

**Disclaimer:** The following document does not offer legal advice nor is it intended to create any additional legal obligations. The document is intended provide guidance on geospatial data asset standards.

The Development Data Library (DDL) is USAID’s public repository of Agency-funded, machine readable data. The draft language below is intended to standardize the submission of geospatial data assets to the DDL. USAID welcomes any feedback or input on this language. Feedback can be shared via [OpenData@usaid.gov](mailto:OpenData@usaid.gov).

**\*\*\* DRAFT \*\*\***

## **Guidance on Submitting Geospatial Data Assets to the USAID Development Data Library**

USAID partners shall submit Geospatial Data Assets to the Development Data Library (DDL). The purpose of this guidance is to define the type of Geospatial Data Assets that should be submitted and how the assets should be submitted.

### **I. Definition of Geospatial Data Asset Types**

#### **A. Geospatial Data Files**

Geospatial Data Files identify the geographic location and characteristics of natural, demographic, cultural, political, or constructed features on the Earth. For example, datasets containing the demographic statistics associated with a subnational administrative unit or the location of schools are considered Geospatial Data Files. Geospatial Data Files include datasets that are stored in file formats traditionally associated with geospatial data, such as a Shapefile (.shp) or GeoTIFF (.tif), as well as other file formats that may not be commonly associated with geospatial data, such as a CSV (.csv). For example, a CSV (.csv) file that includes descriptive information about geographic features or locations is considered a Geospatial Data File. When derived Geospatial Data Files are generated by processing or analyzing primary Geospatial Data Files, both the derived and primary Geospatial Data Files should be submitted.<sup>1</sup>

#### **B. Geospatial Analysis Files**

Geospatial Analysis Files are files that describe or execute data processing or analysis methods. They include files as simple as a text document that describes the steps in a data processing and/or analysis workflow or scripts/tools/models that actually execute a data processing and/or analysis workflow. Geospatial Analysis Files shall be submitted

---

<sup>1</sup> See the definition of “Complete” under the Open Data in section I. Definitions of OMB M-13-13: <https://project-open-data.cio.gov/policy-memo/#i-definitions>

in addition to the Geospatial Data Files because they are required for an end user to understand and replicate the data processing and/or analysis that was performed. For example, a script written in the Python programming language and used to manipulate or analyze data is considered a Geospatial Analysis File and should be submitted.<sup>2</sup>

### C. Geospatial Data Visualization Files

Geospatial Data Visualization Files include files that are used to generate hard copy, digital, or web-based geospatial data visualization products and any other relevant files that are required for an end user to replicate a data visualization product. For example, an Esri Map Document (.mxd), QGIS Project Document (.qgs) or Adobe Illustrator file (.ai) that was used to create a digital map is considered a Geospatial Data Visualization File. Additionally, files written in programming languages that are used to create interactive web-based data visualization products, such as HTML, CSS, or Javascript, are also considered Geospatial Data Visualization Files. In many cases, multiple files are required to make a geospatial data visualization product. In these cases, each component of the data visualization product should be submitted so that it can be fully replicated by an end user.

## **II. Geospatial Data Asset Submission Requirements**

### A. Geospatial Data Files

- 1) Both primary and derived Geospatial Data Files shall be submitted.<sup>3</sup>
- 2) To fulfill the submission requirements, Geospatial Data Files shall be submitted in Shapefile (.shp), GeoTIFF (.tif), CSV (.csv), or GeoJSON (.geojson) file formats.
- 3) In addition to fulfilling the submission requirements noted above, Geospatial Data Files that are stored in other file formats that may be useful to end users can be voluntarily submitted in those file formats. For example, Geospatial Data Files stored in a File Geodatabase (.gdb) or as a Web Map Service (WMS) can also be submitted.
- 4) Geospatial Data Files that are stored in Shapefile (.shp) or GeoTIFF (.tif) formats shall include metadata that follows ISO 19115 using the ISO 19139 XML implementation schema.<sup>4</sup>
- 5) Geospatial Data Files that are stored in a CSV format and contain latitude and longitude coordinates shall include information regarding: 1) the method used to create the latitude/longitude coordinates; and 2) the spatial reference that was used to generate the latitude/longitude coordinates. Additionally, the dataset should include a column containing values that indicate the geographical precision of the latitude/longitude coordinates using the International Aid Transparency Initiative standards.<sup>5</sup> For

---

<sup>2</sup> See the definition of “Described” under the Open Data in section I. Definitions of OMB M-13-13: <https://project-open-data.cio.gov/policy-memo/#i-definitions>

<sup>3</sup> See the definition of “Complete” under the Open Data in section I. Definitions of OMB M-13-13: <https://project-open-data.cio.gov/policy-memo/#i-definitions>

<sup>4</sup> For example, see the instructions on changing the metadata style in Esri ArcGIS Desktop: <http://desktop.arcgis.com/en/desktop/latest/manage-data/metadata/choosing-a-metadata-style.htm>

<sup>5</sup> For example, see the International Aid Transparency Initiative Geographical Precision coding system: <http://iatistandard.org/201/codelists/GeographicalPrecision/>

example, a latitude/longitude coordinate may represent an exact location, the center of a populated place, or the center of subnational administrative unit.<sup>6</sup> This information should be included as an attachment to the actual CSV dataset.

- 6) Geospatial Data Files that are stored in Shapefile (.shp), GeoTIFF (.tif), or GeoJSON (.geojson), JSON (.json) file formats shall use the Geographic Coordinate System World Geodetic System 1984 (GCS WGS 1984) spatial reference. If the use of a different spatial reference was required for an analysis, the spatial reference that was used should be indicated elsewhere in the metadata.

#### B. Geospatial Analysis Files

- 1) Geospatial Analysis Files that are written in programming languages shall include a description of the language they are written in (for example, a Python file would be .py, a Javascript file would be .js, etc.), the intended purpose of the file, use limitations, and execution instructions.
- 2) In addition to fulfilling the submission requirements noted above, Geospatial Analysis Files shall include descriptive comments and relative pathnames to enable efficient and effective use by an end user.
- 3) In addition to fulfilling the submission requirements noted above, if Geospatial Analysis Files are stored in an online repository, such as GitHub, these files can also be submitted as a clone of the repository.

#### C. Geospatial Data Visualization Files

- 1) To fulfill the submission requirements, Geospatial Data Visualization Files created using desktop GIS software shall be submitted in Esri Map Document (.mxd) or QGIS Project (.qgs) file formats.
- 2) In addition to fulfilling the submission requirements noted above, Data visualization products that are finalized in separate desktop software or stored in a different file format can also be submitted if they may be of use to an end user. For example, when a data visualization product is stored as a Map Package or is finalized in Adobe Illustrator, the Map Package (.mpk) or Adobe Illustrator (.ai) file can be additionally submitted.
- 3) Geospatial Data Visualization Files written in a programming language and used to generate web-based data visualization products shall include an indication of the programming language (for example, a Python file would be .py, a Javascript file would be .js, etc.)
- 4) Geospatial Data Visualization Files shall reference relative pathnames to the source of all Geospatial Data Files that are displayed in the data visualization product.
- 5) When products created with the Geospatial Data Visualization Files are included in documents that are submitted to the USAID Development Experience Clearinghouse (DEC), such as a report that includes a map created with a Geospatial Data Visualization File, the URL that links to the location of the document in the DEC should be provided.

**\*\*\* DRAFT \*\*\***

---

<sup>6</sup> See the definition of “Described” under the Open Data in section I. Definitions of OMB M-13-13: <https://project-open-data.cio.gov/policy-memo/#i-definitions>