

## **Chesapeake Bay Program Guidance and Policies for Data, Information and Document Deliverable Submission**

This document describes the guidelines and policies governing the submission of electronic outputs to the Chesapeake Bay Program. The full guidance document "Chesapeake Bay Program Guidance for Data Management" is available at <https://www.chesapeakebay.net/what/data>. Below are excerpts from that guidance.

### Applicability

These guidelines and policies must be followed by all agencies, institutions, and organizations participating in data and information collection, processing, document generation and submittal to the Chesapeake Bay Program under grant or cooperative agreement funding. The Chesapeake Bay Program has adopted these guidelines and policies in order to improve coordination, compatibility, standardization, and information access across all the Bay Program partners.

In addition to these guidelines and policies, any activities funded with Federal Government funds must also adhere to applicable federal guidelines, policies, and executive orders, such as the [Federal Information Processing Standards \(FIPS\)](#) and Executive Order 12906.

### CBP Data/Information Management and Document Outputs Guidelines and Policies

The Chesapeake Bay Program has adopted the following guidelines and policies pertaining to data and information collection, processing, document generation and submittal to the Chesapeake Bay Program under grant or cooperative agreement funding. Any deviations from these guidelines and policies must be documented in the workplan and approved by the EPA Project Officer. Specific guidelines and policies include:

- Data, Information, and Document Deliverable Requirements
- Deliverable Serving vs. Submission Policy
- Locational Data Policy
- Metadata Policy
- Common Station Names Guideline
- Common Data Dictionary Guideline
- Common Database Design Guideline
- Calendar Date Policy
- Common Method Codes Guideline
- Data Reporting Guideline
- ITIS Biological Nomenclature Policy

### *Data, Information, and Document Outputs Requirements*

Recipients are required to submit data, information, and document deliverables in electronic format unless exceptions are specified in the grant or cooperative agreement workplan.

All data, information, and documents funded by the Chesapeake Bay Program whether through direct Chesapeake Bay Program funding or indirect matching funds are public information and shall be made available to the public unless there is a grant/cooperative agreement award condition that specifies otherwise. In addition, source data collected and processed in the creation of a deliverable shall also be submitted when practical. If source data is submitted, it *shall* also be delivered in electronic format. All outputs must have associated metadata.

**The table below describes suggested formats for electronic deliverables to be submitted to CBP. Alternate formats can be accepted with prior approval.**

<b>Document Type</b>	<b>Acceptable Formats</b>
<b>Text</b>	<b>Preferred:</b> Microsoft Word (DOC; DOCX) Portable Document Format (PDF)* ASCII Text Extensible Markup Language (XML)
<b>Spreadsheet</b>	<b>Preferred:</b> Microsoft Excel 2003 or higher Comma Separated Values (CSV)
<b>Database</b>	<b>Preferred:</b> Microsoft Access 2003 or higher Comma Separated Values (CSV) Microsoft SQL Server 2008 or higher Extensible Markup Language (XML) ASCII delimited text files
<b>Graphics</b>	<b>Preferred:</b> PNG TIFF GIF JPEG SVG PDF
<b>Geographic Information System</b>	<b>Preferred:</b> File geodatabase ESRI compatible raster formats, <b>raster formats compatible with open-source GIS software such as QGIS</b> , shape files, GeoTIFF
<b>Code</b>	<b>Preferred:</b> <b>Python and R</b>

*\* Data tables within PDF documents must be delivered in one of the spreadsheet formats.*

### *Deliverable Serving vs. Submission Policy*

Recipients can submit deliverables directly to the Chesapeake Bay Program or provide deliverables from their organizational data/web servers. State grant recipients shall submit non-point source BMP data using EPA's National Environmental Information Exchange Network (NEIEN). Recipients who plan to directly serve their grant/cooperative agreement outputs through their own data server/web site must include relevant documentation, including access instructions, within their workplan.

Recipients intending to deliver applications to be hosted by the Chesapeake Center for Collaborative Computing (C4) must coordinate directly with the CBP Data Center to ensure those applications are within the C4's capacity to operate and maintain. The technology used to develop those applications must be included in the CBP baseline technology portfolio. Recipients delivering computer applications must participate in the CBP data center release and deployment planning boards. This requirement only applies when client/server or web applications are submitted as part of the deliverables with the expectation that CBP will host those applications within C4.

### *Locational Data Policy*

The Chesapeake Bay Program adheres to the EPA's national geospatial data policy that requires consistent use of latitude/longitude coordinates to identify the location of entities. All data, containing spatial and/or specific geographic locations, collected or assembled under a Chesapeake Bay Program grant or cooperative agreement or to be served on the Internet via the Chesapeake Center for Collaborative Computing, must have latitude and longitude information, as well as datum information, for each entity.

### *Metadata Policy*

The Chesapeake Bay Program has adopted the policy, consistent with Presidential Executive Order 12906, that all data generated or collected using federal funds, submitted to the Chesapeake Bay Program, or served on the Internet, shall be accompanied by metadata fully conforms to the [Federal Geographic Data Committee's requirements for metadata](#). Metadata created for Chesapeake Bay Program deliverables may also be delivered to other federal Clearinghouses. Guidance for completing compliant metadata is documented in [Content Standard for Digital Geospatial Metadata](#). **The USGS offers [Metadata Wizard](#), a free tool that can be downloaded for creating metadata.**

### *Common Station Names Guideline*

The Chesapeake Bay Program has adopted the guideline that all data generated, collected for, or submitted to the Chesapeake Bay Program shall utilize a consistent set of common station names for identifying and reporting monitoring station locations. It is the data provider's responsibility to comply with this guideline.

The purpose of this guideline is to create one master table of station names, to the extent possible, to reduce confusion among cooperating agencies. The Station Names table, maintained on the [Chesapeake Bay Program website](#), should serve as the master list. Updates to this table that are required by data submitters shall be coordinated with the CBP Water Quality Data Manager.

#### *Common Data Dictionary Guideline*

The Chesapeake Bay Program has adopted the guideline that all data generated, collected for, submitted to the Chesapeake Bay Program shall utilize the CBP common data dictionary for defining all data elements and units of measure, **including all attributes present in spatial data**. It is the data provider's responsibility to comply with this policy.

The purpose of this guideline is to create one data dictionary, to the extent possible, to reduce confusion among cooperating agencies. Updates required by data submitters to the dictionary shall be coordinated with the CBP Water Quality Data Manager.

#### *Calendar Date Policy*

The Chesapeake Bay Program has adopted the standard that all data generated, collected for, or submitted to the Chesapeake Bay Program shall adhere to the Federal Information Processing Standard, Representation for Calendar Date and Ordinal Date for Information Interchange (FIPS PUB 4-1).

The standard states, "For purposes of electronic data interchange in any recorded form among U.S. Government agencies, National Institute of Standards and Technology (NIST) highly recommends that four-digit year elements be used". The year should encompass a two-digit century that precedes, and is contiguous with, a four-digit year-of-century (e.g., 1999, 2000, etc.). In addition, optional two-digit year time elements specified in ANSI X3.30-1985(R1991) should not be used for the purposes of any data interchange among U.S. Government agencies. Therefore, it is required to report and store all dates using four digits for the year.

#### *Common Method Codes Guideline*

The Chesapeake Bay Program has adopted the guideline that all data generated, collected for, or submitted to the CBP shall utilize the CBP Method Codes tables. The method codes are defined in the [Guide to Using CBP Water Quality Monitoring Data](#). It is the data provider's responsibility to comply with this guideline that requires the use of standardized method codes, to the extent possible, to simplify data coding and sharing.

The methods used by monitoring agencies and analytical laboratories are critical in providing accurate measurements. Therefore, capturing the methods is a high priority during database development. Modifications to the CBP Method Codes shall be coordinated with the Data Center to maintain

consistency in the table contents. If CBP agencies do not have a pre-defined method code that is acceptable for the work being conducted, the grantee/contractor should work with the funding agency to develop method codes that suits the requirements of the work, while maintaining maximum compatibility with other CBP codes.

#### *Numeric Data Reporting Guideline*

The Chesapeake Bay Program has adopted the guideline that all data generated, collected for, or submitted to the Chesapeake Bay Program shall report numeric data elements at the same level of precision as that of the original measurement. The exact precision of recorded values must be maintained. This guideline has a significant impact on data analysis and the decisions made based on these analyses.

Values should not be zero-filled to greater precision than actually recorded. For instance, if the measured value is 0.03, then the reported value should be 0.03 and not 0.030, which would imply precision to the third decimal place. For values that are recorded as below or above detection, a detection flag (in a separate data field) shall be used to identify the value as below or above the detection limit of the method, and the value shall be reported as the detectable limit. Values should be reported as zero, only if the measured or recorded value is zero. Values that are missing shall be reported as missing or null or nil, to identify values that were sampled but no value was obtained. Missing, null, or nil values are different than those that were never sampled, which should be recorded as a blank field, if they are recorded at all. It is the responsibility of the data submitter to record in the metadata, how measurements are coded, as well as the accuracy of the measurements.

It is important to note that some software tools used in data processing may represent the data internally with more precision than the original measurement, and/or may round the value. For instance, even though a value of 0.3 was entered, the value may be stored and reported as 0.299999.

#### *ITIS Biological Nomenclature Policy*

The Chesapeake Bay Program has adopted the policy that all data generated, collected for, or submitted to the Chesapeake Bay Program shall utilize the [Integrated Taxonomic Information System](#) biological names for identifying and reporting species. It is the data provider's responsibility to comply with this policy.

The purpose of this policy is to create one master table of species names, to the extent possible, to reduce confusion among cooperating agencies. The ITIS taxonomy table, maintained on the ITIS web site, should serve as the master list. Updates to this table that are required by data submitters shall be coordinated with the Data Center to maintain one consistent species name list.