A Standard for Units in Heliophysics

Bob Weigel, Baptiste Cecconi, Brian Thomas, Doug Lindholm, Andriy Koval, Jeremy Faden, Lee Bargatze, Bobby Candey, and Shing Fung

We describe an effort by representatives from the metadata standards groups of HAPI, SPASE, and ISTP Metadata Guidelines to develop a recommendation for the representation of scientific units in metadata. The recommendation is

> “Some metadata models in Heliophysics do not have a constraint on unit strings (that is, standard has not been chosen). Our recommendation is that if a standard is used, it should be VOUnits.”

Background

- **ISTP Metadata Guidelines** - Part of the International Solar-Terrestrial Program that began in 1980s, guidelines later in parallel with CDF file format.

SPASE - Space Physics Search and Extract. Began ~2004. More formal metadata model than ISTP and community developed.

- **HAPI** - Heliophysics API. Began in 2016. A standard for serving time series data via an API. Metadata model is not domain specific - allows linking to domain-specific metadata.

Motivation

In the metadata standard for the HAPI API, we have units and unitsSchema attributes, with the optional unitsSchema to be used by units-aware software.

We noted that two Heliophysics metadata standards, ISTP and SPASE, do not have constraints on unit strings (UNITS and Units). ISTP has SI_CONVERSION and SPASE has UnitsConversion, which allow automatic unit conversion.

Primary questions:

1. Is SI_CONVERSION/UnitsConversion enough?
2. If not, what standard should we use for unit strings?

Question: 1. Is SI_CONVERSION/UnitsConversion enough?

Examples of SI_CONVERSION/UnitsConversion:

UNITS = Re SI_CONVERSION = 6371200 > m
UNITS = eV SI_CONVERSION = 1.602e-19 > J
UNITS = eV SI_CONVERSION = 1.60217646E-19 > J
UNITS = deg SI_CONVERSION = 1.0 > degree *or degrees or (degree)*
UNITS = deg SI_CONVERSION = 0.0174532925 > rad

Question: 1. Is SI_CONVERSION/UnitsConversion **enough?**

Advantages:

- Used in Heliophysics community for many years
- Avoids requirement that data providers use a standard style for unit strings (and providers have strong opinions in this regard)

Disadvantages:

- Not used outside of Heliophysics (?)
- Uneven use by (especially older) missions
- Some inconsistency in implementation across missions
- Lack of software using it
- Additional standards needed (e.g., precision on numbers, what terms are valid, e.g., degree or degrees or deg). Current standard states only "... the factor that the variable must be multiplied by in order to turn it to generic SI units"

Question: 1. Is SI_CONVERSION/UnitsConversion **enough?**

It *could* be enough with additional constraints beyond existing specification "… the factor that the variable must be multiplied by in order to turn it to generic SI units".

However, much of the needed addition will overlap with work that has already been done in the development of standards for unit strings.

Conclusion:

We want to allow data providers to express unit strings in a standard form but still keep the keep the SI_CONVERSION/UnitsConversion option. (Which will be useful for validation.)

General agreement that Heliophysics should have a recommendation "… if a standard for unit strings is used it should be X".

Question: 2. If not, what standard for unit strings?

- Studied many standards and associated software libraries, when available: NIST SP 811, VOUnits, UDUNITS2, EDS-RFC, QUDT, UCUM, MMS, Cluster, SMDX, ….
- Also, met with the CODATA working group DRUM (Digital Representation of Units of Measurement) team.

Question: 2. If not, what standard for unit strings?

- Much discussion on why machine readable is important, why “SI strings” is ambiguous, and many other nuances.

- Historically, what was placed as values for unit attributes in NASA missions has included non-unit information such as documentation, labels, and quantities, e.g.,

  1/cc (qual only if scan=2=SW)
  T/F
  1/N SUM |B|, nT
  V-dc

Question: 2. If not, what standard for unit strings?

- UDUNITS-2 and VOUnits used on both boundaries of Heliophysics (Earth Science on inner edge and Astronomy on outer).
- Much debate on this. Flexibility of UDUNITS-2 syntax was attractive given many existing unit strings follow one of the allowed representations (e.g., m^2, m**2 allowed as is N-m, N.m, and N*M.)
- UDUNITS-2 contains XML tables for unit strings and a text description (not formal grammar) for their creation that is distributed with UDUNITS-2 source code. Others have built on UDUNITS-2 (e.g., CF-Conventions).
- VOUnits is “proper” standard (community developed, versioned, and has grammar) and supported by AstroPy and other libs.

Conclusion

General agreement that Heliophysics (via IHDEA) should have a recommendation “… if a standard for unit strings is used, it should be VOUnits”.

Plan

- **HAPI** - already has units and unitsSchema and will add VOUnits as an option for unitsSchema. Some work needed to allow unitsSchema version to change without updating HAPI version.
- **ISTP** - already has UNITS. Will add UNITS_VO or equivalent. Will also update SI_CONVERSION.
- **SPASE** - already has Units. Will add a way to express the schema of units, possibly with an XML attribute on Units.