

# International Solar-Terrestrial Physics (ISTP) Metadata Guidelines Status

```
ISTP Metadata
DEPEND_0: EPOCH
FILLVAL: -1.0E31
0100 UNITS: m/s
1001 FIELDNAM: Ion Velocity
0100 0110 LABLAXIS: Ion Velocity
1001 0101 ...
0110 11101 010011001110001010
1010 0100 1001001010100010011
1110 1001 0110001001001001010
0100 0110 0101001001001001010
1001 0101001001001001011
0110 1110111011111010101
0101001001001001011
1110111011111010101
```

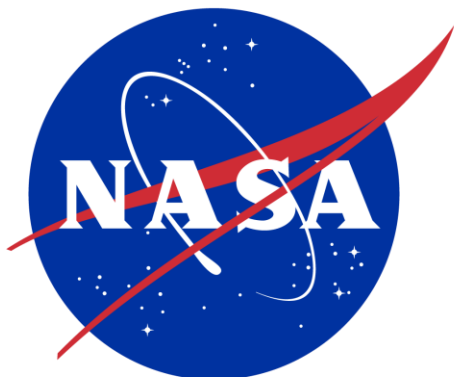
Andriy Koval (NASA/UMBC), Robert M. Candey (NASA), Sarah Fooks (NASA/Adnet), Robert E. McGuire (NASA, retired), Tamara J. Kovalick (NASA/Adnet), Bernard T. Harris (NASA), Lan Jian (NASA), Eric Grimes (NASA/Adnet) and others

Space Physics Data Facility (SPDF)

<https://spdf.gsfc.nasa.gov>

NASA Goddard Space Flight Center

*IHDEA 2024 Oct 18*



# History

- International Solar-Terrestrial Physics (ISTP) for coordinated, simultaneous investigations of the Sun-Earth space environment in 1990s
  - NASA
  - European Space Agency (ESA)
  - Institute of Space and Astronautical Science (ISAS) of Japan
  - Academy of Sciences (Russia)/Rosaviakosmos
- Science community developed the ISTP Metadata Guidelines and other conventions for describing and naming the datasets
- Later adopted by the Interagency Consultative Group (IACG)
- Missions add more attributes: Cluster, THEMIS, RBSP (PRBEM), MMS, etc.
  - [https://github.com/IHDE-Alliance/ISTP\\_metadata/blob/main/v1.0.0/Multi\\_Mission\\_Use\\_of\\_Attributes.md](https://github.com/IHDE-Alliance/ISTP_metadata/blob/main/v1.0.0/Multi_Mission_Use_of_Attributes.md)
  - [https://github.com/IHDE-Alliance/ISTP\\_metadata/blob/main/v1.0.0/Non-ISTP%20Mission%20Global%20Attributes.md](https://github.com/IHDE-Alliance/ISTP_metadata/blob/main/v1.0.0/Non-ISTP%20Mission%20Global%20Attributes.md)
  - [https://github.com/IHDE-Alliance/ISTP\\_metadata/blob/main/v1.0.0/Non-ISTP%20Mission%20Variable%20Attributes.md](https://github.com/IHDE-Alliance/ISTP_metadata/blob/main/v1.0.0/Non-ISTP%20Mission%20Variable%20Attributes.md)
- Widely-used data display and analysis tools depend on the guidelines
- Used to populate SPASE metadata

# ISTP Guidelines Structure and Metadata Concepts

- ISTP/IACG Guidelines (mid 1990s) and subsequent extensions by SPDF define implementation standards for CDFs and NetCDFs
  - Include general file naming conventions
  - Data is time-ordered and time-identified; times vary by record
  - Set of required and suggested metadata (details on next slide)
  - Variable attributes can point to other variables by name and carry arguments
    - Attributes thus carry information about relationships among variables
    - Variables can carry metadata (e.g., labels for dimensional variables)
  - Global attributes provide overall context of the dataset
  - Missions add their own metadata requirements
- **CDAWeb additional concepts: “Master” CDFs and “Virtual” Variables**
  - “Master” CDF is the use of a “skeleton” CDF (structure and metadata but no data) to insert supplemental or updated metadata for CDFs as a dataset
  - “Virtual” variables are computed variables, using specialized CDF attributes to link defined variables and routines within CDAWeb/CDAWlib

# ISTP Metadata Elements

- **Variable attributes required for automated processing:**

- CATDESC for longer variable description
- DEPEND\_0 points to time variables
- DEPEND\_1, 2, 3 point to variables that describe other dimensions
- FIELDNAM short variable name for plots and column headers
- FILLVAL values indicating missing or bad data
- LABLAXIS/LABL\_PTR for axis and column titles
- UNITS/UNIT\_PTR
- VALIDMIN/MAX for valid data range

- **Global attributes required for automated processing:**

- Data\_type identifies the data type of the dataset. Both a long name and a short name are given.
- Descriptor identifies the name of the instrument or sensor that collected the data
- Instrument\_type defines instrument type in CDAWeb
- Logical\_source carries source\_name, descriptor, and data\_type information
- PI\_affiliation and PI\_name for acknowledgements
- Source\_name generally is spacecraft name

# ISTP Metadata Tools and Usage

- Current ISTP/IACG/SPDF Guidelines  
[https://spdf.gsfc.nasa.gov/sp\\_use\\_of\\_cdf.html](https://spdf.gsfc.nasa.gov/sp_use_of_cdf.html)
- Move updates to Github in markdown  
[https://github.com/IHDE-Alliance/ISTP\\_metadata/tree/main/v1.0.0](https://github.com/IHDE-Alliance/ISTP_metadata/tree/main/v1.0.0)
- SKTeditor metadata creation tool <https://spdf.gsfc.nasa.gov/skeditor>  
originally in Java and now in Javascript by Eric Grimes
- PDS added CDF-A as a standard format, which is CDF with ISTP Guidelines and two SPASE attributes, but no compression or sparse variables
- ISTP metadata independent of CDF and easily used in other self-describing science formats like netCDFs used by GOES, ICON, GOLD
- Added SPASE and DOI global attributes to CDAWeb datasets via Master CDFs and shown in CDAWeb interface
- CDF Time variable types, especially CDF\_TIME\_TT2000 nanoseconds from J2000 in Terrestrial Time in 8 byte integer and properly handles leap seconds and is well-defined

# Development

- Formalize Steering Committee
- Define governance and update processes
- Review version on Github, add general dataset creation recommendations and lessons-learned
- Create Github issues for various additions, such as author list for DOIs, DOI, Variable\_display\_order, Variable\_display\_indent\_level, Associated\_parent\_variable, Dataset\_group, Mission\_parent)
- Adopt mission additions like for Cluster/Solar Orbiter: Representation, Tensor\_order, Coordinate\_systems, Rotation\_matrices, Unit\_quaternion
- Add content from **mission-specific documents** that reference the ISTP guidelines
- Consider requirements specific to model results
- Add recommendations for unit representation in addition to current UNITS used for labels, such as UDunits and VOUnits
- Add explanatory material from the CF Conventions (<https://cfconventions.org/>) that also apply in heliophysics
- Add crosswalk with SPASE metadata (Fung et al. (2023) doi.org/10.1016/j.asr.2023.09.066)
- Stay flexible for interactions with missions and enabling framework for CDAWeb services

# ISTP Global Attributes to SPASE

Table B1  
Comparison of global attributes for numerical data between the SPASE metadata model and ISTP Guidelines.

## Metadata Mapping - CDF Global Attribute to SPASE Numerical Data

CDF Global Attribute	SPASE Numerical Data Mapping	Edit?	Non CDF Metadata Sources, Processing Programs, Misc. Notes
LOGICAL_SOURCE	NumericalData/ResourceID		SMWG, Spacecraft lookup table: cdaweb_sc_list.tab, stream editing
LOGICAL_SOURCE	NumericalData/ResourceID, NumericalData/Parameter/Cadence	Yes	SMWG, cdaweb_spase_map_cadence.pro
LOGICAL_SOURCE_DESCRIPTION Title	NumericalData/ResourceHeader/ResourceName	Yes	Hand edits as required
DESCRIPTOR	NumericalData/ResourceHeader/ResourceName	Yes	Hand edits as required
TEXT	NumericalData/Parameter/Caveats	Yes	Hand edits as required
PI_NAME	NumericalData/Parameter/Caveats	Yes	Hand edits as required
PI_NAME	NumericalData/ResourceHeader/Contact/Name		Stream editing
ACKNOWLEDGEMENT	NumericalData/ResourceHeader/Contact/PersonID		Stream editing
ACKNOWLEDGEMENT	NumericalData/ResourceHeader/Contact/Acknowledgement	Yes	Stream editing, Hand edits as required
PI_NAME	NumericalData/AccessInformation/Acknowledgement	Yes	Stream editing, Hand edits as required
Not applicable	NumericalData/AccessInformation/Acknowledgement	Yes	Stream editing, Hand edits as required
Not applicable	NumericalData/AccessInformation/RepositoryID		SMWG
LINK_TITLE	NumericalData/AccessInformation/AccessURL/URL		URLs set to match the SPDF CDF directory tree structure
LINK_TEXT	NumericalData/ResourceHeader/InformationURL/Name	Yes	Hand edits as required
HTTP_LINK	NumericalData/ResourceHeader/InformationURL/Description	Yes	Hand edits as required
PI_AFFILIATION	NumericalData/ResourceHeader/InformationURL/URL	Yes	Hand edits as required
LOGICAL_SOURCE	NumericalData/ResourceHeader/InformationURL/Acknowledgement	Yes	Hand edits as required
MISSION_GROUP	NumericalData/InstrumentID		SMWG
INSTRUMENT_TYPE	NumericalData/InstrumentID		SMWG
Not applicable	NumericalData/MeasurementType	Yes	Stream editing, Hand edits as required
Not applicable	NumericalData/TemporalDescription/TimeSpan/StartDate		Dates set by tracking of the CDAWeb data product CDF file content
Not applicable	NumericalData/TemporalDescription/TimeSpan/[Relative] StopDate		Dates set by tracking of the CDAWeb data product CDF file content
LOGICAL_SOURCE	NumericalData/TemporalDescription/Cadence	Yes	Hand edits as required, cdaweb_spase_map_cadence.pro
CAVEATS	NumericalData/Caveats		Stream editing
TITLE	NumericalData/Keyword		Stream editing
MISSION	NumericalData/Keyword		Stream editing
PROJECT	NumericalData/Keyword		Stream editing
DATA_VERSION	NumericalData/Keyword		Stream editing
DISCIPLINE	NumericalData/Keyword		Stream editing
DATA_TYPE	NumericalData/Keyword		Stream editing
ADID_REF	NumericalData/Keyword		Stream editing
GENERATION_DATE	NumericalData/Keyword		Stream editing
NSSDC_ID	NumericalData/Keyword		Stream editing
MODS	NumericalData/Keyword		Stream editing
SOFTWARE_VERSION	NumericalData/Keyword		Stream editing
GENERATED_BY	NumericalData/Keyword		Stream editing
RULES_OF_USE	NumericalData/Keyword		Stream editing
TEXT_SUPPLEMENT_1	NumericalData/Keyword		Stream editing
LOGICAL_FILE_ID	Used to cross check LOGICAL_SOURCE Metadata		Stream editing

5722

# ISTP Variable Attributes to SPASE

Table B2

Comparison of variable (parameter) attributes for numerical data between the SPASE metadata model and ISTP Guidelines.

**Metadata Mapping - CDF Variable Attribute to SPASE Parameter**

CDF Variable Attribute	SPASE Numerical Data Parameter Mapping	Edit?	NonCDF Metadata Sources, Processing Program, etc.
FIELDNAM	NumericalData/Parameter/Name	Yes	Often custom editing required
DEPEND_0	NumericalData/Parameter/Set		
CATDESC	NumericalData/Parameter/Set		
cdf_variable_info.name	NumericalData/Parameter/ParameterKey	No	Parameter Key populated without using Var. Attr.
VAR_NOTES	NumericalData/Parameter/Caveats	Yes	Hand edits as required
AVG_PTR_1	NumericalData/Parameter/Caveats		Stream editing
AVG_TYPE	NumericalData/Parameter/Caveats		Stream editing
VIRTUAL	NumericalData/Parameter/Caveats		Virtual Variable designation flag
FUNCT	NumericalData/Parameter/Caveats		Virtual Variable support metadata
FUNCTION	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_0	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_1	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_2	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_3	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_4	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_5	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_6	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_7	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_8	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_9	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_10	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_11	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_12	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_13	NumericalData/Parameter/Caveats		Virtual Variable support metadata
Component_14	NumericalData/Parameter/Caveats		Virtual Variable support metadata
MONOTON	NumericalData/Parameter/Caveats	Yes	Rarely needs editing
Not applicable	NumericalData/Parameter/Cadence	Yes	cdaweb_spase_map_cadence.pro via CDF Global Variable LOGICAL_SOURCE
TIME_RES	NumericalData/Parameter/Cadence		Stream editing
RESOLUTION	NumericalData/Parameter/Cadence		Stream editing
UNITS	NumericalData/Parameter/Units		Stream editing
SI_CONVERSION	NumericalData/Parameter/UnitsConversion		Set from UNITS value via stream editing
DICT_KEY	NumericalData/Parameter/CoordinateSystem/CoordinateSystemName		Stream editing
FRAME	NumericalData/Parameter/CoordinateSystem/CoordinateSystemName		Stream editing
COORDINATE_SYSTEM	NumericalData/Parameter/CoordinateSystem/CoordinateSystemName		Stream editing
DICT_KEY	NumericalData/Parameter/CoordinateSystem/ CoordinateSystemRepresentation		Stream editing
FRAME	NumericalData/Parameter/CoordinateSystem/ CoordinateSystemRepresentation		Stream editing
REPRESENTATION_1	NumericalData/Parameter/CoordinateSystem/ CoordinateSystemRepresentation	Yes	Rarely needs editing
DISPLAY_TYPE	NumericalData/Parameter/RenderingHints/DisplayType		
LABLAXIS	NumericalData/Parameter/RenderingHints/AxisLabel		
LABLAXIS	NumericalData/Parameter/RenderingHints/RenderingAxis		
LABLAXIS	NumericalData/Parameter/RenderingHints/Index		
FORMAT	NumericalData/Parameter/RenderingHints/ValueFormat		

5723



# Backup

# Why Metadata Conventions

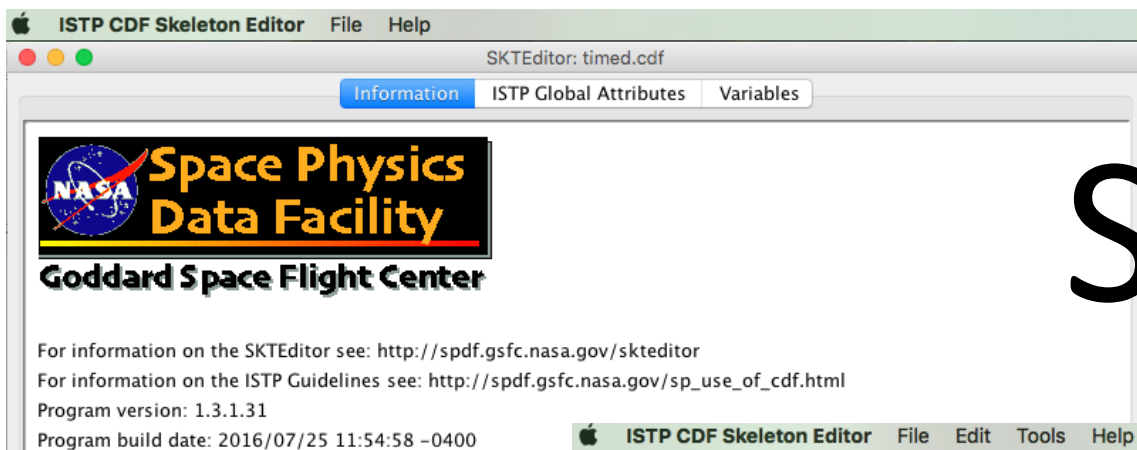
- Leverage standardized self-describing data formats, metadata for datasets and parameters, time conventions, and dataset and filename conventions to enable effective data analysis and browsing using generic easy-to-use software and web services
- Restricting metadata representations limits the number of equivalent possibilities with which software must deal, and thus fosters **interoperability**
- Conventions standardize ways to name things, represent relationships, and locate data in space and time
- Enables developing applications with powerful extraction, regridding, analysis, visualization, and processing capabilities
- Abstracts general data models to represent data semantics
- Embody data provider's knowledge and capture the meaning in data and make data semantics accessible to humans as well as programs
- Provide higher-level abstractions such as coordinate systems, standard names for physical quantities for comparing different data and distinguishing variables

# Dataset creation: Understand the Data to be Loaded

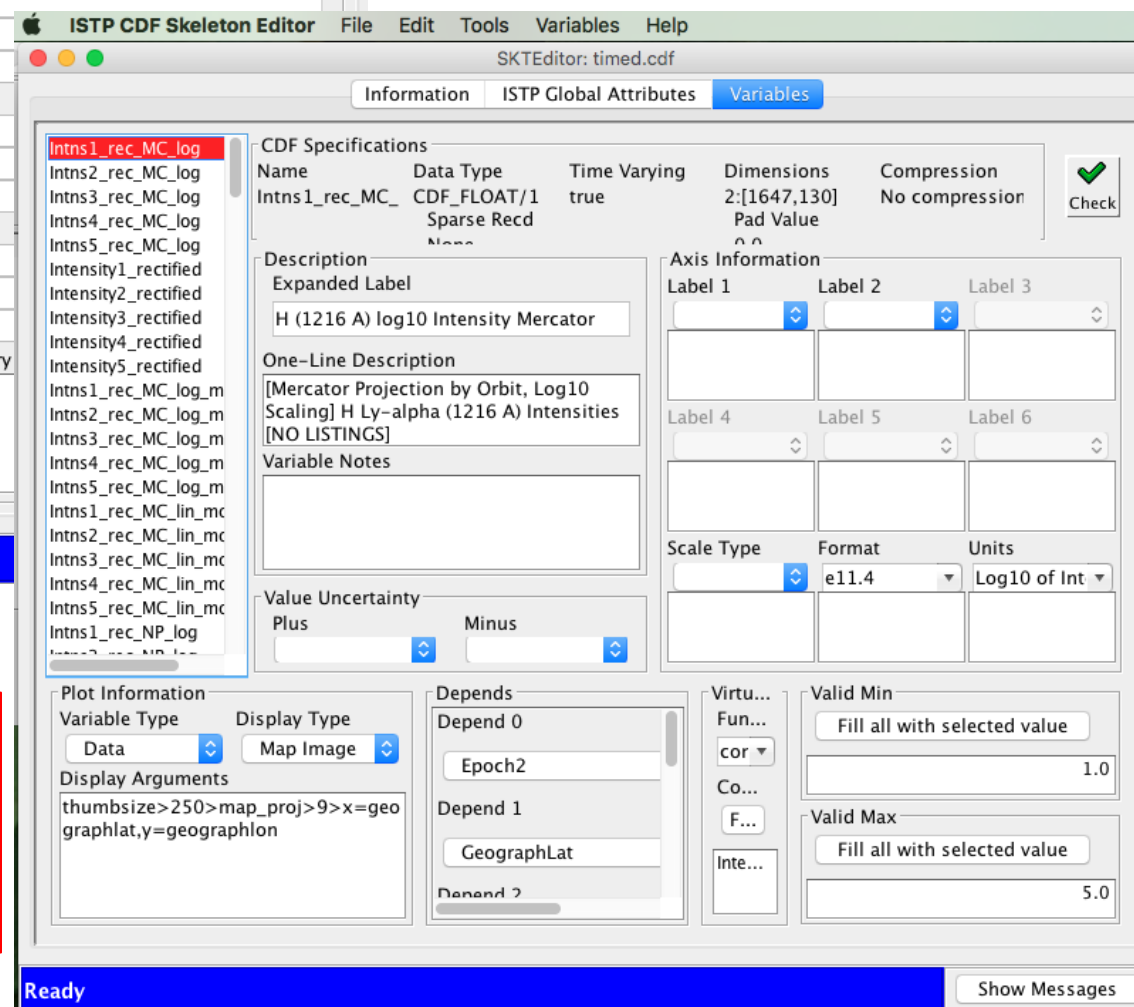
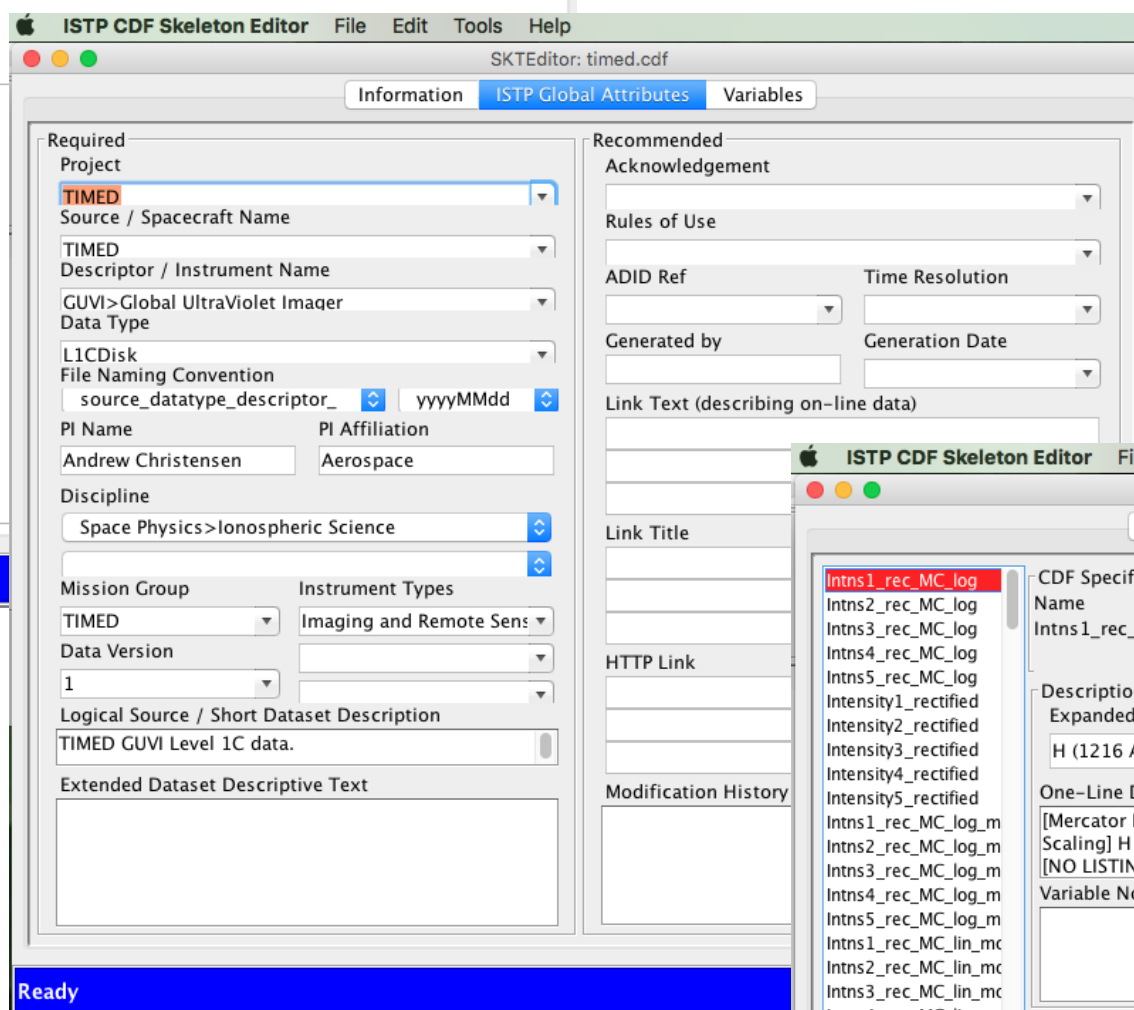
- What are the key data quantities
  - What is their definition/meaning?
  - How are they going to be named?
    - N.B. MMS parameter naming convention: `sclD_instrumentID_paramName`  
»
- Understand (at the dataset level)
  - Dimensionality and dependencies
  - Variance with time and dimension
    - ISTP conventions allow >1 time variable in a file
    - Carry slowly-varying data as variables rather than in attributes
- General rule is to capture relationships in the structure
  - Otherwise capture relationships in variable attributes
  - Want relationships to be logically-structured and machine-readable
    - Available for more general-purpose codes to exploit
- Let CDF/NetCDF deal with mechanics of efficient data storage
  - Once more: lay out data by what's science logical and useful
    - E.g. methods to handle slowly-varying data include setting “`sparse=sRecords.PREV`” in CDFs

# Tool to Create/Edit a CDF/NetCDF File Compliant to ISTP Standard

- SKTeditor is a Java, web-start application, soon to be in JavaScript
  - Guide designers to good choices consistent with ISTP guidelines
  - Create new CDF/NetCDF or check/correct then modify existing skeleton file
- Guided by the interface flow, add or edit
  - Scalar and higher-dimensional variables, multiple time variables
    - Times as `cdf_epoch` or preferably `cdf_time_tt2000`
  - Variable attributes (descriptions, labels, units, `display_type`)
  - Global attributes and file naming
  - Virtual variables (functions in CDAWlib, compute values on-the-fly)
- Checking and validation functions
  - Against ISTP standards
  - For PRBEM, MMS or other specified project compliance reporting
- New JavaScript SKTeditor plans to add capability to add SPASE metadata at the same time when creating a dataset
  - Incorporate Lee Bargatze's ADAPT business logic to reduce effort

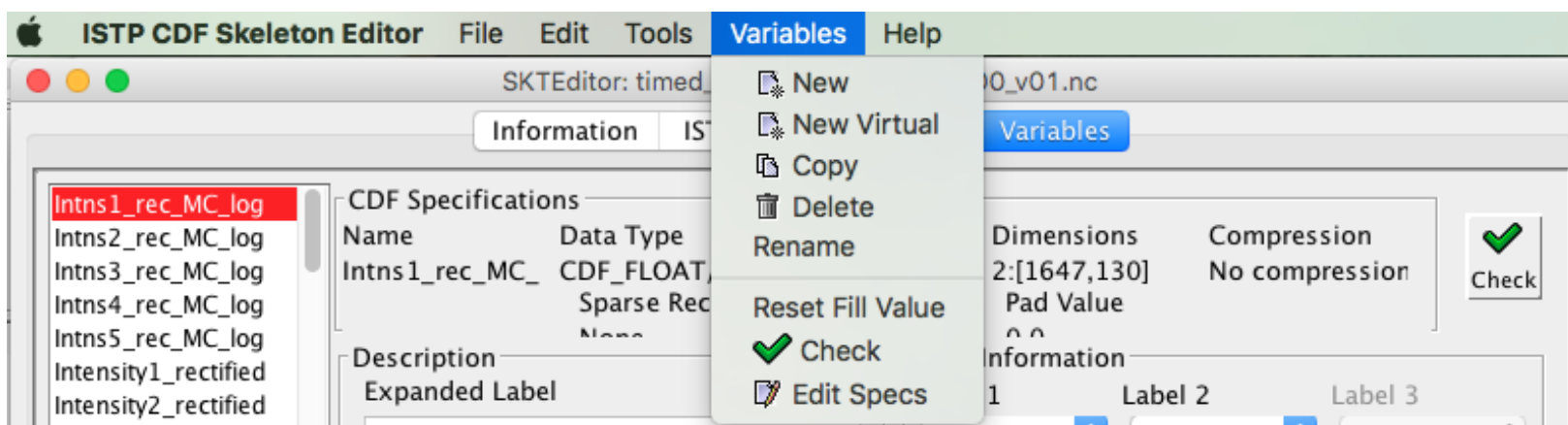
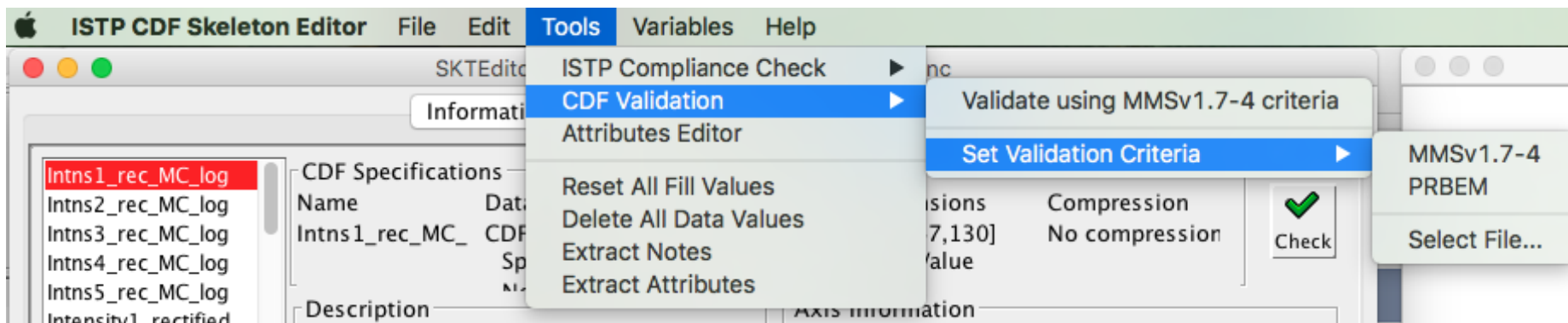
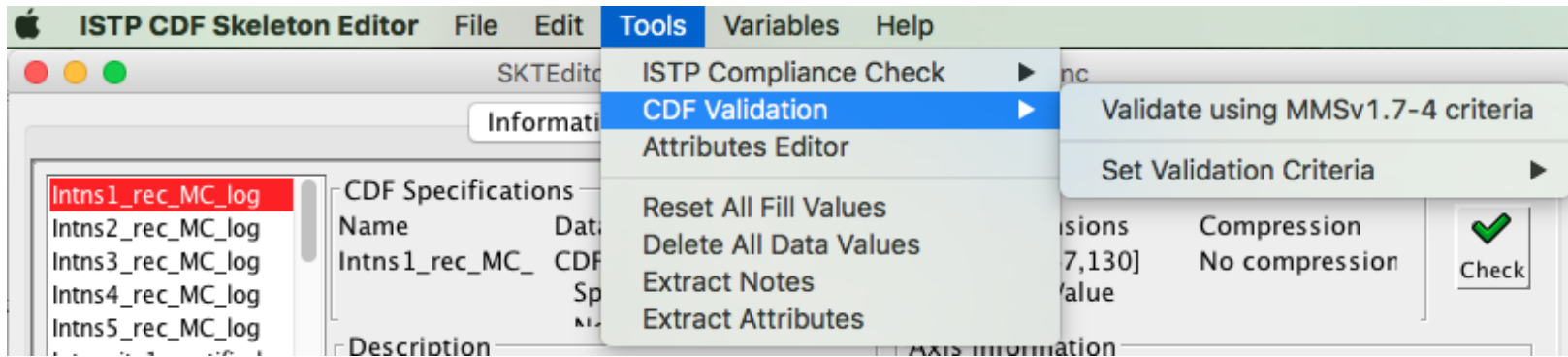
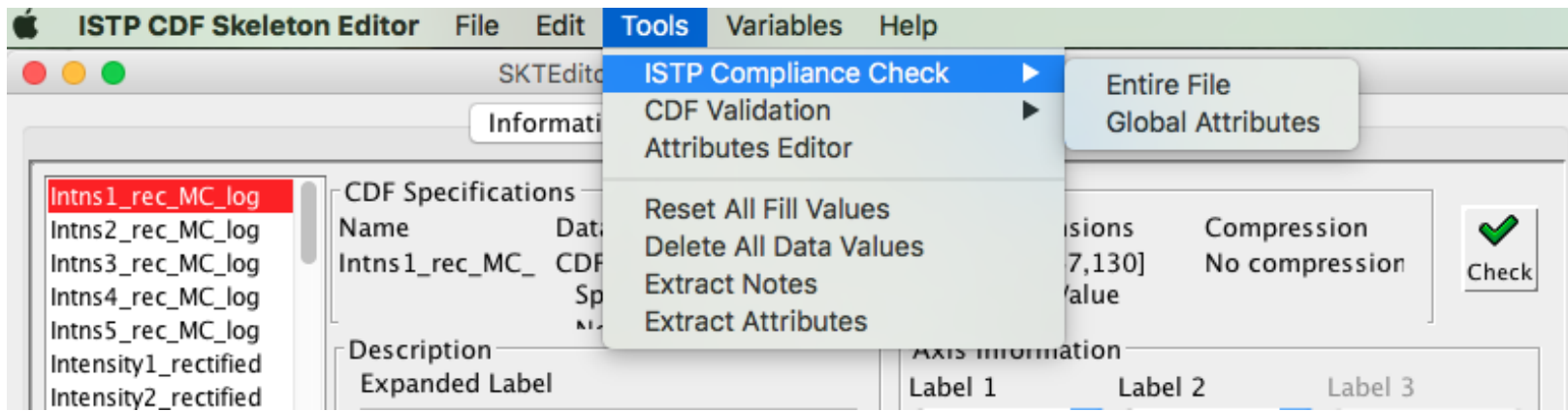


# SKTEditor (Java)



The following variables are not ISTP-compliant:

- Intns1\_rec\_MC\_log
  - DISPLAY\_TYPE attribute value 'map\_image>THUMBSIZE>250>MAP\_PROJ>9>x=GeographLat,y=GeographLon' is not all lower case.
  - DISPLAY\_TYPE attribute value changed to 'map\_image>thumbsize>250>map\_proj>9>x=geographlat,y=geographlon'.
  - DISPLAY\_TYPE error: invalid keyword 'thumbsize'
  - DEPEND\_2 is not 1 dimensional
  - DEPEND\_2 is wrong size
  - DEPEND\_1 is not 1 dimensional



# Development of ISTP Metadata Editor

New JavaScript based web-browser tool to help users create/update CDF datasets with ISTP and SPASE (Space Physics Archive Research and Extract) metadata

**ISTP Metadata Editor** File Tools Help

Information Global Attributes Variable Attributes

**Required**

Project [Project]  
LWS>Living With a Star

Source / Spacecraft Name [Source\_name]  
PSP>Parker Solar Probe

Descriptor / Instrument Name [Descriptor]  
ISOIS-EPILO>Integrated Science Investigation of the Sun, Energetic Particle Instrument Lo

Data Type [Data\_type]  
L2-ic>Level 2 ic

File Naming Convention  
source\_descriptor\_datatype yyyyMMdd

PI Name [PI\_name] PI Affiliation [PI\_affiliation]  
David McComas Princeton University

Discipline [Discipline]  
Solar Physics>Heliospheric Physics  
Space Physics>Interplanetary Studies

**Recommended**

Acknowledgement [Acknowledgement]  
Cite McComas et al (2016), doi:10.1007/s11214-014-0059-1

Rules of Use [Rules\_of\_use]  
See https://spp-isois.sr.unh.edu/ISOIS\_Terms\_of\_Use.html.,Cite as, "McComas et al. (2016)" a

Digital Object Identifier [DOI]

SPASE ID [spase\_DatasetResourceID] Time Resolution [Time\_resolution]  
spase://VSPO/NumericalData/ParkerSolarProt 1 minute to 1 hour

Generated By [Generated\_by] Generation Date [Generation\_date]  
ISOIS SOC, University of New Hampshire 20220129

Links

- :: Data Rules of Use Edit Delete
- :: Instrument paper at Space Science Reviews Edit Delete
- :: Magnetic field data for pitch angle calculation courtesy of the FIELDS team Edit Delete

# NetCDF Issues

- No predefined time variable types
  - Time not always the unlimited dimension
  - CDAWeb adds CDF\_TIME\_TT2000 virtual variables for NetCDF datasets, computed from various time schemes (base time, time units)
- CDAWeb adds missing Fillval, Validmin/max, Var\_type, depend\_0, and other attributes
- NetCDF to CDF converter adds attributes to store version, dimensions, sizes, compression, chunking, and string (not character) information
- Compression requires careful block size determination
- CDF to NetCDF converter converts time variables to binary or encoded string forms
- Supports only NetCDF4 model with no groups or user-defined variable types



# Some Standards and Conventions

- **SPASE** <http://www.spase-group.org> dataset descriptions for easy searching
- **Heliophysics Data Portal** <https://heliophysicsdata.sci.gsfc.nasa.gov>
- **ISTP/IACG/SPDF Guidelines** for global and variable attributes  
[https://spdf.gsfc.nasa.gov/sp\\_use\\_of\\_cdf.html](https://spdf.gsfc.nasa.gov/sp_use_of_cdf.html)
  - SKTEditor metadata creation tool <https://spdf.gsfc.nasa.gov/skteditor>
  - Defining additional standard attributes: Cluster, THEMIS, RBSP (PRBEM), MMS, etc.
- **Dataset naming and file naming** recommendations  
[https://spdf.gsfc.nasa.gov/guidelines/filenaming\\_recommendations.html](https://spdf.gsfc.nasa.gov/guidelines/filenaming_recommendations.html)  
and file naming templates:  
<https://github.com/hapi-server/uri-templates/wiki/Specification> \$Y/data\_\$Y\_\$j\_id\$x.cdf
- **CDF** <https://cdf.gsfc.nasa.gov> scientific data format (including pure Python library  
<https://github.com/MAVENSDC/cdflib>)
  - Time variable types [https://cdf.gsfc.nasa.gov/html/leapseconds\\_requirements.html](https://cdf.gsfc.nasa.gov/html/leapseconds_requirements.html)
- **NetCDF** <https://www.unidata.ucar.edu/software/netcdf/>
- **FITS** <https://fits.gsfc.nasa.gov/>
- **UDunits** [www.unidata.ucar.edu/software/udunits/](http://www.unidata.ucar.edu/software/udunits/)
- **VOUnits** (<https://ivoa.net/documents/VOUnits/>)
- Tools enabled by standards: CDAWeb and CDAWlib IDL/Python library, Autoplot <http://autoplot.org>, SPEDAS <http://spedas.org> IDL/Python library