

## **ESDS–RFC-019 Technical Working Group Final Report**

The ESDS-RFC-019 Technical Working Group (TWG) has completed a review of the International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) file format RFC with the following conclusion that:

**The Standards Process Group should endorse ESDS-RFC-019 (ICARTT file format) as a Recommended Standard.**

### **History**

The TWG conducted a review of the ICARTT file format specification RFC dated September 2009 from the perspective of implementation and operational suitability. The ICARTT file format was initially developed to fulfill the data management and broad collaborative research 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.

A set of review questions was adapted from the HDF5 and NetCDF 3 classic reviews. These can be found on the SPG web site (<http://www.esdswg.org/spg/rfc/esds-rfc-019>).

There were a total of 12 reviews received from the community that included data providers and managers, scientific analysts and programmers, and research scientists. The reviews were characterized with a predominantly positive response to ICARTT file format's ease of usage, well-described and straightforward implementation, ability to provide descriptive metadata along with the actual data in the files, and ensuring interoperability among the users of airborne datasets. Of particular interest, there was also a consolidated response from the NASA Airborne Science program that was positive.

### **Recommendation**

The TWG bases its recommendation on an analysis of the following factors.

#### **Strengths:**

A major strength of the ICARTT file format according to the reviewers was that it has provided an easy to use and standard approach to share airborne data sets to facilitate broad collaborative scientific research among the airborne measurement and atmospheric chemistry modeling and satellite communities. The file format improves upon previously used formats by the airborne science community and has been widely accepted and used

by that national and international airborne science community including NASA, NOAA, British, French and the German airborne science programs. In particular, it is well suited for exchanging and storing aircraft observations. For example, one data manager mentions the following in his review:

**"It is an excellent format for aircraft observations. It is readily readable by eye (ASCII data). It is flexible and can be read by many different applications..."**

Many of the reviewers have directly implemented the specification and all reviews confirmed the completeness of the file format specifications for implementations. Quite a number of reviewers mentioned the availability of tools to verify conformance to this file format and to visualize the data. For example, two independent reviewers comment:

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

**"The ICARTT specification detail has always been sufficient for my purposes"**

Many of the reviewers commented on the ease of the use of the ICARTT file format and the flexibility it provides for aircraft observations by including metadata along with the data itself. The format appears to be valuable for assuring interoperability between different user groups without regard to the sensor performing the measurements. Following are a sample of reviewer comments:

**"For my purposes, it offers a good amount of descriptive metadata without getting too bulky."**

**"The format is easily read automatically ... with all necessary information taken by the software from the file header. This works very well, and means the same code can be used to read files from different instruments etc. without any 'hard-wiring' for different files."**

#### **Weaknesses:**

Weaknesses of the ICARTT file format identified by the community were that 1) As an ASCII based format it is not as efficient as the binary formats such as NetCDF 2) The file format is not suitable for large 3-dimensional data sets such as models although it works very well for time series data and 3) the tools are not widely available as those for NetCDF. For example, according to one reviewer:

**"The format works very well for time-series data, however it is not suitable for use with several dimensions e.g. 3D model output. For this type of data, netcdf format performs best and is also well supported by software packages such as IDL."**

There are some idiosyncrasies in ICARTT's structural conventions, for example some fields are terminated with commas, others by colons or spaces. Date and time formats don't follow widely adopted standards.

Overall, the reviewers found few weaknesses or limitations of the ICARTT file format for its intended use in the earth science community.

### **Applicability:**

The demand for airborne observations continues to increase as atmospheric chemistry models become more complex and satellites observe a broader suite of atmospheric constituents. Such observations fill important needs for validation and assessment purposes as well as process-level understanding. In response, airborne field campaigns have become more complex and encompass a growing suite of observations often including multiple airborne platforms, dozens of instruments teams, and as many as 200 variables. These data are collected by a wide range of research teams and are recorded at different temporal resolutions. The ICARTT file format was initially developed as an airborne data file standard format to satisfy the data archival and exchange needs of a multi-agency, multi-national group of researchers conducting coordinated observations in pursuit of a common research goal. Thus, the design of the ICARTT format was specifically tailored to accommodate airborne observations and arose from a consensus established across the atmospheric chemistry community. The success of the ICARTT study propelled much wider use of the ICARTT data format for subsequent airborne campaigns, such as the Megacity Initiative: Local and Global Research Observations (MILAGRO) study conducted with interagency and international partners in 2006 as well as NASA's Intercontinental Chemical Transport Experiment – Phase B (INTEX-B) campaign. Most recently, the ICARTT file format was adopted for use by international atmospheric chemistry community to support observations during the International Polar Year for the **Polar** Study using Aircraft, Remote Sensing, Surface Measurements and Models, of Climate, Chemistry, Aerosols, and Transport (PolarCAT), including NASA's component - Arctic Research of the Composition of the Tropospheric from Aircraft and Satellites (ARCTAS)- and NOAA's contribution - Aerosol, Radiation, and Cloud Processes affecting Arctic Climate (ARCPAC). Based on the responses of the reviewers, data from thousands of flight hours have been archived in this format. According to the reviews, the ICARTT format will continue to be used in the future campaigns (i.e. **CalNex 2010** - An Air Quality and Climate Change Field Study in California in 2010) addressing chemistry and climate to exchange and archive data.

The ICARTT file format also has many tools developed by multiple-groups for accessing the data. Given the simplicity of the format, the files can be read and created using a single subprogram for multiple types of instruments. Tools and file scanning software are available on a wide variety of platforms and in multiple languages and packages (i.e. IDL, Matlab) that makes this file format very useful for the earth science community.

## **Limitations:**

As noted in the Weaknesses section, the ICARTT file format is not suitable for large 3-dimensional data sets. A recurring theme in the reviews is succinctly summarized by one reviewer:

**“Large datasets would be unwieldy in this format. However, most of the observational datasets are not large”**

While the file format includes metadata specification, the metadata fields themselves do not follow any other standards that may limit its use in large automated systems. Another limitation cited by a reviewer is that the files are not as quick to load as binary files as the ASCII files do not allow random access to data.

One reviewer did feel quite strongly that while ICARTT may serve the airborne science community well, it would be good for that community to consider migrating to more widely used standards.

## **RFC Overview**

The RFC can be downloaded from:

<http://www.esdswg.org/spg/rfc/esds-rfc-019/ESDS-RFC-019-v0.03.pdf>

Evidence of Implementation document:

[http://www.esdswg.org/spg/rfc/esds-rfc-019/EoI\\_ICARTT.pdf](http://www.esdswg.org/spg/rfc/esds-rfc-019/EoI_ICARTT.pdf)

The ICARTT RFC review responses can also be found in this folder:

<http://www.esdswg.org/spg/rfc/esds-rfc-019/implementation-review-comments>