

CONSENSUS METADATA STANDARD:

Identifying Chemical and Physical Variables and Descriptive Metadata

REVISION DATE: October 3, 2001

CHECK FOR LATEST VERSION ONLINE: Before using this document, and periodically thereafter, please check for an updated version at <http://cdiac.esd.ornl.gov/programs/NARSTO/metadatastandards>. The latest version will be called consensus_naming_standard_yyyymmdd.pdf.

SCOPE

Provides the approach for identifying (i.e., naming) chemical and non-chemical/physical measured variables and various descriptive metadata elements

Applies to US EPA PM Supersites projects and NARSTO affiliated projects in general.

SUMMARY

This standard points to references tables of names for chemical and non-chemical/physical variables and various metadata elements to use in the Data Exchange Standard files. Data providers are expected to use these tables to determine the appropriate identifiers for the chemical substances, physical properties, and metadata elements (e.g., date, time, locations) they are measuring or modeling.

SPECIFICATIONS

Identifying Chemical Substances with a CAS¹ Registry Number:

Mandatory requirements

- CAS Number (with "C" prefix)
 - The "C" prefix prevents spreadsheet programs from converting some CAS numbers to dates.
- Chemical name (either CAS-9CI, IUPAC, or other common name)
- See Supersites reference table (Attachment 1)

¹ The CAS Registry Number and the CAS-9CI name (Chemical Abstracts Service, 9th Collective Index Nomenclature) are the copyrighted property of the American Chemical Society. The NARSTO QSSC has the permission of CAS to use this information in NARSTO archive data sets. By extension, EPA Supersites Projects and NARSTO affiliated projects may incorporate CAS numbers and CAS-9CI names into data being processed for NARSTO archiving. Furthermore, the use of CAS numbers and CAS-9CI names is permitted as required in supporting regulatory requirements and/or for reports to Government Agencies and in copyrighted scientific publications when the CAS information are incidental to the publication. Any use or redistribution other than that described here is not permitted without the prior, written permission of the American Chemical Society. Please contact the QSSC if you have any questions about the use of CAS information.

Preferred Nomenclature Standards, General

- CAS-9CI (Chemical Abstracts Service, 9th Collective Index Nomenclature)
- IUPAC Nomenclature
- See Supersites reference table (Attachment 1)

Preferred Nomenclature Standard, Polycyclic Aromatic Compounds (PACs)

- IUPAC name, as implemented by NIST/IUPAC search tool
<http://ois.nist.gov/pah/>
- See Supersites reference table (Attachment 1)

Identifying Chemical Substances/Measurements/Calculated Quantities that do not have a designated CAS Registry Number:

Mandatory requirement

- Formal syntax with key phrase and detailed modifier if needed, separated by a ":"
- Supersites reference table (Attachment 2)

Identifying Physical/Non-chemical Measurements:

Mandatory requirement

- Formal syntax with key phrase and detailed modifier if needed, separated by a ":"
- Supersites reference table (Attachment 3)

Identifying Metadata Elements:

Mandatory requirement

- Formal syntax with key phrase and detailed modifier if needed, separated by a ":"
- Supersites reference table (Attachment 3)
- See the NARSTO Data Exchange Standard (DES) file format template (Attachment 4) to see how these are used..

Reference Tables:

- The Supersites Data Management Working Group will compile and maintain reference tables for chemical substances with CAS numbers (See Attachment 1), chemical substances/measurements/calculated quantities without CAS numbers (See Attachment 2), physical/non-chemical variables (See Attachment 3) and various metadata elements (See Attachment 3) for use by data originators.
- Additions or other changes to the reference tables will be made by consensus of the Data Management Working Group. Refer to "Changes and Additions to the Reference Tables" later in this document.

GUIDANCE

Identifying Chemical Substances with a CAS Registry Number:

Along with the required CAS registry number (with a C prefix), the data provider must provide a chemical name. The preferred nomenclature is either the CAS-9CI or the IUPAC name. A widely recognized other common name would also be acceptable.

The Data Exchange Standard format read-and-verify QA software program will check that the CAS number is valid and in the reference table and compare the provided name to both the CAS-9CI and IUPAC names. If it doesn't match either of those names a warning message (flag) will be generated but processing of the data file will proceed. Those flagged names will be reviewed for any obvious errors.

For Polycyclic Aromatic Compounds (PACs), subject matter experts recommend that the IUPAC name, as implemented by the NIST/IUPAC search tool <http://ois.nist.gov/pah/>, is the preferred nomenclature for PACs and should be the name provided with the data.

Identifying Chemical Substances/Measurements/Calculated Quantities that do not have a designated CAS Registry Number and Identifying Physical/Non-chemical Measurements:

We strongly request that these variable names and units as shown in the reference tables (Attachments 2 and 3, respectively) be used. Proposed additions to the reference tables should follow the same syntax and be referred to your Site Data Coordinator.

Identifying Metadata Elements:

The appropriate identifiers for metadata elements including site information, locations, dates, times, and sampling conditions are as shown in the reference table (Attachment 3).

Metadata elements (e.g., key characteristics) are, for the most part, defined in the Data Exchange Standard (DES) input template and accompanying instructions (Excel workbook, Attachment 4). Values for any project-specific metadata fields in the DES header can be obtained from your Site Data Coordinator. Numerous metadata fields located in the main data table have an associated picklist of controlled values. Data providers should use these values. Proposed changes or additions to the picklists should be referred to your Site Data Coordinator.

Additional guidance on data reporting, metadata, the DES, and the DES template may be obtained from the NARSTO Data Management Handbook (<http://cdiac.esd.ornl.gov/programs/NARSTO/narsto.html#gsmp>)

Changes and Additions to the Reference Tables and Picklists:

Site investigators and data users are encouraged to work with their Site Data Management Coordinators to suggest improvements in and additions to the reference tables. Additions or changes to the reference tables should be recommended by a Site Data Management Coordinator to the Data Management Working Group and QSSC for discussion and consensus. The chemical name table will be kept in sync with EPA Environmental Data Registry (EDR) Chemical Registry System (CRS) tables.

IMPLEMENTATION

Reporting CAS Numbers in a Data File

- CAS numbers with a "C" prefix are used in the archive data files. This prevents Excel from inadvertently converting some CAS numbers to nonsensical dates when a .csv file is opened in Excel by default. The "C" identifies the field as text.
- Exclude leading zeros from CAS numbers.

Examples:

C10028-15-6	Ozone
C106-98-9	1-Butene

No Site Data Coordinator

Should your project not have a designated Site Data Coordinator, or the Coordinator is unavailable, please contact the NARSTO QSSC through the web site referenced below.

Identifying Chemical Substances, Calculated Quantities, and Physical/Non-chemical Measurements that do not have a designated CAS Registry Number:

Attachments 2 and 3 are the reference tables for these variables and contain the valid values in the correct format --with a key phrase and detailed modifier if needed, separated by a ":". For example, PM10: area, PM10: count, PM10: mass, and Temperature: air, Temperature: dew point, Temperature: virtual, Temperature: wet bulb. Additional variables may be added following this format.

Identifying Metadata Elements

The valid values for metadata elements including site information, locations, dates, times, and sampling conditions are as shown Attachment 3. The correct format is a key phrase and detailed modifier if needed, separated by a ":". For example, Date start: local time, Date end: local time, and Latitude: decimal degrees, Longitude: decimal degrees. Additional metadata elements may be added following this format (see Change and Additions.... in Guidance section).

REFERENCES

EPA EDR Chemical Registry System (CRS) (<http://www.epa.gov/crs/index.htm>)

Sander, Lane C. and Stephen A. Wise. 1997. Polycyclic Aromatic Hydrocarbon Structure Index. NIST Special Publication 922 available online at <http://ois.nist.gov/pah/>. Chemical Science and Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.

DEVELOPMENT BACKGROUND

The EPA Supersites program recognized, in March 2000, a need for standardization of key metadata including names for chemical variables and descriptive metadata. The Supersites Data Management Working Group has discussed naming in many conference calls (see <ftp://cdiac.esd.ornl.gov/private/narsto/ssdmwg/minutes>) and sought the advice and consent of the Quality Assurance and Organics Working Groups. The standard was approved 20010731 by the Working Group and sent for PI review on 20010806. Send comments to Les Hook (hookla@ornl.gov)



Attachments

Attachment 1. Chemical Variables with CAS Numbers Reference Table

Attachment 2. Chemical Variables without CAS Numbers Reference Table

Attachment 3. Physical/Non-Chemical Variables Reference Table

Attachment 4. Data Exchange Standard (DES) Input Template

**All attachments are available on the QSSC web site -
<http://cdiac.esd.ornl.gov/programs/NARSTO/> .**

Check for latest version.