

An Open Government Research and Development Agenda Setting Workshop

A Summary Activity Report

**April 27-28, 2011
University at Albany, NY**

June 2011

***Prepared by:*
Meghan Cook
M. Alexander Jurkat**

Introduction

The Open Government Research and Development Agenda Setting Workshop was held on April 27- 28, 2011, at the University at Albany in Albany, NY. The workshop, sponsored and conducted by a collaborative team from the Center for Technology in Government (CTG) at the University at Albany, the Tetherless World Constellation (TWC) at Rensselaer Polytechnic Institute, the Institute for Law and Policy (ILP) at New York Law School, and Civic Commons was organized to outline a research agenda focused on opening up, federating, and using data to improve the lives of citizens (See Appendix A for a listing of the workshop team).

This document presents a non-attribution account of the contributions made at the workshop. The information presented in this report is not synthesized or analyzed, it is strictly an account of the workshop activities and discussions. It begins to shed light on complexities of developing a national open government research agenda. Following the release of this activity report we will focus on the analysis of the results working toward a set of recommendations and action steps.

We look forward to continued discussion of a national open government research agenda as a community of research and practice

Workshop Purpose, Approach, and Design

The Open Government Research and Development Agenda Setting Workshop was organized to bring together leaders from government, academia, and the non profit community to outline a research agenda focused on opening up, federating, and using data to improve the lives of citizens.

Held at CTG, the workshop built on the discussions started at the Open Government Research and Development Summit convened by the Networking and Information Technology Research and Development Program (NITRD) and the National Archives and Records Administration (NARA) on March 21st- 22nd in Washington, DC.

The objective of the workshop was to create an actionable and relevant multi-year open government research and development agenda focused on:

- Identifying critical needs
- Mapping needs to potential solutions
- Identifying legal and policy barriers
- Exploring critical evaluative approaches
- Laying out strategies for obtaining future research funding.

Taking an interactive and interdisciplinary approach, the workshop activities were a mix of plenary and small group sessions facilitated by members of the workshop organizing team. Using a set of established facilitation methods and techniques, invited participants took part in several activities including:

- Introductions and icebreakers
- Identification, affinity clustering, and prioritization of research questions
- Small group discussion of clusters
- Identification of biggest challenges in open government
- Identification, categorization, and discussion of research questions by lens and timeframe
- Reflection and thoughts on moving to a national research agenda
- Identification and discussion of closing thoughts

The sections that follow provide the summary results from each of the from each activity in a non-attribution format.

Activity #1: Introductions and icebreakers

At the opening of the workshop, participants took part in a plenary session where they were asked to do three things:

- Introduce themselves stating their name, organization, and a general statement about their work (see Appendix B for a full listing of participants).
- State one word they would use to describe an ideal research agenda.
- Name their favorite online application or site.

Describing an ideal research agenda.

- | | |
|---------------------|--|
| • Actionable | • Interesting |
| • Authoritative | • Intriguing |
| • Bottomless | • Iterative |
| • Challenging | • Meaningful |
| • Clarified | • Multi-level |
| • Clear | • Open data |
| • Collaborative | • Plain spoken |
| • Concrete | • Problem-focused |
| • Consistent | • Robust, Scalable |
| • Diverse | • SMART (specific, measurable, attainable, realistic and timely) |
| • Executable | • Strategic |
| • Expansive | • Sustainable |
| • Feasible | • Target claims |
| • Focused | • Ubiquitous |
| • Fundable | • Usable |
| • Humble | |
| • Intentional | |
| • Interdisciplinary | |

Participants' favorite applications and websites

- Amazon <http://www.amazon.com>
- Bill Shrink <https://www.billshrink.com>
- Brightscope
<http://www.brightscope.com/>
- Business Card Reader for iPhone
<http://www.shapeservices.com/en/products/details.php?product=bcr&platform=iphone>
- Companionlink
<http://www.companionlink.com/>
- Congress for Android
<http://sunlightfoundation.com/projects/congress-for-android/>
- DPREview <http://www.dpreview.com/>
- Dropbox <http://www.dropbox.com/>
- Epicurious <http://www.epicurious.com/>
- Evernote <http://www.evernote.com/>
- Flashlight
<http://itunes.apple.com/us/app/flashlight/id285281827?mt=8>
- Foursquare <https://foursquare.com/>
- Free Caddie <http://www.freecaddie.com/>
- Gapminder <http://www.gapminder.org/>
- Google Calendar
<http://www.google.com/googlecalendar/about.html>
- Google Docs <http://docs.google.com>
- Google Maps <http://maps.google.com/>
- Google Scholar
<http://scholar.google.com>
- Govloop <http://www.govloop.com/>
- Instagram <http://instagr.am/>
- Kayak <http://www.kayak.com/>
- iPad Game for Cats
<http://www.ipadgameforcats.com/>
- Layar (augmented reality)
<http://www.layar.com/>
- Mindsnacks <http://www.mindsnacks.com>
- myBenefits
<https://www.mybenefits.ny.gov/selfservice/begin>
- Next Bus
<http://www.nextbus.com/homepage/>
- OAuth <http://oauth.net/>
- Pandora <http://www.pandora.com>
- Platform as a Service
http://en.wikipedia.org/wiki/Platform_as_a_service
- Reader Supported News
<http://www.readersupportednews.org/>
- Recalls.gov <http://recalls.gov/>
- Satellite Visibility
<http://www.psychicpsquirrel.com/iphone-apps/satellite-visibility>
- Skype <http://www.skype.com/intl/en-us/get-skype/>
- Trip Advisor
<http://www.tripadvisor.com/>
- Trip It <http://www.tripit.com/>
- Twitter <http://twitter.com/>
- Wikipedia <http://www.wikipedia.org/>

Activity #2: Identify, cluster, and prioritize research questions

This visioning activity allows a group to brainstorm a set of ideas and organize them in a way that promotes a shared understanding and it sets the foundation for subsequent discussions. It also generates discussion and offers a technique to discover similarities and differences in perspectives. The physical product of the activity is a prioritized set of named “clusters” of questions.

Participants were given an overview of the “affinity clustering” process and then were asked to respond to the elicitation question:

- *What is a research/researchable question focused on the use of government information to improve the lives of citizens?*

Each participant wrote their research questions on 8x11 sheets of paper, one question per sheet. Then in a round robin fashion each was asked to read their question aloud to the group. As each question was shared the paper was posted on the wall. As the questions were posted the group began to work together with the facilitators to decide which items were alike or related and to decide placement in emerging cluster. After all the questions were read aloud and posted to the wall, the group worked to refine the clusters. Once all questions were placed into clusters, the group named each cluster with a statement that best described the ideas represented in that cluster. After all clusters were named, participants were each given four sticky dots and asked to vote on the clusters they thought were most important to address in the subsequent discussions at the workshop. The prioritized list of clusters that emerged from this activity are listed below followed by a full list of the items included in each cluster.

Cluster	Votes
The Value of Open Government /Ecosystem of Open Government	19
What Do Citizens Want?	16
Government Capabilities	10
Incentivizing	10
Getting Data Into Citizens’ Hands	9
Building the Toolkit	8
Data Quality/Fitness for Use	6
Standards/Interoperability/Integration	6

Cluster: The Value of Open Government / Ecosystem of Open Government

- How does access to data effect the quality of life for the average person?
- How do we holistically measure total ROI in open government?
- How can open government lead to public accountability?
- What is the best way to analyze the costs and benefits of “smart disclosure” of data on consumer markets?
- What are the goals and purposes of government information transparency in different contexts and across agencies?
- How does the value of government information vary across levels of government?
- Who is going to pay for all of this? What are the best ways to spend the funds?
- Does the publication of data encourage collaboration among stakeholders?
- What is open government and who says?
- What does open government mean?
- Will open government information initiatives actually make government more open? What are the longer term cultural changes?
- Does publishing data improve government performance?
- Does a correlation exist between better information and better service?
- What is the open data ecosystem? What are the relationships among vendors, citizens, journalists, and start-ups?
- What are the concrete suggestions on how to bring coherence to the open data ecosystem?
- Qui Bono? Who benefits the most from open government data?
- What is the business case for open government? Where are the case studies?
- What are the key metrics targets for evaluating if an open government initiative is working?
- What outcomes do we seek from open government (savings, trust, better service, more equal empowerment)?
- How do we measure the success of open government so we can justify the resources spent?
- The economics of open data—what is the value proposition for opening data? Why do it?
- When quantifying the value of open government, what types of value—cost savings (e.g., IT dashboard), economic activity (e.g., brightscope)—exist? How do consumer surplus, service delivery, policy outcomes factor in?
- How do we measure improvement in citizens’ lives?
- How can we evaluate whether open government increases or decreases economic and political inequality?
- Will open government widen the digital divide and redefine it? What are the effects of limits on a citizen’s ability to use and interpret information?
- How can we provide the best value by integrating open data across all levels of government, business, and industry (as appropriate)? How do we address the value to business?
- How do we execute open government at the local level with minimal budgets?

- What are the funding opportunities for open government data?
- How can open data improve the efficiency of government itself (i.e., intragovernment collaboration)?
- What, if any, relationships exist between levels and types of educational expenditures and education outcomes?
- What types of information most effectively enhance existing initiatives of civil society?

Cluster: What Do Citizens Want?

- How do we uniquely identify a citizen and insure his/her well being?
- How does government know (hear about) what citizens need to know?
- How do we know what citizens need and how/if they can access it?
- How do citizens know what they need to know?
- What data is available and useable to help improve citizen health?
- Who is currently using what data sets in data.gov and how do they evaluate the transactions?
- What citizen-generated and citizen-created data is most critical to open, effective government?
- What data sets now available (e.g., in data.gov) will help consumers make better marketplace choices? What are the biggest gaps in available data?
- What information has high consumer demand?
- How do citizens find out information on subjects such as whether they can get their house reassessed for tax purposes (given the real estate downturn) or if their drinking water is safe?
- How do I find out what data is available about me and how do I control access to it?
- Can the public tell us where data isn't available or isn't usable?
- How do we collaborate more openly with citizens to help us identify issues based on government information?
- What type of citizens care about which types of government information?
- What is the politically relevant constituency for technology-enabled open government?

Cluster: Government Capabilities

- How can government best support citizen access and interpretation of government data?
- What government information has the potential to support citizen decision-making on important events?
- How can government capture data about its own website users and use that data to improve its services and foster public engagement?
- How can we identify market failures that are rooted in a lack of publically available data?
- How well do current government organization business processes and practices lead to publically usable data? Can they do it better?
- How can we share and combine public and private data to make it more useful to both governments and citizens?

- To what extent should government provide information rather than data?
- How do we organize present government data to facilitate a more open dialogue around key issues (creating true two-way feedback)?
- What new capabilities do government agencies need to increase openness?
- What policies prevent government from responding to citizen demands resulting from greater openness?
- Which state and federal governments have open government initiatives?

Cluster: Incentivizing

- How do we incentivize agencies/governments to share data/information?
- How do we get the lawyers on-board?
- How do we build and incentivize communities (both real and virtual) to care about and use government data?
- Can the government identify where data is lacking or unusable better than any individual or group can?
- What motivations, agenda, and values do citizens bring to the engagement with a government agency?
- What incentives create public participation (e.g., money, recognition)?
- How do we incent government to release more data?
- How do we minimize the cost of opening up legacy government data silos?
- How do we make open data more sustainable? Can the government charge companies for data?
- Where/how can individuals or organizations decide to opt out of government data sets?
- What are the legal, regulatory, and similar barriers to citizen participation?

Cluster: Getting Data Into Citizens' Hands

- How can I find usable government data on topic X? How can I find apps that use that data?
- How can government use citizen-produced data to be innovative, effective, and efficient?
- How can government help increase digital literacy among citizens so they better use government data?
- At what point or how long does it take for citizens to give up a search for government data online and why?
- How do we provide a proactive service plan for a citizen based on what we already know about him/her?
- Can we push information out based on individual profiles? Do citizens want this?
- How do we define communities to facilitate citizens knowing about happenings in their communities (health, energy, environmental, poverty)? Are citizen definitions feasible?
- How can we use healthcare outcome data more effectively to improve the health of citizens.
- What methods are effective in helping citizens understand and manipulate government data?

- Where do citizens go to find government data?
- How do we make it possible for citizens to compare what their state and local areas pay for the same services/goods? How do we make them understand why differences exist?
- Do we really want one central compendium of data?
- Given the constraints of government marketing and all of the noise online, what is the best way to reach out and inform citizens of open government projects?
- How are governments defining data? Do these definitions exclude citizen access to certain documents?
- What is the impact of mobile computing on citizen use of data? Can citizen use of data be furthered through mobile technology?
- How can the broadest government datasets (such as Social Security and Medicare) be used without violating privacy concerns? Should they be?
- How can we disseminate data about nutrition information from local restaurants? This would aid citizens in making decisions about dietary and medical concerns.
- What data collected by governments is currently available in a best practice usable format?
- What is the full universe of data that government collects? What and how much is siloed away from the public?
- What can Medicaid expenditure data tell us about the variations in medical practice and adherence to best practices for healthy outcomes?

Cluster: Building the Toolkit

- What data is captured and already analyzed, but not made public or easily understood?
- What are the trends in usage of existing data sets and formats in terms of social impact?
- How do we merge information for government and public? Is one website for all information desirable or feasible?
- How much can the formats for data disclosure be standardized across industries and regulatory agencies?
- What is the relevance of government enterprise IT trends (such as consolidation, adoption, of hybrid cloud hosting, shared services) to open gov/open data?
- How can we make it easier for community groups to use government data without having to hire programmers or other IT professionals? Can we make data mashups as easy as blogging?
- How can we promote the inclusion of semantics in structured data to make it accessible via the semantic net?
- What is the role of large scale spatial/GIS data (imagery)?
- How can we do better than keyword searches in finding relevant databases from open government?
- What type of information is used by citizens in third-party apps?

Cluster: Data Quality/Fitness for Use

- How can we support the augmentation of data (adding structuring metadata, derivative data products, correcting) over time?
- How do we find data?
- How do I know what was done to the data before I use it? How do I know how reliable it is for my purposes?
- Can we create different models for providing access to data that doesn't create more confusion (among app developers and users)?
- What are the best criteria for establishing what data should not be publically available?
- How do we work with data that has some restrictions (terms of use rules, privacy, security)?
- What specific datasets are giving us "big" problems?
- What is the framework for choosing and defining high value data sets based on public value?
- What kinds of high value open government information are not in digital form?
- How can we discover the internal biases of the data we wish to work with?
- How can a data user know that the data is fit for use?
- Can we automate publishing to keep public data more timely and useful?
- What can we say about the data quality as a whole? What happens when we mix data sets across different domains?
- What information currently being demanded has low confidence?
- Are governments sharing established information (e.g., GIS data) or in-progress information (real time data about sales tax collections)?

Cluster: Standards/Interoperability/Integration

- What strategies are best for publishing complex unstructured data?
- Can the various types of information be centralized, perhaps in one portal (like Google)?
- What entities should be responsible for open government technical architecture, infrastructure, policies, and standards in our global, civil society?
- Does the United States need its own standards or can the United States leverage existing international standards (metadata)?
- What are the best ways to let outdated policies/standards die?
- What strategies and technical architectures help data sharing and interoperability?
- How do we share (accurate) information across multiple agencies?
- How can we start effectively sharing more data and then analyze it better to identify issues earlier and more frequently that improve the lives of citizens?
- If standards are key to scale and use of data, why do we see resistance to standards?

Activity #3: Small Group Discussion of Clusters

After identifying, clustering, and prioritizing clusters, participants self-selected into small groups each focused on a specific cluster. Each small group was asked to develop a descriptive statement for their cluster and then discuss and record their answers to the following questions:

1. *What are the research traditions, tools, techniques, and theories that must be engaged to develop answers to the questions within the cluster?*
2. *What are the challenges researchers face in addressing the cluster?*
3. *What are the strategies to address these challenges?*
4. *What organizations should be targeted to support or participate in research about this cluster?*
5. *Why is this cluster important to efforts to use government data to improve the lives of citizens?*
6. *What are the implications of not addressing the questions represented in this cluster?*

The answers recorded by each small group and shared with the larger group are provided below. Please note that all clusters identified in the previous activity were discussed by a small group..

Small Group: The Value of Open Government /Ecosystem of Open Government

- *Descriptive Statement:* What is the value of an open data ecosystem? How can we catalyze the imagination of those who would or could benefit from open government?
- *What are the research traditions, tools, techniques, theories that must be engaged to develop answers the questions within the cluster?* The group identified econometrics and public value frameworks as potentially analogous. Quantitative and qualitative measures are necessary to capture total worth.
- *What are the challenges researchers face in addressing the cluster?* The lack of a generally recognized framework for measuring the effects of open government. Quantification measures cannot capture all the benefits, but they are the most influential in establishing value. Tracking information about data usage is unavailable or difficult to access. It is difficult to determine why citizens want certain information.
- *What are the strategies to address these challenges?* Researchers need to define and consult with a variety of stakeholders. They need to formulate guidelines and metrics, not hard-and-fast rules. The results of these studies should fall somewhere between the scattershot response to a Google inquiry and the too-narrow focus of a Wikipedia entry. Case studies should be identified and explored. Efforts should begin at a low, simple level and built over time.
- *What organizations should be targeted to support or participate in research about this cluster?* Citizens, consumers, employees, employers, civil and government authorities, researchers, and civic organizations.

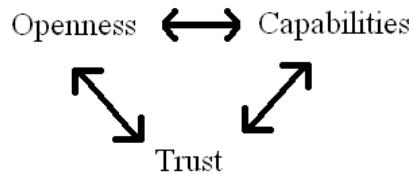
- *Why is this cluster important to efforts to use government data to improve the lives of citizens?* A valuation framework allows projects and resources to be prioritized.
- *What are the implications of not addressing the questions represented in this?* Unless you can prove the worth of open government initiatives, no one will support them.

Small Group: What Citizens Want

- *Descriptive Statement:* How do we figure out what citizens want?
- *What are the research traditions, tools, techniques, and theories that must be engaged to develop answers the questions within the cluster?* User desires can be measured by looking at available metrics for existing data sets, consulting with data and information managers, and employing a variety of tools (surveys, focus groups, usage matrices, online tools, feedback, and crowdsourcing), we can gauge citizen desires.
- *What are the challenges researchers face in addressing the cluster?* The challenges arise because data of potential interest may not be publically available, researchers must separate issues of the usability of data from the value of data, a large number of factors compound the issue, the government lacks an effective ability to survey/measure public, and no methodological consensus exists.
- *What are the strategies to address these challenges?* The challenges can be addressed by employing usability testing, working with diverse stakeholders, maintaining transparency, promoting crowdsourcing, and using feedback tools (creating new ones and analyzing existing feedback tool data). To the extent possible, feedback tools should be automated.
- *What organizations should be targeted to support or participate in research about this cluster?* Supporting organizations could include various local, state, and federal agencies; social data platforms like Socrata; academic centers; news media; and foundations.
- *Why is this cluster important to efforts to use government data to improve the lives of citizens?* By matching people more effectively with public resources, value is added to their lives and support for open government initiatives is maximized.
- *What are the implications of not addressing the questions represented in this?* If citizen desires are not included in the process of open government, resources devoted to this topic may be wasted and government inefficiencies exacerbated.

Small Group: Government Capabilities /Data Fitness for Use

- *Descriptive Statement:* How can open government initiatives drive innovations and improvements in government capability to collect, manage, use, integrate, and share information? The relationships at issue are between government and the private sector/civil society, across levels of government, within a particular government or agency. The following diagram shows a relationship among government capabilities, openness and trust.



- *What are the research traditions, tools, techniques, and theories that must be engaged to develop answers to the questions within the cluster?* The group identified information science/technology and computer, public policy, organization studies/change management, diffusion of Innovation, and usability approaches.
- *What are the challenges researchers face in addressing the cluster?* Researchers must define the question with a realistic view of the answer. They must identify all parts, formulate holistic views that include both sides (government and citizen), be mindful of the level of analysis, and explore cultural resistance.
- *What are the strategies to address these challenges?* The challenges can be addressed through an ecosystem perspective, feedback between research and practice, “depoliticizing” the subject, using best practice research communication and translation, a strong justification for their work, an experimental “sandbox,” and interdisciplinary team structures (government, academic disciplines).
- *What organizations should be targeted to support or participate in research about this cluster?* Organizations at all levels (federal, state, and local) should be consulted.
- *Why is this cluster important to efforts to use government data to improve the lives of citizens?* Efforts in this area are needed to move the discovery, implementation, and evaluation of open government initiatives forward.
- *What are the implications of not addressing the questions represented in this?* Failure to address this area will result in lost opportunities and mismanaged data.

Small Group: Incentivizing

- *Descriptive Statement:* The current government lacks a marketplace for government and citizens to more effectively share and use data. Such a marketplace would provide “carrots” for government decision-makers (data owners, legal departments, leadership) to release government data. It would also serve to identify the data needs of consumers (companies, individuals, and communities). It would create a connection between these two groups (suppliers and consumers) and lower the “cost of entry” for both.
- *What are the research traditions, tools, techniques, and theories that must be engaged to develop answers to the questions within the cluster?* The group listed the study of government transparency from a business perspective (Archon Fung), online communities researchers (in the area of communications and behavioral science), and traditional information scientists from database perspective.
- *What are the challenges researchers face in addressing the cluster?* Challenges lie in the diversity of communities and the lack of a coherent research community in online governance.

- *What are the strategies to address these challenges?* The revitalization of the digital government research community with specific focus on open government. It is critical to engage multi-disciplinary researchers. Finally, researchers need to increase value and reduce effort.
- *Why is this cluster important to efforts to use government data to improve the lives of citizens?* A marketplace for the sharing and use of data would improve citizens' lives and understanding, both directly and indirectly and help generate an innovation economy.
- *What are the implications of not addressing the questions represented in this?* Unless supply and demand is linked in a continual and meaningful way, data will not be used or useful.

Small Group: Building the Toolkit and Standards/Interoperability/integration

- *Descriptive Statement:* How do we create a shared “methodology” for publishing data that facilitates “mashing”? The group indicated that a data publishing methodology could include metadata, standards, dictionaries, provenance, and best practices and would focus on an iterative process of defining the data, prioritizing it, collecting it, publishing it, and using it (analyzing and augmenting). Mashing is a derivative of the software development idea of a mashup—a [Web page](#) or application that uses and combines data, presentation, or functionality from two or more sources to create new services. The term implies easy, fast integration, frequently using [open APIs](#) and data sources to produce enriched results that were not necessarily the original reason for producing the raw source data ([http://en.wikipedia.org/wiki/Mashup_\(web_application_hybrid\)](http://en.wikipedia.org/wiki/Mashup_(web_application_hybrid))).
- *What are the research traditions, tools, techniques, and theories that must be engaged to develop answers to the questions within the cluster?* The group identified master data management, data science, informatics, ontology, standards formation process (e.g., W3C), unofficial emergence (e.g., internet protocols), case studies, and diffusion Theory/Practices
- *What are the challenges researchers face in addressing the cluster?* The challenges identified include the diversity of data techniques and methods, concerns about the integrity and (perhaps hidden) biases of any given data, the existence of proprietary claims that might require incentives for sharing, the natural friction/apathy/inertia prevalent in any organization, and the actual and opportunity cost of devoting resources to making data interoperability.
- *What are the strategies to address these challenges?* Potential strategies to counter the challenge are consensus building on standards, encouraging the frontloading of integrity and provenance information, the predictive association of data, efforts to evangelize the value of open data including an identification of the mutual gain from facilitating sharing, and the promotion of risk taking in this area.
- *What organizations should be targeted to support or participate in research about this cluster?* Support for these efforts might be found in executive and legislative leaders, private/public standards group (OASIS/NITRE), data collecting agencies, and representatives of consumers (assuming such groups exist).

- *Why is this cluster important to efforts to use government data to improve the lives of citizens?* Data interoperability and sharing are key to realizing the greatest possible gains from open government.
- *What are the implications of not addressing the questions represented in this?* If data cannot be used, analyzed in innovative combination, shared or “mashed,” open government is relatively meaningless. Lack of interoperability promotes secrecy, wasted effort, and lost opportunities. Finally, lack of simple interoperability concentrates power (knowledge = power) in the hands of the those with the money and talent to leverage the data that does exist. More and more data will be merged and analyzed as time goes on. The question is who is going to do it and how widespread its benefit will be.

Activity # 4: Considering research questions by lens and timeframe

In this activity, participants were asked to identify research questions based on perspectives; lens and timeframe.

The perspectives are described as:

- *Lens:* Four lenses were proposed as ways to think about open government research. Participants were asked to identify research questions from the perspective of four lenses 1) law and policy, 2) management, and 3) technology and 4) cross-cutting.
- *Timeframe:* Two timeframes were presented as targets 1) Short term- research completed within 18 months and 2) Long term- research completed would take longer than 18 months. What emerged during the exercise was a third category, mid-term, that fell in between short term and mid term- although no specific time was associated with this category.

Workshop participants completed this activity by identifying research questions and writing one question per sheet of paper. Then they read each question aloud and stated in what lens and timeframe it belonged. The following presents the non attribution contributions as categorized by the participants.

Lens: Law and Policy

Short-term

- What kind of data should be disclosed for what purposes? Who should decide?
- What do we know already about how citizens use or respond to different kinds of public data (e.g. campaign finance data, sex offenders, deadbeat parents, product defects, airline on-time data, etc.)?

- How can industry data transparency and disclosure be used most effectively as regulatory tools?
- What are examples of modern sunshine laws? How do we modernize sunshine laws to reflect technology capability?
- How to provide sufficient incentives to agencies to disclose high value data?

Mid-term

- What is the comparative effectiveness of disclosure as a regulatory tool?
- Define meaningful use.
- Decide what data can be shared.
- What, when, where, and how should opacity continue to exist in government?
- What legal policy framework needs to be in place to sustain open government? Are existing open government laws adequate?

Long-term

- Do we, and if so, how do we adequately capture open data in the National Economic Accounts (e.g., GDP)?
- What new information architecture do we need for government data to create optimal openness? Is proactive disclosure possible? Can we share freely so that we request information rarely?
- How does open government change what constitutes good public acquisitions? Can we incorporate dynamic feedback?
- Does open government change role/impact of lobbying/lobbyists?
- Can we use policy/laws to increase the quality of government data? What would that look like?
- Can we create info data/categories/definitions that allow for gradations of exposure/usage/confidentiality?
- How can open data put the science back into political science? Can it lead to innovating on the system itself?
- Does open data lead to reduced corruption? Is it the promised disinfectant?
- Will increasing openness in government eventually require a new constitutional convention?
- What does open government look like in 2018?
- Is government data a form of wealth, a means to economic development? What are the benefits of openness (e.g., GDP, job creation)? For example, tax credits are said to create jobs. Can a similar claim or analysis be attributed to opening government knowledge? Isn't it a huge resource waiting to be tapped? Can we value it?
- What is the impact of international open government? Might it move us away from governance by nation-state?

Lens: Management

Short-term

- What data/information does government have that the public needs? How do we put it into a format that is useable?
- How do we arrive at a set of standards most effectively and efficiently?
- What are the most effective routings for government to establish now that raise the likelihood that open government lives in the future?
- Create prioritization guidelines to determine what data to open up. Downplay factor of “making my job easier.”
- Compare/contrast open government qualities internationally.
- Purchasing partnerships: Is the RFP process preventing governments from adopting open, cost-efficient technology?
- Free access and paid access friction. Government currently sells data. Will this model change? What can fill the money gap?
- Does making government information available inspire citizens to think more carefully about their choices (government & personal)?
- What management philosophies need to change and why?
- How do you measure open government’s success? Can you incubate culture change?
- Regardless of open government initiatives, is there a valid, generalizable method for assessing return of investment?
- How can we create an open government commons to connect research, practitioners, and outcome in real time?
- How do I find testimonials of quantified success stories?
- How does open government help me serve my constituents and value add to their lives?

Mid-term

- Does the public find “contextualized” data more “useful” than raw data? Is there a spectrum of users?
- How do we ensure we meet the needs and demands of society and citizens? Over the long term, such demands will change.
- Do we need to define high value data?
- Government data is owned by the public—they paid for it and should have access. Why should government/academics/others decide what is high value? Even old data can lead to better/new services or delivery methods.
- Test the assumption that existing business process generate usable data of value to others.

Long-term

- Can nationwide information standards open governments and create greater efficiency?
- What skills do governments need to maximize web 2.0 technologies?
- What is the future of the civil service in an era of citizen bureaucracy? What skills are needed? How many people are required to serve what population?
- How does open government help manage and support decision-making?

- How does the structure of government organizations change with the opening and sharing of government information?
- What government business processes can be radically altered or streamlined in future era of free, “mashable,” comprehensive open government data?
- How do we manage data quality and accuracy?
- How do we better proactively embrace change and help bureaucracies – SHIFT?

Lens: Technology

Short-term

- Create toolkits to publish open data (from internal systems).
- How can we quickly start laying down the foundation for open government technologies to make it easier/faster for government agencies to publish accurate fresh and meaningful data?
- In a world where cost is everything, how do you make it nearly costless for government to share data publicly?
- How are business analysts and application developers educated to deliver useful tools and disseminate data within the flow of business processes?
- What basic standards are needed to ensure that machine-readable data from different federal departments and agencies is interoperable?
- How do we enable people to control access to potentially sensitive data?
- What will data access look like (raw, cataloged, something else) when data increases exponentially in size and quantity?
- What metadata should be included with data sets to facilitate reuse, understanding, uptake?
- Create better systems to work with/deal with/integrate large volumes of feedback?

Mid-term

- How do we insure privacy?
- How can we combine/share/relegate/provide a central point for public (local, state, federal, international) and private data to make it easier to access and take action on it?
- What technologies are required to associate similar data in disparate databases (e.g., are John Doe and J.R. Doe the same person)?
- Can learning technology be leveraged as a tool to educate citizens on how to use data to collaborate, crowdsource, and provide oversight?

Long-term

- What does it look like to no longer be publishing data; instead, having data available by default?
- Do information and communications technologies advances (e.g., remote sensing, crowdsourcing) change the need for government?
- How will we find relevant international data in the long term?
- Scaling, metrics, language.

- How do we annotate data and data usage so we facilitate reuse and tracking (as well as maintain credit, authorship)?

Lens: Cross-Cutting

- What new government data do we need to construct performance assessments useful to the public?
- Have developments in open government made it into new curriculum in universities? What are we teaching future leaders about open government?
- Can we get academic consensus on the relationship between intellectual property and open government?
- Clarify definitions within and across the legal, policy, management, and technology areas for terms such as “Data,” “Database,” “Public,” “Private.”
- How does the velocity of the innovation cycle relate to open government?
- What causes government officials to be interested in participating in open government? How do we get broader support from all levels?
- How has the open government ecosystem evolved since 9/11? What are the effects on society?
- How can we balance security and open data?
- When/how does data get beyond boundaries of one “owner” such as government so that the starting point is not that we have a problem/question of openness?
- How will open government change a citizen’s sense of personal capacity within government? Will citizens become more powerful?
- How do we maximize public participation in the broad open government discussions and about individual open government projects? Can we use some form of crowdsourcing with a push?
- How can we define and assess outcomes of open government?
- Can we manage better outcomes? What is the success measure?
- What are the measures of “improving the lives of citizens” and how do we operationalize them?
- How can we turn existing government data into performance assessments that are meaningful to the public/stakeholders?
- How to measure “openness” in government?

Activity #5: Biggest Challenges Faced in Open Government

During the workshop, participants were asked to identify, in a worksheet, their top three challenges in open government. Worksheets were collected and transcribed. Workshop facilitators then grouped common challenges together. The following is a list of the challenges categorized by theme.

Data Available, Usable, Interesting, and Standardized

- Lots of important data are not available because they are not produced, even though type could be. For example, during the recent recession, important national policy (ARRA) was based in part on an understanding that state government spending would decline sharply. But there is no authoritative data on current state spending.
- Most people are more interested in sports or the royal wedding than in government. How can we use the data to “tell stories” that will interest citizens?
- Creating common standards and ways of describing these data sets to foster collaboration between government agencies/departments, stakeholders, and the public.
- Analyzing what data sets already exist in Data.gov and elsewhere that sheds lights on some markets.
- Data semantics and conventions: A clear challenge/barrier exists when groups/individuals outside of a group which produces a given data set tries to use that data set.
- Getting the data in a useable and meaningful format.
- Determine how to make many data sets that are potentially usable to citizens or institutions/organizations that represent important citizen interests.
- How can we design an extensible metadata system that will let us federate catalogues from the many agencies, states, municipalities, etc.?
- Data is just one of the many products of open government. How do we bring more key products, vocabularies, and other (nontext documents) products into the game?
- Publishing accurate, fresh and meaningful data. Also make it easier (automating) etc. to publish this data.
- Establishing standards that will allow consistent use and sharing of information.

Performance Measurement

- One of the most important purposes of open government is to measure performance which will improve services. Very few elected officials care about this, very few advocacy groups promote it, very few citizens know about it.
- Defining the success measure for “open government”.
- Better understand the effects and outcomes of open government
- Develop tools for measurement and evaluation

The Value of Open Government/ The Business Case

- How to measure the total return on investment for open data?
- What are the best practices (architecture, tools, standards) for how to open up government data to maximize benefits and costs?
- Support/priority from executive level to understand the importance of doing “this” and providing them with a means to make a choice or prioritization of initiatives. Cost benefit analysis.
- Identifying social, economic, political problems that open government may help solve. Can open government help reduce gaps in the transportation system? Identify communities with severe social needs not otherwise recognized? Can it reduce energy consumption? And how?
- There are over 300,000 data sets on data.gov, and tens of thousands more from states, cities, and tribes throughout the US and from abroad. How do we find relevant data for the needs of citizens in this firehose of data?
- We need to provide a framework for open government. If we do not, then elected officials, citizens, and business will not have a context for what is made public, why, etc. There is some basic work that needs to be done first. We need to explain what open government is.
- Incentive – solutions are fragmented and tactical because there is no institutionalization of the need or ROI, and so sustained planning is not committed to.
- Identifying how exposure of data can be used to promote collaboration to arrive at governmental solutions. Here, data is not especially numerical data, but complex, unstructured data. BPMN, data flow diagrams, use cases.
- Explain how digital government is important to everyday citizens?
- The information provided is very specific and aggregated because in the raw state it cannot be interpreted consistently. Situation specific data cannot be used to answer broader questions.

Strategy and Implementation

- There needs to be a clear “owner” for the execution and delivery of open government. While this owner is not a gate keeper per se, they would be the technical architect to ensure standards are created, security maintained, scalability ensured and there is some level of consistent user interface and tracking or results/feedback
- Defining the segment of open government that relates to industry transparency and data disclosure as regulatory tools – and defining how it relates to other aspects of open government and open data.
- We need to clearly articulate what our strategy is for open government. What are the intended outcomes? If it is determined that selected data can be sold, then we need a business model for this.
- How to foster cross jurisdictional collaboration and interoperability?
- Better understand incentives for disclosure or lack of disclosure.

- Ingraining open government ideals into processes – making individual “open government” efforts unneeded.

Engagement

- More open collaboration with the citizens, aka, feedback loops, crowdsourcing, etc.
- Finding methods to educate citizens in how to examine data so that they can understand and comment on government operations, to participate in crowdsourcing, and press government on performance metrics
- Showing importance of open government with stakeholders (data owners, public servants, etc.).

Culture and Perceptions

- Acceptance: Cultural resistance to open government initiatives based on a need to see a significant ROI before committing resources makes a big part of the job the marketing of value, rather than working the solutions.
- This could fundamentally change the ways we do business in government. How shall we prepare for it?
- Culture -- remains the biggest challenge because culture change in a large government agency can be very difficult. It's all about incentives.
- “Fear” of putting information out there by elected officials as well as county employees.
- Trying to get an agency that has been traditionally reluctant, and in fact has career employees who have been selected and promoted to be secretive/non-transparent, to embrace Open.data and Open.gov.
- There is misinterpretation of the information, and lower confidence results in less capability to provide information.

Policies

- Policy blockades: Open government initiatives generally push the limit of what has previously been done, in terms of data exchange and collaborative practices/technologies. Most policies were created when an entirely different landscape was in place. The conflict here prevents forward progress.
- Many of our policies are outdated and the process to update them is painful. Additionally, the incentives to fix policy are not aligned with those who would be the most involved.

Activity #6: Moving toward a national research agenda

The workshop participants were asked to reflect on the work they had completed so far and comment on what needs to be done in order to move closer to an open government national research agenda. The following is an account of the comments and discussion points made during this activity.

- There is plenty of information generated here but there is a lot. We need to develop a brief, digestible, and packaged document that communicates the high points.
- Develop an elevator pitch that is theme focused.
- A good summary of the information collected and determine some action steps for after the workshop.
- This work is a good starting point, but do we situate within existing research streams and if not where do we find new funding for this type of multi-disciplinary research?
- In order to get funding this information must be problem-focused, for example we must present a case as why open government is a tool to solve problems in national arenas including transportation, energy and health.
- The clusters need further refinement and possibly be broken down even further. Some things are grouped together and could become additional clusters.
- There needs to be a strategic connection between exposing data and performance, this is important because no one will fund research if that isn't there.
- An identification of the logics of open government would be a good addition to this work. For example, we need to test assumptions on ideas such as 1) when government data is released it will be valuable and good things will happen, and 2) that transparency, efficiency, and accountability are different and that we have a tendency to mash them together.
- A research agenda must be multi-leveled meaning that it must address all levels of government and backgrounds and create a loop that connects academic and practitioners. We have seen that there are many viewpoints and perspectives that must be incorporated into a national agenda and if we tiered an approach then
- We need a roadmap of who is conducting research in open government. If we have no current state of research then when we have completed a research agenda, we don't know where to start.
- We need a large section on how government officials respond to releasing data. This would include understanding the legal and management requirements of releasing data. We also need to understand how different government agencies respond to the exposure to more openness?
- In addition to a research agenda we need to spend time seeing what we already know. Is there a central place where we can list all the research related to open government. That way we can assess what we can get what we already know, together and start synthesizing it.
- Possibly a study on whether government focuses on things it can already disclose

- The focus seems to be on government, we need more studies on citizenry and what it means to be a citizen, to understand the effects of having a lot of data available and how that does or does not reinvigorate citizenry.
- Any national research agenda should be participatory and involve citizens from the start. There will and should be lots of people who influence this agenda.
- We need a focus on long-term research. This means the long haul questions such as what does the world it look like and mean when there are 3 million data sets available? Right now there is 300K datasets available but that is quickly increasing. Its important to know what we can answer today but also important to think beyond that.
- The idea of an open government ecosystem is compelling and should be developed. It could help us juxtapose the ideas and present a large amount of information. We could enlist the help of other organizations such as the Open Forum Foundation and post it to places such as the open government playbook.
- This activities in this workshop are a trigger to see what landscape looks like and start to understand how to put together a coherent agenda, parse the universe and better connect the communities. It will be important to breakdown the complexity and see how each of our works are represented in this ecosystem. A research agenda also needs to capture what was done in the past. It may serve more as a roadmap of both past and present.
- Each person here has been doing open government for quite some time, this will help us decide what each of our next steps are- to help make decisions about the ten things each of us could do.
- We see that there is commonality among the academics and the practitioners. We both want to understand the impacts of the actions we take. For example, we want to learn more about the impacts outside of government in areas such as political or civic implications. Also, the impacts inside government on the process of decision making and the change in output.

Activity #7: Closing Thoughts

At the end of the workshop participants were asked to share final closing thoughts. Some participants passed on this because they stated their thoughts has been captured by others or had been adequately documented in the previous activity. The following is the list of closing thoughts shared:

- Open data can lead to open government. Opening data is at the center of open government.
- Open data and open platforms are the stimulus for economic and government transformation. There can be a reorganization of the relationships between government and citizens.
- Citizens' expectations of government are changing and we must meet those expectations.
- Most important is framing what is open government and understanding the value.
- Open government measurement and evaluation will be imperative to determine.
- Open government requires collaborative methods for governing.
- The rest of the world has changed – government needs to change with them.
- Technology is ahead of government and it's not stopping. We have to catch up and try to get ahead!
- Open government should make us rethink the way government works.
- We first need to develop a common understanding of open government.
- Open government research will have to have a multidisciplinary and multi-sector approach
- While governments are trying to do more with less we could tap the public to take on some responsibilities.
- As government we need to make it easier to publish trusted data.
- We need to understand the effects of transition and institutional design.
- We would like to know more about the engagement gap between data available and used.
- Open government can and will spur economic growth.
- Open government improves government processes.
- There needs to be evidence based open government – don't assume what the effects will be.

Appendix A: Workshop Team

Center for Technology in Government, University at Albany, State University of New York

- Theresa Pardo, Director
- Meghan Cook , Program Manager
- Jana Hrdinova, Project Associate
- Brian Burke, Senior Program Associate
- Alan Kowlowitz, Staff
- M. Alexander Jurkat, Staff

Tetherless World Constellation (TWC), Rochester Polytechnic Institute

- Jim Hendler, TWC chair , Departments of Computer Science and Cognitive Science

Institute for Information Law & Policy, Democracy, New York Law School

- Beth Simone Noveck, Professor of Law and Director of Democracy Design Workshop

Civic Commons

- Andrew Hoppin, Advisor

Appendix B: Workshop Participants

Mr. Brian Burke
Senior Program Associate
Center for Technology Government
University at Albany, SUNY

Dr. Daniel Chan
Acting Chief Information Officer of New
York State
New York State Office for Technology/Chief
Information Office

Mr. Girish Chhugani
Senior Advisor, Technology Strategy
New York City Department of Information
Technology and Telecommunications

Dr. Cary Coglianese
Edward B. Shils Professor of Law and
Professor of Political Science
Director, Penn Program on Regulation
University of Pennsylvania Law School

Ms. Meghan Cook
Program Manager
Center for Technology in Government
University at Albany, SUNY

Dr. Anthony Cresswell
Deputy Director, Center for Technology in
Government
University at Albany, SUNY

Ms. Sarah Crane
Acting Director, USA.gov
United States General Services
Administration

Dr. Sharon Dawes
Senior Fellow
Center for Technology in Government
University at Albany, SUNY

Mr. Dominic DiFranzo
PhD student
Rensselaer Polytechnic Institute

Dr. Thomas Gais
Director, Nelson A. Rockefeller Institute of
Government
University at Albany, SUNY

Mr. Chris Gerty
Open Government Analyst / Systems
Engineer
National Aeronautics and Space
Administration

Mr. Adam Gigandet
Chief Information Officer
New York State Department of Motor
Vehicles

Dr. Stuart Graham
Chief Economist
Office of the Chief Economist
United States Patent and Trademark Office

Mr. Remington Gregg
Senior Advisor
United States Office of Science and
Technology Policy

Mr. Joel Gurin
Chief of the Consumer and Governmental
Affairs Bureau
United States Federal Communications
Commission

Dr. Teresa Harrison
Professor
Department of Communications
Faculty Fellow, Center for Technology in
Government
University at Albany, SUNY

Dr. Jim Hendler
Tetherless World Chair
Departments of Computer Science and
Cognitive Science
Rensselaer Polytechnic Institute

Mr. Sean Herron
Open Government Analyst
National Aeronautics and Space
Administration

Mr. Andrew Hoppin
Partner and Co-founder
New Amsterdam Idea, LLC
Former Chief Information Officer New York
State Senate
Consultant, New York City Department of
Information Technology and
Telecommunications (DoITT)

Mr. John Kaehny
Executive Director
ReinventAlbany

Mr. Todd Khozein, MD
Founder
Second Muse

Mr. Tom Lee
Director of Sunlight Labs
Sunlight Foundation

Dr. Deborah McGuinness
Tetherless World Chair
Departments of Computer Science and
Cognitive Science
Rensselaer Polytechnic Institute

Ms. Kim McKinney
Chief Information Officer
Information Technology Services
Broome County Government

Dr. Ines Mergel
Assistant Professor
Maxwell School of Citizenship and Public
Affairs
Syracuse University

Ms. Lisa Nelson
Program Analyst
Research and Strategic Partnerships,
Citizen Services and Innovative
Technologies
United States General Services
Administration

Dr. Theresa Pardo
Director, Center for Technology in
Government
University at Albany, SUNY
tpardo@ctg.albany.edu

Mr. Robert Samson
Member of NYS Spending and Government
Efficiency (SAGE) Commission
New York State Government

Ms. Sarah Schacht
Executive Director, Founder
Knowledge as Power

Ms. Jennifer Shkabatur
Doctoral candidate
Harvard Law School

Mr. Nicholas Skytland
Open Government Initiative
National Aeronautics and Space
Administration

Mr. Todd Stavish
Technical Account Manager
Socrata

Mr. Robert Vitello
Principal Consultant
v257.org
Former Deputy Commissioner for Planning
and Technology
NYS Department of Labor

Dr. Robert Ward
Deputy Director, Nelson A. Rockefeller
Institute of Government
University at Albany, SUNY

Mr. Derek Werthmuller
Director of Technology Services
Center for Technology in Government
University at Albany, SUNY

Mr. Nick Willett-Jeffries
Sage Policy Analyst
New York State Spending and Government
Efficiency Commission

Appendix C: Workshop Agenda



Open Government Research and Development Agenda Setting Workshop *Using government data to improve the lives of everyday citizens*

April 27-28, 2011

Center for Technology in Government, University at Albany

AGENDA

Wednesday April 27, 2011

10:30 am **WELCOME**

- Overview of the workshop
- Participant introductions

11:00 am **An overview of the NITRD – NARA Open Government Summit**

- Andrew Hoppin, Civic Commons
- James Hendler, Tetherless World Constellation, RPI
- Beth Simone Noveck, Institute for Information Law & Policy, NYL
- Theresa Pardo, Center for Technology in Government, UAlbany

A Research Agenda for Smart Disclosure

- Joel Gurin, Consumer and Governmental Affairs Bureau, FCC

11:45 am **Plenary facilitated discussion**

- What are the questions a research agenda focused on the use of government data to improve the lives of citizens must address?

1:00 pm **LUNCH**

1:30 pm **Small group discussions on the research clusters**

- What are the research traditions, tools, techniques, theories that must be engaged to develop answers the questions within the cluster?
- What are the challenges researchers would face in addressing the cluster?
- What are the strategies to address these challenges?
- What are the target organizations to support or participate in such research about this cluster?
- Why is this cluster important to efforts to use government data to improve the lives of citizens?
- What are the implications of not addressing the questions represented in this?

- 2:45 pm **BREAK**
- 3:15 pm **Plenary report out and discussion**
- 5:00 pm **ADJOURN FOR THE DAY**
- 6:00 pm **Bus to dinner at University at Albany campus**
- 8:30 pm **Bus returns to hotels**

AGENDA

Thursday, April 28, 2011

- 9:00 am **WELCOME**
- Introduce new participants
 - Reviewing Wednesday's discussions
- 9:10 am **Defining our target audience and our workshop product**
- 10:00 am **Plenary facilitated discussion**
- Considering the law and policy, management, and technology perspectives and cross-cutting questions
- 11:00 am **BREAK**
- 11:15 am **Plenary facilitated discussion**
- What are the cross-cutting questions?
 - What's missing?
- 11:45 am **LUNCH**
- 12:15 pm **Plenary discussions**
- What is the most critical message our research agenda should communicate to government practitioners, researchers, civil society, and funding agencies (others...)?
- 1:00 pm **Closing comments and next steps**
- 1:15 pm **ADJOURN**