

CSV Header Definition

CSV Header Definition Proposal Sandbox

We have currently agreed that XML would be the format of choice.

Three existing schemas have been suggested:

- [TransducerML](#)
 - Transducer Markup Language (TML) is a language for capturing and characterizing not only data from transducers, but information necessary for the processing and understanding of that data by the eventual recipient of the transducer data. Both sensors and transmitters can be captured and characterized within TML, leading to the use of the term "transducer" rather than "sensor". TML handles not only static but also streaming transducer data. TML permits the data stream to handle live transducer data both being added to the stream and being deleted from the stream.
- [SensorML](#)
 - SensorML provides standard models and an XML encoding for describing any process, including the process of measurement by sensors and instructions for deriving higher-level information from observations. Processes described in SensorML are discoverable and executable. All processes define their inputs, outputs, parameters, and method, as well as provide relevant metadata. SensorML models detectors and sensors as processes that convert real phenomena to data.
- [netCDF NcML](#)
 - Excellent for time series data and metadata. Could be done real-time over the wire, but not designed for such.

Seems if we want to go with TransducerML, then we should replace our entire CSV scheme with that. I have a difficult time locating good examples, and ones that are a minimal set for both TransducerML and SensorML (better examples for SensorXML exist). Both TransducerML and SensorML seem heavy-weight, NcML is light-weight.

Chris currently proposes using the netCDF NcML schema with some slight modifications.

Based on this CSV string:

```
GARMIN_GPS,2008-09-12T153204,41.482,-105.7331,5800.2
```

Pure NcML example:

```
* <?xml version="1.0" encoding="UTF-8"?>
<netcdf xmlns="http://www.unidata.ucar.edu/namespaces/netcdf/ncml-2.2" format="classic" >

<attribute name="ProjectName" type="String" value="MILAGRO" />
<attribute name="Platform" type="String" value="N130AR" />

<variable name="IDENTIFIER" shape="Time" type="String">
<attribute name="long_name" type="String" value="Keyword identifying this packet" />
<attribute name="units" type="String" value="GARMIN_GPS" />
</variable>

<variable name="Time" shape="Time" type="String">
<attribute name="long_name" type="String" value="time of measurement" />
<attribute name="units" type="String" value="iso-8601" />
</variable>

<variable name="GGLAT" shape="Time" type="float">
<attribute name="units" type="String" value="degree_N" />
<attribute name="long_name" type="String" value="GPS Latitude" />
</variable>

<variable name="GGLON" shape="Time" type="float">
<attribute name="units" type="String" value="degree_E" />
<attribute name="long_name" type="String" value="GPS Longitude" />
</variable>

<variable name="GGALT" shape="Time" type="float">
<attribute name="units" type="String" value="m" />
<attribute name="long_name" type="String" value="GPS Altitude" />
</variable>
</netcdf>
```

- The **netCDF** attribute should probably be replaced.
- The global attributes in this example, [ProjectName](#) and Platform, would be optional.

- Other variable attributes could optionally be included (e.g. missing_value, valid_range, standard_name).
- This could also be used for ASCII data distribution post-flight.

```
<?xml version="1.0" encoding="UTF-8"?>
<csv>
<variable name="IDENTIFIER" shape="Time" type="String">
<attribute name="long_name" type="String" value="Keyword identifying this packet" />
<attribute name="units" type="String" value="GARMIN_GPS" />
</variable>

<variable name="Time" shape="Time" type="String">
<attribute name="long_name" type="String" value="time of measurement" />
<attribute name="units" type="String" value="iso-8601" />
</variable>

<variable name="GGLAT" shape="Time" type="float">
<attribute name="units" type="String" value="degree_N" />
<attribute name="long_name" type="String" value="GPS Latitude" />
</variable>

<variable name="GGLON" shape="Time" type="float">
<attribute name="units" type="String" value="degree_E" />
<attribute name="long_name" type="String" value="GPS Longitude" />
</variable>

<variable name="GGALT" shape="Time" type="float">
<attribute name="units" type="String" value="m" />
<attribute name="long_name" type="String" value="GPS Altitude" />
</variable>
</csv>
```

A post-flight header might look like:

```

* <?xml version="1.0" encoding="UTF-8"?>
<csv>
<dimension name="Time" length="16269" isUnlimited="true" />
<dimension name="sps1" length="1" />
<dimension name="Vector64" length="64" />
<dimension name="Vector31" length="31" />
<attribute name="Source" type="String" value="NCAR Research Aviation Facility" />
<attribute name="Address" type="String" value="P.O. Box 3000, Boulder, CO 80307-3000" />
<attribute name="Phone" type="String" value="(303) 497-1030" />
<attribute name="Conventions" type="String" value="NCAR-RAF/nimbus" />
<attribute name="ConventionsURL" type="String" value="http://www.eol.ucar.edu/raf/Software/netCDF.html" />
<attribute name="Version" type="String" value="1.3" />
<attribute name="ProcessorRevision" type="String" value="3213" />
<attribute name="ProcessorURL" type="String" value="http://svn/svn/raf/trunk/nimbus" />
<attribute name="WARNING" type="String" value="This file contains PRELIMINARY DATA that are NOT to be used for critical analysis." />
<attribute name="DateProcessed" type="String" value="2006-02-22 12:57:06 \-0700" />
<attribute name="ProjectName" type="String" value="MILAGRO" />
<attribute name="Aircraft" type="String" value="N130AR" />
<attribute name="ProjectNumber" type="String" value="145" />
<attribute name="FlightNumber" type="String" value="tf02" />
<attribute name="FlightDate" type="String" value="02/21/2006" />
<attribute name="coordinates" type="String" value="LONC LATC GGALT Time" />
<attribute name="landmarks" type="String" value="19.4383 \-98.9242 MEX,19.1544 \-95.81 VER,22.3053 - 96.123 TAM,19.1636 \-97.6253 PUE,19.735 \-98.981 LUC,19.6847 \-97.0475 T1" />

<variable name="Time" shape="Time" type="int">
<attribute name="long_name" type="String" value="time of measurement" />
<attribute name="standard_name" type="String" value="time" />
<attribute name="units" type="String" value="seconds since 2006-02-21T18:43:35 \+0000" />
</variable>

<variable name="ACINS" shape="Time" type="float">
<attribute name="_FillValue" type="float" value="-32767.0" />
<attribute name="units" type="String" value="m/s2" />
<attribute name="long_name" type="String" value="Aircraft Vertical Acceleration" />
<attribute name="Category" type="String" value="Aircraft State" />
<attribute name="SampledRate" type="int" value="25" />
<attribute name="TimeLag" type="int" value="10" />
<attribute name="TimeLagUnits" type="String" value="milliseconds" />
<attribute name="DataQuality" type="String" value="Preliminary" />
<attribute name="PITCH_BIAS_1" type="float" value="0.0" />
<attribute name="HEADING_BIAS_1" type="float" value="0.4" />
</variable>

<variable name="ALT" shape="Time" type="float">
<attribute name="_FillValue" type="float" value="-32767.0" />
<attribute name="units" type="String" value="m" />
<attribute name="long_name" type="String" value="IRS Baro-Inertial Altitude" />
<attribute name="Category" type="String" value="Position" />
<attribute name="standard_name" type="String" value="altitude" />
<attribute name="SampledRate" type="int" value="25" />
<attribute name="DespikeSlope" type="float" value="" />
<attribute name="DataQuality" type="String" value="Preliminary" />
</variable>
</csv>

```