MEMORANDUM

July 5, 2012

TO: Government Operations and Fiscal Policy Committee

FROM: Dr. Costis Toregas, Council IT Adviser

SUBJECT: FY13 Recommended Work Program for Interagency Technology Policy and Coordination
           Committee (ITPCC)

The following are expected to attend:

Sherwin Collette, MCPS Chief Technology Officer and Chair, Chief Information Officers
    Subcommittee, ITPCC
Gary Thomas, ITPCC Staff
Naeem Mia, Office of Management and Budget (OMB)

Members of the ITPCC and ITPCC CIO Subcommittee may also be available for detailed questions.

Staff Recommendation:

1. **Accept the Work Program and request an implementation roadmap for the next ITPCC worksession.**

Overview

The 2013 work program for the ITPCC was presented to the ITPCC Principals for adoption after the budget process for the County and Agencies was completed and each agency knew what resource capabilities for 2013 were available. In a meeting of ITPCC Principals on June 26, 2012, it was unanimously adopted. This work plan is presented on © 1-2.

Discussion

Several new elements are presented in the 2012-13 Workplan: an Interagency Open Data Initiative, Interagency Mobile Systems/Devices and Applications, and a Strategic Visioning and Planning process.

In addition, a new subcommittee is being created to look at long-term planning and funding strategies for ITPCC projects. All these new elements are a welcome addition to the more traditional (and yet no less important) elements of FiberNet, Health and Replacement Priority of Major IT Systems, and Continuity of Operations Plan (COOP) Development, which have been included in prior year Workplans.

The first important change announced by the ITPCC is the evolution of the Workplan to a two year plan, rather than the traditional annual plan. This is a wise move, as many of the items incorporated are complex and may take time to define, organize, and allocate to responsible organizational structures perhaps not yet in existence. The ITPCC Principals will, of course, be reviewing progress against the goals on a continuous basis.

The Open Data Initiative comes at an opportune time, as the Council's digital lead, Hans Riemer, is actively preparing legislation on the same topic for the County to consider. Current trends in Open Data are well summarized in the Guidelines for Open Data Policies (on © 3-14) just compiled and released by the Sunlight Foundation.

Specific questions that the Committee may want to pose in order to engage with the ITPCC include:

> How will the many desirable and needed projects suggested by the Workplan be funded? (This question is also addressed on Agenda Item #2.)
> Can the Open Data Initiative lead to a Cloud-based solution, under which different agencies will be able to share information in an open manner?
> Will the FiberNet build out support the creation of public WiFi hot-spots, a stated desire of the GO Committee from past worksessions? If so, where might be locations that lend themselves to such a service provision?
> Will the NOC (Network Operating Center) costs be shared by all agencies? And is there an opportunity to lease the network to a private provider of telecommunications services and buy back annual services, including NOC and SLA Management, for a lower cost than that currently absorbed by the County (in other words, a public/private partnership)?
> The COOP solution implementation is now complete in a technical sense; can the same be said for the procedures and human factors of implementation in all agencies? Or are Change Management processes required to be deployed?
> The Security Special Interest Group (SIG) transition to a formal work group is a good first step to improved security; what major new strategies will the work group be able to undertake that the SIG was not able to?

1.0. Digital Citizenship/Digital Montgomery

1.1 Interagency Open Data Initiative
This project will identify how data from across the agencies can be shared efficiently to make better use of information resources to improve services and access to information. The ITPCC member agencies will develop a pilot project that assesses the feasibility of leveraging and expanding the initial MCG Open Data Initiative to include appropriate interagency data; identify and define project requirements, processes, and potential data sets for initial use in an interagency open data environment; create a formal project plan identifying tasks, milestones, and deliverables; and develop a roadmap for future expansion of the pilot project.

1.2 Interagency Web Search Capability
This project will strengthen citizen access to information contained within agency web sites that current search solutions do not easily reveal. The feasibility and implications of leveraging the MCG solution for web crawling and indexing outside agency site information, and a solution for presentation of the data will be assessed. A pilot project to test and assess the viability and benefits for wider use will be developed.

1.3 GIS Data Visualization Project
As part of the broader Digital Community and open data initiative, data visualization, web and mobile device applications that leverage Montgomery County's GIS resources and infrastructure will be examined. The project will examine how to make GIS information and applications available. The existing GIS Policy Group will be utilized for this effort.

2.0. IT Infrastructure

2.1 FiberNet II and ARRA Buildout
In the final phase of the large scale buildout of the FiberNet network through August 31, 2013, the focus will be to maximize the Federal ARRA grant funds and complete connection of all identified sites; to identify long-term operational support and service requirements agreements; and to assess and document budgetary requirements to sustain the FiberNet infrastructure and operations. This includes priority efforts that will:
- Coordinate the FiberNet Hub Fiber Distribution Center Rebuild for all existing FiberNet Hubs to accommodate new ARRA fiber and existing fiber.
- Determine FiberNet service level requirements of each agency and create a new FiberNet Service Level Agreement (SLA)
- Negotiate and execute MOUs for each agency utilizing FiberNet
- Develop a Network Operating Center (NOC) solution for FiberNet that meets agency requirements and recommend solutions to ITPCC
- Document the FiberNet optical plant using OSOInsight
- Comply with the ARRA Grant requirements for open access

2.2 IT Asset Management-Health and Replacement Priority of Major IT Systems
Focus will be maintained on updating the health and replacement priority for major IT systems, including risk assessments for presentation to the Council in FY13.

2.3 Interagency Mobile Systems/Devices and Applications
Interagency expertise will be coordinated to determine the feasibility of leveraging and sharing applications, application development expertise, management of legal issues, and sharing policy

guidelines related to the mobile computing and the use of non-agency issued technology in the workplace.

2.4 COOP Development—Adding Agencies
The CIO Staff Subcommittee will assess the requirements and options for extending interagency planning using best in class tools such as MCG Office of Emergency Management and Homeland Security (OEMHS)'s COOP Automation Solution that won PTI's and NACO's 2012 awards.

2.5 Information Assurance and Risk Management
To strengthen our preparedness and ability to plan for and respond to evolving and emerging risks, the Security SIG will be transitioned to a formal workgroup and tasked with providing information on specified topics in FY13-14 that should help to strengthen agencies' cyber security plans, estimate resource requirements, and identify categories of risk management controls.

3.0.   Strategic Visioning and Planning

3.1 IT Strategic Visioning Retreat
The CIO Staff Subcommittee will plan a day-long interagency group retreat to engage in long range strategic visioning and priority setting for interagency information technologies and solutions as it relates to digital citizenship in a digital Montgomery County.

3.2 Interagency Social Media Communications Pilot
This pilot project would provide a structured examination of the potential uses for social media tools through the creation of a limited scope, special interest group application/project to explore benefits and issues associated with broader utilization of social media tools [LinkedIn, Google+, Facebook, Twitter, etc.] to create and reinforce a 'learning community', enhance communication and information sharing between interagency staff on topics of interagency interest and collaboration initiatives and evaluate the feasibility for expanded uses beyond the pilot phase.

3.3 Special Subcommittees
3.3.1 Major IT Systems Planning and Funding Strategies
The CIO Staff Subcommittee will examine content, organization, and presentation of information regarding the IT infrastructure health of major systems, risks and consequences to the systems, fiscal requirements and strategies, and coordination with established budget and planning processes to enable and improve resource allocation decisions for this critical infrastructure.

APPROVED by ITPCC, June 26, 2012

Guidelines for Open Data Policies

Open Data policies can...

1. Mandate open formats for government data.
2. Mandate the release of specific new government information.
3. Mandate electronic filing.
4. Require any public information to be posted on the Internet.
5. Mandate continuous publication and updates to data.
6. Create permanent, lasting access to government data.
7. Publish bulk data.
8. Create public APIs (Application Programming Interfaces) for accessing information.
9. Remove restrictions for accessing government information.
10. Remove restrictions on reuse of information.
11. Appropriately safeguard sensitive information.
12. Require exemptions to open data policy to be balance-tested against the public interest.
13. Create a portal or websites devoted to specific issues related to data publication or specific policy arenas.
14. Create or explore potential public/private partnerships.
15. Create contests or other events focused on the use of government data.
16. Require digitization and distribution of archival materials.
17. Create processes to ensure data quality.
18. Create a public, comprehensive list of all information holdings.
19. Mandate the use of unique identifiers.
20. Require the publishing of metadata or other documentation.
21. Require the publishing of code.
22. Set appropriately ambitious timelines for implementation.
23. Ensure sufficient funding for implementation.
24. Empower the creation of binding regulations to implement the new policy.
25. Tie contract awards to transparency requirements for new systems.
26. Stipulate that provisions apply to contractors or quasi-governmental agencies handling public data.
27. Create new oversight authority to review implementation of the requirements.
28. Create new legal rights or other legal mechanisms to empower the public.
29. Appeal to values and goals, such as accountability, efficiency, employment and commerce, innovation, civic engagement, and public services provision.
30. Reference and build on existing public accountability policies, like
     i. Freedom of Information Laws
     ii. Open Meetings Acts
     iii. Open Records Acts
     iv. Ethics Protections
     v. Campaign Finance
     vi. Lobbying Disclosure Laws
31. Incorporate public perspectives into policy implementation.
32. Require analytics about the use of open data to be published publicly.
33. Mandate future review for potential changes to this policy or law.

Detailed Information
Guidelines for Open Data Policies

Open Data policies can...

1. Mandate open formats for government data.

The utility, quality, and permanence of information depends on the format in which it's published. "Open" formats are considered best practice by technology and transparency communities because of their versatility: To quote Josh Tauberer[1], open formats "tend to promote a wide range of uses, backward and forward compatibility, and an independence from short-term commercial interests". In other words, these formats are machine-readable (structured), serve searchable, sortable data, and tend to be non-proprietary and/or implemented in open source software. When combined with appropriate methods of distribution, these traits maximize the degree of access, use, and quality of published information. This degree of access and interaction allows citizens and government alike to get the most out of the data.

Specific open data formats include JSON, CSV, and XML (for datasets), and HTML and plain text (which are only semi-structured, but can provide more flexibility for documents). The Open States Project[2] has explored how these formats relate to legislative data in more detail here[3]. More details about file formats and open data best practices can be found in the Open Knowledge Foundation's Open Data Handbook[4], Josh Tauberer's Open Data is Civic Capital[5], the 8 Open Government Data Principles[6], the 10 Open Government Data Principles[7], and The Power of Information report[8].

Open format provisions can be broad or specific in scope. More broadly defined provisions are generally hard to enforce, but can still be helpful as statements of general policy. Provisions that use more specific wording (e.g. those that define both specific datasets and the formats that they'll be released in) are more likely to cause meaningful change but take more effort to craft.

It should be noted that in this context, "data" refers broadly to any information published in electronic formats. This definition refers to a variety of resources, including databases, analytics, documents, transcripts, and audio and video recordings. Although each of these examples represent different kinds of data, all can be published in an open format.

---

[1] Open Data is Civic Capital: Best Practices for "Open Government Data, Josh Tauberer, http://razor.occams.info/pubdocs/opendataciviccapital.html#bestprac

[2] Open States Project, http://openstates.org/

[3] Open States Project Best Practices Wiki, http://code.google.com/p/openstates/wiki/StateBestPractices

[4] Open Data Handbook File Formats Overview, http://opendatahandbook.org/en/appendices/file-formats.html

[5] Open Data is Civic Capital, Josh Tauberer, http://razor.occams.info/pubdocs/opendataciviccapital.html#format

[6] 8 Open Government Data Principles, Public.Resource.Org, https://public.resource.org/8_principles.html

[7] 10 Principles for Opening Up Government Information, Sunlight Foundation, http://sunlightfoundation.com/policy/documents/ten-open-data-principles/

[8] The Power of Information, Ed Mayo and Tom Steinberg, http://www.opsi.gov.uk/advice/poi/power-of-information-review.pdf

④

2. Mandate the release of specific new government information.

Open data laws provide an opportunity not just to update and improve access to information that is already open and/or public, but to specify new datasets and records to be published. Open data policies can create specific mandates about a variety of kinds of data: information ranging from transportation data to lobbying registration databases to the video and audio of public meetings are all fair game (see Provision 1). Careful consideration should be given to the language used to describe what data is affected. Phrases to such as "high-value" or "high priority", when used without direction or indication of how to assign value/priority, can open up loopholes that prevent or slow the release of information desired by the public. It is important, therefore, that the scope of this provision be clearly defined: as with other provisions listed here, the scope of this provision can be broad or narrow, but in the final bill or order, the scope should be explicitly defined, the limitations noted, and the key agencies, committees or other relevant agents are identified. Similarly, policies should be specific about what "new" data can mean: Some policies require new data to be created, collected, and released for the first time, whereas other identify existing datasets and (newly) mandate their release.
Other provisions noted on this page address how to bring the public into the process of determining how datasets can be prioritized for release.

3. Mandate electronic filing.

Many existing disclosure requirements were created as inefficient, paper-based requirements and should be updated to require electronic filing, as long as the filers can be reasonably expected to have access to the necessary technology. Electronic filing requirements save money, make real-time disclosure possible, and allow structured data to be created, while paper filings make reuse and analysis more difficult.

Electronic filing is currently required in various places across the United States Federal Government, including the Federal Election Commission[9] -- where House candidates are required to file disclosures electronically, and Senate candidates are notably exempt[10]. Similar requirements can be found throughout the states: For example, in 2012 Delaware passed a bill (SB 185)[11] mandating that all lobbyist registration and disclosure be filled electronically by default.

4. Require any public information to be posted on the Internet.

The government makes tremendous amounts of information available to the public, but only a small subset is available on the Internet even as more people look online first to find these records. To close this gap, public information should be published online in a timely fashion subject only to common-sense exceptions (such as redacting personally identifiable information in certain contexts.) Implementation could be by legislative action or executive directive. The "Public Online Information Act" has been introduced on the federal level[12] to require public information to be available online.

---

[9] Electronic Filing, Federal Election Commission, http://www.fec.gov/elecfil/electron.shtml
[10] Sunlight on Electronic Filing (E-Filing), Sunlight Foundation, http://sunlightfoundation.com/about/issues/e-filing/
[11] Delaware Senate Bill 185, http://legis.delaware.gov/LIS/lis146.nsf/vwLegislation/SB+185/$file/legis.html?open
[12] The Public Online Information Act, Sunlight Foundation, http://sunlightfoundation.com/policy/poia/

5. Mandate continuous publication and updates to data.

It is not enough to mandate the one-time release of information: data is often created on an ongoing basis and should be released the same way. A one-time release of data is in some sense incomplete as soon as additional information is generated. Therefore, in order to ensure that the information published is as accurate and useful as possible, specific requirements should be put in place to make sure that government data is released as quickly as it is gathered and collected (in "real time"). This kind of rapid publishing becomes less of a burden when combined with others measures for online publishing, such as electronic filing (Provision 3), data portals (Provision 13), and APIs (Provision 8).

6. Create permanent, lasting access to government data.

Information released by the government should be sticky: Once released, it must remain "findable" at a stable location or through archives in perpetuity. Although portals and websites can be vehicles for accessing this data over the long term (see Provision 13), it is critical that the data's permanent release & accessibility is defined so as to apply to the data itself, not just the means of access.

Provisions relating to permanence can also be expanded to relate to updates, changes, or other alterations to the data. For best use by the public, these changes should be documented to include appropriate version-tracking and archiving over time. These provisions should build on the strengths of existing records management laws and procedures.

7. Publish bulk data.

Bulk access is a simple, but effective means of publishing datasets in full, giving the public the ability to download all the information stored in a database at once. Bulk downloads are often the simplest, most direct way of maximizing reuse and analysis of a dataset. Although they aren't absolutely necessary for the release of bulk data, data portals can be helpful indexes of specific sources for bulk data downloads.

Daniel Schuman and Eric Mill explore bulk data further as it relates to legislative data and the federal THOMAS system in this blog post[13].

8. Create public APIs (Application Programming Interfaces) for accessing information.

Government bodies can develop APIs, or Application Programming Interfaces, that allow third parties to automatically search, retrieve, or submit information directly from databases online (See Josh Tauberer's Open Data is Civic Capital[14]). Navigating requirements for bulk data and APIs should be done in consultation with people with technical expertise and also likely users of the information. For a lengthier discussion of the benefits of APIs, see the recently developed Federal

---

[13] Appropriators May Undercut Legislative Transparency, Daniel Schuman and Eric Mill, Sunlight Foundation, http://sunlightfoundation.com/blog/2012/05/30/appropriators-may-undercut-legislative-transparency/
[14] Open Data Is Civic Capital, Josh Tauberer, http://razor.occams.info/pubdocs/ opendataciviccapital.html#bestprac

Web Policy[15], and for a slightly more critical take on APIs (and their relationship to bulk data), see this post by Eric Mill[16] of the Sunlight Foundation.

9. Remove restrictions for accessing government information.

Open data that is out of reach of the public is hardly open. To provide truly open access, you must provide both the right to reuse government information (explored in Provision 6) and remove arbitrary technical restrictions, such as registration requirements, access fees, and usage limitations, among others. Whether these technical restrictions have been specifically put in place (i.e. access fees) or are the accidental result of the choice of data format or software (i.e. usage limits), it is appropriate for an open data policy to address and remove these barriers to access. The aim should be to provide broad, non-discriminatory access so that any person can access the data at any time without having to identify him/herself or to provide any justification for doing so. More detailed exploration of these limitations can be found in Josh Tauberer's book Open Government Data[17].

10. Remove restrictions on reuse of information.

Most government restrictions on the reuse of government information serve no purpose but to restrict the public value of important information. If information is to be truly public, there should be no license-related barrier to the public's interaction with or use of that information.

Outside of data legally exempted from public use or access because of privacy or security restrictions (see Provisions 11 and 12), to be completely "open," public government information should be released completely into the public domain and clearly labeled as such. At a minimum, licenses that grant the right to use, download, and reproduce government data can be applied, and any restrictions should only be created to serve the public interest. The fewer restrictions the better. Opening data into the public domain (or at least into free public use) removes arbitrary barriers to information access (further explored in Provision 9), helps disseminate knowledge, aids in data preservation, promotes civic engagement and entrepreneurial activity, and extends the longevity of the technological investments used to open information in the first place.

11. Appropriately safeguard sensitive information.

Open data policy should be complementary to pre-existing legislation and directives about access to public information (see Provision 30), which means taking into consideration pre-existing protections for sensitive information for privacy, security, or other reasons. While these protections should be upheld, careful thought should be given to the language used to describe what (if any) additional information will be exempt from the policy, as overbroad terms can create loopholes that undermine the soundness of provisions requiring openness.

---

[15] Digital Government, Building a 21st Century Platform to Better Serve the American People, The White House, http://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government.html
[16] Government: Do You Really Need an API? Eric Mill, Sunlight Foundation, http://sunlightlabs.com/blog/2012/government-do-you-really-need-an-api/
[17] Open Government Data, Josh Tauberer, http://opengovdata.io/2012-02/page/5-1-3/principles-universality-use

12. Require exemptions to open data policy to be balance-tested against the public interest.

Exemptions to disclosure are a necessary component of many transparency requirements. Unfortunately, these exemptions are often crafted as blanket categories for entire types of information, without consideration for competing interests. Valid privacy and security concerns should be addressed through provisions that recognize the public interest in determining whether information will be disclosed or not. For example, rather than saying "information relating to X topic are exempt from disclosure", provisions should require that "information relating to X topic are exempt from disclosure if the potential for harm outweighs the public interest their disclosure." Public interest here does not mean public attention, but instead refers to interests like democratic accountability, justice, and effective oversight.

13. Create a portal or websites devoted to specific issues related to data publication or specific policy arenas.

Data portals and similar websites can facilitate the distribution of open data by providing an easy-to-access, searchable hub for multiple datasets. At their best, these portals or hubs promote interaction with and reuse of open data (see the note about bulk data in Provision 7) and provide documentation for the use of information (see Provision 8). Portals can be generalized (such as Data.gov[18], "open data portal") or specific (e.g. a spending or ethics portal), and can vary in terms of their sophistication.

Portals and other related websites also provide governments with the opportunity to go into detail about issues and policies related to its commitment to openness and transparency. To facilitate their "findability" these websites should be allowed to be indexed and searched by third parties (such as search engines).

Some examples of websites created through different kinds of open data policies include:
- San Francisco's Data Portal: https://data.sfgov.org/[19]
- Austin's Open Data and Open Government Hub: http://austintexas.gov/austingo2.0[20]
- Missouri's Accountability Portal: http://mapyourtaxes.mo.gov/MAP/Portal/Default.aspx[21]
- NASA's Open Government Initiative Plan: http://www.nasa.gov/open/[22]
- USASpending.gov: http://usaspending.gov/[23]
- Recovery.gov: http://recovery.gov[24]

14. Create or explore potential public/private partnerships.

Partnerships can be useful in the effort to increase awareness of the availability of open data and in

---

[18] Data.gov, Federal Government, http://www.data.gov/

[19] Data.SFgov.org, San Francisco, https://data.sfgov.org/

[20] Austingo2.0, Austin http://austintexas.gov/austingo2.0

[21] Missouri Accountability, http://mapyourtaxes.mo.gov/MAP/Portal/Default.aspx

[22] NASA, http://www.nasa.gov/open/

[23] Federal Spending Transparency, http://usaspending.gov/

[24] Recovery Act Spending Transparency, http://www.recovery.gov/Pages/default.aspx

connecting government information to that held by non-profits, think tanks, academic institutions and others. Ed Mayo and Tom Steinberg[25] have noted that such partnerships can aid civic participation and help identify the gaps in services delivery, among other benefits. Poorly planned public/private partnerships run the risk of subsidizing private sector actors at the expense of the public. (For example, see the Government Accountability Office's digitization project described here[26].)

15. Create contests or other events focused on the use of government data.

Like partnerships (Provision 14), contests and events (both held in real space and online) are an effective mechanism for generating use of, interest in, and attention to the government's open data resources and repositories. Further, hosting events and facilitating both creative and practical uses of government data can help spur civic innovation and build communities around information. Events can range from barcamps, hackathons, and apps contests to town halls, webinars, and public hearings, with both technical and non-technical communities in mind. Outside of structured, government-run events, participation in developer communities through listservs and meetups could also be explored. (For example, see the Health Data Initiative Forum[27], the Apps for Democracy contest[28], or the Illinois Reform Commission's listening sessions[29].)

16. Require digitization and distribution of archival materials.

Open data policies can address not only information currently or soon to be available in an electronic format, but also undigitized archival material. See for example Vancouver's Open Data motion[30], which critically notes not only the importance of thoughtful digitization of archival information but the imperative to release this data to the public, ideally eventually in the same formats and in the same locations as modern data.

17. Create processes to ensure data quality.

Data quality will not be ensured through data release alone: efforts need to be made to keep the data up to date, clean, accurate, and accessible. In the executive memorandum[31] that established that Washington, DC would share internal data on DC.gov[32], the city specified not only the need to maintain data quality but, broadly, the processes required to do so and responsibilities of the agencies involved. Other approaches to ensuring data quality include assigning specific staff responsible for maintenance (see the Open Government Directive on financial data[33] (Section 2.a.))

---

[25] The Power of Information, Ed Mayo and Tom Steinberg, http://www.opsi.gov.uk/advice/poi/power-of-information-review.pdf

[26] GAO did sell exclusive access to legislative history to Thomson West, Free Government Information, http://freegovinfo.info/node/1798

[27] The Health Data Initiative Forum, http://www.hdiforum.org/

[28] Apps for Democracy, http://www.appsfordemocracy.org/

[29] Illinois Reform Commission, Listening Sessions, http://www.reformillinoisnow.org/townhall.html

[30] Open Data Motion, Vancouver, Canada, http://vancouver.ca/ctyclerk/cclerk/20090519/documents/motionb2.pdf

[31] Executive Memorandum on Data Sharing, Washington, DC, http://www.scribd.com/fullscreen/26442622?access_key=key-20rfsh26eu0ob66xlbmu

[32] DC.gov, http://dc.gov/DC/

[33] Open Government Directive, Section on Financial Data, http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf

and creating other audit processes.

In any case, data quality concerns should not be accepted as an excuse for exempting or restricting the release of information, but a challenge that becomes clearer and easier to address when data is released. Data with serious accuracy and quality concerns should be adequately documented to avoid creating confusion or misinformation.

Similarly, public data reporting streams that are separate from what is used within government should be avoided whenever possible, as redundant or parallel data streams can create opportunities for data quality to suffer.

18. Create a public, comprehensive list of all information holdings.

Government bodies often do not know what information they have. Open data policies should require a full public listing of government information. This comprehensive listing empowers policymakers and administrators to determine whether information is being appropriately managed and empowers the public oversight of those determinations. Publicly accounting for agency information helps ensure that information is managed to benefit the public interest, can create efficiencies among government departments, and empower journalists and policymakers. To provide up-to-date information, agencies can also be required to regularly audit their information holdings.

In an Open Data White Paper[34] released in June 2012, the UK Ministers of State for the Cabinet Office and Paymaster General noted among a list of open data strategies and principles that "Public bodies should maintain and publish inventories of their data holdings"(Section 2.46, Principle 13). The Obama Executive Memo on Regulatory Compliance Data[35] and, in particular, the Department of Transportation's index of major datasets ("Regulatory Enforcement and Compliance Data") are other examples of this provision in action. For more details, see also this blog post from John Wonderlich[36] explaining the need for indexes.

19. Mandate the use of unique identifiers.

Unique identifiers within datasets empower analysis and reuse by allowing disparate datasets to be combined and to be more carefully mapped to real-world entities. Without unique identifiers, some analysis can become difficult or impossible, since similar names may or may not refer to the same entities. Importantly, identifiers should be non-proprietary and public. A typical example of where unique identifiers are often required is found in lobbying disclosure[37]. For more information, see also this list of extensive resources[38] about the need for unique identifiers for corporate entities.

20. Require the publishing of metadata or other documentation.

---

[34] UK Open Data White Paper, http://www.cabinetoffice.gov.uk/resource-library/open-data-white-paper-unleashing-potential
[35] Presidential Memoranda on Regulatory Compliance, http://www.whitehouse.gov/the-press-office/2011/01/18/presidential-memoranda-regulatory-compliance
[36] The Missing Open Data Policy, John Wonderlich, Sunlight Foundation, http://sunlightfoundation.com/blog/2012/03/22/the-missing-open-data-policy/
[37] A State by State Look at Lobbyist Disclosure, Eric Dunn, Sunlight Foundation, http://sunlightfoundation.com/blog/2011/07/19/a-state-by-state-look-at-lobbyist-disclosure/
[38] Six Degrees of Corporations, Sunlight Foundation, http://sunlightfoundation.com/sixdegrees/resources/

Metadata and other documentation about the data provided by the government can be useful to the public and government alike. Notations such as these add helpful context about the data's creation that will aid in the public's use of that information and support current and future archival and data quality efforts. The Open Data White Paper[39] released by the UK's Ministers of State for the Cabinet Office and Paymaster General in June 2012 notes that the UK data portal (www.data.gov.uk[40]) already includes "basic metadata about all its datasets, including timing and geographical scope" as well as "a link to a departmentally supplied description of the data and details of a contact point within the department who data users can ask for further details" (Section 2.46, Principle 14[41]).

21. Require the publishing of code.

Not only the data, but the code used to create government websites, portals, tools, and other online resources can provide internal and external benefits, often as valuable open data itself. Governments should employ open source solutions whenever possible to enable sharing and make the most out of the opportunities provided. The Consumer Financial Protection Bureau (CFPB) began publishing open code on the social code site GitHub in 2012, citing that doing so helped them fulfill the mission of their agency and facilitated their technical work. (More information is available in the announcement blogpost on the CFPB's website[42].)

22. Set appropriately ambitious timelines for implementation.

Setting a clear deadline can demonstrate the strength of a commitment and can help to translate these commitments into results. They can also help to identify failures clearly, opening the door to public oversight. Relevant actors should be given enough time to prepare for the changes brought on by the new open data policy, but not so much time that the policy becomes inoperable. The timeline should be firm, provide motivation for action, and have actionable goals that can be used as a metric for compliance. These goals or checkpoints can include qualitative and quantitative measurements.

23. Ensure sufficient funding for implementation.

Like any other initiative, implementing an open data policy should be done with an eye on long-term sustainability. One way to do this is to consider funding sources for the implementation of the policy as well as its future maintenance. Sufficient funding can mean the difference between successful and unsuccessful policies.

For example, in 2011, the Electronic Government Fund, which supports Data.gov, the IT Spending Dashboard, and USASpending.gov, among other programs, was sliced from over $34 million to $12.4 million. Without the work of the advocacy community, funds would have dropped as low as $8 million. This dramatic change in funding has continued implications for federal data, some of

---

[39] UK Open Data White Paper, http://www.cabinetoffice.gov.uk/resource-library/open-data-white-paper-unleashing-potential

[40] Data.Gov.UK, http://www.data.gov.uk/

[41] UK Open Data White Paper, http://www.cabinetoffice.gov.uk/resource-library/open-data-white-paper-unleashing-potential

[42] Consumer Financial Protection Bureau Source Code Policy, http://www.consumerfinance.gov/blog/the-cfpbs-source-code-policy-open-and-shared/

which can be explored in this tag on the Sunlight Foundation blog[43]. By contrast, in 2012, California approved a bill (SB 1001)[44] to pay for maintenance, repair, and improvements to their Cal-Access public disclosure database by increasing the registration fees for those engaged in lobbying and with political action committees.

24. Empower the creation of binding regulations to implement the new policy.

While some questions may defy easy treatment in the process of creating an open data policy, specific officials can be appropriately empowered to create regulations or guidance to ensure a strong, reliable policy. For example, in the proposed Public Online Information Act[45], a central regulator is empowered to create and set data standards. Similarly, the Dodd-Frank Financial Reform bill empowers regulations to require public reporting of royalty payments made by the extractive industry (see Section 1504)[46]. A similar approach is taken in the proposed DATA Act[47].

25. Tie contract awards to transparency requirements for new systems.

Existing procurement, contracting, or planning processes can be used to create new defaults and requirements for IT systems and databases -- to bake open data requirements into new systems being planned. See for example the White House's Digital Government strategy[48], which proposes the creation of similar new requirements.

26. Stipulate that provisions apply to contractors or quasi-governmental agencies handling public data.

The government often uses third party entities or contractors to handle, research, or generate government information, and the use of outside services should not necessitate sacrificing important public protections.

Similarly, these public protections should generally apply to quasi-governmental agencies and other similar actors, such as multi-state agencies, government-sponsored entities, publicly-funded universities, and self-regulatory organizations (like FINRA[49]).

27. Create new oversight authority to review implementation of the requirements.

Whether through a central authority, review board, or other similar body, creating an oversight

---

[43] SaveTheData Tag, Sunlight Foundation, http://sunlightfoundation.com/blog/taxonomy/term/savethedata/
[44] Yee's Cal-Access Bill Reconsidered, Approved Along With Other Open Government Bills, http://sd08.senate.ca.gov/news/2012-05-31-yee-s-cal-access-bill-reconsidered-approved-along-other-open-government-bills
[45] Public Online Information Act, Sunlight Foundation, http://sunlightfoundation.com/policy/poia/
[46] H.R. 4173 The Dodd-Frank Wall Street Reform and Consumer Protection Act, 111th Congress, http://www.govtrack.us/congress/bills/111/hr4173/text
[47] H.R. 2146 The DATA Act, 112th Congres, http://www.opencongress.org/bill/112-h2146/show
[48] Digital Government, Building a 21st Century Platform to Better Serve the American People, The White House, http://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government.html
[49] FINRA: Open Data not on the table, Nancy Watzman, Sunlight Foundation Reporting Group, http://reporting.sunlightfoundation.com/2012/Finra_open_data/

authority helps ensure compliance with new open data measures. New oversight bodies should conduct their work independently and publicly. See, for example, the Recovery Accountability and Transparency Board[50] established as part of the Recovery Act.

28. Create new legal rights or other legal mechanisms to empower the public.

An open data policy can create mechanisms that allow individual members of the public to play a dynamic role in policy oversight and compliance. For example, the right to sue serves as the ultimate enforcement mechanism of the Freedom of Information Act, some countries (like Canada)[51] have FOI ombudsmen with special legal enforcement powers, and some countries also have special anti-corruption agencies.

29. Appeal to values and goals, such as accountability, efficiency, employment and commerce, innovation, civic engagement, and public services provision.

Publishing open data has many practical and normative implications which can be noted and explored in the text of the open data policy. These values and goals can be noted for the record as part of the policy.

30. Reference and build on existing public accountability policies, like
- **Freedom of Information Laws**
- **Open Meetings Acts**
- **Open Records Acts**
- **Ethics Protections**
- **Campaign Finance**
- **Lobbying Disclosure Laws**

Open data policies should be informed by provisions already on the books, building on precedent for opening information and taking advantage of pre-existing laws, executive orders, and other policies that defend and establish public access, define standards for information quality, disclosure, and publishing.

31. Incorporate public perspectives into policy implementation

Implementing the details of an open data policy will benefit from public participation, especially since open data policies can have effects government-wide and also have consequences for a variety of different stakeholder groups. Formal mechanisms for collaboration can include public hearings, draft proposals, and online resources, like wikis and email lists. For example, in 2012, New York City created a wiki[52] to encourage collaborative input on the open data policies, standards, and guidelines that would be enacted as part of its then-newly passed open data law.

---

[50] The Recovery Accountability and Transparency Board, http://www.recovery.gov/About/board/Pages/TheBoard.aspx
[51] A Citizen's Guide on Using the Freedom of Information Act and the Privacy Act of 1974 to Request Government Records, http://www.gpo.gov/fdsys/pkg/CRPT-106hrpt50/pdf/CRPT-106hrpt50.pdf
[52] New York Open Data Wiki, http://nycopendata.pediacities.com/wiki/index.php/NYC_Open_Data

13

32. Require analytics about the use of open data to be published publicly.

Statistics about the use of and interaction with government data can be mandated as part of an open data policy and can strengthen the goals of the policy. For example, the New York State Senate publishes their monthly website analytics as part of their open data portal (See, for example "NY Senate Web Presence May 2012"[53]).

33. Mandate future review for potential changes to this policy or law.

Just as publishing open data is an ongoing process that requires attention to its quality and upkeep (Provision 5), so too does the policy that establishes it. In order to keep up with the times, current best practices, and feedback from existing policy oversight, open data policies should be written in a way that makes them open to future revision. Open data policies should acknowledge that the context in which they operate is rapidly changing over time, and will likely need sustained attention to remain relevant.

---

[53] New York Senate Web Presence Analytics May 2012, http://www.nysenate.gov/report/ny-senate-web-presence-analytics-may-2012