Open Data Exec Order & Policy: DHS Agency Actions

1. Create and maintain an Enterprise Data Inventory
2. Publish Enterprise Inventory Schedule and Scorecard update each quarter
3. Create and maintain a Public Data Catalog Listing and document if data cannot be released
4. Create a process for identifying and documenting high values datasets in the Enterprise Data Inventory
5. Create a process to engage with customers to help facilitate and prioritize data releases
6. Generate DHS Digital Strategy web page updates

(Enterprise Architecture Information Repository - Data) – DMWG (Components DM) and EDMO maintains data portion – 815 published data assets and growing
EDMO and Digital Government Team
EDMO extracts from EAIR and adds Open Data metadata and reviews any changes with DMWG
EDMO and Web Council work with Components to release new datasets (machine readable format) using OpenFEMA process
EDMO and Web Council work with Components to process requests using OpenFEMA process
EDMO and Digital Government Team
Create an Enterprise Data Inventory

• Ensure the Inventory contains one metadata record for each data asset at your agency.

• Include, at a minimum, all data assets which have been posted on Data.gov by August 1, 2013.

• Submit to OMB the inventory as a single JSON.

DHS Status: Enterprise Architecture Information Repository (EAIR) – 815 Data assets currently identified, focused on support to DHS systems inventory. Does not include all Information Product high value datasets generated by DHS systems.

DHS Status: Continue to track systems in the EAIR that lack data assets on the EDMO Scorecard. Engage with Components to identify high value datasets that should be in the EAIR.
Use the Enterprise Data Inventory to produce the Public Data Catalog

• Extract all dataset information from the EAIR
• Eliminate all datasets with PII, Law enforcement or National Security Sensitive or Trade Secret data (metadata indicators)

Convert DHS metadata mapping to Open Data Common Core metadata

• DHS EAIR metadata is more extensive than Open Data
• DHS currently lacks indicator in the EAIR for sharing with the Public
• Get final Component concurrence on datasets marked as “Public or Non-Public disclosure” - Nov 4

Publish on DHS Digital Strategy Page

• Public Data Catalog data set is converted to JSON for Open Data catalog sweep
Engage with DHS Collaborative Bodies (ITSGB, Web Council ESC, ASC, ISCC)

- Discuss new data dissemination requirements from OMB
- Present overview of process for identifying high value datasets at Components
- Seek approval and feedback
- Enterprise Data Management will track and report progress

Public Dissemination of DHS High Value Data

- All Public dissemination of DHS data must use machine readable formats (HTML, XML, XLS, RDF, etc) – evaluate existing web pages with Components for future use
- Public dissemination of DHS data should use open licenses and standards like Drupal, APIs

FEMA has developed a process for identifying and approving high value datasets

- Incorporates all governance and approval bodies (Web Managers, Privacy, General Council)
- Usable by any Component
- Presented to Web Council ESC and endorsed
Engaging with Customers to help facilitate and prioritize data releases

- Provide “Suggest a Dataset” submission link on DHS Digital Strategy, Developer and Data web pages
- Provide suggestions and feedback to Component POCs and web content managers
- Provide feedback to requestors on receipt of suggestions
- Work with requesting organizations to streamline data sharing where possible

Dissemination of Customer suggestions and feedback

- Track and report number of suggestions to Web Council ESC
- Track and report Component responses to Customer dataset suggestions
- Track and report new dataset additions to EAIR
- Engage with Components on keeping their data current on data.gov

Engage with DHS Collaborative Bodies (ITSGB, Web Council ESC, ASC, ISCC) on quarterly schedule

- Report progress
- Report issues
- Seek feedback
Roles and Responsibilities for Promoting Efficient and Effective Data Release

**Enterprise Data Management (EDMO) and DMWG**

- Assist with identifying and adding new datasets to EAIR
- Use EDMO scorecard to track new datasets added to EAIR by Components
- Provide scorecard results to Collaborative Bodies and Digital Government Team
- Provide quarterly inventory and public data catalog update to OMB

**Component Web Content Managers, Web Council**

- Assist with identifying and submitting new high value datasets to EDMO
- Assist with public dissemination of datasets through web pages approval process
- Assist with improving existing public web page publication of reports and statistics to include machine readable supporting tables

**Component Program Managers**

- Develop dissemination strategy for Programs as required by SELC
- Review existing web pages and datasets for machine readable format compliance
- Provide high value dataset approval documentation to Component web managers
- Include dissemination updates in investment document information sharing questions and Program Reviews

COMPLETE BY NOVEMBER 27, 2013
Overview of Digital Strategy changes
- DHS data asset inventory and inventory schedule overview – **Done**
- Public Data Catalog overview – **Delivered**
- High Value Data sets process overview – **Delivered**
- Customer Suggestion and Feedback process – **Delivered**
- Roles and Responsibilities – **Delivered**

New links for Digital Strategy page – **Published**
- Public Data Catalog [www.dhs.gov/data.json](http://www.dhs.gov/data.json)
- Inventory Schedule – continue to enhance EA Information Repository
- Open Data/Digital Government Scorecard – How is DHS doing?

Updates to Developer and Data pages – **Published**
- Additions to Developer page – waiting on Security approval to FEMA Document API (best practice for handling multiple dataset versions) – **In Progress (Testing and Refining)**
- Continuing additions to DHS Data page - USCIS FOIA logs, Immigration Statistics, FEMA datasets, USCG datasets – **Continuing**

Data.Gov web site has new functionality coming – **Delivered**
- [www.DHS.gov/data.json](http://www.DHS.gov/data.json) Agency public data catalogs will be automatically ingested by Data.Gov scanning process
- New public data catalogs produced on quarterly basis to Digital Strategy page
Executive Order 05092013 and M-13-13

- Requires a transition strategy for exposing more high value datasets
- Requires DHS to provide statistics on identifying high value data sets and making them available
- Requires DHS to develop a plan for exposing an Enterprise Data Catalog
- Work with Data Producers to produce datasets in machine readable format

Information Resource Management Strategic Plan

- Details on DHS approach to make more information accessible and identify high value data
- Policies, standards and processes to make information more accessible
- Focus on Mobile workforce support and information delivery through APIs
- Guidance on improvements to Mobile information presentation and security

Why High Value Data Sets and Must They be Identified?
What is High Value Data

- High-value data is information that can be used to increase agency accountability and responsiveness; improve public knowledge of the agency and its operations; further the core mission of the agency; create economic opportunity; or respond to need and demand as identified through public consultation – Open Government Policy.

DHS Perspective – Data is an Enterprise Asset

- Source Data considered to be key tactical inputs to developing data aggregations, dashboards, data services and common operating pictures
- Information Products that provide critical or situational updates to mission operators
- Information products that provide mission perspective to partners and the public

Public Perspective – Information on Web Pages or Data.Gov

- Data and information products that improve the understanding of various mission outcomes like immigration and border crossing statistics, international and domestic cargo statistics, boating safety statistics
- Data and information products that improve mission transparency by sharing with data.gov communities like Safety
Make open data the new default data standard.

- Machine readable and open formats (HTML, XML, XLS, RDF, etc.) - Open data standards
- Open licenses (DRUPAL, API)
- Government wide Data.gov metadata for publication standard

Collect or create information in a way that supports downstream information processing & dissemination

- Build information systems to support open data needs & maximize interoperability

Incorporate new interoperability and openness requirements into core agency processes like System Engineering Life Cycle (SELC) and resource management
How Are High Value Data Sets used at DHS?

**Information Dissemination**
- Web pages are traditional method of sharing with public (PDFs, search, XLS, HTML, XML)
- All web pages are approved by Component web managers and oversight offices (Privacy, Legal, Public Affairs)

**Discovery and Search**
- EAIR provides metadata descriptions and categorizations that business and technology can use to find the datasets they are looking for

**Segment Architecture Frameworks**
- Used for As-Is and To-Be target architecture surveys
- Used for investment decisions
- Information flows graphically depict the end-to-end information lifecycle of significant mission areas, i.e. screening, vetting, intelligence, etc.
What Support is needed for Open Data?

CIO Council
- Promote improvements to Web pages that share with the public (PDFs, search, XLS) to include machine readable content
- Require programs to provide dissemination strategies that clearly define how data can be made more transparent and what information can be shared with external partners and the public

Information Sharing and Safeguarding Governance Board
- Promote expanded information sharing by requiring programs to make information available using machine readable formats like the National Information Exchange Model (NIEM) as the default data exchange standard

Applications Services Council
- Require information to be treated as an enterprise asset and make information accessible and understandable the DHS information sharing default
- Make machine readable format and open data transparency the DHS defacto data standards for information sharing and exchanges (XML, RDF, HTML, JSON)
QUESTIONS & COMMENTS

CONTACT:
MICHAEL SIMCOCK, Chief Data Architect
DEPARTMENT OF HOMELAND SECURITY
Office of the Chief Information Officer | Enterprise Data Management Office
michael.simcock@hq.dhs.gov
202.447.3740