

## Review of ICARTT file format implementation and operational suitability

NASA's Earth Science Data Systems Standards Process Group (SPG) is considering the International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) file format for adoption as a community standard. The ICARTT file format was developed to fulfill the data management needs for the ICARTT campaign in 2004. This file format is text-based and composed of a header section (metadata) with critical data description information (e.g., data source, uncertainties, contact information, and brief overview of measurement technique), and a data section. Although it was primarily designed for airborne data, the ICARTT format proved to be practical for other mobile and ground-based studies and various data types. The ICARTT file format has since been widely accepted in the airborne field study community and used in recent major airborne studies sponsored by NASA, NSF, NOAA and international partners.

You are invited to review this Requests For Comment (RFC) in the context of your **implementation experience** with this data format specification and its **suitability for operational use**. You only need to answer questions applicable to you. Please send your completed review to:

[spg-rfc-019@lists.nasa.gov](mailto:spg-rfc-019@lists.nasa.gov).

*Implementation Experience questions:*

1. (*Your background*) Describe in a sentence or two your overall implementation experience related to the proposed specification. (*e.g., specification implementer, tools developer, data provider, scientific analyst, science user, etc.*) Have you directly implemented the ICARTT format specification? Did you use a pre-existing software package, and if so, what did you use?

No direct implementation experience.

I have previously reviewed format and find it reasonably complete. The **file format scanning software is great.**

**I prefer a more complete set of accompanying metadata and standard values for metadata fields. See NARSTO Data Exchange Standard <http://cdiac.ornl.gov/programs/NARSTO/> and specifically [ftp://narsto.esd.ornl.gov/pub/DES\\_metadata/var\\_names\\_web\\_sources/NARSTO\\_template\\_atmospheric\\_measurements.xls](ftp://narsto.esd.ornl.gov/pub/DES_metadata/var_names_web_sources/NARSTO_template_atmospheric_measurements.xls).**

2. (*Completeness*) Does the specification provide all the detail you need to implement it in software? (*e.g., to read or write a data file; to implement the specification, a profile or extension; or develop a tool such as a format translator*) If not, describe what is missing in the specification.

Yes.

3. (*Accuracy*) Do any parts of the specification contain inaccuracies, or internal

inconsistencies? If so, please provide details.

I believe comma separators are needed in example:

**Indep\_Var VarName\_1 VarName\_2 VarName\_3 ... VarName\_n**

4. (*Clarity*) Is any part of the specification ambiguous, or poorly explained? If so, please provide details.

**In the 2.3.B. File header information section under "Normal comments (SUPPORTING information):"** I would suggest that more "best practices guidance" or examples could be given.

See additional best practices comments below.

5. (*Usefulness*) How well does this specification meet your information sharing needs? (*e.g., does it work well with the data types and data manipulations in your application? Does it properly represent your datasets? What are the pros and cons of this data format?*)

See comment in 1.

6. (*Implementation*) What implementation challenges does the proposed standard present? Please provide details, if any.

**A more complete set of metadata would ensure that data can be understood well into the future.**

*Operational Suitability questions:*

7. Do you currently use or plan to use the ICARTT format in a production setting? Do you plan to distribute data in this format to science collaborators and other researchers?

No

8. Why do you choose to use the ICARTT format over other data formats for your applications?

Not applicable.

9. Does the ICARTT file format meet your requirements for storing and accessing data?

Not applicable.

10. Have you or your users encountered any difficulty when using some of the data in the ICARTT format? If you have, please provide a brief description of your experience.

Not applicable.

11. What operational challenges or limitations does ICARTT present? Please provide details.

I prefer that dates and times be given explicitly rather than as an offset from a starting point.

12. What benefits does the ICARTT file format present? Are there any drawbacks to using this file format? (*e.g., Does it offer the flexibility you want to package the data types in your applications? Does it facilitate interdisciplinary studies?*)

Not applicable.

13. How much data do/will you provide or archive in the ICARTT format? (*Number of distinct data sets, total data volume, number of files.*)

Not applicable.

14. How many users do you have or expect to have for data in the ICARTT format, and what is your expected user community?

Not applicable.

15. (*User comments*) Any additional comments, observations or criticisms of the ICARTT format and the RFC can be provided here.

Consistent with my comment about offering additional best practices guidance, I would offer some resources that could be included as a link in the “Informative references” section. See, [6] MILAGRO Data Policy and Management Plan: [http://www.eol.ucar.edu/projects/milagro/data/MILAGRO\\_DataPolicy.html](http://www.eol.ucar.edu/projects/milagro/data/MILAGRO_DataPolicy.html) .

As part of the MILAGRO project I developed “Guidelines for Using Consistent Names and Units in ICARTT Format Data Files” ([http://cdiac.ornl.gov/programs/NARSTO/MILAGRO\\_ICARTT\\_Names\\_20060316.pdf](http://cdiac.ornl.gov/programs/NARSTO/MILAGRO_ICARTT_Names_20060316.pdf))

**Note that references to NARSTO could be easily removed. \*\*\***

Specifically, the purpose of these guidelines is to provide a means for ICARTT data file format users to implement consistent variable names and units.

- These names and units (column headings) can be used to report measurements and metadata for various data types.

- These names and units have been constructed to be “friendly” to most data user software and data systems. The combined name and unit column headings will be no longer than 30 characters.
- The chemical names are cross referenced with CAS Registry Numbers and CAS Ninth Collective Index nomenclature when applicable. Names for other chemical and physical measurement types and metadata are cross referenced with NARSTO Data Exchange Standard names. In both instances, this permits a nominal cross reference to the NARSTO DES data file format.
- Cross references via the CAS Registry Numbers to other organizational standards can be added (e.g., EPA AQS Parameters).
- These designations for names and units of chemical and physical constituents and descriptive attributes (metadata) could be maintained by the standard management team.

\*\*\* If the standard management team would like to incorporate these names and units, either as links at their current locations or assume responsibility for the files, I would be available to assist with the transition.