



White House Open Data Policy (ODP) and EPA's DRAFT Environmental Information Management Policy (EIMP)

Overview & Discussion

Joint ENLC and NTB Meeting Agenda

June 19-20, 2013 (1:15 – 3:45)

Churchill Hotel

1914 Connecticut Ave.

Washington, D.C. 20009

Room: Kalorama Room

- White House Open Data Policy (ODP) Purpose
 - ODP Requirements
 - Generally, Open Data is...
 - ODP Timeline: Agency Deadlines
 - Additional Agency Responsibilities for ODP
- EPA's Environmental Information Management Policy (EIMP) Purpose
 - Crosswalk of ODP Requirements & EIMP
- EPA's Current Minimum Metadata
- Evaluation of ODP Progress
- Open Discussion



Open Data Policy (ODP) Purpose

- Executive Order 13642 – “Making Open and Machine Readable the New Default for Government Information”: released May 9, 2013 <http://www.gpo.gov/fdsys/pkg/FR-2013-05-14/pdf/2013-11533.pdf> establishes the Open Data Policy (ODP)
- The purpose of the ODP is to:
 - Establish a framework to institutionalize the principles of effective information management at each stage of the information's life cycle
 - Promote interoperability and openness by applying this framework to all information
 - Mandate that agencies build or modernize information systems in a way that:
 - Maximizes interoperability and information accessibility
 - Maintains internal and external data asset inventories
 - Enhances information safeguards
 - Clarifies information management responsibilities

1. Support downstream information processing and dissemination activities:
 - a) Use machine-readable and open formats when information is collected
 - b) Use data standards
 - c) Ensure information stewardship via use of open licenses
 - d) Use common core and extensible metadata
2. Build information systems to support interoperability and information accessibility
 - a) Design must be scalable, flexible, and facilitate data extraction in >1 format and for >1 use
 - b) Document data schemas and dictionaries
3. Strengthen data management and release practices
 - a) Create/maintain an enterprise data inventory
 - b) Create/maintain a public data listing
 - c) Create a process to engage with customers to facilitate data release
 - d) Clarify agency roles and responsibilities
4. Strengthen privacy and confidentiality measures through risk-based analysis and consideration of the 'mosaic effect'
5. Incorporate these requirements into IRM strategic plan



Generally, Open Data is...

- **Public.** Consistent with OMB's Open Government Directive, agencies must adopt a presumption in favor of openness to the extent permitted by law and subject to privacy, confidentiality, security, or other valid restrictions.
- **Accessible.** Open data are made available in convenient, modifiable, and open formats that can be retrieved, downloaded, indexed, and searched.
- **Described.** Open data are described fully so that consumers of the data have sufficient information to understand their strengths, weaknesses, analytical limitations, security requirements, as well as how to process them.
- **Reusable.** Open data are made available under an open license that places no restrictions on their use.
- **Complete.** Open data are published in primary forms (i.e., as collected at the source), with the finest possible level of granularity that is practicable.
- **Timely.** Open data are made available as quickly as necessary to preserve the value of the data.
- **Managed Post-Release.** A point of contact is provided.



ODP Timeline: Agency Deadlines

- After August 7, 2013 (90 days) EPA must:
 - Develop metrics and milestones to meet the Cross-Agency Priority (CAP) Goal the Chief Performance Officer (CPO) will set in 90 days (by August 7, 2013)
- By November 5, 2013 (180 days) EPA must:
 - Integrate ODP into policies, regulations, guidances, & operational activities
 - Create and maintain a publically available enterprise data inventory
 - Develop process for publically releasing additional data sets
 - Update the Agency IRM Plan
 - Submit a progress report to OMB
- Beginning FY2014/Q2, EPA must report progress quarterly



Additional Agency Responsibilities for ODP

- Conduct analysis of privacy, confidentiality, and security risks into each stage of the information lifecycle to identify information that should not be released (Overseen by Senior Agency Official for Privacy (SAOP))
- Participate in interagency working group to support information interoperability and openness
- Evaluate current processes and identify implementation opportunities that will result in more efficient use of resources and taxpayer dollars

- Progress will be assessed by OMB and the public through:
 - Analysis of the agency's updates to Information Resource Management (IRM) strategic plans (part III, section 5)
 - Completeness of the enterprise data inventory (part III, section 3.a)
 - Data made available in the agency's public data listing (part III, section 3.b)
 - Initial and Quarterly progress reporting



EPA's DRAFT Environmental Information Management Policy (EIMP) Purpose

- EPA's Environmental Information Management Policy (EIMP) implements several information-related initiatives within EPA and across Federal, with the goal of making EPA's information readily discoverable, available, defined and useful to the public and all stakeholders
- The EIMP establishes a framework to ensure that all information more fully supports EPA and partner business processes by:
 - Requiring EPA information to be managed throughout its lifecycle
 - Addressing stakeholder requirements and fostering the availability of information that is appropriate and adequate for its intended use and reuse
 - Outlining processes for facilitating the creation, capture, discovery, access, sharing, reuse, integration and disposition of EPA information
 - Requiring clear documentation and availability for discovery, evaluation, and use in accordance with federal regulations and directives



Crosswalk of ODP Requirements & EIMP

Management Practice	Open Data Policy	EIMP
Addresses all information types	In definitions only (same info definition as in EIMP)-data focus	Yes
Lifecycle management	Yes	Yes
Machine readable	Yes	As appropriate/feasible
Open Formats as appropriate - documented and shared data dictionaries/schema	Yes	Yes in data related procedures/guidance
Data standards	Yes	Yes
Common core & extensible metadata	Yes	Yes
Maintain internal & external data asset Inventory	Yes	In Data Procedures only
Internal & external needs considered throughout lifecycle; including potential uses not in the original design	Yes	Yes
Ensure Privacy and confidentiality are fully protected	Yes (includes risk assessments)	Yes (risk assessment not addressed)
Incorporate new interoperability and openness requirements into core agency processes	Yes	Yes
Coordinate with the Records Management Officer	Yes	Yes



EPA's Current Minimum Metadata

- EPA Environmental Information Management Policy Minimum Fields - Principles
 - All information resources, structured (e.g., data) and unstructured (e.g., documents) warrant some minimum level of metadata.
 - Metadata should help users to find, evaluate and access the information resources that might meet their government business needs.
 - The Dublin Core offers a well established base standard from which to draw a set of minimum fields, definition and common practices.
 - EPA has some additional demands not met by the Dublin Core. These include the need to protect sensitive information and the proper management (i.e., retention periods) of all resources.
 - Some resources (e.g., datasets, web content) are subject to additional Agency, Federal or International standards (e.g., Open Data, ISO). Programs need to assure this minimum set is included with any Agency implementation of external standards.
- Recommended Minimum Metadata Fields
 - Dublin Core: Title, Creator (Contact), Date, Description, Type, and Subject (Keywords or Tags)
 - Recommended to support additional EPA needs: Retention Policy, Sensitivity, and Resource Location



Open Discussion

- Approach
- Next Steps