

REPORT ON THE NEW JERSEY OPEN GOVERNMENT DATA THOUGHT FORUM

Convened by the Public Technology Institute (PTI)
at the request of the Geraldine R. Dodge Foundation

Issue Paper #4 | November 2015
Bloustein Local Government Research Center

Marc H. Pfeiffer, MPA
Assistant Director and Senior Policy Fellow
Bloustein Local Government Research Center
Edward J. Bloustein School of Planning and Public Policy
Rutgers, The State University of New Jersey

Credits

Sponsor: The **Geraldine R. Dodge Foundation** supports leadership, innovation, and collaboration for a better New Jersey. For nearly 40 years, the Foundation has nurtured leaders, ideas, and institutions that transcend self-interest and promote a sustainable future. Dodge focuses on issues critical to its home state of New Jersey, and organizations that have a direct and meaningful impact here. The Foundation believes that philanthropy includes not only providing resources, but also connecting leaders across sectors, sharing expertise, and promoting collaboration to help build movements for change around important issues.

The Dodge Foundation funds arts, education, environment and media initiatives that are innovative and promote collaboration and community-driven decision-making. It listens to the needs of the nonprofit sector by offering sustained and comprehensive technical assistance, and pays careful attention to the public policy environment in which its grantees operate. The Foundation works with nonprofit, community, government, and business leaders to imagine a better New Jersey. www.grdodge.org

Organizer: Located in Washington D.C., the **Public Technology Institute**, a technology organization created by and for cities and counties, actively supports local government executives and elected officials through research, education, executive-level consulting services, and national recognition programs. As the only technology organization created by and for cities and counties, and now including state governments, PTI works with a core network of leading local officials — the PTI membership — to identify research opportunities, provide thought leadership and professional development opportunities, share solutions, provide technology support services, recognize member achievements and address the many technology issues that impact local government. www.pti.org

Convener/Author: On behalf of PTI, **Marc H. Pfeiffer** served as the Thought Forum’s convener, planning and moderating the event and writing this report. In 2012 Mr. Pfeiffer retired from a 37-year career in New Jersey local government administration. He served as a municipal administrator in several municipalities, and served for 26 years in the State’s local government oversight agency, the Division of Local Government Services. He served the agency as Deputy Director for 14 years, and periodically as Acting Director. His work in state government involved a wide variety of public policy, legislative, fiscal, and administrative activities.

Several assignments during his tenure in state government involved public records. He worked on the legislation that became the Open Public Records Act and served as the Government Records Council’s first Executive Director (Acting). He initiated the GRC’s operation and delivered the first OPRA training to hundreds of state and local officials. He also served on the State Records Committee for 10 years and worked with the Division of Archives and Records Management on local and state government records issues for 23 years.

Among other public service activities, he is the Assistant Director of the Bloustein Local Government Research Center blousteinlocal.rutgers.edu, where in addition to research projects, he makes his extensive government experience available as a guest lecturer and as a consultant in other collaborative efforts. He also assists the Rutgers School of Public Affairs and Administration in curriculum development and instruction related to the State’s Certified Public Manager Program.

This report was edited by Debra Meltzer at meltzer.dc@gmail.com.

The author also thanks Deborah T. Poritz, New Jersey’s former Chief Justice (among many other roles that have contributed to New Jersey’s civic life) for her thoughtful comments on aspects of the issues.

Mr. Pfeiffer can be reached at marc.pfeiffer@rutgers.edu.

Contents

Why This Report is Important to New Jersey	4
Executive Summary	5
Part 1 – Overview of the Thought Forum	7
Part 2 – Setting the Table.	9
Part 3 – The Open Government Data Imperative and New Jersey . . .	13
Part 4 – Conclusions and Recommendations	24
Appendix:	
A: List of Open Government Data Reference Sources	28
B: List of Attendees	29
Endnotes	30

Why this report is important to New Jersey

This report examines a growing public imperative — the routine disclosure of government data in New Jersey. As part of a growing national movement, open government data is about taking all kinds of state and local government datasets and making them available and useful to the public.

Opening data enhances government transparency, adds social and business value to data, promotes economic growth, and encourages individuals with a wide range of skillsets to help solve public problems. It relies heavily on new and developing technologies that the public, researchers, businesses, and the media have already started using.

While access to data has become a routine expectation by some in the civic and private sectors, it is, in effect, a new government service. This report observes that open data is affected by uncertainty, evolving technology, and great promise — all elements that can lead to disappointment, frustration, and conflict if the challenges and the opportunities are not well understood.

The report identified six broad challenges that face open government data in New Jersey:

- The impact on government of open data decisions and policies;
- The availability of time, money, and attention to address open data issues;
- The disconnect between public officials and open data advocates;
- The management, presentation, curation, and quality control of data;
- The conflicts between state records laws and personal privacy concerns; and,
- The adequacy of current technology to manage open data.

The challenges are daunting, but opportunities exist for overcoming them, or at least reducing them. This report highlights actions that move government in that direction:

- The impact on government of open data decisions and policies;
- Change happens slowly in government, but it happens, and so it will with open data;
- Government agencies can re-engineer data collection practices;
- Ways to protect personal privacy can be developed;
- Any government agency can start posting basic data;
- State agencies can provide websites for aggregated data; and,
- Support and training can come from the state's higher education, government advocacy, civic, and philanthropic organizations.

The full report is online at www.pti.org/programs/tech/njopendatarep.asp.

Executive Summary

Forty leaders from across sectors in February 2015 gathered at a “thought forum” on the state of open data at all levels of New Jersey government organized by The Public Technology Institute and hosted at the Geraldine R. Dodge Foundation in Morristown. Attendees represented six sectors of the open data community with interest in the subject: media, academia, state government, municipal and county governments, open government advocates, and civic groups

Open government data means taking datasets of all kinds collected and held by government agencies at every level and making them publicly accessible, readable by software, available without cost, and reusable and redistributable without restriction. By opening government data this way, it enhances government transparency, releases the untapped social and commercial value of data, and encourages participation and engagement by citizens, businesses, and organizations to help solve problems and encourage economic growth.

The attendees heard from several speakers: Chris Daggett, President & CEO of the Dodge Foundation, Waldo Jaquith from U.S. Open Data, Mark Headd, Philadelphia’s original Chief Data Officer who is now with Accela Inc., Jeff Jarvis of the City University of New York’s Tow-Knight Center for Entrepreneurial Journalism, and Marc Pfeiffer, the event’s convener and Assistant Director of the Bloustein Local Government Research Center at Rutgers University. Attendees learned about each group’s varying interests and perspectives. They engaged in several exercises to identify the challenges of open government data in New Jersey and suggested actions that could be taken to address them.

Significant challenges to developing an open data culture in New Jersey government were identified:

- **Governance:** how the structure of New Jersey government affects government activities and the behavior of agencies throughout the state and local government
- **Resources:** what levels of time, attention, and money are available and necessary in order to open data?
- **Understanding:** is there agreement on what open government data is, and the responsibility of government to provide it?
- **Data:** how it is managed, presented, curated, and used, and how quality is assured?
- **Records and Data:** The Open Public Records Act was enacted in 2001. OGD is a relatively new concept, and comes from changes in technology and societal expectation. What challenges arise where these two public policies intersect?
- **Technology:** what technologies are necessary to manage open government data and what is the capacity of government agencies to manage them?

The group identified opportunities for addressing the challenges. They included recognition that opening government data in New Jersey would not happen all at once, and would evolve over time as public attention to it grows and the technology used in managing open data improves and becomes more cost effective. It was observed that government agencies that collect and produce data should re-engineer their requirements to reflect current technologies and that the state should maintain a centralized resource for providing data it collects, particularly for data submitted by local agencies.

Government agencies should simply start making their data openly available. They should be encouraged to learn about open data, assess their own practices,

and start institutionalizing measures that allow for open data. This includes inventorying their data, ensuring that contracts for technology services provide for open data, prioritizing what they open within available resources, and managing their technology to facilitate open data. “Starter” data sets for municipalities are suggested in the report, subject to agency and public needs.

It was also recognized that some of the challenges in producing and managing open data are rooted in the state’s Open Public Records Act. There is the overarching issue of how to address citizens’

expectations of privacy and to ensure personal protection in an increasingly networked world. In addition, participants made recommendations for philanthropic, civic, academic, and government-related organizations to partner in order to develop guidance, training, and support material for government agencies that choose to move forward on open data initiatives.

The full report can be downloaded at www.pti.org/programs/tech/njopendatarep.asp.

Part One—Overview of the Thought Forum

The February open data “thought forum,” convened by The Public Technology Institute at the Geraldine R. Dodge Foundation, brought together 40 leaders from sectors across the open data community to discuss open government data in New Jersey and to learn about the concerns of the groups.

Generally, open government data (OGD)¹ is a recent imperative facing government agencies that has evolved from three separate issues: diminishing public trust in government, the availability of technology that simplifies production and use of government data, and the interest of the commercial sector in monetizing public data through value-added technology. Joel Gurin, a leading advocate of open data has described it as “the world’s greatest free resource – unprecedented access to thousands of databases – and it is one of the most revolutionary developments since the Information Age began.”ⁱ A bit hyperbolic perhaps, but it belies the challenge of how to “mine” the resource and make it useful. The thought forum was intended to examine this imperative and its impact on New Jersey government; its goal was to examine the challenges as well as the opportunities it presents.

The attendees at the thought forum represented six sectors of interest in the subject: media, academia, state government, municipal and county governments, open government advocates, and civic groups. While space limited the number of participants and several sectors were not represented (law, business, non-profit), the discussion covered the full range of issues.

Dodge Foundation President & CEO, Chris Daggett, opened the forum by highlighting the organization’s interest in open data as part of its ongoing goal of supporting the New Jersey news ecosystem. The Foundation believes in the role of the press and

importance of accessible government data as keys to a vibrant democracy. Daggett explained Dodge is exploring ways philanthropy might be able to assist in resolving open government data issues.

Three invited experts made presentations that set the stage for the discussions to take place later in the day:

- **Waldo Jaquith** from the US Open Data Initiative led off with an overview of what constitutes open government data. He highlighted examples that illustrated the range and scope of government data, and its current state across the country and throughout the world.
- **Mark Headd**, Philadelphia’s original Chief Data Officer who is now with Accela Inc. (a provider of cloud-based civic engagement solutions for government) could not attend in person, but presented via the Internet. He discussed the tools used to open and present data, and explained some of the challenges that currently face governments creating open data portals. He also described how they could be improved.
- New Jersey’s “Entrepreneurial Journalist” **Jeff Jarvis**ⁱⁱ followed with observations on the importance of open data to transparency, the value that open data can bring the public, and the role of the media as a watchdog. He highlighted the value that can be gained by creating an “ethos of openness” in New Jersey and how the role of journalism can be enhanced by using open data.

Forum convener **Marc Pfeiffer**, closed the presentations with a look at the “darker” side of open data, looking at some of the downsides that need to be considered when implementing open government data programs.

¹ This report will use alternatively use “OGD” and the full phrase throughout this report – no difference is intended and is done to enhance readability.

After the presentations, attendees discussed the concerns, interests, and opportunities they or their organization had about open government data. Participants were encouraged to talk with those whose comments interested them in order to learn more about their issues and concerns.

Attendees broke into small groups made up of representatives from each of the six sectors. They listed the primary challenges of open government data, and as they reported on their findings, common issues and themes were recorded and the full group shared ideas on how these issues could be addressed .

This report integrates the comments of the speakers, the activities of the event, and the supplemental research that ensued. The analysis and recommendations are those of the author and do not necessarily represent the Geraldine R. Dodge Foundation or the Public Technology Institute.

Many thanks to the attendees and speakers for their contributions – this report could not have been written without them. This report is online at www.pti.org/programs/tech/njopendatarep.asp.

Part Two—Setting the Table

I. THE OPEN GOVERNMENT DATA IMPERATIVE

Generally, three forces have given rise to where we are today: steadily diminishing public trust in government, the availability of technology that simplifies production and use of government data, and the interest of the commercial sector in monetizing public data through value-added technology. Once released, the value of OGD goes beyond these elements and expands the universe of those who use data to help solve public policy problems. Among other concerns, it raises the unintended consequence of compromising personal privacy when information on individuals collected for government purposes is disclosed.

Considered more broadly, open government data is part of the “open knowledge” movement. Open knowledge is any content, information or data that people are free to use, re-use, and redistribute — without any legal, technological or social restrictions.ⁱⁱⁱ

The movement is international in scope, currently reflected at all levels of government in different places in different degrees. It is supported by a diverse group of organizations around the world that span the political spectrum. It is an element of larger efforts whose goal is to develop more responsive and accessible (open) governments with political mechanisms that are transparent and flexible.

More specifically, as described in a recent Pew Research Center report,^{iv} “Open Data and Open Government are related notions and often thought to be interchangeable. Yet they are not quite the same, and indeed open government is about more than open data. One (Open Government) is an end and the other (Open Data) is a means to that end. Open Government is a policy posture of the Obama Administration as well as many governors and mayors of all political persuasions. The notion is to use modern technology and other tools to help citizens better understand how government works, engage more

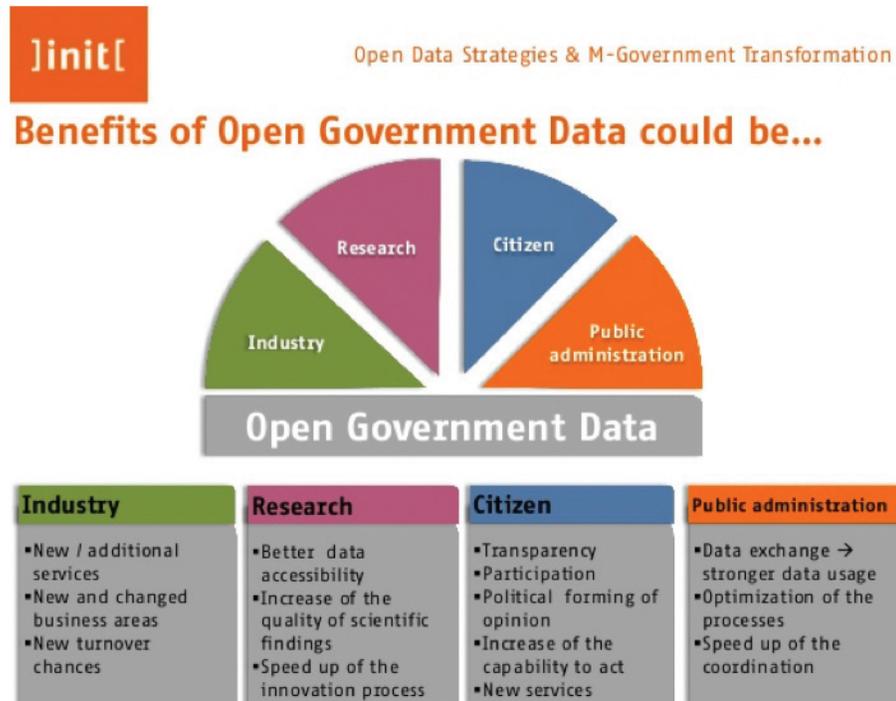
effectively with government, create economic value for businesses, and improve the delivery of government services.” This is the “open government data imperative.”

What is this “data” we are discussing?

In this context, open government data is: data produced or commissioned by government or government controlled entities.^v The open data world uses several definitions of “data”, some more expansive than others. For the general purposes of this report, we are differentiating between *open records* as defined by New Jersey’s Open Public Records Act (OPRA), and *open data* as a subset, in which “data” means “geospatial, tabular, textual, legislative, and source code that is maintained in an electronic, digital or optical format” and held or maintained by a government agency.^{vi}

Why should data be open? The answer, of course, depends somewhat on the type of data. However, there are common reasons such as:^{vii}

- **Transparency:** In a well-functioning, democratic society citizens need to know what their government is doing. To accomplish that, they must be able to access government data and information freely, and share that information with other citizens. Transparency isn’t just about access; it is also about sharing and reusing it. Often, to understand material it needs to be analyzed and visualized; this requires that the material be easily accessible so that it can be used and reused without restriction.
- **Releasing social and commercial value:** In a digital age, data is a key resource for social and commercial activities. Everything from finding your local post office to building a search engine requires access to data, much of which is created or held by government entities. By opening up data, government can help drive the creation of innovative businesses and services that deliver social and commercial value.



- **Participation and engagement:** Access to data encourages citizens, businesses and organizations to actively participate in government processes. Much of the time, citizens are only able to engage with their own government sporadically — maybe just at the voting booth every two or four years. By opening up data, it enables citizens to be much more directly informed and involved in decision-making. This is more than transparency: it's about encouraging a community of active participants, citizens who not only know what is happening in the process of governance, but are able to contribute to it.

The federal government developed “Project Open Data” which argues that government agencies should apply the following principles toward data:^{viii}

- **Public:** Agencies should have a presumption that their data should be open, subject to laws and imperatives of privacy, confidentiality, security, or other valid restrictions.
- **Accessible:** Data should be machine-readable and easily retrievable and searchable. Open data formats such as csv, json, odf, and xml are preferred over proprietary vendor formats (i.e., pdf, xls, etc).^{ix}

- **Described:** Open data should be clearly described so that users can understand how to use them, as well as the data's strengths and weaknesses.
- **Reusable:** Under open licenses, there should be no restrictions on use.
- **Complete:** Data should be published in primary forms.
- **Timely:** Data should be released as quickly as necessary to preserve its value.
- **Managed post-release:** Agencies should designate a point-of-contact for data releases to respond to queries.

Presenter Waldo Jaquith put the definition succinctly, suggesting that open government data is:

- Accessible publicly
- Readable by software
- Without cost
- Reusable and redistributable without restriction

Generally, in society there are many areas in which open data has potential uses and applications:

- **Culture:** Data about cultural works and artifacts — for example titles and authors — and generally collected and held by galleries, libraries, archives and museums.
- **Science:** Data that is produced as part of scientific research from astronomy to zoology.
- **Finance:** Data such as government accounts (budgets, debt, expenditures and revenues) and information on financial markets (stocks, bonds etc).
- **Statistics:** Data produced by statistical offices such as the census and key socioeconomic indicators.
- **Environment:** Information related to the natural environment such as presence and level of pollutants, the quality and rivers and seas.
- **Transport:** Data such as timetables, routes, on-time statistics.

To be pragmatic, why are we discussing open government data in New Jersey? Advocates make the following points that OGD:³

- Can reduce an agency's OPRA load
- Can increase sharing between agencies, levels of government, and academic and non-profit researchers and improve policy making
- Can facilitate innovation in the private sector
- Can improve public trust and confidence in government decision-making.
- Or, you are doing it because somebody is making you.

II. BIG PICTURE OBSERVATIONS ABOUT OPEN GOVERNMENT DATA

Ultimately, the elements of open knowledge and open government data are issues that affect how we govern and who participates. They create new challenges and opportunities for elected and appointed government

officials who manage day-to-day operations. It is likely that they will irrevocably change how public administration works and create new forms of administrative practice.

For most agencies, OGD is a new government activity or service. Like any new service, it can provide both challenges and opportunities. It requires setting



priorities. Much the same as the introduction of websites fifteen years ago, it requires new resources of time, attention, and ongoing expenditures by public officials and their organizations.

OGD challenges conventional beliefs about access and use of government data: Should all taxpayers be compensated by charging businesses for the commercial use of data (and how is that defined)? Should journalists have added access (and how is journalist defined)? Is there a threshold of overuse that should be applied to citizens with a cause who access data regularly? Should there be changes to the political process based data use (or abuse)? When does “data” blend into “records” and how are they differentiated? It is further challenged by the expectations of open data advocates, and the technological capacity and will of government agencies of varying sizes to adopt and institutionalize open data policies.

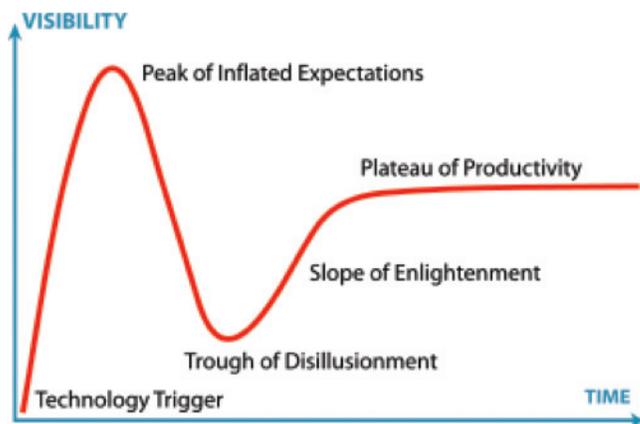
It raises complicated questions: Does the government's responsibility stop at providing the data? Is there also a responsibility to provide tools to view, analyze, and display data for those unable to afford their own data management tools? With the advent of technology that enables sophisticated data matching, where do privacy concerns start and end, how are they addressed, and can privacy concerns result in less data being open? Are the resources spent worth the value gained from open data efforts or are those resources better spent on other initiatives? Is there a point of diminishing returns, where too much “sunshine” on data can “burn”, or does debate surrounding data (data analysis can provide contradictory conclusions) result in an unintended consequence, i.e., the inability to make decisions?

Like many issues in a democratic society, the answers to these questions are that “it depends” on the place,

the people, and their competency at self-government. However, the convergence relating to the public’s lack of trust, the emergence of new technologies, and the value of data have presented new challenges to governance and public decision-making.

III. INFLUENCE OF THE NEW TECHNOLOGY ADOPTION CURVE

The challenge of open government data is akin to a concept in technology that describes how new technology is adopted. This concept is represented as a curve (deemed the “hype cycle” by its originators, the Gartner Group).^{xi}



While some of its terms are pejorative, it viscerally graphs the challenges of new technologically based services.

Viewing the development of open government data over the last few years and comparing it to where it is today, it is arguably approaching the “Peak of Inflated Expectations” in context of the private-sector phases described in the table.^{xii} These phases can also be applied in the public sector.

This implies that early adopters will lead the process as technology is refined based on their experiences; this in turn will lead to a new generation of tools, practices, and lower costs that will heighten expectations. More and more agencies will adopt open data practices and their adoption will be easier. For example:

- We see it in the federal government through its Open Data Project, the growth of www.data.gov as a starting

No.	Phase	Description
1	Technology Trigger	A potential technology breakthrough kicks things off. Early proof-of-concept stories and media interest trigger significant publicity. Often no usable products exist and commercial viability is unproven.
2	Peak of Inflated Expectations	Early publicity produces a number of success stories—often accompanied by scores of failures. Some companies take action; many do not.
3	Trough of Disillusionment	Interest wanes as experiments and implementations fail to deliver. Producers of the technology shake out or fail. Investments continue only if the surviving providers improve their products to the satisfaction of early adopters.
4	Slope of Enlightenment	More instances of how the technology can benefit the enterprise start to crystallize and become more widely understood. Second- and third-generation products appear from technology providers. More enterprises fund pilots; conservative companies remain cautious.
5	Plateau of Productivity	Mainstream adoption starts to take off. Criteria for assessing provider viability are more clearly defined. The technology’s broad market applicability and relevance are clearly paying off.

point to find government data, the use of GitHub by agencies developing new applications (see sidebar on Open Code Sharing), and the 18F project for improving digital delivery of federal services.

- We see it in state data policies and portals in Connecticut, Hawaii, Delaware, and New York.
- We see it in open data policies in large cities such as Boston, Los Angeles, New York, Philadelphia, and Seattle and in smaller places like Montgomery County, MD and Madison, WI as well as in regional efforts in the Pittsburgh and Hartford regions.
- We see it in applications developed by profit and not-for profit organizations such as the US Open Data Initiative, Accela, ESRI, Google, Sunlight Foundation, the Bloomberg “What Works Cities” project, funding

efforts of the Knight Foundation and the Alfred P. Sloan Foundation, the advocacy work of law.resource.org^{xiii} and others.

We also see that the commercial sector is developing applications and marketing solutions that build expectations. They are promoting the experience of early adopters as easy solutions to open data opportunities. This gets amplified by the development of civic sector solutions and together they create a cacophony of claims, solutions, and answers to problems that can confuse and

puzzle government officials who did not know they had a problem to solve.

As a completely new field of endeavor and outside of dedicated IT staff, most government administrators are currently not prepared for the open data imperative nor have they allocated the resources needed for it. The new technology adoption cycle applies here; more needs to be done for open data policies to become mainstream. In the meantime, the outliers on the leading edge will muddle through.

Part Three—The Open Government Data Imperative and New Jersey

The open government data imperative issues raised at the Thought Forum break down into six challenges and their attendant opportunities. They are:

- I. Governance: how does the structure of New Jersey government affect government activities and the behavior of agencies throughout state and local governments?
- II. Resources: what levels of time, attention, and money are available for opening data?
- III. Understanding: is there agreement on what open government data is and what is the responsibility of government to provide it?
- IV. Data: how is data managed, presented, curated, and used, and how is its quality is assured?
- V. Records and Data: The Open Public Records Act was enacted in 2001. OGD is a relatively new concept, and comes from changes in technology and societal expectation. What challenges arise where these two public policies intersect?
- VI. Technology: what technologies are necessary to manage OGD and what is the capacity of government agencies to manage it?

I. GOVERNANCE

State government technology (Executive Branch)

New Jersey's central technology agency, the Office of Information Technology, manages most of the Executive Branch's mission critical databases as well as its administrative and finance systems. It maintains state-developed software for many agencies and maintains other systems using third-party software. In recent years, it has started using cloud and hybrid cloud services. It also maintains the state's primary network, the Garden State Network, used by most agencies for internet and telecommunications access. It obtains most of its spending authority from user fees assessed to agency customers.

Many state agencies also maintain and support their own technology resources, including smaller, specialized applications, cloud-based applications, internal networks, and device support.

The head of OIT (the State's Chief Information and Technology Officer) does not oversee or manage all information technology resources, hardware, software, and data for state government. Individual agencies have the responsibility to arrange, either through OIT or their own resources, for their technology needs. OIT has

adopted standards. However, there are exceptions to most rules, these are in part driven by the diversity of agency technology. The state maintains two data-related sites:

1. Early in the Christie Administration, the State Treasurer's office coordinated the development of the State "Transparency Center" (<http://www.yourmoney.nj.gov/>). This site which includes the Governor's Performance Center, posts state agency performance reports quarterly (regrettably as PDFs). The Transparency Center is otherwise rich with fiscal data and has good search tools. It is updated regularly, and provides download capabilities for much of the data. Most of the data is available in an open format (though PDF-based data lingers in some areas).
2. In early 2015, OIT deployed a Socrata-hosted data site (www.data.nj.gov) to serve as a hub for state agency data. The state also recruited a Chief Data Officer to coordinate data management efforts within OIT and across agencies. Efforts have been made to ensure consistent data standards, formats and practices across agencies. However, a scarcity of resources, OIT's lack of enforcement authority, and agency resistance has long hindered a cohesive and comprehensive approach to data management in the state enterprise. That said, the site is available to agencies to post data using Socrata's open standards, as long as the agencies also post the same data on their own websites.

There are other state sites that hold datasets. Many agencies have their own web pages that link to agency-related data (some may be hard to find unless the user knows where they are). Others maintain applications that can be used to find, download, or display data. Two of the latter are GIS based, the general N.J. Geographic Information Network (<https://njgin.state.nj.us>), maintained by the OIT Office of Geographic Information Systems, and the N.J. GeoWeb maintained by the Department of Environmental Protection (<http://njwebmap.state.nj.us>). Other sites include the Department of Education's NJSmart Education Data System (www.nj.gov/education/njsmart/) and

Department of Health's N.J. State Health Assessment Data System (www26.state.nj.us/doh-shad/home/Welcome.html), among other systems; both agencies were represented at the Forum.

Regrettably, the state has not undertaken what would be a complicated inventory of data held by agencies. It also has not undertaken the simpler (but still not trivial) task of creating and maintaining a single website index of data that could be posted on agency websites. A legislative proposal that would adopt an "open data lite" policy, which included the requirement to undertake a full inventory, was proposed in the Legislature in 2013 (and reintroduced in 2014 as A-2071/S-3137), but has not yet received a committee hearing.

Other state government agencies

There are three other participants in New Jersey government data activity: the Legislature, the Judiciary, and independent agencies (a.k.a., in-but-not-of agencies). The Legislature and Judiciary manage their own information technology and make their own decisions regarding open government data. Both branches of government have been active in advancing open data, but have been limited by available resources and, in some cases, policy decisions (i.e., the Legislature exempts itself from the public access requirements of the Open Public Records Act).

Independent agencies, while subject to OPRA, set their own technology standards and priorities. They are subject to direction and guidance from the administration, either directly or through their governing boards, whose appointees come through the Executive Branch. Some agencies, New Jersey Transit for example, are embracing open data standards for some of their activities, while other, less prominent ones may not be inclined to treat open data as a priority.

State agency data activities

Historically, state agencies have set standards and requirements for public and private entities to keep records as required by law as well as for the submission of reports related to regulatory or statutory requirements. This includes employee record keeping, the way in which

financial records are kept, environmental permits and reports, the submission of tax and fee reports, and a myriad of matters in which entities and individuals living or working in the state must take action according to state laws.

Many of these submissions started as paper forms. Over time and with the development of technological tools, more and more of these activities have been modified to allow electronic recordkeeping and submissions. In these cases, state agencies took what were fields on paper forms and converted them to data using standard field names, types, lengths, etc. Organizations engaged with these agencies were required to adopt their standards as appropriate to the purpose.

While the state collects an increasing volume of data electronically, there are still many activities that have not been converted to data standards with the associated electronic submissions. Furthermore, there are public sector activities for which the state does not collect data, but does set standards for recordkeeping; many of those standards are still rooted in their original paper-based environment.

In other cases, electronic submission of information may be required, but not collected as data, e.g., municipal audit and related financial documents may be submitted as PDFs, which are quite useless for data analysis purposes. There are other cases in which electronic documents are collected, but the collecting agency has not had the necessary resources to create datasets that meet the standards of open data.

From a practical standpoint, after almost a decade of shrinking numbers of employees working in state government coupled with mostly stagnant budgets, many agencies lack the resources to engage in improvements in the absence of third-party funding (i.e., federal aid) or policy directives to focus on a specific initiative. That said, there is evidence of change as demonstrated by new data sites managed by OIT and the Departments of Education and Health, and as retiring managers and employees are replaced by ones more appreciative of the opportunity.

Local unit activities

New Jersey is the country's most densely populated state, and has over 1,500 local units (municipalities, counties, local authorities, and boards of education). All these agencies derive their legal authority to act from the actions of state government through legislation and its supporting regulations.

The state's local units are a diverse group of agencies, however, and each must be considered based on its population, geographic size, the number of services it delivers, its management capacity, technology maturity, political dynamics, and for our purposes, by how each one manages its data.

Only the largest and most sophisticated units have programmers to develop and maintain application software; most local unit applications are provided by third-party vendors and run on the agency's servers or are moved to the cloud. At this time, there are no standards or requirements for how the data in those systems is managed for open government purposes. Unless a state agency has required standards for recordkeeping purposes, developers use their own business judgment for data design and reporting standards.

Alternatively, some local units provide a shared service environment where, for example, a county government will host cloudlike services for their constituent municipalities. This can include specific applications (Sussex County), GIS systems (Morris County), or records management services (Monmouth County).

Most local units are resource challenged. While there are some exceptions, state policies that limit increases in spending and property tax levies make the allocation of resources for new projects difficult. If state government wanted to require property-tax-based agencies (municipalities, boards of education, and counties) to engage in open government data practices, a state constitutional provision^{xiv} that limits the ability of the state to mandate new services without providing revenue to the local unit to pay for it could limit its effectiveness.

Because the circumstances of each agency vary widely along with its technological capacity, there are additional

challenges to developing standard patterns and practices among these groups concerning open government data. To the extent that there is a public policy goal of developing consistency in open government data across government agencies, there are very few mechanisms currently in place to reach that goal. However, in the absence of state-based public policy, a small number of local units are taking first steps toward implementing open government data policies and practices (e.g., Newark and Jersey City).^{xv}

If authorities wanted to ensure consistency in property-tax based local units across the state, and if state oversight entities were to establish standards for a mandatory, common set of open data for local units to post, such a policy would likely implicate the State Mandate/State Pay law. Consideration of alternative procedures would be a better approach to reaching this kind of broad, all-encompassing goal.

II. RESOURCE CHALLENGES

Appreciating that open government data is a new government service, the following resource related considerations emerge:

- The time and attention of management allocated to making decisions and of staff to executing those plans is variable and often scarce;
- Funding has not been traditionally allocated for the purpose opening data. Resources for staffing, organizing and curating data, and for technology systems may be limited and require new resources;
- Staff members within the agency may have limited technology skills and supplemental talent may not be affordable. This may require the agency to seek external guidance to implement the service;
- Public support for OGD spending may be limited, especially in the context of budgets and priorities; for property-tax-based local units, these limitations may include issues related to levy or expenditures laws.
- Making the proposition more challenging is that

potential savings from OGD are unclear and not quantified. This may be offset by public advocacy for the presentation of at least minimum data sets to demonstrate political or policy intention in order to meet open data goals.

Do all parties understand the resource requirements? A recent blog post^{xvi} by Arnaud Sahuguet, the chief technology officer of New York University's GovLab listed some of the factors that can hide the true cost of developing open data:

- Unexpected startup costs if data is kept in a legacy computer system that requires reformatting;
- Quality-related costs to keep open data fresh and up-to-date;
- Legal costs to comply with open data legislation;
- Liability costs in case something goes wrong, such as publication of nonpublic information; and
- Public relations costs that can occur when a jurisdiction generates bad press from open data about poor performance metrics or workforce diversity problems.

III. DATA CHALLENGES

About datasets

The general public is not aware of the range of datasets held by a given government agency or that similar agencies often hold common data (i.e., municipal budget data). Alternatively, proponents of open data, business interests, and policy advocates will have knowledge of datasets that relate to their concerns. There are also government datasets designed for internal use that may have value if they are publicly opened. Given policy dynamics concerning open data, agencies will need to set priorities for opening and maintaining data (and possibly develop a ranking process to limit subjective decisions).

Arguably, where local datasets are collected by oversight agencies (mostly state agencies) the agencies should be responsible for aggregating and posting them to provide

a one-stop location for the entire data set. For many local unit activities, standard data formats do not exist, as oversight agencies have not promulgated or required electronic data standards. While nothing prevents agencies that want to adopt open data policies from doing so, the lack of data standards for datasets that span multiple agencies presents medium- to long-term challenges.

Datasets are not always clean, and may include bad data. This can lead to faulty reporting or policy analysis. (See the Sidebar on “Reasons for Bad Data” for more on this topic.)

Once datasets are slated to be “open,” the sets must be prepared, meta-data created, the validity of the data checked, and its formats confirmed. In addition, once opened, the datasets need to be curated and kept up-to-date. Resources are required to do this along with personnel who are assigned the responsibility for these tasks.

During the Forum attendees suggested that policies and priorities should be adopted that favor opening datasets that save money and bring in revenue. However, this would face challenges because of the OPRA default requirement that electronic data in standard formats be disclosed at no cost.

Government agencies that use third-party vendors to manage their data may face additional challenges. Vendors have been known to charge additional costs when specialized reports are needed or to impede requests by setting up expense and process obstacles when full data sets are requested.^{xvii} Contracts with such vendors that do not provide for access to the government-owned data need to include provisions for downloads of datasets as an agency may need them. Added costs should not be a barrier to opening data that is owned by the government but held or managed by a third party.

Smaller government agencies (at all levels) will likely need guidance on how to best manage data. The value of providing institutional support and guidance to government agencies and possibly the public should not be overlooked. This could be accomplished by creating an open data ombudsman or by support from advocacy organizations. Alternatively, similar agencies or their

Key Reasons for Bad Data^{xxix}

- Systems are too old and data entry too time consuming, or too difficult to extract and analyze
- Error-prone data entry systems
- Ineffective system controls (inadequate separation of duties/ access controls)
- Un- and under-trained workers,
- Bad/inconsistent definitions
- Siloed systems
- Lack of centralized control of data
- Problems with data collected by private-sector contractors - poor formatting, costly programming

DJ Patil, the recently appointed Chief Data Scientist of the White House, summarizes the data problem well, noting that “you have to start with a very basic idea: Data is super messy, and data cleanup will always be literally 80 percent of the work. In other words, data is the problem.”

professional organizations could work together to create open data consortiums. Lacking such support, these agencies will need to fend for themselves. This comes with a price: the duplication of efforts, inconsistency across agencies, and public confusion.

Open records and personal privacy

Data privacy issues must be carefully considered. The technology and the science behind “big data” permit otherwise unconnected datasets to be merged. This can result in potential privacy challenges (i.e., time and location data from the automated license plate readers mounted in police vehicle merged with vehicle ownership files). In addition to names and addresses other data elements can be added to the mix as well. The increased reliability of facial recognition software suggests more complicated challenges for government entities operating security video cameras.

A new science of “data re-identification” is developing. These techniques result in users being able to assign personal identifiers to what was intended to be anonymous data. This is a significant societal shift that touches broad issues of what privacy means in today’s internet- and social-media-enabled world, and has implications that may affect the disclosure of government data. This is a highly complicated and challenging issue (and is further discussed below in Section V, “Public Records”).

Data standards

Inconsistent data and formatting can result in a duplication of efforts, limit the ability to aggregate and compare data across agencies, and ultimately increased costs to the public.

To improve competition between vendors that provide services to government agencies, it would be advisable to require vendors to meet data standards. This can serve as means of differentiating them and can incentivize compliance with standards when they exist. While the notion of encouraging vendor compliance has high value, allowing (as was suggested) compliant vendors to avoid procurement competition create challenges to principles of open and fair competition.

Technology system design plays a role in how data becomes accessible. In addition to meeting technical standards for data, systems need to have the capacity to output data that meets open data standards (as described earlier).

For some agencies, the requirements of federal agency regulations may cause them to triage its data management efforts, as different reporting models are needed for different entities.

There is considerable value in providing sets of data over time. However, it should be recognized that there are obstacles and cost factors that inhibit this, i.e., earlier data storage models that require updating and the conversion of data to useful and more modern formats.

Other elements

Geographic Information Systems (GIS) provide powerful tools for data analysis and visualization. They are an

opportunistic data challenge, as GIS tools have high value to researchers, businesses, and policy analysts. There are, however, wide variations in system capacity that vary with individual agencies (if they have GIS systems to begin with). GIS applications also possess a range of technological capabilities that can enable open data (or not), thus creating a disparity in how data is presented to the public.

The newest opportunity (with its accompanying challenges) is the so-called “Internet of Things” (IoT); this is the use of low-cost sensors that constantly produce data. This information is supported by specialized technology tools that facilitate the analysis of “big data.” “...there is enormous potential in the number of sensor and data collection channels already in place in government, such as traffic signals, telecommunications towers, video surveillance cameras, citizen-connected data such as “smart” water meters, or even citizen-owned devices like smartphones. Adding computing capability and advanced analytics to existing infrastructure could make adoption of this trend much more manageable.”^{xviii} This new technology is already generating unprecedented amounts of data; technologists, advocates, and researchers are only starting to address its impact on security and privacy issues.

Finally, there is the conceptual challenge of designating third parties to process and format information and then allowing them to charge for it vs. requiring the government to pay for it up front and making it available to interested parties for free. New Jersey state agencies have generally avoided this problem, as in those data elements affected by this (legislative documents, the compilation of laws, the publication of rules and court cases) are generally available in their native forms at no cost to the public, although third parties that manage them may add value and then charge for the added services.

Exception to this are the state’s construction codes; they require the use of detailed technical manuals published by a national standards organization that charges for copies. Locally, municipal law codes (ordinance books) are managed by third parties under contract to the municipality and are accessible through municipal

websites, although download capability may be limited. The national advocacy movement at “[Public.Resource.org](#)”^{six} focuses on this issue with their axiom “Law is the operating system of our society ... So show me the manual!” as the basis for opening up these resources.

IV. UNDERSTANDING THE IMPERATIVE: THE DISCONNECT

“ But in the 21st century, the “available upon request” model of transparency doesn’t quite cut it anymore. Digitally-connected citizens are not satisfied waiting weeks for a response to a request for information. They expect governments to be “transparent by default” and be able to find answers to their questions in a quick search.”

—*From Solving Government Data Challenges* ^{xx}

There is a disconnect between open data advocates (civic and business advocates) who want data, and managers of government agencies who hold it along with taxpayers who would pay for the new service. It highlights a lack of agreement on the value of investing in OGD and a lack of understanding of what the open government data imperative means.

This is reflected by the concerns and questions that elected, appointed, and administrative government officials express:

- Elected officials are concerned that data may be misrepresented and possibly altered for political purposes. This can result in lack of buy-in and political will to develop OGD.

- Political leadership and public administrators are not often synchronized – the two sides do not perceive the values of OGD at the same levels. They will ask, “Why should we do this when we have so many other priorities to meet?”
- There are also the challenges of committing resources for the development and maintenance of OGD (time, attention, money).

It is very clear that there are trust issues between government officials and journalists. Regrettably many government officials and administrators feel that the press intends to use data to create “gotcha” stories.

While valid to a degree, there is culpability on both sides. Some New Jersey government officials have deserved (or undeserved) reputations that often lead journalists to distrust them. In effect, data can become a battleground, something government officials will look to minimize.

These issues are amplified by a generation gap, where millennials, who make up the bulk of today’s journalists have a different view of government than the predominant decision-makers who are baby boomers and their predecessors. This also underscores the reality that many government organizations are trying to run 21st century government using 20th century standards, something that these younger “digital natives” do not understand.

There is a gap between expectation and reality among those seeking data. The opening of some datasets may lead to questions for which data does not exist, or the data received is not what was expected and therefore does not answer the question(s) being asked. In addition, data dumps of raw data, while fulfilling policy, make it challenging for end users to find the value in it. This can lead to additional questions and challenges that require additional resources to address. Furthermore, data users may make assumptions about what data government collects and the forms in which they maintain it; the reality may be inconsistent with what those requesting information may expect. This is a misalignment of what government can do, versus what citizens expected it to do.

Government agencies at all levels vary in their technological capacity; open data advocates that expect

consistency are often disappointed and critical when they find that organizations vary in their approach to data. These inconsistencies between and within state agencies reflect our disjointed technology planning practices (or lack thereof). The differences also reflect New Jersey's interpretation of "home rule," the ability of local units to control most aspects of government service delivery and local decision-making. This conflict can raise issues of trust between government officials and open data advocates.

Governments that decide to open their data must be prepared to address how they prioritize the data they open. There will be questions about the reliability and veracity of "selected" data being opened and the challenges of why given sets are opened while others not. If the goal is opening and posting all data instead of just what is requested, what happens if the cost of posting everything is not deemed affordable? Answers to these questions can rest on values, costs, technological capacity, existing formats, and policy/political implications. Governments can address these issues by engaging interested parties in the decision-making processes concerning open data.

Waldo Jaquith has summarized the significant conundrum of open data.^{xx1} To reflect the context of this report, text in *italics* has been added or ~~struck out~~ from his original blog post:

There is no incentive model that makes it rational for government ~~employees~~ to publish open data. The safe thing to do is to publish bland, unobjectionable, low-value data.

One solution to this is probably the hardest possible solution: *CULTURE CHANGE*. So long as there are legislators *or journalists* who are willing to seek out small mistakes to pounce on to score political ~~or journalistic~~ points, and so long as agencies don't support...*open data initiatives*..., then it will be irrational for government ~~employees~~ to publish datasets that could be used to make them look bad.

The Sunlight Foundation has developed guidelines on policy setting and public engagement that may be of value to agencies considering their open data policies.^{xxii}

One approach to overcoming the understanding gap would be to encourage cultural and practical training for government officials and their staff on the roles and responsibilities of journalists in a democratic society as well as training programs for journalists on how government works. Regardless, there will always be challenges to address.

V. PUBLIC RECORDS ISSUES

Open data as a cost efficiency

Treating OPRA requests and responses as data is commonly recommended by open government advocates; they posit that an agency that automatically posts the requests it receives and the information given in response will reduce its OPRA resource load. This warrants consideration but needs to be considered just as any other open data priority would be. It has technology and resource demands, as well as privacy implications. This is different from the concept of treating data commonly requested as open data and posting it so that interested parties do not have to ask for it.

Notwithstanding privacy issues, ideally there would be a single system all government agencies could use to record OPRA requests. Such a system is used by Executive Branch departments, though not posted as open data.

The argument that posting OPRA requests and their responses will reduce costs of OPRA compliance is considered speculative and is unproven. It may be the case that if technological capacity exists, it should be done. The absence of technology is not seen as a reason to invest in technology with an assumption that savings will follow.

OPRA issues overlap many open data issues and can create complications, as well as compliance advantages. For example:

The challenges of time to comply, technological capacity and limits on cost reimbursement for narrow, but legal data requests. Under current law, the act of posting data which an individual

The Challenge of Open Code Sharing^{xxx}

can then search for the desired, specific information does not remove the requirement for the custodian to comply directly to the request, particularly in the case of individuals who cannot or will not use online tools. A recent Government Records Council (GRC) case, however, concluded that directing someone with a request to a specific online record can serve as a valid response. This issue is complicated by outliers who request information on a frequent or regular basis (they may be looking for a “gotcha” moment or are driven by narrow issues) and whose detailed requests are expensive and time consuming. (Up to this point, courts have been reluctant and inconsistent in allowing the special service charges allowed under the law.)

There is a question of compliance under OPRA that relates to data sets and requests; is data a “set” or an individual record, and how must it be disclosed (as data or in as a report)? Also, does an agency have the responsibility to provide different versions of data that is already opened if the individual requests already posted data in a different form; this is problematic and time consuming particularly in light of limits on special service charges.

There is a penchant of many agencies to engage in legal reviews of OPRA requests (usually related to finding exceptions to disclosure). These costs are then charged to the person filing the request and can affect the availability of data sets as well.

Privacy and data

OPRA’s policies regarding an individual’s right to privacy versus the release of electronic data need to be updated, as this issue was of less importance when the law was enacted in 2001. Outside of filing a denial of access charge with the court or the Government Records Council, there is no mechanism for agencies to obtain binding and consistent guidance on the “reasonable expectation of privacy” of individual data. Denial of access complaints are

There are two elements to “open” when it comes to government-developed software (applications). The first is the use of open-source tools to develop new software. There are many such tools (and many have private sector proprietary counterparts, some of which may bring added value). However, more often than not, open-source development tools can meet the needs of government agencies. In addition, these tools can be shared by multiple organizations, thus providing value to all and dividing the cost among many. This benefits taxpayers by lowering the overall price tag of the project. There is a culture of open-source advocates that make very compelling arguments for developing new government applications with open-source tools. <http://opensource.com/government/13/5/top-5-misconceptions-open-government>

The second is the principle that when government develops software, it should be developed in an open manner that lets interested parties contribute to the process, and suggest improvement to the code, i.e., open-source development. The corollary to this is that once developed (and paid for by taxpayers) it should be available to any other organization that wants to use it. These principles are put into practice by using such collaborative hosting services as GitHub^{xxxi} as a software repository and development management tool.

adjudicated based only on data presented in a specific case, often without expert testimony on how data can be used in light of contemporary re-identification technology.

State law allows access to government records under the traditional common law as well as under OPRA. While an OPRA request may be denied due to a specific exception to disclosure, an appeal to the courts under the common law may allow disclosure under specific circumstances. This creates confusion and inconsistency in records administration throughout the state.^{xxiii}

From a public policy perspective, when OPRA was enacted, public access to many government records were limited by the functional obscurity of government data, in that the technology of the day effectively limited access to data. In an era of “big data,” new and more powerful technology tools and the ever-increasing ability of non-

government organizations to manipulate and correlate data sets, personal privacy can unknowingly be placed at risk from an otherwise ordinary data set. Since OPRA was enacted, access to once-unavailable data is now provided without consideration of the implications of disclosure on privacy.

Agencies and the public need a clear and consistent processes to determine personal privacy issues, in advance of an OPRA “failure to disclose” case being filed (and the potential fees that may be awarded at public expense to plaintiffs). Furthermore, there are substantial issues surrounding the notion that “because government collects personal information, it must be disclosed,” and these issues warrant more detailed consideration.

The state once studied issues of privacy. The Privacy Study Commission, was created as part of the OPRA law, but was dissolved after it submitted its report in December of 2004.^{xv} The Commission was prescient, as it foresaw many of the issues we are facing today. Regrettably, the report resulted in no legislated or government policy action, and its recommendations were virtually disregarded.

Records retention

Government records custodians must also consider records retention issues. As agencies move from paper-based records to digitally stored data, records retention schedules, based in part on the ability to store paper records warrant reconsideration. It may be that with the lower costs of storing digital records, retention times should be increased. These standards are set by a separate agency, the State Records Committee, and changes are often driven by the custodian agencies. The Committee and its support staff it has undergone recent changes in management and personnel, and are currently hard-pressed to manage a thorough overhaul, as they are involved in other projects involving new records storage technologies and standards.

Fear of format

A long-standing concern of officials who post data in closed formats (e.g., PDF image files) is to ensure that users would not be able to “alter” the data. While

superficially well-intended, the concern is misplaced. As long as the originating agency maintains the “single version of truth” on its own website or in other forms, someone who alters government data runs the risk of having the data compared to the original “version of truth.”

There is a similar concern regarding the opening of documents that contain an individual’s signature and the fear that this digital signature could be used in a forgery. This concern is also misplaced. As long as an image of a signature is posted, regardless of its format, it can be forged. Even if redacted, the signature exists on the original documents which are subject to disclosure. The reality is that such forgeries are irrelevant for law enforcement purposes, and have little value in the real world from a fraud basis. If they are used, they are minor elements in a larger unlawful act. Still, this issue can be mitigated by the use of a “s/_____” in a facsimile of a signed document with the signer’s name inserted as text.

VI. MANAGING THE TECHNOLOGY

Do all government agencies possess the technological capacity to develop and manage open data? The answer is, “No, it varies, a lot.”

In the local government environment, 90% or more of the applications used by municipalities are commercial packages or online services. They have varying degrees of flexibility and many place limits on user control over data output and formatting. Such applications include those used for payroll, police dispatch and record keeping, financial management, GIS, and code administration. Some software systems make using data sets and extracting data easy, others do not.

State government uses a wide variety of applications whose code is maintained and supported directly by OIT and individual agencies. In addition, many units use commercial software that runs on state servers as well as online services administered by third-party providers (in the “cloud”). While the state-administered applications may lend themselves to providing open data, the commercial and online services may yield a different result.

Agencies need guidance to deal with copyright, reuse, licensing, etc. to simplify the process for government and to meet open data standards. While this is a small issue, agencies need authoritative guidance. Wikipedia notes that NJ has copyright rules about its government (and local data):

“It is the policy of the state of New Jersey that all documents originating from web sites of executive departments and non-independent agencies are ‘available to the public and anyone may view, copy or distribute State information found here without obligation to the State’ unless the document specifically states otherwise. Likewise, all records obtained from state, county, or local government entities in New Jersey via the state’s Open Public Records Act (OPRA), per a 2009 decision of the New Jersey Supreme Court, may be reproduced including for commercial purposes.”^{xv}

A remaining challenge is more pragmatic. New technologies and new research will cause the open data environment to constantly evolve; government agencies need to be ready and able to adapt to these as yet unanticipated expectations.

Procurement of technology goods and services

The restrictions posed by government laws and rules that concern how agencies are allowed to procure technology goods and services is a common source of frustration for government administrators and technology managers. This is not unique to New Jersey; it is a national problem. The need for a different, more flexible model challenges long-held principles of public procurement that award contracts to the lowest responsible bidder or to the proposal that is the most advantageous, price and other factors considered. A failure to follow procedures to the letter often results in vendor protests that delay projects while adding costs or result in goods and/or services that do not fully meet the needs of the agency.

The current procurement model was developed in the 1950s based on commodity purchases of clearly specifiable goods or services that changed little over a multi-year contract. Over the years, the process has also

become encumbered by other well-intended policies relating to issues and practices intended to promote social agendas (e.g., a contractor’s adherence to the MacBride Principles as they relate to doing business in Northern Ireland). Other guidelines were established to prevent collusion and corruption. Enterprising agencies have used innovative legal maneuvers to work around some of those problems, including the use of cooperative contracts (i.e., piggybacking on contracts awarded by other agencies in New Jersey or those used by other government entities across the country). They may also use blanket, cost-plus contracts for software and specialized services. However, these solutions do not always result in the best price for the product or service; they result in a price that meets the requirements of a law.

There are solutions that can be applied to this area, but they must be carefully considered because of their potential to result in unintended consequences, and they must be enacted through legislation.^{xvi}

The Value of Coders

Skilled coders have the ability to solve almost any problem with sufficient time and funding. If there are coders and money to pay them, opening governing data become less cumbersome.

The need for skilled coders and project managers is critical to understanding the processes that enable data to be opened. Regrettably, in New Jersey government this is not a trivial problem. Many agencies have no reason to employ programmers. For those that do, government hiring and benefit practices make it difficult to compete with the private sector. When third-party software is involved, vendors have free rein to charge for new services.

Some advocates believe that all New Jersey government agencies have up-to-date coders and/or technology administrators who are easily able to make government data open and available. However, opening and maintaining data are not minor undertakings. In reality, many agencies do not have technology specialists on staff, and even for those agencies that do, managing data will require considerable time, effort and money.

Part Four— Concluding Thoughts and Recommendations

“ Moving New Jersey governments to an open government data model will not happen all at once— nothing does. ”

Rapidly improving technology is enabling and encouraging government agencies to adopt open data policies and to implement technology-based solutions to open data. The effort is supported by evolving standards, software, and service solutions that reduce the resources needed to implement open data policies.

The new technology introduction curve may be instructive in assessing adoption of open data practices. There are early adopters using technology that requires a high resource commitment; others are starting smaller, with limited or first-step implementations of open data projects using a variety of formats as part of a commitment to open government.

As technology becomes simpler and increasingly user friendly, more places will adopt practices that are consistent with public expectations, agency priorities, and resource availability. That said, state government agencies may need to consider appropriate legislation as a way to begin meeting open data goals. While legislation has been introduced, it should be carefully considered in light of the challenges described above.

For local units, legislation may not be appropriate given the limitations imposed by State Mandate/State Pay provisions and the challenges described above. The notion of one-size-fits-all may not be appropriate for the range of local units. If the Legislature wants to develop policies for local units, it should proceed carefully, review options and consider policies that encourage local creativity and needs.

RE-ENGINEERING OF DATA REQUIREMENTS IS WARRANTED

N.J. government data is a combination of data originated and used locally, and data collected locally and sent to state agencies for use pursuant to state requirements. Originators who want to open their data should be able to link to the state disclosure site, and the state maintain a centralized resource for providing data it collects, particularly for data submitted by local agencies. This would eliminate duplication of effort by individual agencies and give data users a form of “one-stop data shopping,” with data from different sources aggregated in one place.

In cases where the state is the “data manager” it should provide the hosting, access, and tools, or engage in partnerships with other agencies or third parties to do so. Some third-party organizations may be able to host and monetize data by adding value services, but they would also have to provide basic Open Public Records Act (OPRA) custodian services.

This would require that legacy paper-based systems incorporate digital standards and that existing digital standards be reviewed to reflect open data policies. Changes can be implemented in existing IT systems over time (or all at once, if well planned); oversight agencies can impose requirements for third-party software to adopt new standards during upgrades or over a limited period of time.

This proposal for the state to assume a broader data management role has OPRA implications:

- It needs to address OPRA access issues when local data is held by the state; the principle of “custodianship” should transfer to the state for those datasets.
- It needs to be able to handle proper requests by a person without technical skills or tools who “doesn’t do data,” but wants information.

START SOMETHING!

In the absence of state government decrees or the imposition of mandates, all agencies (state and local) can begin developing open data on their own; starting slowly and taking small steps is wise. Incremental progress and change can help work through any questions surrounding resources, data, and understanding. The diversity of New Jersey government agencies will result in many and varied approaches to making data more available. Agencies can start the process and increase its development over time using the following guidance:

- **Learn** more about open data. Use the resources cited in this report to learn about open data policies and technology, and determine how they can apply to the organization.
- **Assess** the organization's current practices and begin institutionalizing open data practices. Start with the low hanging fruit to get a feel for what works for the organization and its clients, constituents, and customers.
- **Inventory** data to whatever extent is feasible in order to see what is available.
- **Ensure** contracts for technology services (i.e., payroll, financial administration, police records, GIS services) provide:
 - o The capability to download agency data in open formats on demand and at no additional cost;
 - o Routine reporting capability of data in open and agreed-upon presentation formats generated locally or on request and at no additional cost;
 - o That all agency data held by and processed by the vendor is owned by the agency;
 - o For the return of data in an open format when the contract is ended, upon request and at no additional cost;
 - o That any fees for development of specialized reports or queries are at a reasonable rate to ensure agency and public access to information.
- **Prioritize** what to post – create a process to determine which datasets to post, and how they will be

posted. Consider common OPRA requests as a guide to what can be posted, and balance:

- o Value to users as well as management concerns (consider consulting with interested parties to make the assessment).
- o Resources needed to provide useful metadata and policies for keeping the data current.
- **Manage** the technology and develop administrative processes to post data:
 - o Adopt a set of open standards that is practical for the organization.
 - o Review the options for posting data, e.g., on a local website, through a specialized commercial data portal, or in coordination with other agencies by using aggregated data sites (county, regional or agency types).
 - o For local units that use a third-party web service, encourage providers to develop and provide access to pages or sites that focus on data presentation.
 - o Ensure that data is indexed and locatable on whatever website or portal is used.

Possible municipal government starter data (and non-data), i.e., low hanging fruit:

- Budgets and financial information (bill lists, employee salaries, etc.) in open data formats: csv, json, or xls instead of (or in addition to) PDFs, images, or proprietary formats
- Minutes of governing body meetings in searchable formats
- Data that is regularly requested through OPRA requests
- Crime reporting data as appropriate to the community
- Restaurant inspections
- GIS-based zoning information, voting locations, public facility locations
- Property assessment and tax data (MODIV tax lists)
- Performance data captured by the agency
- Statistical data the organization uses as part of its mission

OPRA AND PRIVACY ISSUES

- Create a statutory process to determine and reconcile privacy issues using a special master appointed by the Judiciary, or create of a type of Institutional Review Board with the authority to make binding and precedential decisions to deal with privacy issues. This model is currently used by higher education institutions when dealing with human subject research; the concept may be relevant to privacy decisions. A new concept is that of “data ethics,... an expansive set of practices and behaviors grounded in a moral framework for the betterment of a community (however defined).”^{xvii} This approach may have value in the way that open government data is treated and used.
- Another approach might be to re-establish the state’s Privacy Study Commission as a means to further explore these issues, with the goal of legislative recommendations that address the opportunities.
- Other OPRA issues raised in this report will need to be addressed through amendments to the law. However, prior to deciding what changes need to be made, a process should be established to assess the issues in order to develop appropriate and balanced amendments to the law.^{xviii}
- The State’s Records Management Service, the agency responsible for setting records retention standards, should be granted additional resources to address the impact of technology on records retention and management.

THIRD-PARTY ACTIVITIES

Philanthropic and civic organizations could collaborate to help develop guidance and support to assist government agencies looking to establish open data policies and practices. They could also provide training programs for local officials.

The Sustainable Jersey program can provide useful guidance as part of its certification activities (in part, already supported by philanthropic funding).

Organizations that provide programs for government employees and officials could develop continuing education, academic certification and degree courses that infuse open-data policy practices into them as appropriate. These include:

- Rutgers Center for Government Services (certification programs and continuing education programs for local government officials)
- Rutgers School of Public Affairs and Administration; the school houses the state’s Certified Public Manager program, the Center for (Government) Technology Leadership, and Institute on Transparency in Governance, all of which could contribute to training managers in the development of open data principles.
- Rutgers Center for Executive Leadership in Government (certificate programs for local government executives)
- Institute for Information Policy and Law, part of the Rutgers School of Law (Camden)
- The state’s higher education graduate and undergraduate public policy and public administration programs
- Law school programs (the newly combined Rutgers and Seton Hall law schools)
- Continuing education programs for government officials through the N.J. State League of Municipalities, N.J. Association of Counties and other organizations that represent government agencies and

provide continuing education programs. Examples include: the Municipal Clerks Association of N.J., Association of Environmental Authorities, NJ-GMIS (government technology managers), CARMA/ NJ (County Archives and Records Management Association of NJ), the Constitutional Officers Association (County Clerks), N.J. Institute of Continuing Legal Education, and others.

Other third-party activities might include:

- Developing an initiative to consider and advocate for legislative and policy changes that will lead to the implementation of the recommendations in this report.
- Creating a lab or research center to support and advocate for agencies looking to implement open data actions; collaborate with some of the third-party partner organizations noted above (perhaps a New Jersey Open Data Institute).
- Establishing thought forums to work through related issues, such as the aforementioned OPRA challenges, and those of government technology procurement.

Appendix

A. OPEN DATA REFERENCE SOURCES

- <http://opengovernmentdata.org/>
- www.sunlightfoundation.com
- www.usopendata.org
- <http://opendatahandbook.org/guide/en/how-to-open-up-data/>
- www.opendatafoundation.org/
- www.opendata.guide
- <http://us-city.census.okfn.org/faq>
- <http://opendatamanual.readthedocs.org/en/latest/index.html#>
- www.socrata.com/open-data-field-guide/

B. THOUGHT FORUM ATTENDEES

Category	Invitee	Organization
Academic	Kelly Robinson	Rutgers, NJ DataBank (SPAA)
Academic	Angie McGuire	Rutgers, NJ Data Book (CELG)
Academic	Matthew Hale	Seton Hall Univ. Public Administration Program
Advocacy	Mark Headd	Accela Inc.
Advocacy	Waldo Jaquith	U.S. Open Data
Advocacy	Ellen Clarkson	Citizens Campaign
Advocacy	Anna Lukasiak	Code for America: Open JC
Advocacy	Alex Torpey	Rethink Leadership/Veracity Media
Advocacy	Lauren Skowronski	Sustainable Jersey
Civic Sector	Deborah Poritz	Fund for New Jersey
Civic Sector	Chris Daggett	Geraldine R. Dodge Foundation
Civic Sector	Molly De Aguiar	Geraldine R. Dodge Foundation
Civic Sector	Ingrid Reed	Ingrid Reed
Local Govt.	William Kearns	General Counsel, NJ League of Municipalities
Local Govt.	Katya Wouk	Montclair Twp. Webmaster (NJ GovWPro)
Local Govt.	Sean Canning	Mount Olive Twp. Chief Administrative Officer
Local Govt.	Allen Weston	N.J. Association of Counties
Local Govt.	Patricia Tumulty	N.J. Library Association
Local Govt.	Justin Heyman	Technology Management – NJ-GMIS
Local Govt.	Matthew Watkins	West New York, Chief Administrative Officer
Media	Jeff Jarvis	BuzzMachine.com/CUNY Tow-Knight Center
Media	Shannon Mullen	Gannett New Jersey
Media	Steve Leibman	NJ Advance Media
Media	Debbie Galant	NJ NewsCommons/Hyperlocal Media
Media	Lee Keough	NJ Spotlight
Media	Augustin Armendariz	NY Times
Media	Nancy Solomon	WNYC-FM
State Govt.	Michelle Smith	Judiciary - Clerk of the N.J. Superior Court
State Govt.	Bari Erlichson	N.J. Department of Education
State Govt.	Cathleen Bennett	N.J. Dept. of Health
State Govt.	Colette Lamothe-Galette	N.J. Dept. of Health
State Govt.	Joseph Donahue	N.J. Election Law Enforcement Commission
State Govt.	Andrew Pratt	N.J. OIT - CIO's Office
State Govt.	Kathleen Smith	N.J. OIT - Data Management
State Govt.	Tommi Povia	N.J. OIT – Data Management
State Govt.	Katrina McCarthy	Rowan Univ. GIS Services – NJMAP
Event Convener	Marc Pfeiffer	Public Technology Institute
Event Recorder	Debra Meltzer	Public Technology Institute

Endnotes

- i <https://www.safaribooksonline.com/library/view/open-data-now/9780071829779/>
- ii Mr. Jarvis comments on media issues at www.buzzmachine.com and is Director of the Tow-Knight Center for Entrepreneurialism at CUNY.
- iii This section is generally sourced from <https://okfn.org/opendata/>
- iv American's Views of Open Government Data: www.pewinternet.org/files/2014/10/PI_OpenData_072815.pdf (page 11)
- v <http://opengovernmentdata.org/#sthash.zS4vrzj.dpuf>
- vi From the City of Boston's Open and Protected Data Policy: <https://data.cityofboston.gov/download/2rjs-rb6r/application/pdf>
- vii <http://opengovernmentdata.org/>, with minor edits for clarity
- viii There are several descriptions of open data most of similar principles. We have chose the principles from the federal governments "Project Open Data" at <https://project-open-data.cio.gov/> with some minor clarifying additions.
- ix See <http://opendatamanual.readthedocs.org/en/latest/appendices/file-formats.html> for explanations of open data formats
- x Adapted from <http://civic.io/2013/06/22/why-publish-open-data/>
- xi https://en.wikipedia.org/wiki/Hype_cycle
- xii https://en.wikipedia.org/wiki/Hype_cycle#Five_phases
- xiii <https://law.resource.org/> is the site that advocates for privately published documents that are used as law by government agencies be made open and freely accessible
- xiv Colloquially known as "State Mandate/State Pay" and administered by the Local Mandates Council www.state.nj.us/localmandates/
- xv See <http://data.ci.newark.nj.us> and <https://data.openjerseycity.org>
- xvi From: www.govtech.com/data/Managing-the-Hidden-Costs-of-Open-Data.html
- xvii The ongoing case of *Gannett v. Raritan Borough* is an extreme case of this challenge. See <http://scarincilawyer.com/high-profile-opra-case-results-in-542000-in-legal-fees-for-nj-city/> for an impartial overview.
- xviii Excerpt from www.deloitte.com/us/techtrends2015 "Tech Trends 2015, the Fusion of business and IT, A public sector perspective"
- xix <https://public.resource.org/> is the parent of the site at note x.
- xx From Socrata, <http://discover.socrata.com/rs/851-SII-641/images/%5BeBook%5D%20Opening%20the%20Books%20-%20Solving%20Government%20Challenges%20with%20Open%20Financial%20Data.pdf>

- xxi From <https://usopendata.org/2014/11/12/culture-change/>
- xxii From <http://sunlightfoundation.com/opendataguidelines/>
- xxiii From <http://njfog.org/2015/07/10/the-common-law-right-of-access-to-public-records-and-the-cost-of-the-fight/>
- xxiv The Privacy Study Commission Report is online at http://www.state.nj.us/privacy/prc_final_report_v21.pdf
- xxv https://en.m.wikipedia.org/wiki/Copyright_status_of_work_by_U.S._subnational_governments#New_Jersey
- xxvi An example of procurement challenges and ideas for solutions are at <http://civic.io/2013/10/14/five-ways-to-make-government-procurement-better/>
- xxvii This concept of data ethics is discussed at <https://www.accenture.com/ke-en/insight-outlook-case-data-ethics>
- xxviii A thought forum on OPRA issues might be an appropriate starting point to move OPRA amendments forward
- xxix <http://firstround.com/review/everything-we-wish-wed-known-about-building-data-products/>
- xxx <http://ben.balter.com/2014/08/03/why-isnt-all-government-software-open-source/> and <https://government.github.com>
- xxxi Github is an open source software project that manages and stores revisions of software and other IT and data projects. It provides access control and several collaboration features, such as a wikis and basic task management tools for every project. <https://github.com/about>

The Bloustein Local Government Research Center

New Jersey is served by more than 1,500 distinct local government agencies: municipalities, school districts, utilities, counties, and more. Yet, even with this wealth of opportunity, precious little substantive research has been done within the local government environment to inform some of our state's most pressing policy issues.

The **Bloustein Local Government Research Center**, or **Bloustein Local** <http://blousteinlocal.rutgers.edu/>, serves as a focal point and engages in a range of services, including:

- Encouraging and conducting applied and academic research on local government fiscal and administrative issues, emphasizing application and support to New Jersey local government.
- Developing resources that can assist others in conducting research and analysis.
- Organizing and hosting conferences and symposia on New Jersey local government fiscal and administrative issues.
- Supporting New Jersey local government fiscal and administrative policy development, implementation, and analysis through contract research and on-call advice for organizations and institutions that engage in local government policy setting and policymaking.
- Promoting and increasing public understanding of local government issues by partnering with and supporting civic and media organizations that inform and educate the public on local government matters.

A list of the Center's current projects may be found online at <http://blousteinlocal.rutgers.edu/projects/>.

RUTGERS

Edward J. Bloustein School
of Planning and Public Policy

Bloustein Local Government Research Center
Rutgers, The State University of New Jersey
33 Livingston Avenue
New Brunswick, N.J. 08901

p. 848-932-5475 blousteinlocal.rutgers.edu

© 2015, Rutgers, The State University of New Jersey