

OPEN DATA – MANAGING INFORMATION AS AN ASSET

1. **REASON FOR ISSUE:** This handbook establishes guidance regarding the Department of Veterans Affairs (VA) data asset collection efforts and procedures for managing information as an asset. Open Data promotes openness and interoperability while properly safeguarding information technology systems and information within the context of all applicable VA standards, during the implementation of the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act), Title II Open, Public, Electronic and Necessary (OPEN) Government Data Act, and Office of Management and Budget (OMB) guidelines. The policies in this directive apply to the entire Department but are non-binding guidelines for the Office of Inspector General.
2. **SUMMARY OF CONTENTS/MAJOR CHANGES:**
 - a. This is a new handbook that covers procedures to ensure that Administrations and Staff Offices effectively contribute to VA's need to publicly disseminate information. It is also intended to meet the Evidence Act requirements which call for each Federal agency to manage data as a strategic asset and make its data holdings available to the public without compromising privacy and security or VA standards. This handbook addresses guidance issued by OMB related to Open Data, which includes the quarterly submissions and other requirements for establishing, populating, and maintaining an enterprise-wide data asset inventory according to quarterly milestones.
 - b. This handbook codifies in policy the Department's Open Data Working Group (ODWG), and its relationship to the Data Governance Council.
 - c. This handbook sets forth the roles and responsibilities of VA's Chief Data Officer (CDO), Open Data Lead, ODWG, and the Open Data Liaisons and Data Stewards in Administrations and Staff Offices.
3. **RESPONSIBLE OFFICE:** Office of Enterprise Integration (OEI) (008), Office of Data Governance and Analytics (008B).
4. **RELATED DIRECTIVE:** VA Directive 0900, VA Enterprise Data Management (VADM), dated December 8, 2020.
5. **RESCISSION:** None.

CERTIFIED BY:

**BY DIRECTION OF THE SECRETARY
OF VETERANS AFFAIRS:**

/s/
Shana Love-Holmon
Acting Assistant Secretary
for Enterprise Integration

/s/
Shana Love-Holmon
Acting Assistant Secretary
for Enterprise Integration

DISTRIBUTION: Electronic Only

OPEN DATA – MANAGING INFORMATION AS AN ASSET

TABLE OF CONTENTS

1. PURPOSE. 4

2. SCOPE. 6

3. RESPONSIBILITIES. 6

4. PROCEDURES. 12

5. REFERENCES. 14

6. DEFINITIONS. 15

APPENDIX A – BUREAU CODES AND PROGRAM CODES 29

APPENDIX B – FREQUENTLY ASKED QUESTIONS..... 34

APPENDIX C – THE VALUE OF OPEN DATA 40

OPEN DATA – MANAGING INFORMATION AS AN ASSET

1. **PURPOSE.** This handbook provides Department-wide policy for ensuring VA disseminates appropriate machine-readable data and metadata to the public at VA's Open Data website (at: www.va.gov/data) and the Federal Data Catalog (at: www.data.gov). The value of Open Data is:
 - a. Transparency to improve public trust with internal and external stakeholders;
 - b. Modernization to provide comprehensive, timely, aggregated statistical data using stakeholder feedback on usage, enabling evidence-based advocacy on behalf of Veterans, their families, caregivers, and survivors; and
 - c. Data availability to inspire innovation.
 - d. Open Data enables Federal agencies to post machine-readable data on agency websites so that researchers, academia, and entrepreneurs can mine the data and build new products and services, while improving government efficiencies so that agency data is in one place.
 - e. Making data open and accessible in a standard, machine-readable format by default can have significant productivity and cost savings for agencies. When conducting a cost-benefit analysis to determine whether and to what extent to modify existing datasets and systems, the following potential benefits can be considered:
 - (1) Save time and money responding to Freedom of Information Act (FOIA) requests - When data is open by default, the public can access the information it seeks directly, freeing the time and cost related to responding to FOIAs;
 - (2) Avoid duplicative internal research -transparency into the total universe of data held by VA helps prevent the possibility of wasting funds re-collecting data simply because a particular program or department is unaware of that data's existence. Further, it may be possible to reduce the scope and cost of new collections based on the ability to re-use and/or pair with existing data. Maintaining a central data catalog for VA makes it easier to understand what information is currently available, and reviewing this catalog prior to the start of any new data collection is a recommended best practice;
 - (3) Discover complementary datasets held by other agencies - the benefits of transparency into VA's own datasets are amplified when every agency maintains its own standardized data catalog. Programs may realize that some or all of the data they need are already held by one or more other agencies, or that more powerful conclusions can be drawn from combining existing agency-held datasets with additional data across other agencies;

- (4) Empower employees to make better-informed, data-driven decisions - published details about each dataset owned by VA in a specific format will power a central search engine at [Data.gov](https://data.gov) that every single Federal employee (and member of the public) can use to easily locate data held, owned, and/or created by the Federal Government. Making it easier to find existing data is key to being able to then incorporate that data into VA's everyday decision-making processes;
 - (5) Engender trust through transparency; and
 - (6) Generate revenue and create new jobs in the private sector - McKinsey estimates that open health data alone adds over \$300 billion to the economy each year, according to [Business case for open data | resources.data.gov](https://www.businesscaseforopendata.com/resources). For example, entrepreneurs and non-profits integrate existing open government datasets with web apps that connect you with the nearest hospital in case of an emergency, according to [Business case for open data | resources.data.gov](https://www.businesscaseforopendata.com/resources). Making more VA data publicly available in standards-compliant, machine-readable formats makes it easier for private sector companies and entrepreneurs to create new innovations fueled by VA's data.
- f. Open Data guidance was officially disseminated on May 9, 2013, with the release of Executive Order 13642, Making Open and Machine Readable the New Default for Government Information and OMB M-13-13, Open Data Policy- Managing Information as an Asset. OMB policy was codified in the Evidence Act. Title II: OPEN Government Data Act also requires Federal agencies to develop and maintain a comprehensive data inventory that accounts for all data assets created by, collected by, under the control or direction of, or maintained by the agency and issued their own implementing guidelines.
- (1) This handbook contains mandatory VA procedures and responsibilities for holders of data to include:
 - (a) Administrations and Staff Offices approval of appropriate machine-readable data and metadata used in managing information created, collected, and disseminated to the public;
 - (b) Processing requests for correction of disseminated information that does not meet the established VA standards; and
 - (c) Complying with the OMB and Federal Data Strategy requirements.
 - (2) Since Administrations and Staff Offices disseminate a wide variety of data, the standards contained herein cannot be implemented in the same manner consistently by every organization. The handbook standards are designed to be generic enough to cover all disseminated data in printed, electronic, or other forms of media. The idea of allowing flexibility is that it shall allow each organization to have the opportunity to adapt the

standards to their existing information resource management and administrative practices while ensuring that product evaluations abide by the same considerations of technical and reproducibility aspects as to the Veteran or stakeholder point of view.

- (3) The information included in this handbook covers existing products and new products, however, it is important that when evaluating work products, the same consideration is given to the technical and the reproducibility aspects as to the context from the Veteran or stakeholder point of view.
 - (4) Under Secretaries, Assistant Secretaries and other key officials are encouraged to inform other applicable Administrations and Staff Offices throughout VA having substantial interest or expertise in the material proposed to be disseminated.
2. **SCOPE.** The handbook is limited to Open Data. The handbook supplements policies and responsibilities prescribed in VA Directive 0900, VA Enterprise Data Management (VADM).
 3. **RESPONSIBILITIES.** Additional responsibilities can be found in Directive 0900.
 - a. **Assistant Secretary for Enterprise Integration (A/S OEI)** shall: Provide resources for execution of the Open Data program.
 - b. **Under Secretaries and Assistant Secretaries and Other Key Officials** shall:
 - (1) Identify strategies, leveraging Administrations and sub-organizations including data governance organizations, to drive value in programs through the use of Open Data, establishing goals and objectives to realize Open Data's value in programs individually and collectively through governance and the Data Governance Council. By publishing Open Data, the opportunity to improve the quality of the data and peer review of the data exists. Execution of integrating Open Data to create value is the responsibility of program leaders.
 - (2) Appoint primary Open Data Liaison and alternate.
 - (a) The appointee should:
 - i Be a person that can speak authoritatively on behalf of the Department's data management practices;
 - ii Be a current VA employee;
 - iii Be aware of the OMB's Open Data efforts;
 - iv Be available to the public for data asset inquiries or coordinates public responses within the organization; and

within their purview especially the High Value Assets included in the Enterprise Information Model;

- (2) Establish Department-wide policies and procedures to ensure appropriate machine-readable data and metadata are posted at VA's Open Data website and that posted data follows appropriate use and disclosure standards;
- (3) Assign an Open Data Lead to serve as the program manager for the Open Data program;
- (4) Request designation of a primary and alternate Open Data Liaison from each Administration and Staff Office to participate in the ODWG; and
- (5) Establish and maintain a mechanism for tracking and responding to public requests.

e. **OEI Open Data Lead** shall:

- (1) Manage the Open Data Program and provide day-to-day support to organizations as it pertains to the Open Data Program and respond to and coordinate Open Data public requests while coordinating all available data in the Department, as applicable, is published in the Open Data website;
- (2) Serve as VA's liaison officer with OMB and other Federal agencies regarding management and operations of the Open Data program;
- (3) Coordinate and facilitate meetings and events in support of the Open Data program and CDO or designee, including the quarterly VA ODWG meetings and updates to the Data Governance Council and special events;
- (4) Collaborate and coordinate across the enterprise to support the implementation of the VA Data Strategy and Evidence Act requirements as related to Open Data; and
- (5) Provide recommendations to CDO on VA Open Data policy, program priorities, requirements based on stakeholder needs, legislative, OMB, Government Accountability Office (GAO), and Congressional requirements and reports.

f. **Data Governance Council members** shall:

- (1) Ensure the Open Data liaison list is maintained, including an email address for the public to use to submit requests, at VA's Open Data website;
- (2) Ensure that data stewards follow the appropriate Departmental data stewardship guidance, and are current VA employees, directly responsible and knowledgeable for the data asset associated with their name, aware of

OMB's Open Data efforts, fully understand their roles and responsibilities should anyone in the public domain contact them regarding their data asset which is listed in the public data asset listing. Organizations can identify an organizational data asset coordinator, if needed, to support an organizational lead;

- (3) Ensure data stewards are appointed and notify CDO with updated list on an annual basis; and
- (4) Provide support to Open Data Liaisons to implement strategy to drive value in programs through the use of Open Data, implementing goals and objectives to realize Open Data value in programs individually and collectively through governance and the Data Governance Council.

g. **Open Data Working Group** shall:

- (1) Set forth Open Data Policy, guidance, and training for the Department;
- (2) Oversee and contribute to VA's expansion, enhancement, and enrichment of the enterprise data inventory with ongoing release and updates of machine-readable data and metadata to meet quarterly OMB, Federal Data Strategy, and Evidence Act requirements;
- (3) Participate in ODWG, Data Governance Council meetings and support Open Data Program events;
- (4) Respond to and coordinate organizational public requests;
- (5) Provide recommendations to CDO and Open Data Lead on VA Open Data policy, program priorities, technology investments, as well as OMB, GAO, and Congressional reports; and
- (6) Provide feedback/guidance to CDO.

h. **Administration and Staff Office Open Data Liaisons** shall:

- (1) Participate in the ODWG and lead the Department's efforts to ensure all the data within the Administration/Staff Office they represent, as applicable, is published in the Open Data website;
- (2) Issue Administration and Staff Office communications, training and guidance, outlining the data collection effort requirements;
- (3) Form an internal committee or team consisting of representatives with the following subject matter expertise, including but not limited to: Privacy Act, Artificial Intelligence and other Data Users, Ethics, FOIA, and Information System Security Officers (ISSO). The internal committee's purpose is to review data assets that meet Open Data criteria, namely: data assets are

machine readable; data assets do not violate VA policy or privacy and security laws; access levels are assigned appropriately to data assets; FOIA exemptions are appropriately identified by data asset and exempted data assets are identified for redaction from the enterprise data inventory; and work with data stewards and notify CDO in OEI of any redacted data assets and metadata fields, citing appropriate FOIA exemptions;

- (4) Ensure certain data is de-identified by appropriate VA standards including unintended means such as the Mosaic Effect (Reference Glossary);
- (5) Identify possible locations of known data assets holdings (e.g., data warehouses, FOIA reading rooms, Enterprise Mission Assurance Support Service (eMASS) database, intranet, and internet webpages, etc.) and provide accurate estimates of percent of data available in VA's Open Data webpage versus the organization/Staff Office's holdings;
- (6) Examine Privacy Act Systems of Records (https://www.oprm.va.gov/privacy/systems_of_records.aspx) and PIAs and coordinate with the VA Privacy Office for any potential impact to Open Data assets;
- (7) Take Open Data training provided by VA, OMB, and/or General Services Administration (GSA) as required and made available through ODWG meetings;
- (8) Ensure organizational internet sites reflect appropriate uniform resource locators (URLs) for all metadata documents (e.g., data dictionaries, Privacy Act Systems of Records, and the occasional supporting technical documents classified as public). This location shall permit the public to access the documents via a generated hyperlink/URL;
- (9) Utilize hosted data asset collection and data story/data visualization tool at the following site: www.va.gov/data; take the training and review the documentation at: www.va.gov/data;
- (10) Search for existing and additional data assets, link data assets to provide more valuable data assets, and prioritize data assets according to value based on customer feedback and risk;
- (11) Review possible sources for data asset nominations, for example:
 - (a) Existing collections (e.g., National Center for Veterans Analysis and Statistics' website content tracker, etc.);
 - (b) Customer feedback (from events, Data.gov Help Desk requests, data user surveys, GitHub); and

- (c) VA Strategic Plan Crosscutting functional best practices (e.g., Office of Management, Office of Human Resources); OIT scan of the external facing VA internet website.
- (12) Assign data stewards/point of contact for each entry and have that individual/organization responsible for completing metadata content and validating existing content.
- i. **Data Stewards** shall:
- (1) Respond to Open Data Liaison/lead and public requests, including Geospatial data asset releases;
 - (2) Participate on an internal organizational committee consisting of representatives from various business lines and/or subject matter expertise including POs, FOIA staff, Ethics and ISSOs;
 - (3) Follow the appropriate Departmental data stewardship, Open Data guidance and VA data standards;
 - (4) Respond to public domain inquiries and complaints regarding data assets under their purview which are listed in the public data asset listing;
 - (5) Work with the organization's Open Data Program to ensure that the data assets under their purview are published according to the predetermined schedule;
 - (6) Ensure organizational internet sites reflect appropriate URLs for all metadata documents (e.g., data dictionaries, Privacy Act Systems of Records, and the occasional supporting technical documents classified as public). This location shall permit the public to access the documents via a generated hyperlink/URL;
 - (7) Utilize hosted data asset collection and data story/data visualization tool at the following site: www.va.gov/data; take the training and review the documentation at: www.va.gov/data; and log into <https://www.data.va.gov/login>;
 - (8) Secure necessary level of reviews and approvals within organization;
 - (9) Complete approved metadata content and validate existing content;
 - (10) Be current VA employees, directly responsible and knowledgeable for the data asset associated with their name, aware of OMB's Open Data efforts, and fully understand their roles and responsibilities should anyone in the public domain contact them regarding their data asset which is listed in the enterprise data inventory;

- (11) Work with Open Data liaisons and notify CDO in OEI of any redacted data assets and metadata fields, citing appropriate FOIA exemptions;
- (12) Ensure data assets and associated metadata are valid at least twice per year, verifying the data steward's name and organization are properly identified in the data asset collection system; and
- (13) Ensure that all data assets under their purview are included in the enterprise data catalog and the enterprise information model with the necessary metadata; the Open Data designations; and are published in the Open Data platform as permitted by the asset designation in the timeframe prescribed.

j. **Privacy Officers** shall:

- (1) Review the listed data asset nominations on a running basis for any privacy concerns or issues that include findings of personally identifiable information (PII) and protected health information (PHI); and consult with the ISSO to identify and resolve any associated information security concerns;
- (2) Review the potential for reidentification of de-identified data and consider the Mosaic Effect; and
- (3) Identify redacted data assets and metadata fields, citing appropriate FOIA exemptions.

4. **PROCEDURES.**

a. The procedure for the Department's Quarterly data asset collection shall be as follows:

- (1) **Collection Activities.**
 - (a) Search for existing and additional data assets; link data assets to provide more valuable data assets; and prioritize data assets according to value based on customer feedback and risk. One of the objectives of Open Data is to prioritize and release valuable data through public engagement. This objective should be balanced by the obligation to prevent inappropriate data release.
 - (b) Under Secretaries, Assistant Secretaries and other key officials should also focus on the following:
 - i Data assets discovered in the two-way customer feedback process and at events;

- ii Contribution of data assets in all categories (public, restricted public, and non-public);
 - iii Privacy Act System of Records;
 - iv Data that may be given to the public via FOIA requests;
 - v Data collections;
 - vi Expanded metadata;
 - vii Keywords;
 - viii Application Programming Interfaces;
 - ix Machine readable;
 - x Real-time delivery;
 - xi Worldwide domain;
 - xii Data quality improvements;
 - xiii Public access to previously restricted public data;
 - xiv Free access to previously charged for data;
 - xv OIT inventories (such as administrative and operational); and
 - xvi Data dictionaries.
- (c) Possible sources for data asset nominations include:
- i Existing collections (e.g., National Center for Veterans Analysis and Statistics' website content tracker, etc.);
 - ii Customer feedback (from events, Data.gov Help Desk requests, data user surveys, GitHub);
 - iii VA Strategic Plan;
 - iv Crosscutting functional best practices (e.g., Office of Management, Office of Human Resources); and
 - v OIT scan of the external facing VA internet website.
- (d) Assign data steward/point of contact for each entry and have that individual/organization responsible for completing metadata content

and validating existing content. Organizational leads are responsible for confirming that their organization's outlined data stewards are:

- i Current VA employees;
- ii Directly responsible and knowledgeable for the data asset associated with their name;
- iii Aware of OMB's Open Data efforts; and
- iv Fully understand their roles and responsibilities should anyone in the public domain contact them regarding their data asset which is listed in the enterprise data inventory.

- (e) All collected data assets are not official inventory entries, but merely nominations, which shall be included once final approval is secured.

b. Post-Collection Analysis.

- (1) Have designated Program Office POs and ISSOs as well as other designated personnel review the listed data asset nominations on a running basis for privacy and information security concerns to include findings of PII and PHI as well as the reidentification of de-identified data via the mosaic effect. Identify redacted data assets and metadata fields, citing appropriate FOIA exemptions.

c. Conduct Necessary Level of Reviews within Organization.

- (1) Final Internal Review and Submission
 - (a) Submit approved data assets to organizational lead for submission to Open Data Liaison. Open Data Liaison notifies OEI of any redacted data assets and metadata fields, citing appropriate FOIA exemptions.

5. **REFERENCES.** Since May 2013, the Executive Office of the President has released a series of guidance regarding its Open Data Policy. Note: Components of OMB M-13-13 have changed since its release in May 2013. Contents of these documents have in part supported the construction of this handbook.

- a. Executive Order 13642, Making Open and Machine Readable the New Default for Government Information (May 9, 2013). [2013-11533.pdf \(govinfo.gov\).](#)
- b. OMB Memorandum 13-13, Open Data Policy-Managing Information as an Asset (May 9, 2013). [M-13-13 \(archives.gov\).](#)
- c. Foundations for Evidence-Based Policymaking Act of 2018, P.L. 115-435, 132 Stat. 5529 (Jan 14, 2019). <https://www.govinfo.gov/content/pkg/PLAW-115publ435/pdf/PLAW-115publ435.pdf>.

- d. Supplemental Guidance on the Implementation of M-13-13 “Open Data Policy – Managing Information as an Asset”. [Supplemental Guidance on the Implementation of M-13-13 “Open Data Policy – Managing Information as an Asset” | resources.data.gov](#)
- e. Project Open Data, Open Data Policy- Managing Information as an Asset. [Redirecting to https://resources.data.gov/ \(cio.gov\)](https://resources.data.gov/ (cio.gov)).
- f. A repository of Federal Enterprise Data Resources. [Home | resources.data.gov](#).
- g. Federal Data Strategy, Leveraging Data as a Strategic Asset. [Welcome - Federal Data Strategy](#).

6. DEFINITIONS.

- a. **Enterprise Data Inventory:** The major output of the Open Data effort shall be the generation of the Department's enterprise data inventory.
 - (1) The inventory is no more than a comprehensive listing of select data assets and associated characteristics known as metadata.
 - (2) It is available to the public in its entirety with appropriate redactions given applicable FOIA exemptions.
 - (3) Under no circumstance does metadata include CUI, PII PHI, nor conflict with VA policy, and standards for data access and use. Nor shall any databases or data sets containing CUI, PII, PHI, or that are inconsistent with VA policy and standards for data access and use, be uploaded.
 - (4) Typical metadata fields include items such as title, description, and data steward/point of contact information. Some types of data assets to include statistical or analytical reports shall require far less.
 - (5) The public shall have access to the entire redacted enterprise data inventory. Where necessary, FOIA exemptions shall be cited and affected data assets shall be redacted. Data stewards in each Administration and Staff Office are responsible for determining appropriate redactions citing appropriate FOIA exemptions and submitting those to the OEI.
 - (6) The public shall see entries classified with a "public" access level via the submission of a hyperlink. This subsection of the inventory is known as the public data listing.

Source: (M-13-13 (archives.gov))

- b. **Data Asset:** According to [PUBL435.PS \(govinfo.gov\)](#), a data asset is "a collection of data elements or datasets that make sense to group together." For

VA purposes, a data asset is a work product generated by a VA employee or VA-affiliated entity which possesses one or more of the following attributes:

- (1) Contains a collection of record-level data (data about an individual) and/or aggregated data (a form of information) generated in support of the Department's mission;
 - (2) Structured in composition and done so in a purposeful and meaningful way;
 - (3) Encased in one of many file formats to include but not limited to.xls,.csv,.xml, json, or an Application Programming Interface (API). (Portable Document Format (PDF) should not be used);
 - (4) Presented in one of many mediums to include but not limited to electronic or digital;
 - (5) Composed primarily of numerics (record-level data or aggregated information) and/or numerics with descriptive narratives such as in a report;
 - (6) Physically housed either within VA's domain (Intranet, data warehouse, etc.) or in the public domain (Internet);
 - (7) May or may not have a data dictionary or be associated with a Privacy Act System of Records;
 - (8) May or may not be releasable in part or in whole under provisions of the FOIA and/or Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, the Privacy Act; or Executive Order 13556, Controlled Unclassified Information;
 - (9) May require permanent information security safeguards and dissemination restrictions;
 - (10) Can consist exclusively of internal VA data/information or partial VA data/information merged with another agency's records and;
 - (11) Was not generated in its entirety by another Federal agency or other government entity; and
 - (12) Is consistent with all VA standards for data access and use.
- c. **Collection Boundaries.** The enterprise data inventory consists of two major components: a public data listing and a restricted listing. A data asset's inclusion in either listing is based upon the access level (public, restricted public, or non-public) which is determined, assigned, and submitted into the inventory's metadata by the Administration and Staff Office of jurisdiction. The public shall

have access to the entire enterprise data inventory with appropriate redactions given applicable FOIA exemptions.

(1) **Public Data Asset Listing.**

The public data listing should only contain vetted and official data assets generated by each Administration and Staff Office and assigned a “public” access level. These products may or may not be sourced entirely with agency information. Data assets generated by VA-affiliated entities (i.e., contracted vendors) are permissible so long as approved for release by the organization.

(a) Included Data Assets.

- i Non-re-identifiable aggregated and statistical data;
- ii Aggregated/statistical reports;
- iii Analytical reports;
- iv Geographic and Information System (GIS) maps (no latitude and longitude data, except as required by Administration and Staff Office);
- v Master non-re-identifiable aggregated/statistical files; and
- vi Fully vetted/official data (non-draft) and information.

(b) Excluded Data Assets.

- i Record-level data;
- ii PII data;
- iii PHI data;
- iv GIS products with latitude and longitude data (except as required by Administration and Staff Office);
- v Classified data;
- vi National security data;
- vii Re-identifiable data;
- viii Sensitive data;
- ix Data extracts;

- x Data snapshots;
- xi Non-vetted/non-official data and information;
- xii Created exclusively by another Federal agency or other government entity; and
- xiii Any data asset placed in the restricted public data listing.

(2) **Restricted Data Asset Listing.** The restricted data asset listing should only contain vetted and official data assets generated by each Administration and Staff Office and either assigned a “restricted public” or “non-public” access level. These products may or may not be sourced entirely with agency information. Data assets generated by VA-affiliated entities (i.e., contracted vendors) are permissible so long as approved for release by the organization.

(a) Included Data Assets

- i Record-level data;
- ii PII data;
- iii PHI data;
- iv GIS products with latitude and longitude data (except as required by Administration and Staff Office);
- v CUI data;
- vi Classified national security data;
- vii Re-identifiable data;
- viii Master data files; and
- ix Vetted/official data and information.

(b) Excluded Data Assets

- i Non-re-identifiable aggregated and statistical data;
- ii Non-aggregated/statistical reports;
- iii Non-analytical reports;
- iv Master non-re-identifiable aggregated/statistical files;
- v Data extracts;

- vi Data snapshots;
- vii Non-vetted/non-official data and information;
- viii Any data asset placed in the public data asset listing;
- ix Created exclusively by another Federal agency or other government entity; and
- x CUI data.

(3) **Metadata** For VA purposes, the data asset collection tool breaks down the 21 metadata fields into two distinct buckets: “Required” fields and “Required-If-Applicable” fields. The two buckets align with OMB definitions:

(a) VA Metadata Buckets

- i Required (15)
- ii Required – If-Applicable (6)

(b) OMB Metadata Definitions

- i A critical metadata field which is mandatory for all data assets regardless of the designated access level.
- ii A critical metadata field which is mandatory should the data asset have the outlined metadata properties.

Definitions for each metadata field follow. *Note:* Some business lines may elect to capture additional information above and beyond the OMB requirements below.

Reference: DCAT-US Schema v1.1 (Project Open Data Metadata Schema) at: [DC AT-US Schema v1.1 \(Project Open Data Metadata Schema\) | resources.data.gov.](https://resources.data.gov/AT-US-Schema-v1.1-Project-Open-Data-Metadata-Schema/)

REQUIRED FIELDS:

#	Field-OMB Tag	Label	Definition
1	title	Title	Human readable name of the asset. This should be in plain English and include sufficient detail to facilitate search and discovery; acronyms should be avoided.
2	description	Description	Human readable description (e.g., an abstract) with sufficient detail to

#	Field-OMB Tag	Label	Definition
			enable a user to quickly understand whether an asset is of interest; no more than six to ten sentences.
3	keyword	Tags	Keywords that help users discover your dataset; terms used by technical and non-technical users should be included.
4	modified	Last Update	Most recent date on which the datasets were changed, updated, or modified.
5	publisher	Publisher	The name of the publishing entity and optionally their parent organization(s).
6	contactPoint	Contact Name and Email	Contact person's name and email address for the asset, including first name and last name.
7	identifier	Unique Identifier	The data collection tool shall automatically generate a unique identifier.
8	accessLevel	Public Access Level	The degree to which this data asset could be made publicly available, regardless of whether it has been made available; options are: "Public" (data asset is or could be made available to all without restrictions); "Restricted Public" (data asset is available under certain use restrictions); and "Non-Public" (data asset is not available to members of the public). Reference Attachment 4 for detailed definitions.
9	rights	Rights	This may include information regarding access or restrictions based on privacy, security, or other policies. This should also serve as an explanation for the selected

#	Field-OMB Tag	Label	Definition
			“accessLevel” including instructions for how to access a restricted file, if applicable, or explanation for why a “non-public” or “restricted public” data asset is not “public,” if applicable. In addition, if the data asset contains one or more exemptions pursuant to a FOIA exemption, an additional explanation is required for each redaction(s). The field is text, limited to 255 characters.
10	bureauCodeUSG	Bureau Code	The Federal agency bureau code from OMB Circular A-11, Appendix C. Format is XXX:XX. Reference the bureau codes located in Appendix C.
11	programCodeUSG	Program Code	The OMB-defined code related to the data asset. Format XXX:XXX. Reference the program codes located in Appendix C.
12	conformsTo	SchemaVersion	The version of the Project Open Data schema being used (for example, 1.1).
13	dataset	Dataset	A container for the array of Dataset objects.
14	isPartOf	Collection	The collection of which the dataset is a subset.
15	license	License	VA uses the standard value found at https://project-open-data.cio.gov/v1.1/schema/#license . The license each data asset (dataset, API, and select other) is published with. URL Copyright and other rights pertaining to U.S. work.

REQUIRED-IF-APPLICABLE FIELDS:

#	Field-OMB Tag	Label	Definition
16	distribution	Distribution	Holds multiple download URLs for datasets composed of multiple files and/or file types.
17	spatial	Spatial	The range of spatial applicability of a data asset (dataset, API, and select other). This could include a spatial region such as a bounding box or a named place.
18	temporal	Temporal	The range of temporal applicability of a data asset (dataset, API, and select other). For example, the start and end date for applicability for the data.
19	downloadURL	DownloadURL	URL providing direct access to a downloadable file of a dataset.
20	mediaType	Media Type	The machine-readable file format of the distribution's downloadURL.
21	accessURL	AccessURL	The alternative landing page used to redirect user to a contextual, agency-hosted "homepage" for the data asset (dataset, API, and select other) when selecting this resource from the Data.gov user interface.

- d. **Access Levels:** A metadata field assigned by the data steward which signifies the level of availability as well as release restrictions for an identified data asset in the enterprise-wide data asset inventory. The three access levels are as follows:
- (1) **Public:** Data asset is or could be made publicly available to all without restrictions.
 - (2) **Restricted Public:** Data asset is only available under defined certain use restrictions that shall vary based upon the content and be the subject of a data use agreement. Example 1: Data asset that can only be made available to select researchers under certain conditions because the data asset contains sufficient granularity or linkages that make it possible to re-identify individuals, even though the data asset is stripped of PII and PHI. Example 2: Data asset that contains PII and is made available to select researchers under strong legal protections. This category includes some but not all data assets designated as Controlled Unclassified Information CUI, consistent with Executive Order 13556.
 - (3) **Non-Public:** Data asset is not available to members of the public. This category includes data assets that are only available for internal use by the Federal Government, such as by a single program, single agency, or across multiple agencies. Example 1: Data assets that are only available for internal use by the Federal Government, such as by a single program, single agency, or across multiple agencies. Example 2: Some but not all data assets designated as CUI, consistent with Executive Order 13556 -- Controlled Unclassified Information | [whitehouse.gov \(archives.gov\)](https://www.whitehouse.gov/archives). Some non-public data assets may still potentially be available to other intra-agency operating units and/or other government agencies, as discussed in OMB M-11-02: Sharing Data While Protecting Privacy at Memorandum to the Heads of Executive Departments and Agencies ([archives.gov](https://www.archives.gov)).
- e. **API:** An Application Programming Interface (API) is a set of definitions of the ways one piece of computer software communicates with another. It is a method of achieving abstraction, usually (but not necessarily) between higher-level and lower-level software. Source: ([Glossary \(opendatahandbook.org\)](https://opendatahandbook.org)).
- f. **Catalog:** A catalog is a collection of datasets or web services. The Comprehensive Kerbal Archive Network (CKAN) is a popular, open-source data catalog that can be installed independently or run as a service. CKAN provides a solution for maintaining an internal or external catalog. In addition to providing an effective means of organizing the agency's data catalog, CKAN provides several opportunities for generating the needed data.json file. CKAN provides a convenient CSV export that can then be imported into the Catalog Generator for conversion into an appropriately formatted JSON file. Alternatively, CKAN provides a JSON API of the data catalog which can then be altered to provide an automatic update for the appropriately formatted JSON file. Source:

<https://www.govinfo.gov/content/pkg/PLAW-115publ435/pdf/PLAW-115publ435.pdf>

- g. **DKAN:** DKAN is an open-source data management platform that treats data as content to publish, manage, and maintain Open Data. Granicus Open Data is based on DKAN. Source: DKAN Open Data Platform | DKAN Open Data Platform (getdkan.org).
- h. **Inventory.Data.gov:** Inventory.data.gov is the GSA hosted data collection tool previously used by VA Administrations and Staff Offices to submit approved data assets and associated metadata.
- i. **Data:** A value or set of values representing a specific concept or concepts. Data become "information" when analyzed and possibly combined with other data to extract meaning, and to provide context. The meaning of data can vary depending on its context. Data includes, but is not limited to, 1) geospatial data 2) unstructured data, 3) structured data, etc.
- j. **Data Asset:** For VA purposes, a work product generated by a VA employee or VA affiliated entity which contains a collection of aggregated data and was generated in support of the respective business function and the Department's overall mission. The definition is meant to be broad; a data asset can possess one or more of the following attributes:
 - (1) contains a collection of record-level data (data about an individual) and/or aggregated data (a form of information) generated in support of the Department's mission;
 - (2) structured in composition and done so in a purposeful and meaningful way;
 - (3) encased in one of many file formats to include but not limited to.xls,.csv, or.xml;
 - (4) presented in one of many mediums to include but not limited to electronic, digital, or paper;
 - (5) composed primarily of numerics (record-level data or aggregated information) and/or numerics with descriptive narratives such as in a report;
 - (6) physically housed either within VA's domain (Intranet, data warehouse, bookshelf etc.) or in the public domain (Internet);
 - (7) may or may not have a data dictionary or be associated with a Privacy Act System of Records ;
 - (8) may or may not be releasable in part or in whole under the FOIA and Privacy Act;

- (9) may require permanent information security safeguards and dissemination restrictions; or
 - (10) can consist exclusively of internal VA data/information or partial VA data/information merged with another agency's records and was not generated in its entirety by another agency. Additionally, the asset's aggregated data makes sense when grouped together, such as, in an analytical report. A data asset's formatting is not limited to one type, but rather to multiple types to include but not limited to.csv and.xml. Common work products are statistical or analytical reports, databases, or APIs.
- k. **Dataset:** A dataset is an organized collection of data. The most basic representation of a dataset is data elements presented in tabular form. Each column represents a particular variable. Each row corresponds to a given value of that column's variable. A dataset may also present information in a variety of non-tabular formats, such as an extensible markup language (XML) file, a geospatial data file, or an image file, etc.
 - l. **Database:** A collection of data stored according to a schema and manipulated according to the rules set out in one Data Modelling Facility.
 - m. **Data Publication Process:** A process, developed in consultation with VA Office of General Counsel (OGC), to determine whether data assets have a valid restriction to release. The objectives of this process are to: review information for valid restrictions to public release to ensure proper safeguarding of privacy, security, and confidentiality of government information; document reasons why a data asset or certain components of a data asset should not be made public at this time; consult with VA OGC and Office of Information Technology Security regarding the barriers identified. A general overview of this process is published on the www.va.gov/digitalstrategy page. The overview includes information on the actual process by which data is determined to have a valid restriction to release and examples of what kinds of characteristics a data asset has that leads to a determination to not release. If data cannot be released, Senior Leadership shall ensure a complete analysis of issues that could preclude public disclosure of information collected or created. If it is determined that the data should not be made publicly available because of law, regulation, or policy or because the data are subject to privacy, confidentiality, security, trade secret, contractual, or other valid restrictions to release, reasons shall be documented in consultation with VA OGC. For each data asset listed in the inventory, one of three "access levels" shall be designated: public, restricted public, and non-public.
 - n. **Enterprise-wide Data Asset Inventory:** Within VA, the major output of Open Data shall be the generation of a comprehensive listing of all data assets called the Enterprise-Wide Data Asset Inventory. By using data asset collection tool, this inventory shall serve as the repository for all identified VA data assets. Each Administration and Staff Office should populate metadata for its data assets.

- o. **Esri:** Esri is a company which provides GIS software, location intelligence, and mapping. Esri provides ArcGIS, a powerful GIS technology that offers tools to capture, view, edit, manage, analyze, and share data in the context of location. It includes access to thousands of curated datasets and maps that can be explored and leveraged for analysis and insight. ArcGIS can be used in the cloud, on mobile devices, and on desktops to create maps, apps, dashboards, 3D scenes and models, and data science notebooks. ArcGIS is comprehensive and scalable digital mapping and analytics software. Source: [About Esri | The Science of Where; Esri Technology | GIS Software & Location-Based Technology.](#)
- p. **JSON:** JSON (JavaScript Object Notation) is a lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate. It is based on a subset of the JavaScript Programming Language, Standard ECMA-262 3rd Edition - December 1999. JSON is a text format that is completely language independent but uses conventions that are familiar to programmers of the C-family of languages, including C, C++, C#, Java, JavaScript, Perl, Python, and many others. These properties make JSON an ideal data-interchange language.
- q. **Machine-Readable File:** Refers to information or data that is in a format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost.
- r. **Metadata:** To facilitate common understanding, a number of characteristics, or attributes, of data are defined. These characteristics of data are known as "metadata", that is, "data that describes data." For any particular datum, the metadata may describe how the datum is represented, ranges of acceptable values, its relationship to other data, and how it should be labeled. Metadata also may provide other relevant information, such as the responsible steward, associated laws and regulations, and access management policy. Each of the types of data described above has a corresponding set of metadata. Two of the many metadata standards are the Dublin Core Metadata Initiative (DCMI) and Department of Defense Discovery Metadata Standard (DDMS). The metadata for structured data objects describes the structure, data elements, interrelationships, and other characteristics of information, including its creation, disposition, access and handling controls, formats, content, and context, as well as related audit trails. Metadata includes data element names (such as organization name, address, etc.), their definition, and their format (numeric, date, text, etc.). In contrast, data is the actual data values such as the "US Patent and Trade Office" or the "Social Security Administration" for the metadata called "Organization Name". Metadata may include metrics about an organization's data including its data quality (accuracy, completeness, etc.).
- s. **Mosaic Effect:** The mosaic effect occurs when the information in an individual dataset, in isolation, may not pose a risk of identifying an individual (or threatening some other important interest such as security), but when combined

with other available information, could pose such risk. Before disclosing potential PII or other potentially sensitive information, agencies shall consider other publicly available data -in any medium and from any source-to determine whether some combination of existing data and the data intended to be' publicly released could allow for the identification of an individual or pose another security concern. Source: OMB M-13-13, [M-13-13 \(archives.gov\)](#).

- t. **Open Data:** For the purposes of OMB's Open Data Policy, Open Data refers to publicly available data structured in a way that enables the data to be fully discoverable and usable by end users. In general, Open Data shall be consistent with the following principles:
- (1) **Public:** Consistent with OMB's Open Government Directive, agencies shall adopt a presumption in favor of openness to the extent permitted by law and subject to privacy, confidentiality, security, or other valid restrictions.
 - (2) **Accessible:** Open Data are made available in convenient, modifiable, and open formats that can be retrieved, downloaded, indexed, and searched. Formats should be machine readable (i.e., data are reasonably structured to allow automated processing). Open Data structures do not discriminate against any person or group of persons and should be made available to the widest range of users for the widest range of purposes, often by providing the data in multiple formats for consumption. To the extent permitted by law, these formats should be non-proprietary, publicly available, and no restrictions should be placed upon their use.
 - (3) **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. This involves the use of robust, granular metadata (i.e., fields or elements that describe data), thorough documentation of data elements, data dictionaries, and, if applicable, additional descriptions of the purpose of the collection, the population of interest, the characteristics of the sample, and the method of data collection.
 - (4) **Reusable:** Open Data are made available under an open license that places no restrictions on their use.
 - (5) **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 and permitted by law and other requirements; derived or aggregate Open Data should also be published but shall reference the primary data.
 - (6) **Timely:** Open Data are made available as quickly as necessary to preserve the value of the data. Frequency of release should account for key audiences and downstream needs.

- (7) **Managed Post-Release:** A point of contact shall be designated to assist with data use and to respond to complaints about adherence to these Open Data requirements.

Source: OMB Memorandum 13-13, Open Data Policy - Managing Information as an Asset (March 9, 2013).

- u. **Project Open Data:** "Project Open Data," an OMB/GSA resource, is an online repository of tools, best practices, and schema to help agencies adopt the framework presented in this guidance. Project Open Data can be accessed at <http://project-open-data.github.io>. Project Open Data shall evolve over time (to Resources.data.gov at: [Home | resources.data.gov](http://resources.data.gov)) as a community resource to facilitate adoption of Open Data practices. The repository includes definitions, code, checklists, case studies, and more, and enables collaboration across the Federal Government, in partnership with public developers, as applicable. Agencies can visit Project Open Data for a more comprehensive glossary of terms related to Open Data.
- v. **Socrata:** Socrata is a data storytelling and insights platform built to add value to Open Data efforts by enabling and empowering content creation in an easy-to-use, streamlined way. To achieve this, Socrata uses a combination of data visualization tools and integrations to create Data Stories and Dashboards that deliver the message behind data in a user-friendly interface. Built on traditional data portal features, Socrata offers a market standards-compliant experience that increases the user value of any Open Data effort.

APPENDIX A - BUREAU CODES AND PROGRAM CODES.

1. For each data asset, there is a metadata field that deals with bureau codes. (VA's equivalent of a bureau is either an Administration or Staff Office.) These codes were established by OMB and are used to partially evaluate participation and progress with the data asset inventory effort.
2. For entry into data asset collection tool, enter the Agency Code, a colon, and then the Bureau Code, for example: 029:15.
3. For each data asset, there is a metadata field that deals with program codes. These codes were established by OMB and are used to partially evaluate participation and progress with the data asset inventory effort.
4. In the instance where an existing Administration or Staff Office program isn't listed or doesn't properly correlate to the specific OMB terminology, please default to the organizational program code.
5. Note: Agency and Bureau codes have been cross referenced with the Office of Management.

VA Bureau Codes

OMB Nomenclature	Agency Code	Bureau Code
Department of Veterans Affairs	029	00
Veterans Health Administration	029	15
Benefits Programs	029	25
Departmental Administration	029	40

6. Reference: OMB Circular A-11, Appendix C, Page 7 [a11_2016.pdf \(archives.gov\)](#).
7. VA Bureau Codes with Supporting OMB-Identified Programs According to OMB records, the designated sub-components of each OMB nomenclature are outlined below. For those Staff Offices not identified below, please default to "Department of Veterans Affairs" nomenclature.
 - a. Veterans Health Administration.

Veterans Health Administration	Bureau Code:Program Code
Inpatient Care	029:040
Ambulatory Care	029:041
Mental Health Service, General Outpatient Care	029:042
Mental Health, Intensive Recovery-Oriented Program	029:043
Mental Health, Inpatient Care	029:044
Mental Health Residential Rehabilitation Treatment Programs (MH RRTP)	029:045
Readjustment Counseling	029:046
Pharmacy	029:047
VHA Research and Development	029:048
Spinal Cord Injury and Disorders (SCI/D) Service	029:049
Blind Rehabilitation Service (BRS)	029:050
Dental Care	029:051
Prosthetics and Sensory Aids Service (PSAS)	029:052
Women Veterans Health Care	029:053
Community Living Centers (CLCs)	029:054
Community Nursing Homes	029:055
State Veterans Home Program	029:056
Adult Day Health Care	029:057
Home Based Primary Care	029:058
Homemaker and Home Health Aide Care	029:059
Hospice and Palliative Care	029:060

Veterans Health Administration	Bureau Code:Program Code
	029:061
Respite Care	029:062
Skilled Home Health Care	029:063
Program of Comprehensive Assistance for Family Caregivers	029:064
Telehealth Care	029:065
Civilian Health and Medical Program of the Department of Veterans Affairs (CHAMPVA)	029:066
Foreign Medical Program (FMP)	029:067
Spina Bifida Program	029:068
Children of Women Vietnam Veterans (CWVV)	029:069
Department of Housing and Urban Development-VA Supportive Housing (HUD-VASH)	029:070
Grant and Per Diem (GPD) Program	029:071
Health Care for Homeless Veterans (HCHV)	029:072
Supportive Services for Low Income Veterans and Families (SSVF)	029:073
National Call Center for Homeless Veterans (NCCHV)	029:074
Veterans Justice Outreach (VJO) Program	029:075
Homeless Veterans Supported Employment Program (HVSEP)	029:076
National Homeless Registry	029:077

b. Veterans Benefits Administration.

Benefits Programs	Bureau Code:Program Code
Veterans Benefits – Administration of Benefits	029:003
Burial Benefits	029:004
Veterans Disability Compensation	029:005
Survivors Compensation	029:006
Dependency Indemnity Compensation (DIC)	029:007
Filipino Veterans Compensation	029:008
Veterans Pension	029:009
Survivors Pension	029:010
Fiduciary Services	029:011
Veterans Mortgage Life Insurance	029:012
United State Government Life Insurance	029:013
National Service Life Insurance	029:014
Veterans' Special Life Insurance	029:015
Veterans Reopened Insurance	029:016
Service-Disabled Veterans' Insurance	029:017
Servicemembers' Group Life Insurance (SGLI)	029:018
Family Servicemembers' Group Life Insurance (FSGLI)	029:019
Servicemembers' Group Life Insurance Traumatic Injury Protection (TSGLI)	029:020
Veterans' Group Life Insurance (VGLI)	029:021
Housing Guaranteed Loan	029:022
Acquired Direct Loan	029:023

Benefits Programs	Bureau Code:Program Code
Vendee Direct Loan	029:024
Guaranteed Loan Sale Securities	029:025
Native American Direct Loan	029:026
Specially Adapted Housing (SAH) & Special Housing Adaptation (SHA)	029:027
Vocational Rehabilitation Loan	029:028
Transitional Housing Loan	029:029
Post-9/11 GI Bill	029:030
Montgomery GI Bill – Active Duty	029:031
Montgomery GI Bill – Selected Reserve	029:032
Survivors' and Dependents' Educational Assistance	029:033
Reserve Educational Assistance Program	029:034
Veterans Retraining Assistance Program	029:035
Post-Vietnam Era Veteran's Educational Assistance Program	029:036
Vocational Rehabilitation and Employment Program	029:037
Vocational Rehabilitation and Employment – VetSuccess On Campus (VSOC)	029:038
Vocational Rehabilitation and Employment – Independent Living (IL)	029:039

APPENDIX B - FREQUENTLY ASKED QUESTIONS.

1. **What is Open Data?** An actionable government-wide initiative aimed at having agencies manage information as an asset across all executive departments and agencies. By managing government information as an asset, the desired outcome is to increase operational efficiencies, reduce costs, improve services, support mission needs, safeguard personal information, and increase public access to government information. Open Data guidance was officially disseminated with the release of Executive Order 13642, Making Open and Machine Readable the New Default for Government Information (May 9, 2013), and OMB M-13-13, Open Data Policy- Managing Information as an Asset (May 9, 2013). Subsequent guidance is routinely released via a combined OMB and GSA resource known as Project Open Data (<http://project-open-data.github.io/>).
2. **What shall Open Data do for VA?** Open Data shall collect or create information in a way that supports downstream information processing and dissemination. This shall be accomplished in part because VA data assets shall be in machine-readable and open formats, use enterprise-wide data standards, and have common core metadata for all newly created information. Additionally, Open Data shall ensure information stewardship through the use of open licenses and the review of information for privacy, confidentiality, security, or other restrictions on release. Finally, Open Data shall build or modernize its information systems so that interoperability and information accessibility is maximized, an inventory for all internal and external data assets can be maintained, information safeguards are enhanced, and information management responsibilities are clearly defined.
3. **What is VA's Data Asset Collection effort?** An internal VA effort to identify, collect, record, and review all of the Department's data assets and associated metadata using a standardized OMB reporting format. All collected data assets shall be stored in VA's enterprise-wide data asset inventory and uploaded to OMB's Government-wide inventory. To the extent permissible by law, VA data assets shall be available to the public via the inventory's public data asset listing.
4. **What is a data asset?** The term "data asset" has several meanings. As defined by OMB's Project Open Data, a data asset is essentially "a collection of data elements or datasets that make sense to group together." All data is to be machine readable. For VA purposes, a data asset is essentially a work product generated by a VA employee or VA affiliated entity which contains a collection of aggregated data and was generated in support of the respective business function and the Department's overall mission. The definition is meant to be broad; a data asset can possess one or more of the following attributes:
 - a. contains a collection of record-level data (data about an individual) and/or aggregated data (a form of information) generated in support of the Department's mission;
 - b. structured in composition and done so in a purposeful and meaningful way;

- c. encased in one of many file formats to include but not limited to.xls,.csv, or .xml;
 - d. presented in one of many mediums to include but not limited to electronic, digital, or paper;
 - e. composed primarily of numerics (record-level data or aggregated information) and/or numerics with descriptive narratives such as in a report;
 - f. physically housed either within VA's domain (Intranet, data warehouse, bookshelf etc.) or in the public domain (Internet);
 - g. may or may not have a data dictionary or be associated with a Privacy Act System of Records ;
 - h. may or may not be releasable in part or in whole under the FOIA and Privacy Act;
 - i. may require permanent information, security safeguards and dissemination restrictions;
 - j. can consist exclusively of internal VA data/information or partial VA data/information merged with another agency's records; and
 - k. was not generated in its entirety by another agency. Machine readable format is in a standard computer language (not English text) that can be read automatically by a web browser or computer system. (e.g. .xml). Traditional word processing documents, hypertext markup language (HTML) and portable document format (PDF) files are easily read by humans but typically are difficult for machines to interpret. Other formats such as extensible markup language (XML), (JSON), or spreadsheets with header columns that can be exported as comma separated values (CSV) are machine readable formats. It is possible to make traditional word processing documents and other formats machine readable, but the documents shall include enhanced structural elements.
5. **What is a two-way public feedback mechanism?** VA's Customer Feedback Process is a mechanism for engaging Veterans, family members, and the public. This engagement helps VA's Open Data team prioritize data releases. It drives and facilitates the use of VA data assets by the public for the creation of informative and/or useful products that add value to the lives of Veterans and their supporters. Feedback is collected through digital engagement and stakeholder outreach (partnerships, in-person events), including White House events on topics such as aging and mental health. Open communication, responsiveness, and personal relationships have been, and shall continue to be, central to VA's ability to connect with and understand the needs of the end users of VA's Open Data. VA's Open Data Feedback (GitHub) mechanism enables direct two-way feedback between the public and VA's Open Data team. As an example, through VA's Open Data Feedback mechanism, there was much discussion about improving the content of the VA facilities data file. As a result, VA updated facilities data releases. In

addition, data stewards identified as the point of contact through Data.Gov engage in direct communications with requestors. VA's Open Data Lead interacts with requestors through Data.gov. In addition, VA continues to submit data release proposals based on user feedback through e-mails from data users sent to The National Center for Veterans Analysis and Statistics (NCVAS). VA receives close to 1,000 email inquiries on VA data every year. VA responds to each person individually. We learn from them, and we try to make our data products better based on their input.

6. **What is the Enterprise-Wide Data Asset Inventory?** Within VA, the major output of Open Data activity shall be the generation of a comprehensive listing of all data assets called the Enterprise-Wide Data Asset Inventory. By using the Granicus data asset collection tool, this inventory shall serve as the repository for all identified VA data assets. Select metadata for each Administration and Staff Office's data asset shall populate the fields contained there within.
7. **Where shall the Enterprise-Wide Data Asset inventory be stored?** Both VA and OMB shall have copies of the Inventory. VA shall maintain a copy of all information that shall be submitted to OMB via quarterly data collection uploads. The VA uploads shall be appended with other Federal agency submissions. The GSA shall host and manage OMB's Government-wide inventory known as the data catalog at: <http://www.data.gov>.
8. **Shall the public have access to the entire Enterprise-Wide Data Asset Inventory?** The public shall only have access to the inventory with appropriate redactions for FOIA exemptions as identified by Administrations and Staff Offices in the data discovery and data publication process.
9. **Shall OMB have access to the entire Enterprise-Wide Data Asset Inventory?** Regardless of the access level designation, both OMB and other Federal employees shall have access to the entire inventory.
10. **Shall databases or data sets be included in the Enterprise-Wide Data Asset inventory?** Under no circumstances shall the metadata contain URLs to record-level data (information about a Veteran). That said, characteristics (metadata) about a database or data set that contains record-level data shall be provided. The composition and structure of the inventory prevents document and file uploads.
12. **Does the metadata contain a field to discriminate the availability of the data asset to the public?** Yes, the Inventory calls for the designated point of contact to complete a field known as "access level" which represents the degree to which a data asset could be made publicly available, regardless of whether it has been made available. Options are as "Public" (data asset is or could be made available to all without restricts); "Restricted Public" (data asset is available under certain use restrictions); and "Non-Public" (data asset is not available to members of the public).

13. **Does the actual file for a data asset described in the public data asset listing need to be made publicly available?** Yes, the goal of the public data asset listing is to make these data assets available on the internet and thus eliminate the need for a formal request such as a FOIA request.
14. **Shall data assets containing record-level data or data about an individual be uploaded and/or stored in the Enterprise-Wide Data Asset Inventory?** Under no circumstances shall this type of sensitive data be accessible to anyone regardless of designated access level. The design of the Inventory prevents such uploads. What is permissible is for select documents – only after having been reviewed and approved for release via multiple layers of concurrence - to be available on the internet made available to the public. These documents shall be of the sort that describes the data asset to include a data dictionary, relevant Privacy Act System of Records, and/or a technical document.
15. **Shall documents or files be uploaded into the Inventory?** No. That said, documentation such as data dictionaries and salient technical documents (if necessary) shall be made available via URLs. Who creates the necessary URL links? Each Administration and Staff Office shall be responsible for generating URLs.
16. **What is machine-readable and open format?** Information or data that is in a format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost. Source: [Home | resources.data.gov](#).
17. **What are the best formats to use?** Most generic formats provide an acceptable level of openness, including but not limited to.csv;.xml; and.json. While ISO 32000 is an open standard, PDF does not achieve the same level of openness and should not be used.
18. **Can the entire inventory be released under FOIA?** Yes, the inventory listing can be released. The individual data assets shall be evaluated on a case by case basis. The entire inventory is located on the internet at www.va.gov/data.
19. **Who shall be responsible for responding to the received FOIA?** The responsible program as designated by the data asset's program code.
20. **What organization is responsible for coordinating the Department's Open Data efforts?** OEI is responsible for coordinating the Department's efforts. The effort is being facilitated by the ODWG which is operating under the auspices of the Department's Data Governance Council.
22. **Does OMB provide direction?** In addition to Executive Order 13642 and OMB M-13-13, VA receives updated initiative information via the Project Open Data website and participation in OMB/GSA Open Data meetings. Additionally, the effort is discussed at a strategic level during the respective sessions of the CDO Council, Chief Financial Officer Council and Chief Information Officer Council.

23. **How does OMB measure Agency performance for Open Data policy implementation?** The goal is to provide more clarity to Agencies and the public on the OMB evaluation process and progress on the implementation of OMB M-13-13, Open Data Policy-Managing Information as an Asset. The Project Open Data Dashboard compares Agency performance of Open Data policy implementation across the Federal Government. The Project Open Data Dashboard is located at: <http://labs.data.gov/dashboard/offices>. The OMB Leading Indicators Strategy Rubric that informs the quarterly Project Open Data Dashboard is online at: <http://labs.data.gov/dashboard/docs/rubric>. This rubric provides additional transparency on the OMB staff review process.
25. **When an individual data asset is identified, where do I store it?** Each Staff Office and Administration shall individually input the data asset and associated metadata in the data asset collection tool. This shall require that each Staff Office and Administration secure "edit" access.
26. **What are my office's data quality responsibilities?** Each Staff Office shall be responsible for the quality of its metadata, to include, but not be limited to, a quarterly assessment of posted links.
27. **Rather than individually inputting the data asset, can my organization upload a quarterly spreadsheet to data asset collection tool?** Each identified data asset shall be individually inputted into data asset collection tool.
28. **Who do I contact if I have a technical problem with Open Data input?** Please send an email to OITOpenData@va.gov.
30. **Shall there be multiple JSON files or one JSON file?** One JSON file composed only of the public data asset listing shall be posted to www.va.gov/data.json . The data asset collection tool shall automatically generate an updated file daily.
31. **Does my organization need to generate a quarterly JSON file?** No, individual Administrations and Staff Offices shall not be generating a JSON.
32. **Does my organization need an account to access data asset collection tool?** An account is required. Please send an email to OITOpenData@va.gov.
33. **What is a bureau code?** Where can I find my bureau code? Definitions for a bureau code and definitions of VA's bureau codes are in Metadata.
34. **What is a program code?** Where can I find my program code? Definitions for a program code and VA's program codes are in Metadata.
35. **Can I input my own code?** Since all codes are OMB-prescribed, no other inputs are permissible.
36. **What if I can't find a precise program code?** While Administrations for the most part have detailed program codes, there may not be a 1 :1 match between

Administration program and OMB program code. If the outlined codes are not suitable for your program, please default to the program codes for your parent organization.

37. **What is CKAN?** An OMB and GSA-provided inventory tool suitable for customization which is available to agencies for their Open Data inventory efforts. Usage of this is optional, not mandatory.
38. **Is VA using CKAN?** VA was using CKAN until August 31, 2017, at which time VA migrated to Granicus Open Data, based on DKAN, an open source data management platform. CKAN allowed VA's data asset entries to populate GSA's hosted CKAN data asset collection tool. In 2020, VA migrated DKAN to the Socrata platform.
39. **What is Socrata?** It is a data storytelling and insights platform built to add value to Open Data efforts by enabling and empowering content creation in an easy-to-use, streamlined way, including the automated creation of documentation for Application Programming Interfaces. To achieve this, Socrata uses a combination of data visualization tools and integrations to create Data Stories and Dashboards that deliver the message behind data in a user-friendly interface. Built on traditional data portal features, Socrata offers a market standards-compliant experience that increases the user value of any Open Data effort. Socrata is a Software-as-a-Service (SaaS) tool.
40. **What is DKAN?** DKAN is an open-source data management platform that treats data as content to publish, manage, and maintain Open Data.
41. **What is Data.gov?** Data.gov is the complete data asset listing for all Federal Agencies. Anyone in the world may access content through a web browser. Data.gov provides search functionality to help people find data assets. For more information on Data.gov, please see: [How to get your Open Data on Data.gov | resources.data.gov](#).
42. **Shall there be two concurrent data asset repositories: Open Data and Data.gov?** Open Data is the concept the Federal Government is implementing, and it is hosting all Federal agencies' asset inventories at: www.data.gov. VA's Open Data page is located at: www.va.gov/data or <https://www.data.va.gov>. VA's Open Data page includes data stories, data visualizations, and user training.

APPENDIX C - THE VALUE OF OPEN DATA.

[The Value of Open Data](#)

[Department of Veterans Affairs Open Data Portal](#)