



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON, D.C. 20503

September 4, 2009

MEMORANDUM FOR AGENCY CHIEF INFORMATION OFFICERS

FROM: Vivek Kundra VK
US Chief Information Officer

SUBJECT: Evolution and Growth of Data.gov

Data.gov allows the public to easily find, download, and use datasets and data tools that are generated and managed by the Federal Government. As a priority Administration initiative, the vision for Data.gov was encapsulated in the President's January 21, 2009 Open Government and Transparency memorandum where he states that information should be disclosed "rapidly in forms that the public can readily find and use."

Since the May 21, 2009, launch of Data.gov, we have seen an exceptional leap forward in the visibility of, and focus on, high quality, policy-compliant data. Data.gov generates over 150,000 hits per day and drives traffic to agency websites to obtain the datasets and more detailed information about the datasets. Thanks to the supportive participation of numerous Federal Executive agencies, as of August 25, 2009, Data.gov includes 579 datasets in the Raw Data Catalog; 208 feeds, widgets, and tools in the Tools Catalog; and over 110,000 entries in the Geodata Catalog. Continued expansion of datasets and tools on Data.gov is dependent on active and continuous participation by all Federal agencies.

With respect to Data.gov, agencies have the following responsibilities:

1. Create and manage their own datasets and tools,
2. Document metadata for datasets and tools, leveraging the Data.gov metadata template,
3. Understand the information products disseminated by the agency to the public,
4. Determine which datasets and tools are authoritative and appropriate for exposure through Data.gov, and
5. Ensure authoritative data are available in appropriate formats that are machine readable.

I am asking each agency Chief Information Officer to review the responsibilities above and, in cooperation with the agency's Data.gov Point of Contact (POC) (please see attachment A for the current POC roster), ensure that appropriate datasets and tools that are authoritative, meet information guidelines, and are suitable for public disclosure are submitted for inclusion on Data.gov.

In an effort to simplify the dataset and tool submission process, we are releasing the Data.gov Dataset Management System (DMS). The DMS is an online capability that will allow your data stewards to submit datasets and tools for inclusion on Data.gov through your Data.gov POC, maintain the associated metadata, and track the progress of submissions from your agency (with a dashboard). For more information on the DMS, please refer to your agency's Data.gov POC; or email Marion Royal at marion.royal@gsa.gov.

Lastly, I am distributing a draft of the Data.gov Concept of Operations for your review and comment. The Concept of Operations is an overview document that describes the intent, vision, strategies, solutions, and path forward for Data.gov. The Concept of Operations is designed to provide an operational overview of Data.gov as appropriate for a Chief Information Officer, an agency Data.gov POC, or a Data Steward. The Concept of Operations includes some future looking visioning as well as an operational overview of the current Data.gov environment. The U.S. Department of the Interior and Environmental Protection Agency CIO's serve as co-leads for the Data.gov Program for the Federal CIO Council with a Government-wide team including your agency's Data.gov POC. Together, we have successfully implemented the first phase of the initiative in Fiscal Year 2009 and are anxious to implement the next phase of the program (FY 2010 and beyond) based on a well-understood, agreed-upon Concept of Operations.

Accordingly, I am asking that you review and provide feedback on the Concept of Operations draft by September 18, 2009. To aid your review, please see Attachment B for a list of questions on which we are particularly interested in receiving comments. Please email feedback or questions directly to Kshemendra Paul, Federal Enterprise Architect, at kpaul@omb.eop.gov with a copy to feapmo@omb.eop.gov.

Thank you for your continuing support of this singular Presidential initiative.

Attachments

Attachment B: Questions to Consider When Reviewing the Data.gov Concept of Operations Draft

Strategic Vision

In section 1 “Data.gov Strategic Intent”, the Concept of Operations lays out a strategic vision for Data.gov. We believe that Data.gov has transformational potential for Federal data dissemination. We welcome specific suggestions for transformational initiatives and how such initiatives could be achieved through Data.gov.

Performance Measures

In section 1.4 “Measuring Success”, Data.gov aims to help agencies increase their compliance with Federal information dissemination policies. We seek comments on how best to measure the success of Data.gov, possible ways to benchmark implementation by agencies, and methods to measure how Federal dataset dissemination has been/could be improved by Data.gov.

Key Partnerships

In section 2.4 “Who Makes Data.gov Work”, the roles and responsibilities of the Chief Information Officer (CIO) and the CIO appointed agency “point of contact” (POC) are outlined as they relate to the implementation of Data.gov. We are interested in your thoughts about how to organize this effort within agencies and how to energize support from the program offices that supply the data.

Automating Dataset Submission

Our goal is to leverage technology as much as possible to get datasets available on the site as quickly as possible, while maintaining the agency’s role as the “authoritative source” and establishing a process to collect metadata on the various datasets. We seek your support for this model and whether it is likely to ease the workload of providing completed metadata templates for Data.gov. We are interested in comments regarding how such an automated process as that described in the Concept of Operations can be integrated into agency processes.

Modernization

In section 3.2 “Modernizing Agency Data and Metadata Publishing”, the Concept of Operations presents a “publish and subscribe” approach for providing metadata to Data.gov. We are interested in agencies’ ideas concerning the resource implications of this effort as well as how we can develop an even more streamlined and automated approach for harvesting data for Data.gov. Finally, we are interested in how agencies view the potential of incorporating the semantic web into their data dissemination strategies.

General

Are there any other general comments on the Data.gov Concept of Operations or how best to disseminate Federal Government datasets?