



# ENTERPRISE DATA PUBLICATION PROCESS

*Department of Housing and Urban Development*

*Office of the Chief Information Officer*

**PAGE INTENTIONALLY LEFT BLANK**

## Revision History

Date	Version	Summary
11/01/2013	1	Initial Schedule of HUD Enterprise Data Publication Process.
6/13/14	2	Update to incorporate Administrative and Statistical Data per OMB M-14-06
2/20/15	3	Updated to include OGC in the review process on page 6.

## INTRODUCTION

This data publication process documents how HUD identifies datasets for publication.

HUD currently has an incomplete Enterprise Inventory and it will need to engage each of its Program Offices to determine what each Office's data assets are and if the data can be made available to the public. For data that is permitted to be shared with the public, metadata will be created to describe what the data is and how it can best be used by the public. This Inventory Schedule will identify these tasks and set milestones and dates to when these activities will take place. How HUD will ensure that it has identified its data assets is detailed in the HUD Method for Identifying Data Assets section

### Purpose

To develop a clear and comprehensive understanding of what data assets HUD possess, we are required to create an Enterprise Data Inventory (Inventory) that accounts for all data assets created or collected by HUD. This includes, but is not limited to, data assets used in HUD's information systems. The Inventory must be enterprise-wide, accounting for data assets across programs and bureaus, and must use the required common core metadata.

OMB will assess agency progress toward overall maturity of the Enterprise Data Inventory through the maturity areas of "Expand," "Enrich," and "Open."

**Expand:** Expanding the inventory refers to adding additional data assets to the Inventory. HUD should develop their own strategy to expand the inventory and break down the work according to defined classes of data. HUD needs to communicate plans for expanding the Inventory in the Inventory Schedule.

**Enrich:** To improve the discoverability, management, and re-usability of data assets, HUD should enrich the Inventory over time by improving the quality of metadata describing each data asset.

**Open:** HUD should implement tools and processes that will accelerate the opening of additional valuable data assets by making them public and machine-readable, while ensuring adequate policy, process, and technical safeguards are in place to prevent against the release of sensitive data. Agencies are required to increase the number of public data assets included in the Public Data over time.

## Scope

By November 1, 2013 HUD is required to develop and submit to OMB an Inventory Schedule.

Below are the minimum requirements to meet:

- Describe how the agency will ensure that all data assets from each bureau and program in the agency have been identified and accounted for in the Inventory, to the extent practicable, no later than November 1, 2014.
- Describe how the agency plans to expand, enrich, and open their Inventory each quarter through November 1, 2014 at a minimum; include a summary and milestones in the schedule.
- Publish the initial Inventory Schedule on the [www.HUD.gov/digitalstrategy](http://www.HUD.gov/digitalstrategy) page by November 1, 2013 and continue to update it on a quarterly basis.

HUD is also required to maintain this Enterprise Data Inventory after the initial November 30, 2013 submission, continuing to expand, enrich, and open the Inventory on an on-going basis.

## HUD Process for Publishing Data Assets

This section describes the Department of Housing and Urban Development (HUD) process to identifying data assets for possible publication to meet the requirement of the Open Data Policy M-13-13, and the Guidance for Providing and Using Administrative Data for Statistical Purposes M-14-06. HUD realizes the importance of “administrative data”, which is data the government collects for programmatic and regulatory purposes. Administrative data helps aid in providing statistical benefits by giving the agency a means for measuring how well its programs are achieving the performance outcomes and goals. The OCIO works closely with its partners in the Office of Policy Development and Research (PD&R) and the Office of General Council (OGC) to identify administrative data for statistical purposes. The link and information below describe some of the key guidance material that HUD has developed to steer this process.

First, datasets are manually identified through multiple avenues:

1. The Enterprise Information Management Program staff will review the HUD Enterprise Information Management repository to identify datasets that were identified in past Data Quality Assessments, and Data Management Maturity Assessments; and look at System of Record Notices (SORN’s);

2. Members of our Data Steward Advisory Group (DSAG) will review HUD's Internet and Intranet for their respective Program Area. The DSAG Member will provide lists of datasets that are available, and discusses internally with their respective Program Area Team members about their availability to be provided to the public. Membership of the DSAG is composed of all HUD major business areas and the ten (10) Regional Offices;
3. The DSAG Member documents datasets that are going to be made available publicly and submits the datasets to the DSAG Chair, Co-Chair, and Chief Data Steward;
4. Once approved the DSAG Chief Data Steward shares the list of datasets with the HUD CISO and SAOP; and OGC's, John Shumway ([John.B.Shumway@hud.gov](mailto:John.B.Shumway@hud.gov)), Administrative Law Division.
5. The Chief Data Steward/HUD Datagov POC publishes the dataset to the Public HUD Dataset Listing at [www.HUD.gov/data.json](http://www.HUD.gov/data.json), and submits the dataset to the OMB MAX via a JSON inventory file.
6. Leveraging the Agency Suggestion spreadsheets provided by Data.gov to determine if the suggested data sets from Data.gov site visitors are viable candidates for publishing; and
7. Posting a "Nominate a dataset" link on our HUD Intranet (<http://hudweb.hud.gov/po/i/edm/news.cfm>)

Those datasets are then entered and tracked on our internal HUD Dataset Status Log. This log exists on our internal SharePoint. Once the dataset has been approved for publication and there are no privacy related issues, the information is then posted on our [data.hud.gov](http://data.hud.gov/) web page, in addition a JSON file is created and also uploaded to this page, and our digital strategy webpage [http://portal.hud.gov/hudportal/HUD?src=/Digital\\_Strategy](http://portal.hud.gov/hudportal/HUD?src=/Digital_Strategy), where OMB harvests those files for uploading to Data.gov.

If datasets cannot be published, they are documented in our Data Not Released List. This list is maintained by our Chief Data Steward in our DSAG SharePoint Site. In addition, they are listed in our Enterprise Data Inventory, which is uploaded to OMB MAX as updates are identified.