]

Very often Congress will ask the Executive Branch for reports on specific programs and activities, which are not easily understood by the present state-of-the-market reporting methods. Further Congress and their corresponding accountants and analysts at the Government Accountability Office (GAO) must link this and other data along with spending reports against the relevant section of the bill, Act or Appropriations. Using today's manual processes, it is

[1] http://www.whitehouse.gov/the_press_office/Transparency_and_Open_Government/
[2] http://www.pewtrusts.org/news_room_detail.aspx?id=48156

estimated that a minimum of $25,000 in staff time is required for each mark to determine if the grantee was qualified to receive a mark, to provide oversight, to audit the expenditure, perform follow up, etc. With around 9,000 marks in the 2009 spending bill alone[3] (not including the American Recovery and Reinvestment Act of 2009), the time required equates to a minimum of $225,000,000 of taxpayers' money.

RECOMMENDATIONS TO MOVE FORWARD



The Administration and Congress share a goal to provide improved accountability and transparency as world class mangers of our nation's finances. We propose the implementation of a state-of-the-art, secure electronic financial framework that provides a better, more cost effective, consistent and interoperable solution. The framework will provide a pathway to reconcile budget, appropriations, and execution data without having to fundamentally change each underlying financial system. It will improve the way Congress and the Executive Branch communicate data with greater precision and visibility by leveraging current technical and financial infrastructures -- resulting in a better, more cost effective, consistent and interoperable solution to this long standing problem.

[3] http://www.heritage.org/research/budget/wm2318.cfm

Near-Term Solutions: Capacity-Building and Pilots

As this figure demonstrates, the Budget, Appropriations, and Execution process all rely on information about dollars that are requested by the Administration in the President's Budget, enacted by Congress in Appropriations law, and spent by Agencies through authorized programs that deliver results. Although each entity in this chain tracks the same dollars for the same programs and purposes, there is no common standard to "follow the money" and track results.

The current electronic financial framework is close to being able to support this beginning-to-end ability but lacks a common computerized standard that would allow linking of information between disparate systems and enable enhanced interoperation. Adopting and applying proven, open financial data standards applying best practices machine readable structured data formats to the existing infrastructure will yield the granularity and integrity necessary to meet the following four goals using this new digital framework to provide:

1: Greater precision and timeliness in budget development and execution reconciliation,
2: Improved understanding of programs to enable better management decision-making,
3: Better information to better enable auditable financials, and
4: Enhanced transparency and accountability.

Specifically, over the past decade both technical and legislative steps to modernize data communications within the Federal Government have significantly improved the process. Congress now generates most authorizing legislation using industry-standard electronic formats (Extensible Markup Language or "XML"). By making incremental enhancements to this existing framework, the electronic creation and communication of the corresponding Appropriations legislation and expenditure guidance would enable an enhanced and improved electronic financial supply chain, using the very same industry standards and proven tools. Most of the enabling steps and resources are already in place including key laws and statutes (e.g.: The Government Paperwork Elimination Act of 1998, the federal E-Signatures in Global and National Commerce Act of 2000, and the state Uniform Electronic Transactions Act of 1999). In addition, much of the work established by the "Recovery.gov" efforts can be applied to this initiative

We propose to leverage the XML schema used for current legislation by adding "tags" that would be associated with appropriated dollars anywhere they flow in the spending chain. By starting with the production of appropriations bills from the Government Printing Office (GPO), the Administration could then accept those tags – connecting a modern "Golden Spike" with the Congress -- and require that financial and performance data transfer and reports map to those same tags to ensure that they flow with the money. Once this framework is in place, we would envision moving toward full end-to-end reporting along the lines of agreed-upon data standards, starting with the preparation of the President's budget and ending with performance reporting that is associated with appropriated funds.

This approach carries significant and proven benefits. Governments in Europe and Asia are adopting not only single electronic data standards to improve efficiency and precision, but employing powerful analytic tools and capabilities to be better stewards of public funds. For

example, The Netherlands estimates the country will save 25 percent[4] of its budget preparation and execution costs in the first year alone of implementing an interoperable data standard solution using machine readable structured data -- such as what Congress presently utilizes for drafting the majority of the "words" for legislation based on Extensible Markup Language (XML). In the United States, projecting similar results would equate to estimated savings of at least $2 billion per year as well as providing better information "on demand" across all stakeholders – including the American tax payer – while adopting global best practices in electronic financial data standardization.

The Constitutional Imperative

Driving this "Golden Spike" to create a fully electronic financial framework enables this Administration and Congress to deliver on the promise of a more accountable and transparent Government in a constitutionally consistent manner. The Constitution (*Article 1, Section 7*) already provides a consistent framework to implement an electronic enhanced appropriation:

> ***All bills for raising revenue shall originate in the House of Representatives; but the Senate may propose or concur with amendments as on other Bills. Every bill which shall have passed the House of Representatives and the Senate, shall, before it become a law, be presented to the President of the United States; if he approve he shall sign it, but if not he shall return it, with his objections to that House in which it shall have originated, who shall enter the objections at large on their journal, and proceed to reconsider it. If after such reconsideration two thirds of that House shall agree to pass the bill, it shall be sent, together with the objections, to the other House, by which it shall likewise be reconsidered, and if approved by two thirds of that House, it shall become a law.***

The Constitution already provides for the basic design requirement both directly and explicitly: ***Present the law***. This presentation from Congress to the President can and could be done electronically. The proposed solution is consistent with the current processes and procedures. What is done today with paper and many manual steps, will be achieved far more efficiently and effectively using an aligned electronic approach.

Longer Term Goal: Achieving A Fully Electronic Financial Framework

Establishing a working partnership by and between stakeholders inside and outside the government -- including the Public Interest, Academic, Legislative and Executive Branches -- that forms a community of interest to catalyze and develop the framework to leverages the systems and inherent capabilities currently available in the Federal Government with coupling components to help to develop the electronic "rails"; considers the legislative, management process and technological aspects of the requirements; and fashions an actionable approach which can be quickly and adroitly executed to build a solution.

A standards-based electronic financial framework is the core element toward enabling an open, accountable and transparent Government at both federal and state/local levels; reaching that goal will be prohibitively difficult without the ability to seamlessly transfer financial data automatically (and on request) between authorized entities. Ultimately, this framework can enable:

[4] http://www.xbrl-ntp.nl/.../Holland_Takes_a_Different_Tack_on_XBRL_article_Compliance.pdf

- More precise and timely understanding of “execution to budget” that affords management better decision-making information.
- Greater precision in budget development, drafting and coordination at all levels throughout the budget-execution lifecycle.
- Improved situational awareness, providing greater insight for more timely and effective strategic and tactical decision-making.
- An improved framework for enabling the ability to create auditable financials throughout the enterprise.
- Better use of automated management and visualization tools for improving transparency and accountability to the American public.
- Electronic audit capability for federal, state and local Government linked expenditures and programs.
- More efficient use of resources and scarce senior management bandwidth; reducing costs and enabling greater staff productivity.
- Improved effectiveness of public resources, which is especially critical for Defense, Healthcare and other large and complex programs.

Steps for Integrating Tax Expenditures into the Budget Process and Increasing Transparency

Center for American Progress

Tax Expenditures Are Spending

Tax expenditures are a form of government spending that are implemented through the tax code. They may feel a bit different than direct spending, but for the most part, bear close resemblance to mandatory and discretionary spending. Integrating tax expenditures into the budget process and increasing transparency of tax expenditure spending is critical for more accountable spending.

The following steps outline a possible framework for incorporating tax expenditures into the budget process.

- **Categorization.** A key part of integration would be categorizing tax expenditure spending like direct spending into discretionary and mandatory spending categories.

- **Length of sunsets could reflect the category.** Tax expenditures that are subject to sunsets resemble discretionary and mandatory non-entitlement programs, depending on the length of the sunset. Those that are subject to annual sunset are most similar to discretionary programs, while those that are subject to longer term sunsets are more like mandatory non-entitlement programs (however, unlike mandatory programs, spending amounts are not capped). Tax expenditures that are not subject to sunset resemble mandatory entitlement programs—both are uncapped and insulated from the authorization as well as appropriation process.

- **Integration.** Discretionary tax expenditures would be subject to annual authorization (which automatically triggers appropriation), and worked into the appropriation process under the budget rules (specifically, the § 302(a) allocation specified in the Congressional Budget Act of 1974). Mandatory tax expenditures would not be authorized annually, but may be subject to reauthorization in a given year.

Tax Expenditures and Transparency

Tax expenditure spending should also be more transparent. The IRS collects a vast amount of useful information that can inform analysts and policymakers both inside and outside of government. However, meaningful data on the distributional and economic impacts of tax expenditure spending is limited. This makes assessment of tax expenditure spending challenging. Overcoming obstacles posed by the IRS' confidentiality rules, scope of data, and research capacities, will be key to developing a central system that can inform analysts and policymakers both inside and outside of government.

- **Data Collection.** Tax data is universal – all individuals and businesses must file tax returns. This means the tax system captures important data beyond traditional tax concepts. Important measures, like firm entries and exits as well as employment totals are compiled through employment tax returns filed by businesses. Other useful data points include socioeconomic measures such as marital status, education expenditures, major medical expenses; disability status; charitable contributions; home ownership patterns; family demographic information; and so on.

- **Electronic Filing.** Electronic filing provides the potential for improved data quality since it reduces error associated with transferring data from hardcopy tax returns to an electronic format. Currently, select data from hardcopies is recorded electronically, which leaves vast amounts of potentially informative data inaccessible. The existence of e-file data increases the potential for generating more information, and for manipulating data in ways that can shed light on the distribution and effectiveness of tax expenditures.

- **Confidentiality Rules.** Federal tax information is treated homogenously because of the IRS' confidentiality rules. These rules do not allow the IRS to distinguish among types of data, even with regards to age. This means that information about all taxpayers, whether an individual's gross income or a large corporation's financial information, receive the same type of protection. These rules impact how the IRS, which is required to regularly produce statistical information, aggregates and organizes publicly released data. Variations on this data, including different classification categories or special studies, can be requested by outside users but resource limitations limit the IRS' ability to fulfill these requests. Additional data beyond what is released can only be accessed by an outside entity, including a Federal agency, if a statute provides such entity explicit authorization.

  However, new technology coding systems may make it possible for the IRS to protect taxpayer anonymity while also providing data for assessment. The current standard for confidentiality requires that data be released in a manner that prevents outsiders from using "reasonable means" to identify a taxpayer. It does not specify how such data be released. New systems for data sorting would ideally take advantage of the flexibility allowed for presentation, without requiring the IRS to change its confidentiality rules. For example, coding microdata with various levels of confidentiality may prevent identification of a taxpayer while also preserving data relations at a micro level which are important to producing analysis useful for decisionmakers.

- **Coordinated Research Center.** A coordinated center for assessing and inputting data might be critical to maintaining the balance between taxpayer privacy and data utility. This could potentially be housed in the IRS, since IRS employees have access to raw tax information under the confidentiality rules.

Federal Spending Transparency: Unlocking the Power of Abstraction

Presented to OMBWatch's
Conference to Develop a Long Range Vision for Federal Spending Transparency
May 20, 2010

Jim Harper
Director of Information Policy Studies, The Cato Institute
and Webmaster, WashingtonWatch.com

Abstraction is the idea of considering the general characteristics of things apart from concrete, specific instances. Abstraction is a powerful tool for comprehending our world—and for remaking it.

Consider the abstractions that we use every day, things like language, numbers, and time. Language is the use of uniform sounds to describe things. The invention of language (and the logic that underlies it) is a big part of what separates humans from other members of the animal kingdom. Writing is a further abstraction on language that improved humans' abilities to catalog, comprehend, and improve their lives.

Numbering is a similar abstraction that allows people to measure and compare quantities of things, a huge innovation from early in human history. Without it, life would be unbearable by today's standards.

Time is another abstraction. It describes the movement of the earth with respect to the sun. The uniform system we use for describing "time" lets people synchronize their activities with others, a tremendous aid to living in an organized society.

More recent advances in abstraction show how it continues to improve our ability to work with the world around us. Expressing letters, numbers, sights, sounds, and symbols as 1s and 0s—digitization—is a way of abstracting information that was central to the invention of computing and the Internet. Without digitization, us policy wonks would still be waiting for Xerox copies of bills to be delivered to us by mail or courier.

The Internet protocol (TCP/IP) is an abstraction that rode on the invention of digitization. It's an abstract way for computers to talk to one another. The Internet protocol paved the way for html (hypertext markup language) and the World Wide Web. These invented abstractions deliver information in readily usable form to computers and connected devices across the globe.

Abstraction is powerful because it allows people to work together, using agreed standards for communicating information, to solve the problems in their lives. It can definitely help solve problems in the area of federal spending.

Earlier this year, a small group of earmark transparency activists put together an abstract model for describing earmarks. Our work is presented at Earmarkdata.org/schema and it is attached to this paper as an appendix. The goal, simply put, is to get information about earmarks from Congress in an abstract, and thus useful, form.

Getting earmark data will allow web sites, researchers, reporters, political scientists, and the public to manipulate and use earmark information in any way they choose. People will be able to learn more about whatever they are interested in: They will be able to more easily compare earmark requests and awards with campaign contributions, political party, or seniority, for example. They will be able to make whatever arguments they want to about particular earmarks, earmarking processes, or the practice of earmarking itself. The great thing about abstraction is that it permits new and innovative uses of data—uses we won't know about in advance.

The earmark data model consists of three basic elements:

- **Entities**: "Entities" are *things*. An entity might be an elected official, a bill, or specific provision of a bill. Agencies, programs, contractors, and grantees are all likely entities to describe in a data model.

- **Properties**: Entities are made up of a collection of *properties*—the characteristics that make an entity what it is. The properties of an elected official entity would include things like first name, last name, state and district represented, and so on. One important property of a bill is its bill number (e.g., H.R. 123, S. 1020). Other properties of bills might be the committee(s) they are referred to, their stage in the legislative process, their texts at different stages, and vote tallies on them.

- **Entity-Properties**: An entity-property is a characteristic—a property—that is another defined entity. When a bill (an entity) has as a property that a particular elected official (also an entity) co-sponsored it, that's an entity-property of the bill. (Likewise, the bills an elected official co-sponsors can be entity-properties of the official.)

Entity-properties allow data users to weave together the "stories" they're interested in. What bills (entities) did a given senator (an entity-property of bills) introduce? And how are these bills' other properties similar—the subjects they affect, their passage rates, and so on? The answers tell us things about the senator who introduced them.

Adding new entities, to be used as entity-properties, expands the range of questions that analysts can answer. Did elected officials (entities) with a common donor (entity-property) vote disproportionately for a given bill (entity)? Did bills (entities) passing through a particular committee (entity-property) favor a particular agency (entity) more than other committees did?

These are rough cuts at using a formally organized schema to talk about familiar issues in public policy. Abstracting the policy process into a well organized "language" like this, then getting public policy information in machine-readable formats consistent with this language, will allow advocates, researchers, reporters, web sites, and the public to investigate questions like these, and many, many more. With structured data, the answers to public policy questions will be much easier to come by, and they will have a more solid grounding in facts and statistics than the answers we work with today.

Two Years to a Federal Budget Data Schema

To get federal budget data in useful formats, subject matter experts should work to express the things they talk about in the abstracted language of "data" described above. Each of the things involved in the

federal budget and spending process should be reduced to description as “entities” with “properties”—with the relationships among them signaled by “entity-properties.”

The entities to be described will probably include elected officials, budget documents and line-items in budget documents, agencies and sub-agencies, government programs, bills and line-items in bills, earmarks, contracts and grants, government contractors, grantees, and so on.

Each entity will have a set of characteristics or properties that go into the entity description, with defined ways of referring to these characteristics. U.S. federal legislation, for example, has relatively standard naming conventions: a designation of the bill’s type and the house in which it was introduced followed by a number: H.R. 1234 or S. 1020, for example. A bill has only one such bill number. Likewise, bills have as a characteristic the one, and only one, Congress they were introduced in: presently, a number between 1 and 111. Was a bill also introduced in earlier or later Congresses? That’s a characteristic of a bill: having a predecessor or successor in a different Congress.

Bills may be referred to multiple committees. Thus they have as a characteristic a number of referrals defined by the set of congressional committee entities. Bills may have as a characteristic an unlimited number of “supporters” or “opponents” (entity-properties) outside the formal legislative process, though determining support or opposition is likely to be a subjective judgment compared to the objective signal given by legislators’ votes.

These are brief examples of the structured ways to describe parts of the U.S. federal legislative process. This kind of thinking should go into structuring all federal budget and spending processes over the next two years.

There are many details to think through so that the schema can capture all relevant information while maintaining flexibility and openness to extension so new, related uses of this information scheme can be adopted.

Ten Years to Transparent Federal Budget Data

Over the coming decade, the transparency community should work with government agencies to ingratiate this kind of data-oriented reporting into government operations. Budgets, bills, agencies, and programs should be reported in a structured way, producing data that is consistent, machine-readable, and amenable to data processing by all segments of the policy community.

Indeed, it’s not just reporting. Each piece of the policy making process—the budgets, bills, votes, etc.—should *originate* as structured data, feeding directly into the information infrastructure that the transparency community creates. A budget should come out not just in paper and PDF version, but as a data set containing all the meaning that exists in the physical documents.

The many different parts of the policy process can adopt data-oriented reporting and origination at their own pace. As more of them do, the pieces can be woven together if an overarching schema gives the field of policymaking some consistency. There is much to be done, but the work will be easier and easier as examples of structured policymaking data come into existence and flourish.

Asking government agencies for “transparency” is a little bit like going to a delicatessen and asking for “a sandwich.” It begs lots of questions: “What kind of bread? What kind of meat—or veggie? Lettuce, tomatoes, onions? Oil and vinegar? Mustard? Hot or cold? Maybe toasted? Chips or a pickle? For here or to go?”

The transparency community should “place our order” by building the intellectual infrastructure for federal spending transparency, making clear exactly what information we want and in what form. This work is a small investment compared to the large dividends it will pay in decades to come. It will eliminate a current impediment to transparency by telling the information originators in government what the transparency community wants.

Data-oriented reporting of government activity will make it easier for all interest groups and actors to involve themselves in public policy and to advocate their views. Abstracting public policy processes as described here will help to create an information platform for all interests to use. Ideology, popularity, and rhetoric—the stuff of old-fashioned politics—will always have its place, but the debates that collectively determine the public interest will be better if they rest on good information made available to all.

Appendix — The Earmarkdata.org Data Schema

What follows is the earmark data schema proposed by the Earmarkdata.org project.

Earmark Requests

An earmark request is any communication from a senator or member of Congress to a congressional committee requesting legislative provisions that set aside funds for a specific program, project, activity, institution, or location. These measures normally circumvent merit-based or competitive allocation processes and appear in spending, authorization, tax, and tariff bills.

An earmark request is uniquely identified by its **earmarkrequestid** taken together with its **fiscalyear**.

The properties of an earmark request entity are:

- **earmarkrequestid** *(required)* A monotonically-increasing number unique in a given fiscal year. This number must be assigned by a central authority. A value of "president" is reserved and not legal in earmark request entities. (This is discussed further under Earmark entities.)
- **fiscalyear** *(required)* The fiscal year for which the allocation of funds is requested.
- **projectname** *(required)* The name of the project on which the funds are requested to be spent.
- **amount** *(required)* The amount of money requested.
- **description** A description of the purpose of the earmark requested.
- **date** The date of authorship of the request letter.
- **source** If available, a URI identifying an internet-accessible source for the information contained in the properties of this entity. Ideally, this should be the complete text of the letter.
- **[earmarkrequester]** *(required)* The entity who wrote the request letter (see below). This is always a sitting senator or representative.
- **[beneficiary]** *(one or many required)* One or more recipient beneficiaries (see below). A beneficiary is the entity for whom the funds are allocated.

Earmarks

An earmark is a legislative provision that sets aside funds for a specific program, project, activity, institution, or location. Earmarks may be included in appropriations or authorizations bills. *(See* **citation** *below)*

An earmark request does not always become an earmark, but all earmarks should be associated with an earmark request. If an earmark is not associated with a request, that will be either an undisclosed earmark or a presidentially requested earmark. An earmark will have many properties similar to an earmark request, but these properties are not redundant since they belong to an earmark *as included in legislation*, rather than *as requested*.

The properties of an earmark:

- **earmarkid** *(required)* A monotonically-increasing number unique in a given fiscal year. This number must be assigned by the data manager.
- **fiscalyear** *(required)* The fiscal year for which the earmark directs allocation of funds.
- **type** *(required)* The type of allocation: either "appropriation" or "authorization".
- **[earmarkrequest]** *(zero or one or many)* If the earmark originated from one or more earmark requests, these must be identified by the value of their earmarkrequestids. The referenced

earmark request entity must have been issued in the same fiscal year as the earmark entity. Where this property is present but empty, this earmark is understood to be an **undisclosed earmark**. Undisclosed earmarks are provisions that meet the definition of an earmark but that aren't expressly disclosed in report language or legislation as an earmark. If an earmark was requested by the president, the reserved earmarkrequestid "president" must be used; in this case the value does not indicate a particular Earmark Request entity but merely asserts that the president was among those who requested the earmark.

- **norequesterreason** *(required if* **earmarkrequestid** *has zero values)* If the number of earmark requests named in the earmark is zero, this property *should* be present and explain why no requesters are named or any circumstances surrounding an undisclosed earmark.
- **projectname** *(required)* The project as named in the earmark.
- **amount** *(required)* Amount of the earmark.
- **amountafteradjustment** This amount is the net amount after any congressionally-mandated adjustments such as across the board reductions.
- **description** A description of the purpose of the earmark.
- **[citation]** *(one and only one required)* An entity describing the legislation containing the earmark and its location. An allocation of funds for the same purpose or project in a different piece of legislation is a different earmark.
- **[beneficiary]** *(one or more required)* The organization receiving the earmark (see below).

Citation

A citation is an entity describing in an unambiguous way the legislation or legislative document containing the earmark. This entity must contain all information necessary for a member of the public to locate the document and the part of the document that includes the earmark. As far as possible, this entity should have format properties or include additional properties to facilitate machine-processing so that these citations can be easily identified and cross-referenced.

- **earmarkbill** *(one and only one required)* A consistent and unambiguous identification of the bill that contains the earmark, e.g. an appropriations bill. Suggested is congressional session, followed by Thomas bill database bill type code, followed by bill number, e.g. 111-H-2997. This property is to aid machine processing and cross-referencing earmark allocations with bills.
- **earmarkreport** This property is required if the earmark is contained in a report accompanying the bill. The property should a consistent and unambiguous identification of the report that contains the earmark (e.g, H, Rept. 111-123, S. Rept. 110-987, Conf. Rept. 109-231, J. Rept. 111-212). This property is to aid machine processing and cross-referencing earmark allocations with bills.
- **location** Specific location of the earmark within the document indicated by **earmarkbill** or **earmarkreport**. Includes information that enables the public to locate the earmark in the document (e.g. page number, line number, section number, subsection, paragraph, etc).
- **excerpt** Relevant excerpts of the actual language creating the earmark. This field should include language incorporating by reference any committee or conference reports or other legislative documents.
- **link** *(zero or one or more)* URLs of resources specific to this earmark.

Earmark Requesters

An earmark requester is a senator or representative who submits a particular earmark request or earmark.

Note that if a **govtrackpersonid** property is not present, there is no guarantee that the remaining properties will be sufficient to uniquely identify the requester.

Aside from the **govtrackpersonid** property, all properties should be interpreted as applicable to that person at the time indicated by the **date** property of the earmark request.

- **govtrackpersonid** An id number from the govtrack database uniquely identifying the person. It is *highly recommended* that this property be present, since it is a well known machine-processable, unique, and temporally-unambiguous identifier of a member of Congress. If this property is present, no other properties of this entity are necessary. Conversely, if this property is absent, all other properties of this entity are required.
- **type** The type of congressman (senator or representative) of this requester.
- **state** Two-letter state code identifying the state which this requester represents.
- **districtorclass** The interpretation of this property depends upon the **type** property. For members of the house of representatives, it is the district (or 0 for at-large) which the requester represents. For senators, it is their election class (1, 2, or 3).
- **firstname** The first name of the requester.
- **lastname** The last name of the requester.

Beneficiaries

A beneficiary is an organization to receive funds as proposed in an earmark request or as directed in an earmark.

- **duns** A Dun & Bradstreet Data Universal Numbering System (DUNS) number uniquely identifying the recipient. If present, many other properties of this entity become optional because their values are inferred from the information present in the DUNS database.
- **name** *(optional if* **duns** *present)* The name of the state, locality, business, nonprofit, or other organization receiving funds as proposed in the earmark request or as directed in the earmark.
- **address** *(optional if* **duns** *present)* Street address of recipient/beneficiary.
- **city** *(optional if* **duns** *present)* Locality of the recipient/beneficiary.
- **state** *(optional if* **duns** *present)* The state, province, or territory of the recipient/beneficiary.
- **zip** *(optional if* **duns** *present)* The zip code or postal code of the recipient/beneficiary.
- **country** *(optional if* **duns** *present)* Country of the recipient/beneficiary.

If both the **duns** property and the additional optional properties are present but the entry in the DUNS database contains information which does not correspond to the values of the optional properties, this specification does not define which set of information is more authoritative.

1616 P Street NW, Suite 210
Washington, DC 20036
202-232-1616 fax: 202-232-6680
www.goodjobsfirst.org

JOB REPORTING AS A COMPONENT OF FEDERAL SPENDING TRANSPARENCY

By Philip Mattera, Good Jobs First

May 2010

A substantial portion of federal spending results in the creation of jobs in the private and non-profit sectors of the economy. Some forms of spending are explicitly intended to generate jobs, such as the American Recovery and Reinvestment Act (ARRA) and economic development programs such as Community Development Block Grants. Other programs such as federal highway spending are often justified in terms of their employment impact. Federal contract awards of all kinds end up creating large numbers of jobs.

The federal government has generally not put much emphasis on measuring the direct employment effects of its spending. Agencies such as the Council of Economic Advisors and the Bureau of Economic Analysis use macroeconomic models to assess (among other things) the overall employment impact of government as well as private spending in assessing the state of the entire economy, but they do not look at specific jobs.

ARRA, on the other hand, has given rise to a system in which tens of thousands of contract and grant recipients must report the number of jobs being supported by Recovery Act funds. The system is not flawless, but it is an important step forward in disclosing a key impact of federal spending.

The ARRA system can serve as a stepping stone to broader job reporting for federal spending. The goals of such a system would include:

- An enhanced ability to evaluate program performance;
- A higher level of accountability among contract and grant recipients; and
- An ability to do equity analyses.

To achieve these goals, a broad job reporting system would have to provide even more job data than is currently mandated by ARRA. The main categories would be:

- Data on the number of jobs created by direct recipients of contracts and grants;
- Data on the quality of those jobs, especially average/median wage rates and the portion of workers receiving key benefits such as healthcare coverage; and
- Demographic information on the workers hired for those jobs, including gender, race and age as well as nine-digit residential Zip code (with the privacy of individuals protected).

This kind of reporting system is meant to cover only direct employment impacts at companies and other organizations receiving contracts and grants. It would not deal with indirect job creation at firms that supply goods and services to contract or grant recipients, nor other jobs resulting from the overall improvement in economic conditions that the contract or grant may help bring about.

The system, to the extent possible, should build on existing reporting requirements faced by employers, such as the Bureau of Labor Statistics Establishment Survey and the Equal Employment Opportunity Commission's EEO-1 Survey.

KEY STEPS

I. *Enhancing ARRA job reporting*. Because many thousands of employers are already reporting job numbers to the federal government under ARRA, it makes sense to start with expanding this system to cover job quality and worker demographics. OMB could revise its ARRA guidance to require these additional data fields beginning with one of the upcoming quarterly data rounds (of which there are five remaining).

II. *Extending to economic development programs*. Over a period of about two years, the federal government could adapt the ARRA reporting system to key economic development programs such as Community Development Block Grants, the Commerce Department's Public Works and Economic Development Program, Empowerment Zones and the New Markets Tax Credit.

III. *Extending to all direct contracts and grants to employers*. Over a longer period of five to ten years, the federal government could extend the system to all for-profit and non-profit employers who receive direct contracts and grants.

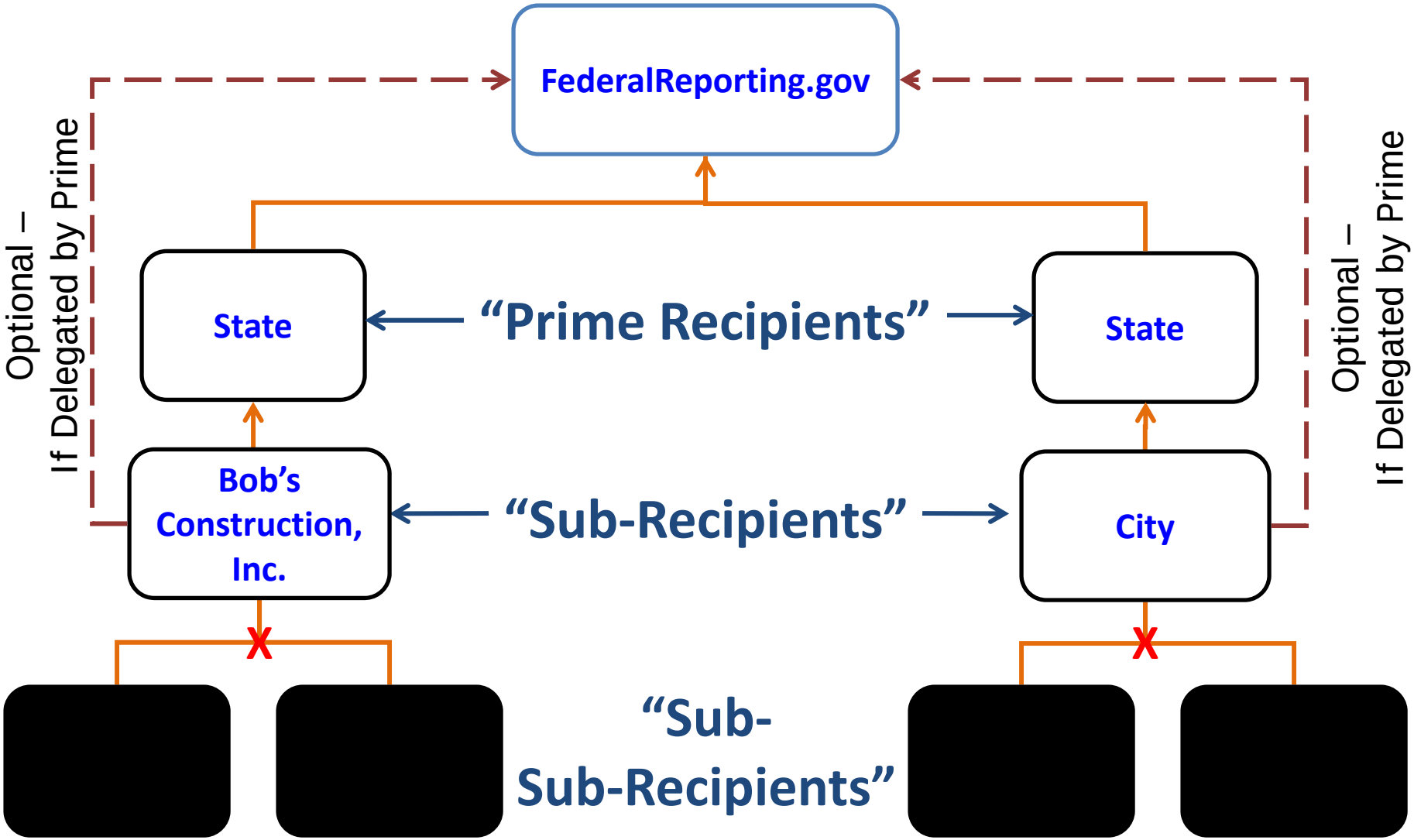
Reporting & Public Access to Federal Spending Information

May 20, 2010

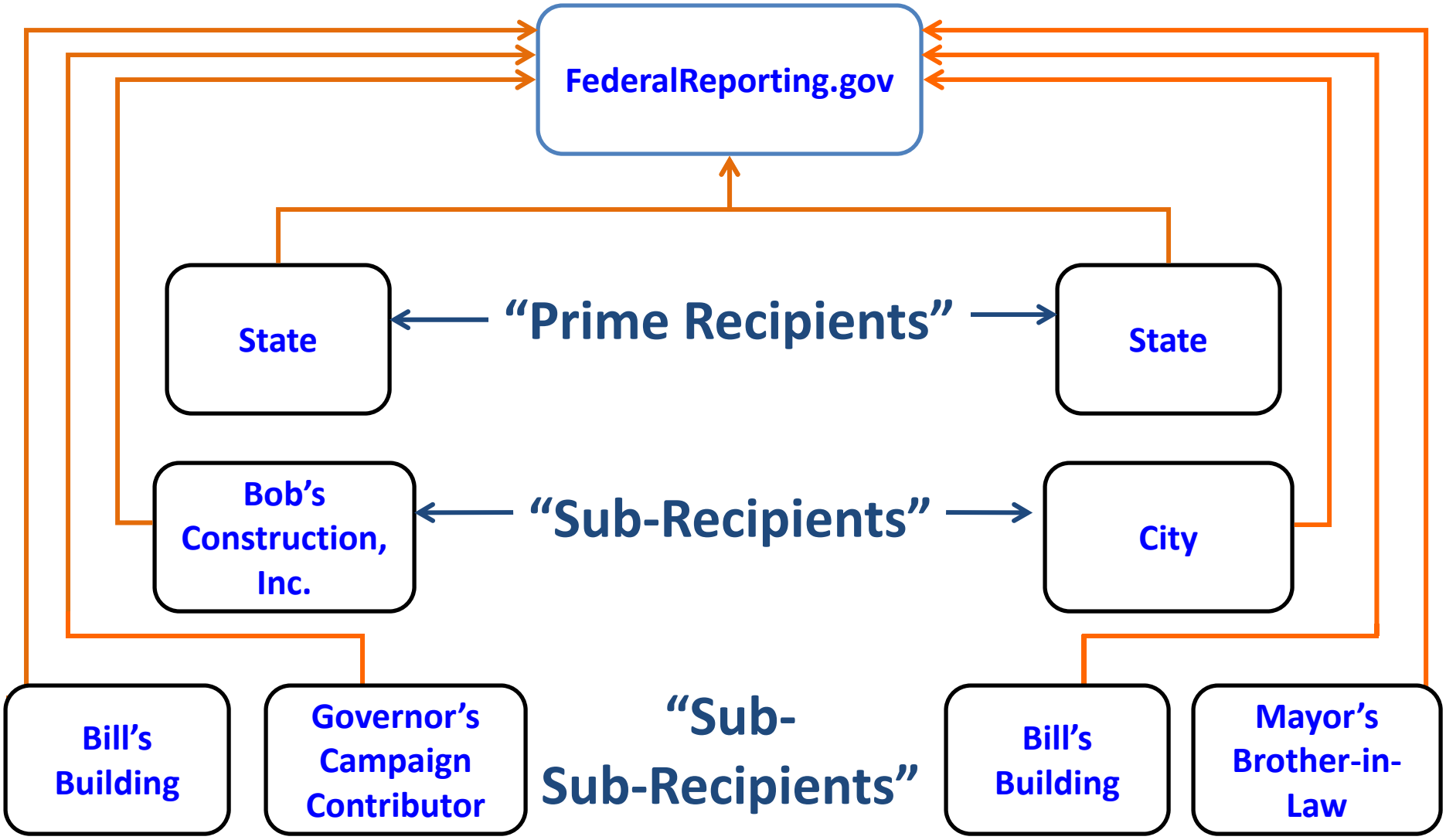
USAspending.gov Today

- Recovery Act Obligations Available
  - Difficult to match obligations to recipient reports on Recovery.gov; automation impossible
- No Cross Check with Treasury Data
  - Treasury checks should be compared to agency obligation data
- Matching obligations to statute virtually impossible
  - Need crosswalk between appropriation and Treasury account symbol
- 1st Tier Sub-Recipient Data Only (Oct. 2010)
  - Per April 6 OMB memo, data to be collected

Current Recipient Reporting: Recovery Act & USAspending.gov



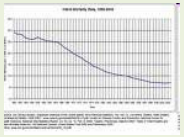
Reporting Requirements: An Alternative Approach



FederalReporting.gov: A New Paradigm for Reporting

All recipients report spending & performance data directly

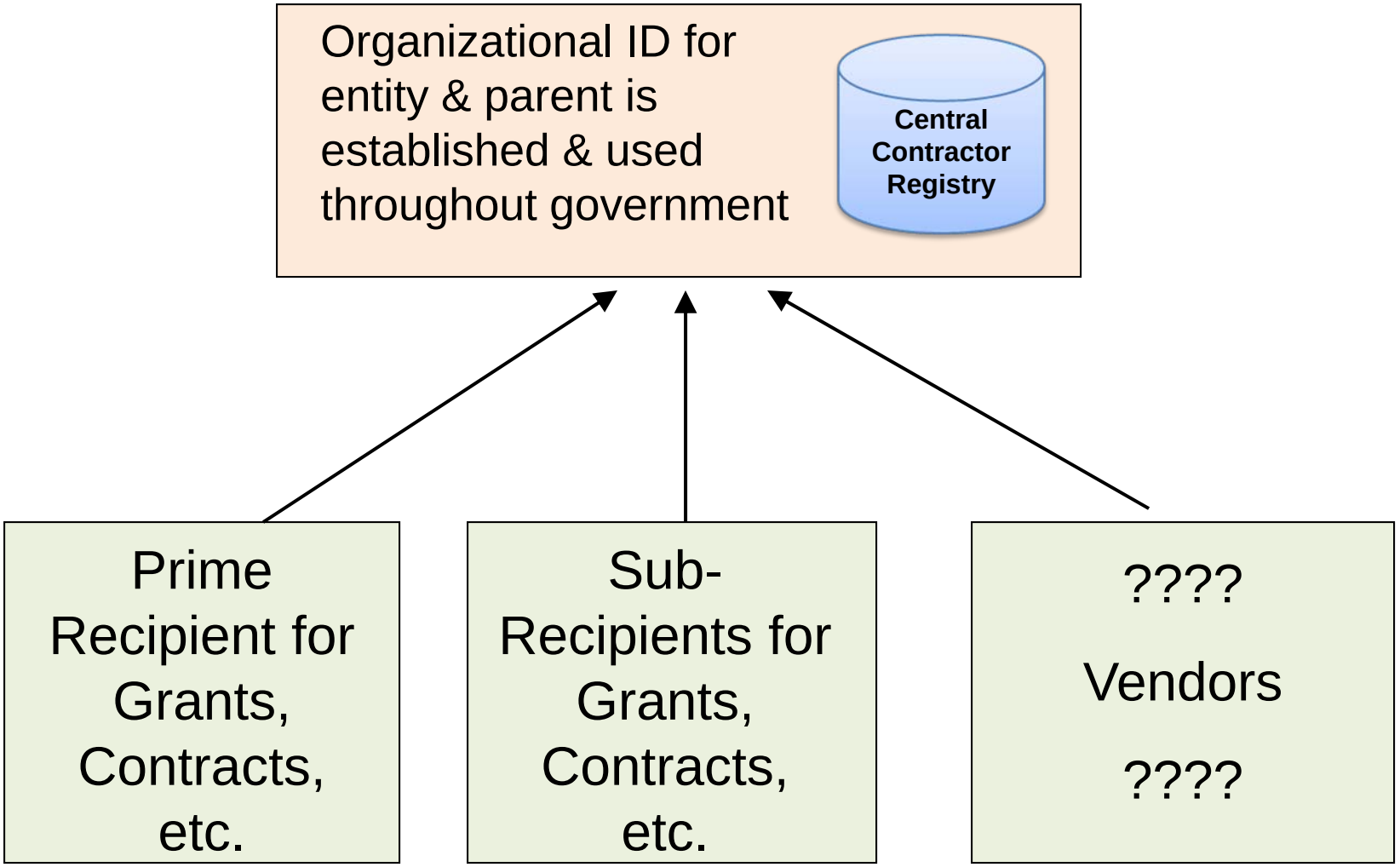
Spending Data



Universal Indicators (e.g., jobs)

Agency/Program Performance Measures

A Comprehensive Registration System: Central Contractor Registry



Proposed Model for Reporting & Disclosure

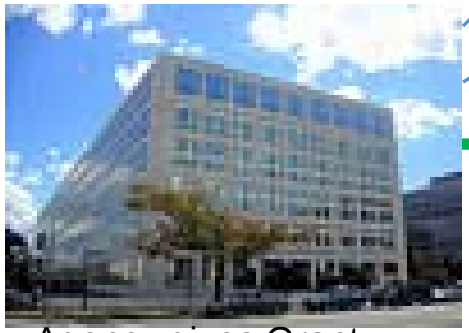
REGISTRATION

Potential recipients & sub-recipients must register.

Central Contractor Registry

An organizational ID for entity & parent is established & used throughout government.

FEDERAL AWARD



Agency gives Grant, Contract, etc.

Treasury writes check & reports data

Reconcile Data from Three Sources

1st Tier Recipient

All recipients report spending & performance data directly

Federal Reporting.gov

Public Access to Web

Machine Readable Formats for Developers

2nd Tier Sub-Recipient

3rd Tier Sub-Recipient

4th Tier Sub-Recipient

Legend

- Funding
- Reporting Data
- Public Access

USAspending.gov
Next Steps

- Recovery Act Recipient Reports
  - Relatively small data set with clear IDs (DUNS)
  - Project identifiers in both sets of data
- OMB Earmark Database
  - Move from OMB site to USAspending.gov
- Obligations matched to statute
  - Treasury account symbol crosswalk

USAspending.gov
Two-Year Horizon (Cont'd)

- Treasury Data
  - Connect outlay data to obligation data to recipient data via DUNS
- Ultimate Recipient Data
  - Federal Reporting.gov works; n^{th} tier sub-recipient data not collected because of policy, not technology
- Contractor Performance Data
  - Move existing data into public view
  - Link to recipient reports & obligations via DUNS
- Agency Budget Execution Data
  - The underlying data that culminates in reports to USAspening.gov

USAspending.gov
Two-Year Horizon (Cont'd)

- IRS Tax Expenditure Data, Summary Level
  - No need/can't connect to recipients (yet)
- President's Budget
  - Standalone info (not linked via unit of analysis yet)
- Appropriations (and Accompanying Reports)
  - Standalone info (not linked via unit of analysis yet)
- Needs Data
  - Using income, race, gender, and other data to contrast with spending decisions
  - To inform spending decisions

USAspending.gov
Long Term

- Flow of funds from budget request to execution to performance
  - President's budget + approps bills + obligation + outlay + recipient report + performance data
  - Need to establish unit of analysis throughout government
- Performance Data
  - Spending and performance data on the same screen
  - Political consideration for program bashing vs. program improvement
- Campaign Contributions
  - Need connecting IDs
  - What about individuals?
- Congressional Lobbying Data
  - Need connecting IDs
- Agency Lobbying Data
  - Need connecting IDs
- SEC Data
  - Link publicly-traded companies to receipt of federal funds
  - Link publicly-traded companies to tax breaks

Adding Data Sources to USAspending.gov

Congress

- Appropriations Bills
  - With associated reports
- Tax Code
  - Specifically, tax expenditures
- Lobbying
- Earmarks
- Earmark Requests

White House / OMB

- President's Budget Request
- Needs Data
  - Unemployment, SES, etc.

Treasury

- Outlays (checks written)
- Tax Expenditures

Adding Data Sources to USAspending.gov

Federal Agency

- Budget Request Documents
- Budget Execution Materials
  - SF133s, etc.
- Contracting
  - RFPs, Bids, and Contracts
- Performance
  - Program Outputs
  - Outcomes Related to Programs
- Procurement Lobbying

Recipient (& Other External Entities)

- Recipient Reports
  - Award receipts
  - Award performance
  - Contractor performance
- Lobbying
- Campaign Contributions

Connecting Data: An Example

Lockheed Martin Corporation, Fiscal Year 2010

Contracts (Agency Reports)

Total dollars: **$7,526,064,979**
The amount for this search is 7.3% of all awarded dollars for the fiscal year.
Total number of transactions: **3,158**
Get list of transactions

Contracts (Recipient Reports)

Total dollars: **$7,125,099,901**
Total number of reports: **2,998**
Get list of recipient reports

Contractor Performance

Number of incidents on all federal projects: **22**
Number of incidents on this project: **8**
Get contractor reports

Tax Forgiveness

Total dollars: **$1,526,064,500**
The amount for this search is 0.2% of all tax breaks for the fiscal year.
Total number of deductions claimed: **189**
Total number of credits claimed: **245**
Get list of credits
Get list of deductions

Lobbying

Total dollars: **$6,064,500**
Total number of legislative lobbying contacts: **215**
Total number of executive lobbying contacts: **130**

Connecting Data: An Example

Joint Strike Fighter Program, Fiscal Year 2010

Contracts (Agency Reports)

Total dollars: **$35,158,556,001**
Total number of prime contractors: **14**
Total number of sub-contractors: **258**
Total number of transactions: **958**
Get list of contractors

Contracts (Recipient Reports)

Total dollars: **$35,158,556,001**
Total number of prime contractors: **9**
Total number of sub-contractors: **216**
Get list of recipient reports

Outlays

Total dollars: **$5,158,556,001**
Get list of transactions

Authorizing Legislation

Total dollars appropriated: **$40,526,064,500**
View statute

Contractor Performance

Total number of contractors: **272**
Number of incidents on all federal projects: **22**
Number of incidents on this project: **8**
This is a high-risk project: View details
Get contractor reports

Performance Based Accountability and Budgeting

Paul Posner, George Mason University

Public managers throughout the world have sought to strengthen public decision-making and performance by establishing strategic goals and performance measures as the touchstone for accountability. The shift to a results-oriented framework promises to improve program and organizational effectiveness and public transparency by systematically linking administrative activities to the performance outcomes that matter most for the organization and its publics.

Performance-based reforms have in fact had a long history in the United States, at all levels of government. Often led by state and local initiatives, public administrators have become gripped by waves of reforms intended to improve performance and enhance public confidence in government. Ushered in with great expectations, reforms such as Planning-Programming-Budgeting, Zero Based Budgeting, and Total Quality Management achieved significant improvements but are widely acknowledged to have fallen well short of their goals of institutionalizing a sustainable focus on performance within government over the longer term.

Many factors accounted for the checkered history of performance reforms in American government, but one lesson learned is that the ultimate success of performance reforms will be predicated on their integration with the most important process engaged in by public managers every year – the budget process.[1] Performance plans and metrics will be consigned to a future of irrelevance if they are not linked to the way we allocate and manage scarce resources. Most importantly, the failure to integrate performance into the budget process will continue to frustrate the ability of various publics to understand and participate in resource allocation decisions.

What is performance-based budgeting?

While the linkage of performance plans and metrics with budgeting was viewed as critical to the success of performance management, few really examined what this meant. Indeed, most simply treated performance budgeting as a proverbial "on-off switch" - you either did it or you didn't. In reality, performance budgeting is more like a dimmer switch, with a continuum of different strategies to link budget decisions to performance data:

- **Presentations** – Budgets are infused with information discussing the performance consequences of budget decisions. The presentations can be linked at both the individual account level in the budget or at the aggregate performance plan goal level.

[1] U.S. General Accounting Office, Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation (Washington, D.C.: GAO, 1997)

- **Budget Restructuring** – The fundamental basis of budget decisions – budget accounts – are reorganized to reflect performance goals. This can reinforce the shift in focus of budgeting from inputs to performance outcomes.
- **Performance reviews and assessments** – Formal assessment processes, such as PART, evaluate how well programs and operations are meeting performance goals and outcomes.
- **Performance targets** – Agencies have targets that are set in either outcome or output terms for performance for the coming year which are integrated with the budget request and appropriation.
- **Performance linked funding** – Under this approach, resource allocation decisions are driven in some mechanical way by performance levels and comparisons. Agencies achieving greater efficiencies by producing at higher levels get higher payments, while those falling short experience budgetary reductions. This approach is what many advocates of performance budgeting envision, but is typically the most difficult to implement.
- **Outcome based budget formulation** – Several jurisdictions have sought to go beyond the confines of traditional agency-centered models of budget formulation to establish outcomes rather than agencies as the primary decision unit for the budget process. The state of Washington notably was among the first to initiate this reform, with the help of David Osborne and other consultants with Public Strategies Group.[2]

Regardless of the form that performance budgeting takes, it is important to have realistic expectations about what such reforms can accomplish. The integration of performance information cannot, and should not, be expected to "take politics out of budgeting" or to supplant the judgment of performance evaluators and analysts for elected officials. There are too many other important criteria that properly belong in budgetary debates, including judgments about equity, needs and the relative priorities given to competing claims.

Some argue for a mechanical model of performance budgeting -- if performance goes up, the agency or staff get rewarded with increased resources, if it goes down, they get penalized. While appealing on first glance, such a model presupposes that there is a single budgetary answer to performance trends, when in fact there are many. Thus, for instance, if the number of drug abusers goes up, it is unlikely that we would penalize drug programs with a loss of funds. In fact, we may find that increased funding is necessary to bring about performance improvements, along with other needed management and program reforms.

The goal of performance budgeting should not be to provide the **answers** to inherently political choices in the budget process but rather to provide a new set of **questions**. The shift in the agenda for the budget process could be expected to bear fruit in a more

[2] This reform formed the basis for David Osborne's book with Peter Hutchinson, The Price of Government (New York: Basic Books, 2004)

informed debate that adds performance goals and results to the other important issues addressed in annual resource allocation debates.

Challenges in Linking Resources to Results

As the foregoing suggests, one could reasonably hope that the debate over resources would not only focus on dollars and inputs but on whether proposals can or have met discrete performance goals and targets. To bring this about, agencies and their stakeholders must step up their efforts to develop a credible supply of information on performance for their programs and operations. Performance can rise to become a new input to the inherently political process of resource allocation only when the performance goals and metrics are perceived to be fair, balanced, comprehensive and accurate.

Once this foundation has been established, however, major shifts must occur in the way that budget choices are framed and presented. Federal agency budgets have generally focused attention on the input side of the budget equation through a focus on the objects of expenditures, e.g. personnel hired, contracts issued, facilities constructed. Introducing performance as a new complementary focus entails a new unit of analysis for budgeting. Ultimately, budget presentations and budget accounts and activities will have to be restructured to highlight their implications for key performance goals.

Budget Account Structure

The account structure used to appropriate funds plays a formative influence in resource allocation. They structure the units of analysis that will form the basis for allocating scarce resources among competing purposes and form the basis for determining what activities are compared. It makes a real difference if marginal decisions focus on tradeoffs among object classes or tradeoffs among programs. If consistent in defining both the scope and costs of program activities, budget officials could compare claims on a level playing field.

Budget accounts and activities also serve as the structure for budgetary accountability and execution. Agency officials must typically ensure that they implement the budgets by being faithful to the allocations agreed to by public officials in the budget process. While agencies may attempt to track and monitor other dimensions such as performance goals, the primary emphasis in administration will be on the budget accounts and program activities in the budget.

Federal budget accounts and presentations have developed over time to reflect various orientations that are based on activities that can be defined by objects of expenditure, performance goals, and programs:

- Budgets with an orientation on objects of expenditures are referred to as line item budgets.
- Budgets with a focus on organizational units

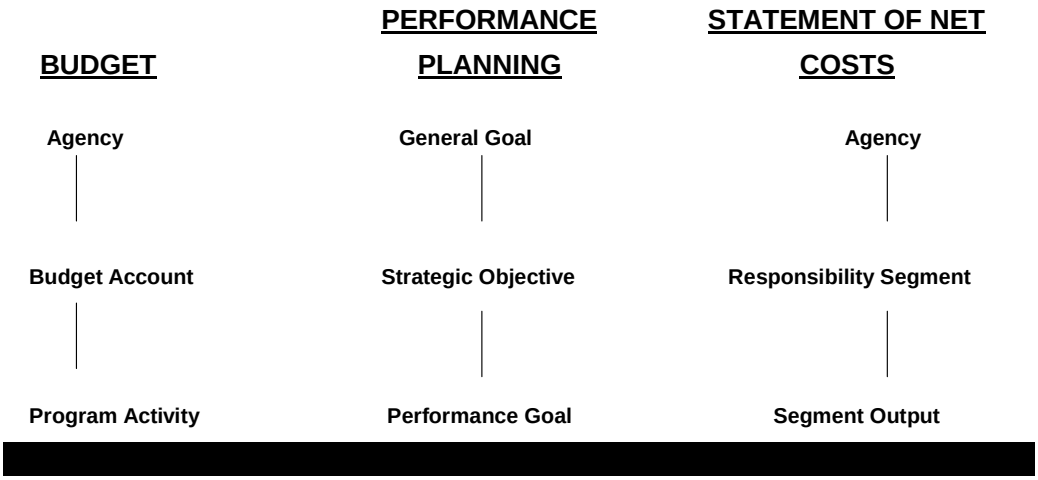
- Budgets with an orientation on programs are referred to as program budgets.
- Budgets with an orientation on performance goals are referred to as performance (or performance-based) budgets. Its primary organization feature is outputs and/or outcomes.

There are over 1,100 appropriations accounts in the federal budget and most accounts have subsidiary program activities that allocate budget authority to more specific levels of inputs, outputs, or outcomes. One informal staff estimate concluded that there were 9,000 program activities – subsidiary units under the account level. However, in many cases, even these PPAs are not "programs." Rather programs are often subsidiary to PPAs and can number in the thousands. The most comprehensive document setting forth the program elements at the most detailed level can be found in each agency's Justification of Estimates provided to the congressional appropriations committees after the President's budget goes forward.

Budget account structures are highly disparate and inconsistent even within the same department. The present budget account "structure" was not created as a single integrated framework but rather developed, for the most part, as separate budget accounts over time to respond to specific needs. Viewing these individually developed accounts collectively discloses not only the variety within the current structure but also its complexity. Not only are appropriations accounts disparate, but so are the program activities within the accounts. Mirroring the disparate nature of budget structures, there is also no common definition of a "program" in the federal government.

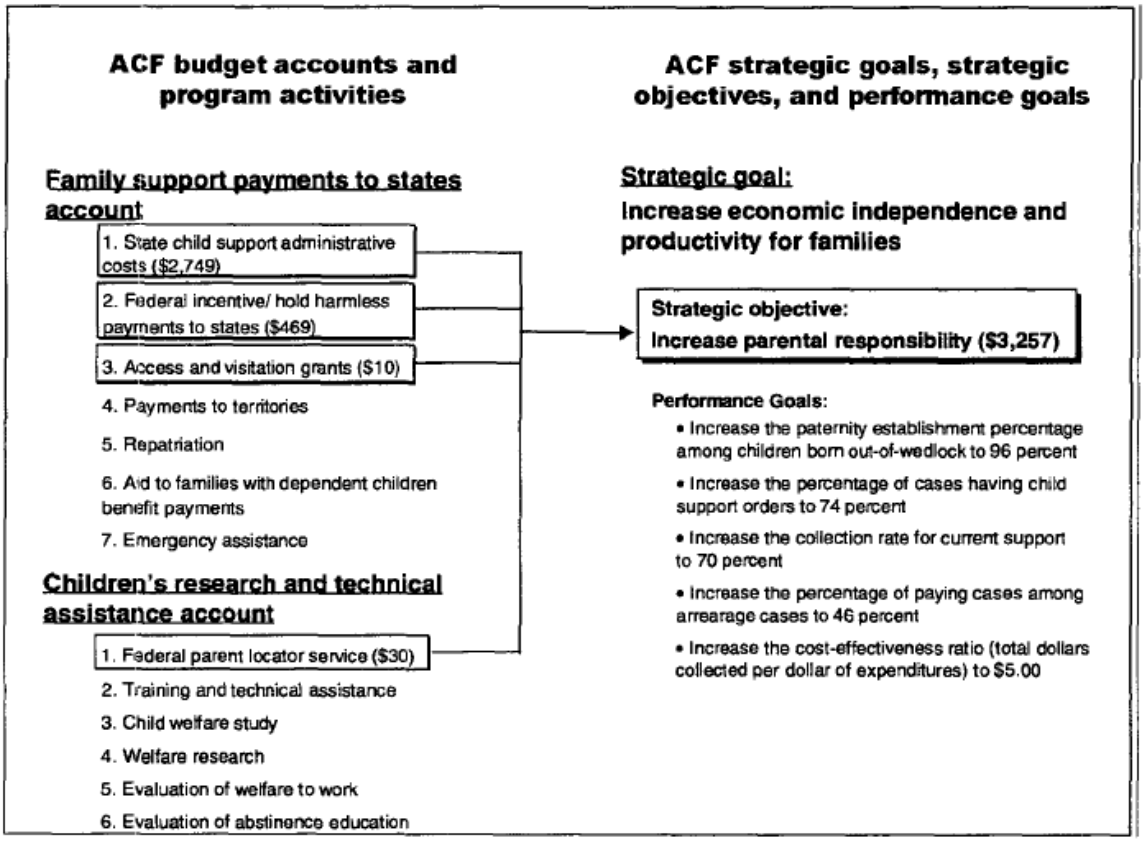
Importantly, the orientation of budget accounts typically differs substantially from the unit of analysis used in both performance plans and in financial statements in federal agencies. As the chart suggests, the focus, unit of analysis and purpose of each of these systems is necessarily different.

Budget, Performance &
Cost Integration Model



Source: GAO.

Performance budgeting principles are difficult to operationalize in our current budget accounting structure. Agencies preparing strategic plans with outcome oriented performance goals face profound obstacles in translating these visions to the budget when the account structure is oriented to line items or organizations. Agencies have had to undertake complex crosswalks that make it difficult to use the performance goals as the basis for resource allocation decisions. The following chart from the Agency for Children and Families in HHS provided a crosswalk showing how programs from several budget accounts contributed to the strategic goals of this bureau.



Note: Dollars in millions. Numbers may not add due to rounding.

Source: GAO analysis based on the Administration for Children and Families' fiscal year 1999 performance plan and Budget of the United States Government Fiscal Year 1999—Appendix.

Along these lines, the federal budget already includes one overarching crosscutting unit of analysis which can serve as the unit of analysis for a reinvigorated performance based budgeting process. These are 19 budget functions and 80 subfunctions in the budget which group programs and agencies around common missions and purposes. Examples include

- Energy function
    - Energy supply subfunction
    - Energy conservation subfunction
    - Emergency energy preparedness
    - Energy information, policy and regulation
- Natural resources and environment function
    - Water resources
    - Conservation and land management
    - Recreational resources
    - Pollution control and abatement
    - Other natural resources

The subfunctions come close to providing a common mission based unit of analysis to capture the many activities of the federal government across a consistent policy frame. It has the advantage of being tied to all of the budget accounts so that the data is directly derived from the budget rather than through a standalone, supplemental analysis. And it has been in use for many years so that trends can be tracked over time.

What has been missing is the use of subfunctions as a unit of analysis by OMB and Congress to make resource allocation decisions. It has largely served as an analytic supplement but could play a more central role in the future should policymakers be convinced that such a crosscutting focus is important.

Consequences

The failure to develop a consistent framework for defining budget choices in performance terms has significant consequences for the public debate. The current system frustrates public oversight and transparency.

- **Transparency -** It is difficult for the public to readily understand how resources are allocated among related programs when structures differ so widely. Food safety, for instance, is split among 15 different federal agencies, all with different budget structures and accountability measures. The lack of a consistent definition of program across the federal establishment as the building block for budgeting is among the most significant barriers to better understanding and tracking.

- **Comparisons** - It is difficult to make comparisons among related programs across different bureaus and agencies without a single consistent unit of analysis cutting across the entire federal program universe. The costs of a single program can sometimes be split among multiple accounts, such as accounts for salaries and expenses and accounts for other expenditure items such as capital or construction. For example, the budget resources used to achieve VA's burial program performance goals are funded by six appropriations accounts spread across separate volumes of its congressional budget justification. Conversely, in other agencies, multiple programs are often included in a single account or program activity.

- **Crosscutting emphasis** - important goals and objectives of policy cut across the narrow confines of budget accounts, bureaus, and departments at the federal level. For instance, nearly one half of FY 2009 budget authority for homeland security is provided by numerous agencies outside the Department of Homeland Security. Indeed, most of the major missions of government transcend the boundaries of federal departments and agencies. The federal budget has a standard set of budget functions and subfunctions that can be used to organize reviews of common programs and missions across agencies, but the budget process remains largely stove-piped by agency.

- **Monitoring implementation** – It is difficult to track implementation of programs without a consistent unit of analysis that flows from budget formulation to execution. Federal agencies and their state, local, nonprofit and private partners will of necessity have to primarily focus on the unit of analysis that governs their budgets and appropriations. Reporting systems and supplemental analyses can be superimposed on the agencies to create a post hoc unit of analysis, but this creates numerous and burdensome adjustments and costs in a fragmented system. Moreover, the accuracy of the crosswalks between the budget unit of analysis and supplemental programmatic displays can be questioned, as agencies struggle to apportion and allocate costs across disparate reporting systems.

- **Accountability** – It is difficult to hold agencies accountable for performance outcomes and results when budget structures don't emphasize performance goals. While supplemental analyses and crosswalks can provide some useful information, agencies must focus primarily on those units of analysis that form the basis for their appropriations.

Remedies

There is a need for a consistent unit of analysis to form the backbone of federal budget formulation and execution. Ideally, the federal budget should be restructured along performance-oriented lines to enable agencies and Congress alike to have greater incentives to consider performance in resource allocation. One GAO study found that those federal agencies which restructured their budgets were able to more easily understand how disparate programs, projects and activities contributed to common goals, even when those programs were in different organizational locations throughout the agency.[3]

As the building blocks for resource allocation, a more cohesive, consistent and performance-oriented account structure is critical to enable both Congress and the executive to consider competing claims on the same terms. A consistent program structure will also facilitate the tracking and monitoring of federal spending across disparate agencies and programs. It will also enable comparisons to be more easily made across related programs focusing on common recipients and goals.

Achieving this vision is complicated and will take years to bring about in our system. Congress and the executive are co-owners of the federal budget and previous efforts at budget account restructuring have foundered when appropriations committees were not engaged as partners. Moreover, establishing common units of analysis across the many departments, agencies and programs is a long-term project as well. Care must be taken to ensure that new program structures and definitions appropriately reflect the purpose of the underlying activities they encompass - apples should be grouped with apples and not

[3] U.S. Government Accountability Office, Performance Budgeting: Efforts to Restructure Budgets to Better Align Resources With Performance (Washington, D.C.: GAO, 2005) GAO-05-117SP

pears, so to speak. And it is critical that these program structures and definitions are in fact tied to the budget itself.

The budget subfunctions discussed above provide a good starting point for such a common unit of analysis. The subfunctions could become the spine of a budget decision-making process that cuts across disparate agencies and programs to focus on common outcomes and missions in budget formulation and budget execution. They have the advantage of being linked to the budget accounts themselves and having trends on spending over many years.

The adoption of a subfunction as a central unit of analysis does not mean that the job is finished. Beneath this high-level mission focus, there would still be the need to develop more consistent definitions of budget accounts and program activities that can be used at the micro level as well.

In sum, the following agenda is presented:

- Congress and OMB should engage in a joint initiative to undertake comprehensive reform of the budget account structure

- The reformed structure should permit and encourage performance-informed budgeting and tradeoffs among related programs

- The reformed structure should add to and not subtract from the information currently provided to Congress for appropriations decisions. Accordingly, even while changing account orientations, agencies should be required to continue to provide other information needed by appropriators to effectively perform their formative role in holding agencies accountable in a democracy

- The budget subfunctions constitute an appropriate unit of analysis which is already tied to the budget accounts and can be readily used to summarize and compare programs across agencies sharing common purposes. Consideration should be given to using the subfunctions as the basis for budgetary tradeoffs within the President’s budget formulation process

Exposing Corruption Exploring Solutions
Project On Government Oversight

Contract Spending: Escaping the Dark Ages

In 2009, the federal government awarded more than $523 billion in federal contracts—contracts for goods, including complex weapons systems, and services.[1] Following the life of a federal contract is difficult, if not impossible in the current labyrinth of government databases. In fact, only recently the government hinted that it might post contracts online.[2] Cynics argue that the government's hide the ball mentality in federal contracting is intended to prevent the public from truly knowing about government spending and the contractors that receive taxpayer dollars. It is very difficult to believe differently until the government decides that absolute openness in federal contract spending is vital to improving competition, building trust in government policies and programs, establishing integrity in the contracting system, and ensuring that contractors are responsible and performing in the best interest of taxpayers.

There are three main issues to focus on when considering contracting transparency. First, the public is not receiving a comprehensive and accurate picture of federal contract spending. Second, the government's stovepipe approach to creating, organizing, and maintaining federal contract data resources unnecessarily confuses the public. Third, a long list of impediments that remove contract data from public view hinders transparency. This white paper discusses concerns in those areas, and provides recommendations to shine a bright light on federal contract spending.

Nominal Public Disclosure

Despite numerous advances in technology, the system for disclosing contract information is barebones to say the least. In fact, the system as it currently stands only provides scant summaries of how the federal government is spending over $500 billion for goods and services each year. For example, USAspending.gov provides addresses, company profile information, good or service codes, some dollar and competition figures, place of performance, and "check the box" information about a contractor's socio-economic factors. The public cannot access copies of solicitation notices, contracts, delivery and task orders, amendments and modifications, responsibility and performance information, and other source selection materials that would provide a more complete picture of federal contract spending.

The current Administration is pressing agencies to do their part in actively releasing information to the public, shifting the presumption under the Freedom of Information Act (FOIA) back to disclosure and using technology to provide searchable, sortable, and downloadable contracting information to the public. No matter how

[1] USAspending.gov, "Top 100 Recipients of Federal Contract Awards for FY 2009," Last updated on May 3, 2010. Available at http://www.usaspending.gov/fpds/tables.php?tabtype=t2&subtype=t&year=2009
[2] 75 Federal Register 26916, May 13, 2010. Available at http://edocket.access.gpo.gov/2010/pdf/2010-11381.pdf. *See* POGO Blog, "Contracts Online -- OMG!," May 13, 2010. Available at http://pogoblog.typepad.com/pogo/2010/05/contracts-online-omg.html

vigorously the Administration promotes openness, however, it must also ensure that information is accurate and timely. The upgrades to the contracting system have been constructive, but the data being released is rudimentary and errors plague the system. For example, past and current versions of the top 100 contractor lists include questionable entries, such as subsidiaries being listed independently of their parent companies, companies being assigned multiple rankings, the listing of federal agencies, and even a bizarre listing called "Government of the United States."

Stovepipe Approach

The current contracting oversight landscape includes approximately nineteen resources, databases, or informational sites. The following is a non-comprehensive list of government web sites that the public should have access to in order to track contract spending and federal contractors:

1. Central Contractor Registration (CCR)
2. Online Representations and Certifications Application (ORCA)
3. Federal Agency Registration
4. Federal Business Opportunities (FBO)
5. GSA Schedule Library
6. Federal Procurement Data System-Next Generation (FPDS-NG)
7. USAspending.gov
8. Excluded Parties List System (EPLS)
9. Past Performance Information Retrieval System (PPIRS)
10. Federal Awardee Performance and Integrity and Information System (FAPIIS)
11. Electronic Subcontracting Reporting System
12. GAO Bid Protests
13. DoD Revolving Door Database
14. SEC Filings (EDGAR)
15. Senate and Executive branch lobbying reports
16. Public Financial Disclosure Report for Executive branch officials (SF-278)
17. FEC PAC and individual contributions
18. DOL violations
19. Wage Determinations On-Line (WDOL)

What is severely lacking is a comprehensive, user-friendly "one-stop shop" for contracting information. Unfortunately, attempts to move legislation through Congress have failed.[3]

The General Services Administration (GSA) appears to finally have listened to POGO, OMB Watch, the Sunlight Foundation, and many others who are frustrated with the current user-unfriendly systems.[4] According to the GSA,[5] nine databases will be integrated by IBM— at a cost of $74.4 million. Absent from the list is USAspending.gov "Version 2.0" (although FPDS-NG, the feeder system into USAspending.gov, is included)

[3] "Strengthening Transparency and Accountability in Federal Spending Act of 2008," S. 3077, 110th Congress. Available at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_bills&docid=f:s3077is.txt.pdf

[4] GSA Press Release, "GSA to Consolidate Contract Performance Databases," February 18, 2010. Available at http://www.gsa.gov/Portal/gsa/ep/contentView.do?pageTypeId=10430&channelId=-24825&P=&contentId=29131&contentType=GSA_BASIC

[5] The nine data sets include: 1. FedBizOpps, 2. Wage Determinations On-Line (WDOL), 3. Central Contractor Registration (CCR), 4. Federal Agency Registration, 5. Online Representations and Certifications Application (ORCA), 6. Past Performance Information Retrieval System (PPIRS), 7. Excluded Parties List System (EPLS), 8. Federal Procurement Data System-Next Generation (FPDS-NG), 9. Electronic Subcontracting Reporting System.

and the recently available Federal Awardee Performance and Integrity Information System (FAPIIS),[6] which will allow government officials to review a contractor's record of integrity or business ethics as required by law. The possible integration of the above data sets is wonderful, but due to the fact that PPIRS is not publicly available, this integrated contracting resource likely will only be available to government officials and it is far from being comprehensive.

Impediments on Contract Spending Disclosure

In addition to the deficiencies caused by the release of barebones contracting information and the lack of integration of systems, other factors may also prevent public access to some of the most useful contracting information. One factor is time. Although contract information can be requested through the Freedom of Information Act (FOIA),[7] it takes months or years to receive an agency response. The public and the media have essentially abandoned using FOIA to request contracting information because of the unreasonable delay.

The second factor is FOIA fees. Exorbitant FOIA fees create additional obstacles that prevent the public from obtaining vital contracting information. Converting to an electronic system will allow documents to be immediately posted on an agency's web site and will reduce many of the burdens of requesting, processing, and delivering requested information under FOIA. Unfortunately, those improvements are slow in coming because the culture within agencies many times disfavors disclosure, which is often considered to result in the punishment of a program, agency, or contractor.

Even if FOIA is used, numerous laws, regulations, and an executive order limit public access to comprehensive contract information. Specifically, public access to contracting information is impeded by Executive Order 12600,[8] source selection information restrictions,[9] the Procurement Integrity Act,[10] the Privacy Act,[11] trade secrets provisions,[12] Controlled Unclassified Information (CUI) designations,[13] and numerous FOIA exemptions, most notably 5 U.S.C. 552(b)(4), which protects "trade secrets and commercial or financial information obtained from a person and privileged or confidential."

Some of reasons for protecting certain information are legitimate. The government and contractors have a right to limit public access to genuine national and homeland security information, information that provides an unfair competitive advantage to other companies, private personal information, and company trade secrets. The problem is that through the years there seems to be a movement toward non-disclosure that only recently has showed signs of shifting in the opposite direction. Urging the Obama Administration to find the proper balance is the goal of today's conference, and is a message that we must deliver to the public, Congress, and the White House.

---

[6] FAPIIS is modeled after POGO's Federal Contractor Misconduct Database and requires, prior to awarding a contract, federal officials to review and consider past performance information with respect to the offeror when making an evaluation of their responsibility and past performance. Pub. Law 110-417, Sec. 872, October 14, 2008. Available at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ417.110.pdf

[7] 5 U.S.C. § 552.

[8] In 1987, President Ronald Reagan signed EO 12600, which states that pursuant to a FOIA request, a contractor "may object to the disclosure of any specified portion of the information and to state all grounds upon which disclosure is opposed." EO 12600 Sec. 4. Available at http://www.archives.gov/federal-register/codification/executive-order/12600.html

[9] FAR Subparts 2.101, 3.104-3, 3.104-4, and 42.1503(b) and (d). Available at https://www.acquisition.gov/Far/current/pdf/FAR.pdf

[10] 41 U.S.C. § 423. Available at http://frwebgate.access.gpo.gov/cgi-bin/usc.cgi?ACTION=RETRIEVE&FILE=$$xa$$busc41.wais&start=960618&SIZE=26916&TYPE=PDF

[11] 5 U.S.C. § 552a.

[12] 18 U.S.C § 1905.

[13] Controlled Unclassified Information is not considered classified, but "the Government must be able to prevent the public disclosure of information where such disclosure would compromise the privacy of American citizens, national security, or other legitimate interests." White House Memorandum for the Heads Of Executive Departments and Agencies, "Classified Information and Controlled Unclassified Information," May 27, 2009. Available at http://www.archives.gov/cui/documents/2009-presidential-memo.pdf

Looking at one example, contractor performance information (compiled in PPIRS and FAPIIS) is considered protected source selection information that **cannot** be released to the public. This is analogous to a parent who isn't permitted to see their child's report card. As a result of this non-disclosure, the public and government appear to have very different opinions about what constitutes acceptable contractor performance.

Performance information can be thought about in many ways, including award or incentive fees or percentages, actual performance grades given during the performance of a contract, a contractor's ability to remain on budget and on time, a contractor's performance record being considered during source selection, a contractor's performance record on non-government contracts, a contractor's record on labor and environmental issues, a bidder's compliance with tax laws, and a company's performance on Wall Street. Performance information might also be considered based the effectiveness of federal agencies in meeting their missions.

All of those criteria are useful to government officials and the public, but there seems to be a wide difference of opinion as to what constitutes adequate or poor contractor performance. For example, the general public has relatively low opinion of KBR, Halliburton, and Blackwater (currently known as Xe) based on media accounts of their performance in Iraq and Afghanistan. Repeated allegations of overbilling, fraud, bribery, poor performance, legal infractions, and excessive use of force against civilians have convinced the public that the government should terminate these contracts and not award any new contracts. Some people are demanding that the companies and executives be held criminally liable for their alleged misdeeds.

However, the government seems to think that KBR, Halliburton, and Blackwater are performing very well in Iraq and Afghanistan. Contracting experts have claimed that the State Department is very happy with Blackwater's performance because all of its officials have been protected, traveling safely in and out of many dangerous sectors of Iraq and Afghanistan. Many of the criticisms of KBR and Halliburton have been dismissed as a cost of doing business with the federal government. As a result, there hasn't been any movement to suspend or debar those contractors, although some individual employees have been held accountable.[14] Instead, the government continues to award billions in contracts to them. Either the government's performance criteria differ greatly from that of the public, or the government has no choice but to continue doing business with risky contractors.

Equally troubling is the degree of influence contractors have over their performance evaluations. POGO has been approached on several occasions by contracting officers who stated that they feel compelled to provide a performance evaluation that will not be challenged by the contractor.

FAPIIS was also crafted with contractors in mind. Contracting officers must give offerors the opportunity to provide additional information that demonstrates their responsibility before the contracting officer makes a nonresponsibility determination based on relevant information from FAPIIS.

It's easy to accept a certain level of due process, but at the same time, the public should be able to see performance information and ensure that the government is making decisions that are in the interest of taxpayers.

[14] See POGO's Federal Contractor Misconduct Database (FCMD):
KBR (http://www.contractormisconduct.org/index.cfm/1,73,221,html?ContractorID=29&ranking=13)
Halliburton (http://www.contractormisconduct.org/index.cfm/1,73,221,html?ContractorID=149&ranking=101)
Blackwater/Xe (http://www.contractormisconduct.org/index.cfm/1,73,221,html?ContractorID=123)

Recommendations

To regain public faith in the contracting system, the government must provide the public with open access to federal contracting information. POGO provides the following recommendations to make government contract spending more transparent:

1. The government should require more electronic filing of information to allow for timely online posting. The result will be a decrease in FOIA requests and the time required to process them.

2. USAspending.gov must become the "one-stop shop" for government officials and the public to access comprehensive, accurate, and timely contracting information. This includes actual copies of each contract, delivery or task order, modification, amendment, other transaction agreement, grant, and lease. Additionally, proposals, solicitations, all award decisions and justifications, audits, performance and responsibility data, and other related government reports should be incorporated into USAspending.gov.

3. Contractor performance metrics must be established and defined and incorporated into contracting information that is publicly available. PPIRS and FAPIIS should be available to the public.

4. The government's use of non-disclosure practices (CUI, FOIA exemptions and unreasonable fees, contractor claims of proprietary data or trade secrets), must be minimized.

5. The Defense Department's revolving door database should be publicly accessible.

6. Executive branch lobbying records should be publicly available, including detailed lobbying reports that describe persons present at a meeting and the policies and programs discussed.

Some Issues for Discussion

Beryl A. Radin
Burt Barnow

Conference to Develop a Long Range Vision for Federal Spending Transparency

Washington, DC May 20, 2010

Federal spending transparency requires answers to two basic questions:

- What do you want to know?
  - If the money gets to the appropriate source
  - How money is spent
  - Effects of the expenditures
  - Level of detail
- Who is “you”? (*See the next slides*)

Cast of characters potentially involved

- OMB/White House
- Department/agency heads
- Senior managers
- Program managers
- Congressional actors
  - Authorizing committees
  - Appropriating committees
  - Political agendas

Cast of characters potentially involved (continued)

- Interest groups – multiple perspectives
  - Advocacy groups
  - Representatives of providers
  - Potential competitors
  - Private sector
- Intergovernmental actors
  - State, substate, and local groups
    - Mirror the federal actors in terms of roles
    - Have their own systems, typologies, goals

Each set of actors has its own needs for information

- Each operates with its own responsibilities and goals
- Differences between macro perspectives and micro perspectives
  - At federal level program managers are concerned about details of program implementation while top officials more likely to look for broader patterns
  - Similar differentiation at the state and local level
- The legislative actors have variable information needs
- Non-government actors will have still different needs—often narrower in focus and possibly their goals will conflict with public sector views

Can one unit of analysis meet all these needs?

- Federalism principles provide different degrees of legitimacy for states and localities to define their own goals
- Congress generally looks at the big picture, but sometimes focuses on specific issues
- Outside groups may want to look at information in still different manner

Some additional questions on spending

- What about the spending do parties want to learn?
  - Program outputs
  - Program outcomes and impacts (Not at all easy to do!)
  - Decision processes
- How far down the chain can we expect to get information?
  - Contracts, subcontracts, subsubcontracts, etc.
  - Problems of information overload
- Can we answer these questions on a consistent government-wide basis?
  - Differences among programs in level at which spent, use of funds
  - Programs have varying interdependencies at both the federal and state/local levels

Tracking the spending is not simple

- Who actually knows how funds are spent?
- How do we find out if reporting is not required?
- Who collects the data now?
- If the data is not collected, how can it be collected? (e.g. who pays for it)
- Who cross checks the data?
  - Multiple sources of information
  - Costs involved

Tracking spending is not simple: An example for Workforce Programs

- One-Stop Career Centers established by the Workforce Investment Act (WIA)
- Over a dozen mandatory partners at One-Stops and often many optional partners (TANF)
- One-Stop infrastructure costs sometimes shared, sometimes paid by WIA
- Spending on participants hard to track due to co-enrollment and sequential enrollment
- Support also sometimes comes from other programs such as Pell Grants, vocational rehabilitation, Veterans' programs

Linking performance to budget process

- Measured performance not the same as assessment of impact/outcomes
  - Performance measures generally short-term and may be inputs, processes, outputs, or outcomes
  - Impact estimates provide data on program effects and take longer
- What does information tell you in terms of budget process?
  - Should we put more or less money into poor performing programs"
  - Promoting efficiency is sometimes counter to promoting equity
- Role of Congress vs. Executive Branch
- Difficulty linking performance information with sanctions
  - What is a sanction in the budget process?
  - More money, less money?

Experience with PART and GPRA: so many perverse impacts

- Examples from workforce programs: fear of application of sanctions
  - Measures based on post-program outcomes lead to cream skimming—most vulnerable not served
  - Cost measures led to providing cheap services rather than intended training—Congress barred their use
  - State and local programs behave strategically and "game" the system—try to look good instead of "doing good"

Given multiple levels of decisionmaking:

- Who should be held accountable?
  - State government, local government, vendors? All?
  - What form should rewards/sanctions take?
    - More/less funds for program?
    - Rewards/sanctions?
    - Should incentives be passed on to implementers?
- How to raise social equity concerns
  - Variation among programs
  - Availability of data
- Appropriateness of punitive action
  - Does it punish the guilty or the innocent?
  - How do we avoid punishing people for results beyond their control?
- Avoid one size fits all
  - Make sure the measures are appropriate
  - Adjust the standards when circumstances vary

What should be done?

- Short term efforts:
  - Give agencies the opportunity to sort program areas by levels of ease in meeting basic transparency requirements
    - OMB should play a facilitating not controlling role
    - Identify problem areas in terms of types of programs
- Long term efforts:
  - Rethink form and substance of sanctions
  - Create typology of different types of programs
  - Create multi-program groups sorted by program type (e.g. block grants)
    - Identify commonalities, differences

Title:
Improving Federal Spending Transparency: Lessons Drawn from Recovery.gov

Author:
Yee, Raymond, UC Berkeley
Kansa, Eric C, UC Berkeley
Wilde, Erik, UC Berkeley

Publication Date:
05-25-2010

Publication Info:
School of Information, UC Berkeley

Local Identifier:
Report 2010-040

Abstract:
Information about federal spending can affect national priorities and government processes, having impacts on society that few other data sources can rival. However, building effective open government and transparency mechanisms holds a host of technical, conceptual, and organizational challenges. To help guide development and deployment of future federal spending transparency systems, this paper explores the effectiveness of accountability measures deployed for the American Recovery and Reinvestment Act of 2009 ("Recovery Act" or "ARRA"). The Recovery Act provides an excellent case study to better understand the general requirements for designing and deploying "Open Government" systems. In this document, we show specific examples of how problems in data quality, service design, and systems architecture limit the effectiveness of ARRA's promised transparency. We also highlight organizational and incentive issues that impede transparency, and point to design processes as well as general architectural principles needed to better realize the goals advanced by open government advocates.

Improving Federal Spending Transparency: Lessons Drawn from Recovery.gov

Raymond Yee, Eric C. Kansa, and Erik Wilde
School of Information, UC Berkeley

UC Berkeley School of Information Report 2010-040
May 2010

Available at http://escholarship.org/uc/item/7tw2w9wx

Abstract

Information about federal spending can affect national priorities and government processes, having impacts on society that few other data sources can rival. However, building effective open government and transparency mechanisms holds a host of technical, conceptual, and organizational challenges. To help guide development and deployment of future federal spending transparency systems, this paper explores the effectiveness of accountability measures deployed for the *American Recovery and Reinvestment Act of 2009* ("Recovery Act" or "ARRA"). The Recovery Act provides an excellent case study to better understand the general requirements for designing and deploying "Open Government" systems. In this document, we show specific examples of how problems in data quality, service design, and systems architecture limit the effectiveness of ARRA's promised transparency. We also highlight organizational and incentive issues that impede transparency, and point to design processes as well as general architectural principles needed to better realize the goals advanced by open government advocates.

Contents

1 Introduction

In this document we present some general principles for making the Federal Budget and spending more intelligible and available for public review and debate. The significance of budgetary transparency is difficult to overstate. Information about federal spending can affect national priorities and government processes, having impacts across our society that few other data sources can rival. Budgetary transparency is therefore central to the whole program of "Open Government." Because the information systems that support budgetary disclosures are so important, they must be carefully designed to offer effective and meaningful forms of transparency.

Key aspects for enabling budgetary transparency include:

- *Open Machine-readable Data in Non-proprietary Formats.* Information should be published in ways that maximize the potential for reuse, aggregation, and analysis on a variety of computing platforms. Non-proprietary formats are essential because they enable consumers of information to use open source or commercial computing applications of their choice. Non-proprietary formats also increase the longevity of information since these formats can be best supported by archival systems.

- *Identifiers to Linked Data.* Budgetary datasets will reference many different codes and coding systems. Unfortunately, entities identified by these codes are often difficult to find. If they do exist on a public website, explanation for codes often exist in difficult to parse PDF documents. Such explanatory data should be published in more reusable formats and services that provide contextual information and help make budgetary data intelligible. Ideally, URIs (*Uniform Resource Identifiers* [2], most commonly HTTP Web links) should be the main identifiers used in budgetary disclosures. URIs can identify linked Web resources that contain important information that give context and meaning to budgetary data. Linked data will make budgetary data easier to understand by clarifying the rationale and purpose behind budgetary appropriations and expenditures. Linking data will also enable agency officials and lawmakers to better gauge program performance and effectiveness.

- *Services for Dynamic Data.* Many "Open Government" and "Open Data" advocates will want to see bulk-download features for budgetary data. While bulk-download features are important and should be implemented, we stress that bulk-downloads are *not sufficient*. Bulk-downloads are most suitable for relatively static datasets. For the Federal Budget, we anticipate that different federal agencies and other government monitoring offices will be continually releasing relevant data. Some of this information will include ongoing spending as new expenditures, grants, and contracts are authorized. In other cases, data may be continually collected on program metrics and reports. New programs may be announced (along with identifiers for these new programs), and other programs may close down. Thus, the information landscape of budgetary information, including key data required to make sense of the Federal Budget will be constantly evolving. Since bulk-download approaches typically deliver data well "after the fact" they must be complemented by other approaches that offer more real-time access. Because of the dynamic nature of budget disclosures, Web services (such as those we proposed at `http://recovery.berkeley.edu/` [8, 9]) need to be implemented. Web services make it easier to incrementally retrieve relevant data, retrieve and monitor updated data, and if well designed, make these data easier to use on a wider variety of platforms. Thus, Web services not only facilitate independent analysis, visualization, and interpretation of budgetary data, but they also enable these independent efforts to be conducted in real-time, based on continually streaming data.

- *Data Archiving and Versioning.* Building trust is a critical element in making transparency work. The public needs to trust that information published today will not be gone tomorrow. Data archiving and curation services are an essential component to transparency. The public must also see guarantees that information will not be altered without notification or explanation. Update and notification services

are essential, as well as version tracking, so that the public can see the history of changes on sometimes critical data. Furthermore, budgetary information should be citable; that is, different versions of budgetary disclosures should be clearly identified and reliably retrievable into the future.

2 Case Study: American Recovery and Reinvestment Act Transparency Measures

The accountability measures around the *American Recovery and Reinvestment Act of 2009* ("Recovery Act" or "ARRA") provide an excellent case study of the interplay of the general principles outlined above with transparency effectiveness. We can glean many lessons for short and long term visions of federal spending transparency from Recovery.gov.

An *Office of Management and Budget (OMB)* memo (April 6, 2010) providing guidelines for the immediate future of federal spending transparency explicitly cites ARRA fund transparency, stating that the memo builds on "achievements of and the lessons learned from implementing the American Recovery and Reinvestment Act" [10]. Recovery.gov and Federalreporting.gov are likely models for what we will see in future systems, especially because policy makers have singled out these systems (especially Recovery.gov) as exemplars. Transparency advocates should therefore pay close attention to how these disclosure systems work in practice. By examining these real world systems, we can learn about advantages and problems of different design approaches.

Specifically, we can look at ARRA transparency measures through the criteria specified in the OMB memo: *accessibility, completeness, accuracy, and usability*. In each of these critical dimensions, we should have expectations for the Recovery Act. President Obama made transparency a centerpiece of this legislation promising that Americans would be able "to see how every penny in this plan [the Recovery Act] is being spent."[1] Such absolute precision may be beyond the capability of any accounting system, especially given ARRA's vast scale and complexity. However, we can examine Recovery.gov and associated systems to see whether we can gain a reasonably intelligible and comprehensive understanding of ARRA spending. Given this more reasonable expectation for ARRA transparency, how effective is Recovery.gov in promoting public understanding and accountability?

3 Missing End-to-end Transparency: Guesswork and the ARRA Jigsaw Puzzle

To begin with, a fundamental requirement for budgetary transparency is the ability to relate actual spending with the intentions of Congress as expressed in legislation. Ideally, we should be able to trace *the full lifecycle of federal spending*, from the agency's budget request to congressional appropriation to disbursement by federal agencies. For example, we should be able to easily track fund allocations from the Congressional record, which documents various versions of the bill, to the final enrolled bill PL-115.[2] that the Congressional Budget Office (CBO) scored (the source of the $787 vs $792 billion figure). Once the bill becomes law, we should then pick up the trail for how the Treasury implemented ARRA provisions and allocated money to different accounts, as identified by *Treasury Account Symbols (TAS)*. From these accounts, we should see how spending unfolded across various agencies and then to the primary and secondary recipients. At each stage, we should have a clear tally of the total number of all relevant entities (e.g., recipients, awards, treasury

[1] "Organizing for America — OFA Blog: Message from President Obama: 'What recovery means for you'." http://my.barackobama.com/page/community/post/obamaforamerica/gGxHGc (Accessed May 13, 2010)

[2] "Public Law 111 — 5 — American Recovery and Reinvestment Act of 2009." http://www.gpo.gov/fdsys/pkg/PLAW-111publ5/content-detail.html (Accessed May 13, 2010)

accounts) and expected sums of funding allocation and spending, as well as an account of the data sources and reasons for any uncertainty.

Unfortunately, reality deviates from this ideal of end-to-end transparency. ARRA data is more like a jigsaw puzzle than a clear tabulation and accounting of spending. Members of the public must invest great effort in making even partial sense of the whole picture, and this process involves a great deal of guess work. Some difficulties should be expected since ARRA, by its very nature, is a complex and sprawling piece of legislation. Investing transparency measures dynamically "on the fly" while coordinating across so many players will naturally lead to gaps and problems. However, these gaps and problems persist and continue to stymie ARRA transparency. It is simply too difficult to know with any real confidence how "every penny in this plan is being spent."

These difficulties should give policy makers pause when they look to the ARRA as a model for making the overall picture of the federal budget and federal spending more transparent. Below we describe some of the specific problems and gaps in ARRA transparency and some strategies to improve transparency effectiveness.

3.1 PL 111-5 and Making Sense of the Legislative Process

Ideally, open government transparency measures should offer the public channels to understand, evaluate, and impact key democratic processes. Understanding the legislative process and how the legislation gets implemented in the Executive branch of government should be a major high level goal. Unfortunately, ARRA transparency measures were scoped too narrowly and largely ignored the legislative process.

Current Problems

- *Need for Machine-Readable Legislation:* The run-up to the final piece of legislation was very confusing; every financial analysis used a different way of catgorizing key budget items. It is very difficult to understand spending levels proposed in different versions of the legislation and how these ultimately became law. One could compare the House and Senate versions of the Recovery Act vs the final piece of legislation, though it is a non-trivial task to do a detailed comparison. There was a big effort by news organizations such as the *New York Times*[3] and *ProPublica*[4] while the final details of the bill were being passed to parse the bill. Since the bills were not in machine-readable form, journalists had to manually transcribe numbers into spreadsheets and then come up with separate tallies. It was next to impossible to get a definitive analysis and compare the various analyses to see whether they actually agreed and where they differed.

- *Need Transparency in CBO Analyses:* Moreover, the CBO released preliminary and final analyses of ARRA in PDF form but not a detailed analytic breakdown on how it arrived at the numbers (including simple matters like adding up all the line items and showing how they do in fact add up to what the analyses came up with). In other words, the CBO analyses are effectively a black box, whose accuracy is difficult to assess and therefore challenge or correct. This is especially problematic in fast moving situations like the passage of ARRA.

The overall point of this discussion is to highlight how transparency in government spending is not just an issue for the executive branch of government; it also places requirements on legislative transparency. Ideally, transparency should promote accountability. Through greater visibility in how laws are created, debated, and enacted, reformers can better identify weakness in political and government processes and build support for improvements.

[3]Hossain, Farhana, Amanda Cox, John McGrath, and Stephan Weitberg. n.d. "The Stimulus Plan: How to Spend $787 Billion — The New York Times." http://projects.nytimes.com/44th_president/stimulus (Accessed May 13, 2010)

[4]Grabell, Michael, and Christopher Weaver. n.d. "The Stimulus Plan: A Detailed List of Spending — ProPublica." http://www.propublica.org/special/the-stimulus-plan-a-detailed-list-of-spending (Accessed May 13, 2010)

Proposed Solutions:

1. *Clear Identifiers Needed for Legislative Line Items:* Funding provisions in all bills should carry clear and unambiguous identifiers expressed in machine-readable formats. These identifiers can enable tracking of funding items through CBO analytical processes and later into the Treasury's accounting systems.

2. *Open Data for CBO Analyses:* The supporting data behind CBO analysis should be made public along with the final report. Datasets showing how line-items actually add up would be very helpful, and these can be linked back to specific provisions in bills to show their context in large and complex bodies of legislation.

3. *Open Access to CBO Models and Analytic Formulas:* Since budgetary analysis represents much more than adding up several line items, the CBO needs to provide formulas and financial assumptions, as well. For example, the tax implications of ARRA spread over ten years and can involve complex computations. Sharing these analytic tools can lead to their refinement overtime, especially if it becomes easier to relate actual budgetary data back to CBO projections.

3.2 Linking Data: Identifiers, End-to-End Tracking, and Comprehensiveness

Just as it is difficult to trace the spending estimates used in the legislative process, it is nearly impossible to know how actual spending patterns mapped to Congress's wishes as expressed in the passed legislation. How do spending provisions in law lead to allocations of money in specific programs, accounts, and projects? How are legislative decisions implemented across various agencies and administrative divisions?

Current Problems

- *Missing Connections between Legislation and Executive Processes:* Through some effort, one can look at items in the original legislation and see how they may fund different programs (TASes) in ARRA. However, there is no official and unambiguous way to tie a given line item in PL 115 to a specific award flowing from that item.

- *Lack of Comprehensive Catalogs of Identifiers:* There is no definitive, public list of Recovery-related TAS identifiers. [5] Without this list it is impossible to get basic understanding of how the Treasury responded to this legislation. It is also impossible to understand the degree to which Recovery.gov data represents the complete and comprehensive picture. We have no way of knowing what should be expected and what data may be missing.

The degree to which reported data represents a complete and comprehensive account of spending must be communicated and demonstrated. As discussed above, there is no definitive list of *all* TAS identifiers used in the Recovery act. Without this list, it is impossible to know whether recipient reporting is 20%, 50%, or 90% complete and comprehensive. Moreover, the Recovery Act authorized exceptions on detailed reporting. In how many instances did such exceptions apply? How much money flowed through channels that did not require detailed reports?

Proposed Solutions:

1. *Publish Machine-Readable Catalogs of Identifiers:* There are many units of analysis and other entities that can directly or indirectly help explain ARRA reporting data. These include data about recipients,

[5] Yee, Raymond. 2009. "ARRA Treasury Account Symbols: the outcome of our FOIA request." Data Unbound. http://blog.dataunbound.com/2009/11/23/foia-outcome/ (Accessed May 13, 2010)

relationships among recipients (contractor-subcontractor or subsidiary relations), federal agencies and bureaus involved, and information about administrative divisions (states, municipalities, congressional districts, counties). Authoritative services to share updated lists of these identifiers would give a reliable basis for understanding if reporting measures gave complete and comprehensive accounts of government spending. The catalogs should also be available in machine-readable formats to facilitate software applications.

2. *Define URI/URL Templates:* Identification systems and coding systems need to be easy to resolve on the Web. Well defined templates for looking up Web resources for an entity associated with a specific identifier or a code improves the usefulness of coding systems. The templates themselves should be easy to read and understand for a layperson, and in doing so, they will be easier to understand for third-party software developers. To improve the longevity of Web identifiers, URI/URL templates should also avoid dependencies on the specifics of back-end implementations such as scripting languages (such that ".php," ".aspx," ".net" should not be in Web addresses).[6]

3. *Enable Linking of Data Across Agencies:* Much budgetary data only makes sense through reference to entities that are not directly described in a given dataset. The use of clear identifiers (ideally Web URIs) can provide key information to help link administrative processes within these various sectors of the Executive Branch with the wishes of Congress as expressed in final legislation. Spending data on different programs can be better contextualized with the use of Web identifiers. We should be able to follow links from spending data to see documentation about a program, relevant metrics reporting on the goals and achievements of that program, and who is responsible for overseeing that program. We also should be able to follow links from spending disclosures on a specific program back to the original legislation authorizing funding for that program.

3.3 Data Consistency, Quality, and Data Integrity Enforcement

Current Problems Measures to enforce data quality and consistency are critical to making data intelligible and usable. This is especially true given the scale and complexity of Federal spending. For the ARRA, there are several data quality problems that have hampered the utility of data and have cast doubt on the overall reliability and trustworthiness of disclosure methods.

- *Confusing Terminology, Unevenly Applied:* Transparency measures will fail if no effort goes toward making basic concepts and data intelligible to the public. Educational aids need to be offered so the public gain grasp the basics of how federal budgeting and spending works. For example, public discussion of ARRA referenced two different grand totals, $782 billion and $797 billion, numbers differing by $15 billion dollars. One estimate came from the CBO as a "budget authority" number, while the other came from the same CBO estimate as an "estimated outlay." Yet the rationale and meaning of these two different numbers remains obscure to the public.

- *Lack of Controlled Vocabularies:* Cities are fields in ARRA recipient reporting. Unfortunately, the public datasets have poor-quality city data. For example, there are numerous misspellings of "Los Angeles" making it difficult to classify and summarize financial reports by cities.

- *Key Data Missing or Omitted:* In many cases, ARRA recipient reporting has critical gaps that short-circuit understanding. In some examples, there are reports of sub-recipients without a corresponding primary recipient. In other cases, there are award keys with more than one primary recipient. In other examples, the local amounts add up to more than the award amount. Sometimes the TAS is only in

[6] For a comprehensive discussion, see: Berners-Lee, Tim. 1998. "Hypertext Style: Cool URIs don't change." http://www.w3.org/Provider/Style/URI (Accessed May 13, 2010)

the form of a placeholder ("91-XXXX"). These missing data can be found in many recovery reports without explanation.

- *No Notification of Updates:* Citizens can download the recipient data from http://www.recovery.gov/FAQ/Pages/DownloadCenter.aspx, but it is very difficult to tell what version of the data is currently available for downloading. There are no services (such as Atom feeds) for sharing update notices. There is no provision for obtaining archived earlier versions of data.

To mitigate some of these problems, spending reporting systems must take a variety of measures. The list below is not exhaustive, but represents some basic approaches to tackling some of the critical quality and integrity issues that can harm transparency efforts.

Proposed Solutions:

1. *Use Existing Controlled Vocabularies:* Reporting systems need to make better use of controlled vocabularies in data entry. The US Census already produces a list of cities, which should be reused in this context. Enforcing a controlled vocabulary for congressional districts would have prevented the political embarrassment over non-existent congressional districts. Use of controlled vocabularies should not be too rigid, since it is sometimes important to accommodate fuzziness and "edge cases." In such cases it is better to have an "other" option, so that there is some assurance that the responder is stating explicitly he or she is opting out of the choices and not making an entry error. A catalog of controlled vocabularies that is readily available and understandable to implementers of federal spending applications would help tremendously.

2. *Enforce Data Model Constraints:* Since Recovery.gov and/or Federalreporting.gov creates an award_key identifier, it can certainly ensure that not more than one primary recipient can be associated with that key. The existence of award keys with more than one primary recipient calls into question the integrity of data collection efforts.

3. *Reuse Existing Data to Promote Quality:* One would assume (perhaps incorrectly) that Federalreporting.gov should know what recipients should be reporting. After all, awards to the primary recipients are coming from the agencies. These agencies can give Federalreporting.gov an authoritative list of recipients. Hence, if a sub-recipient reports to Federalreporting.gov, that sub-recipient should be able to choose from a pre-defined list of awards. Using such measures, there should not be an award_key without a primary recipient.

4. *Business-Rules for Error Checking:* Invite the community to collaborate on a data diagnosis dashboard. The *Recovery Act Transparency Board (RATB)* and public interests groups can define certain business rules that should be in place to test data quality (e.g., one-to-one correspondence between award_key and primary recipient, award_amount must equal total of all local_amount, etc). These tests can help improve quality of data entry and monitor the quality of data once reported.

5. *Report Errors and Version Data:* The open source software development community has well developed practices for identifying problems and publicly track issues. An issue repository or analogous system can help officials, public interest groups, and members of the public identify problems, track progress on fixes, and propose solutions.

6. *Data Citation:* A critical need for making Recovery.gov more trusted and reliable is to make its content citable. One needs solid assurances that the expected version of a dataset will be obtained at an expected Web URI. These kinds of measures complement community versioning and issue tracking discussed above.

3.4 Summary: How Data Silos Hamper ARRA Transparency

One of the largest and most systemic obstacles to greater government transparency comes from the "siloed" mentality of many agency officials and government contractors. "Data silos" are systems defined by their inability to share common types of information. When key data are trapped in individual systems, inconsistencies emerge across these different systems. It becomes much harder to coordinate key government processes, assess performance, or account to the public.

The problems of data silos and the administrative dysfunctions that lead to such silos can be seen in ARRA reporting and disclosure systems. Many of the shortcomings in ARRA transparency come from lack of coordination across different sources and types of government information. As discussed, there is no way to simply connect the intent of the original legislation and the actual spending reports of ARRA funding recipients. Data about the Congressional record and ARRA simply lack the unambiguous identifiers (or "hooks") needed to establish these clear ties.

The lack of apparent coordination between the Treasury and OMB exemplifies our concerns over silos. We have yet to find an authoritative crosswalk between TAS/OMB coding systems. We were able to obtain such data through a *Freedom of Information Act (FOIA)* request, through collations developed by ProPublica, and in some "easter-eggs" inadvertently hidden in Excel spreadsheets provided by Recovery.gov. Obviously, this sleuth work to obtain incomplete "guesstimates" of how TAS amd OMB track money does not represent an ideal picture of transparency. Since OMB and the Treasury both maintain key identifiers needed to track and understand budgetary processes across the Federal Government, it is absolutely critical that they offer a publicly accessible system to reconcile their coding systems.

Similarly, there are telling discrepancies and problems in coordinating information across different sources such as *USAspending.gov* and *Recovery.gov*. Doing this alignment should be easy but is non-trivial because of subtle ways in which DUNS numbers (a unique identifier developed by Dun & Bradstreet and used by the federal government to identify award recipients) and award keys that are supposed to be stable identifiers are not actually always the same between systems. Bridging across these systems is important, because with comparison with USAspending.gov, it is difficult to understand how patterns in Recovery.gov spending compare with the overall picture of the Federal budget. In addition, USAspending.gov offers APIs (services) giving updated streams of data, while Recovery.gov offers bulk downloads. Ideally, we should have both of these capabilities in both systems.

Without specific policy efforts to work against data silos, the problems of reconciling data from different systems will only get replicated. In the future, Congress may enact additional legislation like ARRA that will require specialized websites analogous to Recovery.gov for reporting. However, the key issues for all such specialized sites will be how they relate to the larger context of federal information, including other budgetary disclosure systems such as USAspending.gov.

4 Lessons from ARRA: Agencies Need Incentives to Publish Useful Data

The above discussion of ARRA highlighted several shortcomings, many of which relate to systemic problems of "data silos" in government contexts. This point illustrates how technology choices take place in organizational and political contexts. Budgetary transparency measures will fail if agencies do not have a vested interest in opening data. Without clear incentives, officials will make poor and constraining technology choices, and the quality of implementations will suffer. Currently, officials see little reason to look beyond their own agency needs. This helps motivate the creation of "silos" that trap data within a limited and narrow range of interfaces.

Agencies need to be rewarded for providing data, via reliable services, that see reuse, aggregation, and integration with other government and citizen applications. Agencies must also be responsive to requests

from members of the public and from other parts of the government to make data available in formats and services that maximize convenience and uptake. Hiring a contractor that holds a few conference calls among transparency advocates does not demonstrate meaningful commitment to public collaboration. Consultation processes need to have real impact. Outside agencies and the public interest groups need ways to articulate specific requirements and evaluate deployments to guide revisions that better meet public needs. In other words, meaningful consultation, issue tracking and improvements need to be continual, not just at the initial planning stages. To effectively harness consultation processes to design and build systems, agencies should adopt participatory design methodologies. If agencies saw clear rewards and incentives for providing data and services that see reuse, they will invest the effort and thought needed to support interoperability.

The contracting process, including the *Request for Proposals (RFP)* and the Smartronix bid for building Recovery.gov was less than ideal for building public trust in ARRA transparency. The RFP saw little public input or consultation, and the Smartronix bid itself was only released in a heavily redacted form. These government procurement procedures do not lead to optimal outcomes, and the problems we see in Recovery.gov stem from the deficiencies in the development process. Thus, the process of building transparency systems should be regarded as important as the outcome. Collaborative and iterative processes of design help build the trust, accountability, and shared problem solving that transparency advocates seek. Therefore, participatory design methods should be considered essential for building transparency systems.

5 Citation, Trust, Curation, and Integrity

In bridging across silos, agencies need to consider measures to make their data more reliably identified and permanent. Citation is absolutely integral to understanding. It enables and expresses collaborative knowledge production across space and time, and it is the foundation on which evidence and arguments are identified, assembled, reused, and critiqued. This is true in both formal professional settings of academic communication, and in public debates expressed online in weblogs, forums, and social networks. On the Web, citations usually take the form of a hyperlink (URI/URL) that identifies a specific information resource and, while doing so, makes its retrieval fast and efficient.

For budgetary transparency, citation issues represent a key concern. As already discussed, citable, linked data is required to help provide needed context to budgetary data. However, citation also involves issues of reliability and trust, making data version control and archiving critical issues. These requirements represent key challenges, especially since important information for making sense of the Federal Budget will be distributed across many different agencies and data systems. To help address these issues, designers of systems that support budgetary transparency need to consult the digital library community. This community has developed extensive expertise in designing technologies, business processes, and policies to safeguard data and its integrity, identification, and reliable retrieval [1]. The organization *Citability.org*[7] also provides invaluable guidance for making Web based government information more trustworthy and easy to reference. This group has explored important issues about granularity and specificity in citation and retrieval.

In some ways, the citability and long term curation of different versions of datasets may represent a more important priority than "signing data." Large and complex datasets will likely contain some errors, even with good data validation practices. If we required agency officials to "sign data" as completely accurate, we may create incentives to withhold key information for fear that their accuracy will never be 100% perfect. Instead, we should reward agencies that make their data cleansing and error correction processes more open and transparent. Version tracking and citation can help foster a more collaborative and less adversarial relationship than some signing proposals.

[7] http://citability.org

6 Getting the Big Picture: Metrics, Classification and Discovery

Classification will underpin many uses for budget data. Datasets will need different forms of administrative, technical, and descriptive metadata to be managed. Given the scale and complexity of the federal government, budgetary metadata requirements will likely be similarly complex. The *National Information Exchange Model (NIEM)* may provide a good foundation for meeting these government-wide standards needs. The *Extensible Business Reporting Language (XBRL)* represents another potentially applicable standard that may be useful for describing many financial transactions in government.[8]

Standards and classification concerns are not only technical matters; they are also important issues impacting the usability of data. Searching and retrieving relevant information from large bodies of complex data is a challenge for many information services. Keyword searches are common solutions to this problem. However, keyword searches often yield incomplete and ambiguous results. The "hit or miss" nature of keyword searches can limit the effectiveness of transparency. Poor search capabilities can lead to omissions of relevant data. In turn, this can reduce trust in a disclosure system. Information retrieval systems need to be more robust, reliable, and predictable if they are to support transparency objectives and earn public trust.

To avoid some of the difficulties associated with keyword searches, metadata standards can be used to support "faceted navigation" systems appropriate for exploring and retrieving spending data. In faceted navigation applications, users can leverage sophisticated classification systems using simple and intuitive "point and click" selections. Users progressively home in on more specific information from a larger collection. Because filters are applied across an entire collection, users have greater certainty in the comprehensiveness of their results than with keyword searches. Feedback, in the form of subtotals for the numbers of items that fall under each available facet, guides users in the selection of additional filters [6]. This feature offers users important information cues about the size and composition of the collection they are searching. Dynamic feedback also helps users understand sometimes very sophisticated classification systems.

While classification can be an important enabler to understanding, classifying information across the vast Federal bureaucracy is no easy matter. It can and likely will be debated and contested. Classification exists in social and organizational contexts. Some types of classification may work well in one context but may work poorly in another context. Moreover, classification is not necessarily an objective process. It is shaped by the assumptions and goals of people and organizations. These worldviews and goals often see disagreement and evolve over time.

Thus, not all metadata, including performance metrics, will be appropriate in all situations. Failing to pay adequate attention to context can be harmful, especially with top-down imposed schemes. For example, without careful alignment to an organization's mission, poorly applied metrics may create incentives for the wrong outcomes. Agencies, programs, and the public need to come together to shape classification systems so that they reflect both the need to understand general government-wide processes and the more particular concerns of a specific program or agency. Similarly, metadata systems need to be responsive to evolve to better meet both existing needs and those not yet anticipated. While there is a clear need for government-wide standards, efforts to make the Federal Budget intelligible must also accommodate the great diversity of program goals, needs, and contexts. Just as certain metrics need to be carefully shaped to be appropriate for a given context, classification and description of budgetary data must also allow for enough nuance to facilitate and not obscure understanding.

[8] XBRL is an open data standard only in terms of access to the equivalent of its 'source code'. However, the governance structure of the XBRL consortium differs greatly from open source approaches. Paid membership and a focus on transacting business at physical conferences deviate from open source practices, and represents severe barriers to entry in participating in XBRL standards development. Because of these issues, it may be better not to use only XBRL, and consider NIEM and other standards as well.

7 Looking to the Future: Web Technologies

It is important for transparency measures to be open and easily used by the broadest constituents possible. This overarching need should guide decision-making on various implementation styles and technologies. However, many constituencies on the Web have different opinions on how best to deploy Open Government systems. Different technical styles will be promoted by different vendors, academic experts, and other stakeholders. Agency officials and systems architects will no doubt face many different sales pitches and expertly informed opinions. For example, proponents of the Semantic Web [3] (or "Linked Open Data", which is another term for the same set of technologies) often promote a specific suite of technologies centered on the *Resource Description Framework (RDF)*. RDF tools and technologies can be both powerful and very sophisticated. They enjoy widespread interest in many academic, research, and some commercial communities. In other cases, enterprise computing vendors may promote their specific vision of "Service Oriented Architectures" (usually highly planned and elaborate distributed computing frameworks). On the other hand, developers oriented toward the public Web often favor more "document-centric" and/or *Plain Web* [7] technologies. These other constituents typically favor a different suite of technologies, including Atom feeds and other RESTful services, together with XML [4] and JSON [5] for expressing structured data. Each technical style has its own set of proponents, advantages, and disadvantages, and the landscape constantly shifts as Web technologies continually evolve.

The current Web sees many different approaches working in parallel to meet different needs and solve different problems. Good architectural decisions, especially with regard to Web identifiers (URIs), will enable the Federal Government to best serve the needs of multiple communities, including both Semantic Web and Plain Web proponents. Ideally, budgetary disclosure systems should not favor one contested approach over another or pick individual technology winners and losers. Government transparency efforts therefore should not emphasize one technical style at the expense of another, but should expose information to multiple kinds of interactions. Agency officials should evaluate systems design proposals on the basis of how well they support multiple approaches to using data. Participatory and collaborative design methodologies that involve users (including third-party programmers) and other stakeholders will help agency officials prioritize technology choices that meet actual needs.

8 Conclusions

Policymakers need to understand that the key problems in making government data more open and useful come mainly from organizational challenges and less from technology. Without concerted attention on incentives, administrative processes, and institutional concerns, technologies such as "Service Oriented Architectures," the "Semantic Web," or "Cloud Computing" will remain empty buzzwords rather than real design solutions. A single system or technological choice will not automatically yield meaningful forms of transparency. Rather, officials need to identify strategies to better align agency priorities toward greater openness and collaboration between each other and the public. By institutionalizing and rewarding collaboration, information access, and data portability, agency officials will have the right processes and incentives to find appropriate technologies that best support greater transparency.

References

[1] Micah Altman and Gary King. A Proposed Standard for the Scholarly Citation of Quantitative Data. *D-Lib Magazine*, 13(3/4), March 2007.

[2] Tim Berners-Lee, Roy Thomas Fielding, and Larry Masinter. Uniform Resource Identifier (URI): Generic Syntax. Internet RFC 3986, January 2005.

[3] Tim Berners-Lee, James A. Hendler, and Ora Lassila. The Semantic Web. *Scientific American*, 284(5):34–43, May 2001.

[4] Tim Bray, Jean Paoli, C. Michael Sperberg-McQueen, Eve Maler, and François Yergeau. Extensible Markup Language (XML) 1.0 (Fifth Edition). World Wide Web Consortium, Recommendation REC-xml-20081126, November 2008.

[5] Douglas Crockford. The application/json Media Type for JavaScript Object Notation (JSON). Internet RFC 4627, July 2006.

[6] Marti A. Hearst. Clustering versus Faceted Categories for Information Exploration. *Communications of the ACM*, 49(4):59–61, April 2006.

[7] Erik Wilde. The Plain Web. In *First International Workshop on Understanding Web Evolution (WebEvolve2008)*, pages 79–83, Beijing, China, April 2008.

[8] Erik Wilde, Eric Kansa, and Raymond Yee. Proposed Guideline Clarifications for American Recovery and Reinvestment Act of 2009. Technical Report 2009-029, School of Information, UC Berkeley, Berkeley, California, March 2009.

[9] Erik Wilde, Eric Kansa, and Raymond Yee. Web Services for Recovery.gov. Technical Report 2009-035, School of Information, UC Berkeley, Berkeley, California, October 2009.

[10] Jeffrey D. Zients. Memorandum for Senior Accountable Officials Over the Quality of Federal Spending Information. Office of Management and Budget Memorandum, April 2010.