CAPTURING RAPIDLY CHANGING METADATA (RCM): CURRENT STATUS AND NEXT STEPS:

TIM AHERN
IRIS/EARTHSCOPE EMERITUS
EARTHSCOPE-OCEANS STEERING COMMITTEE

JOEL SIMON
EARTHSCOPE-OCEANS
PRINCETON UNIVERSITY
• WG-5 agreed to evaluate GeoCSV for RCM using the FDSN Framework Process
• WG-5 formed a Proposal Review Team to Evaluate the Framework Proposal and Develop a Recommendation
• The Review Team consisted of
  • Asia – Chin-Jen Lin Academia Sinica/IES
  • North America – Adam Ringler, USGS (chair)
  • Europe – Joachim Wassermann, LMU
• The Review Team Recommended the Adoption of the GeoCSV concept with additional suggestions related to incorporation of the RCM into StationXML
EXPANSION BEYOND “MOVING STATIONS”

• Examples of changing values not easily accommodated in SEED
  • Moving stations (latitude, longitude, elevation, depth) (i.e. MERMAIDs, ice sheets, glaciers, creeping slopes, etc.)
  • Variable sample rates
  • Frequently changing gains (i.e. H_2O Observatory)
  • Different algorithmic estimates of calculated sensor orientation (OBS)

• GeoCSV is flexible enough to allow capture of all relevant Rapidly Changing Metadata we are aware of
GOALS OF THE GEOCSV FOR RCM FORMAT

• A standard way of capturing metadata in an easy-to-understand manner supporting FAIR principles
• Easy to generate in a field setting by non-software experts
• A format that captures information not now in StationXML
• The ultimate goal is to include this metadata in StationXML
IDENTIFIED GEOCSV COMPONENTS

• GeoCSV Header
• Column Identifiers
• A series of Elements as needed
  • Time and SNCL Information (normally necessary)
    • Start time, network, Station, Location, Channel
  • Positional Information
    • Latitude, Longitude, Elevation, Depth
  • Sensor information
    • SensorDescription, Scale, Scale Frequency, Scale Units
  • Timing Information
    • SampleRate, TimeDelay, TimeCorrection
  • Orientation Information
    • Dip, Azimuth, Uncertainties
GeoCSV HEADER

- GeoCSV Header
  - #dataset: GeoCSV 2.0
  - #created: 2021-07-28T23:25:20Z
  - #delimiter: ','
  - #lineterminator: '\n'

<table>
<thead>
<tr>
<th>Column Row Headers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time + SNCL</td>
</tr>
<tr>
<td>Positional</td>
</tr>
<tr>
<td>Sensor/Gain</td>
</tr>
<tr>
<td>Timing</td>
</tr>
<tr>
<td>Orientation</td>
</tr>
<tr>
<td>As Needed</td>
</tr>
<tr>
<td>degrees_north</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>float</td>
</tr>
<tr>
<td>latitude</td>
</tr>
</tbody>
</table>
TIMES AND SNCL ELEMENT

- Start Time
- End Time (opt)
- Station
- Network
- Channel
- Location
## POSITIONAL ELEMENT

- Latitude
- Longitude
- Elevation
- Depth

### GeoCSV Header
- Time + SNCL
- Positional
- Sensor/Gain
- Timing
- Orientation
- As Needed
SENSOR AND SENSOR GAIN ELEMENT

- Sensor Description
- Scale Factor
- Scale Frequency
- Scale Units

GeoCSV Header
Column Row Headers
Time + SNCL Positional Sensor/Gain Timing Orientation As Needed
TIMING ELEMENT

- Sample Rate
- Time Delay
- Time Correction Applied
ORIENTATION ELEMENT

- Dip
- Azimuth
- Azimuthal Uncertainty

GeoCSV Header

<table>
<thead>
<tr>
<th>Column Row Headers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time + SNCL</td>
</tr>
</tbody>
</table>

12
When a new element is needed, these elements can be added with the approval of WG II.
## EXAMPLE FOR CAPTURING DRIFT OF THE ROSS ICE SHELF

### Experiment XH 2015-2017

<table>
<thead>
<tr>
<th>Method/Identifier</th>
<th>Start Time</th>
<th>Network</th>
<th>Station</th>
<th>Location</th>
<th>Channel</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation</th>
<th>Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS Q330 GPS Clock</td>
<td>2014-12-31T23:00:40Z</td>
<td>XH</td>
<td>DR01</td>
<td>*</td>
<td>*</td>
<td>-77.77508</td>
<td>178.34172</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>GPS Q330 GPS Clock</td>
<td>2015-12-31T03:10:28Z</td>
<td>XH</td>
<td>DR01</td>
<td>*</td>
<td>*</td>
<td>-77.76594</td>
<td>178.34611</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>GPS Q330 GPS Clock</td>
<td>2016-01-20T01:08:44Z</td>
<td>XH</td>
<td>DR01</td>
<td>*</td>
<td>*</td>
<td>-77.75806</td>
<td>178.34989</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>GPS Q330 GPS Clock</td>
<td>2014-12-31T23:30:38Z</td>
<td>XH</td>
<td>DR05</td>
<td>*</td>
<td>*</td>
<td>-78.64047</td>
<td>-179.09994</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>GPS Q330 GPS Clock</td>
<td>2015-12-31T22:50:24Z</td>
<td>XH</td>
<td>DR05</td>
<td>*</td>
<td>*</td>
<td>-78.63156</td>
<td>-179.09239</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>GPS Q330 GPS Clock</td>
<td>2016-11-16T17:53:01Z</td>
<td>XH</td>
<td>DR05</td>
<td>*</td>
<td>*</td>
<td>-80.86433</td>
<td>178.43481</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>GPS Q330 GPS Clock</td>
<td>2014-12-31T23:43:19Z</td>
<td>XH</td>
<td>RS01</td>
<td>*</td>
<td>*</td>
<td>-78.18889</td>
<td>169.96239</td>
<td>14.3</td>
<td>0</td>
</tr>
<tr>
<td>GPS Q330 GPS Clock</td>
<td>2015-12-31T17:58:39Z</td>
<td>XH</td>
<td>RS01</td>
<td>*</td>
<td>*</td>
<td>-78.18333</td>
<td>169.965</td>
<td>1.4m</td>
<td>0</td>
</tr>
</tbody>
</table>

Experiment XH 2015-2017 Actual number of rows is >12,000
## EXAMPLE OF POST DEPLOYMENT OBS ORIENTATIONS

<table>
<thead>
<tr>
<th>Method/Identifier</th>
<th>StartTime</th>
<th>Network</th>
<th>Station</th>
<th>Location</th>
<th>Channel</th>
<th>dip</th>
<th>azimuth</th>
<th>azimuthal uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLOpy</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL38</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>179.37</td>
<td>2.20</td>
</tr>
<tr>
<td>STACH</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL38</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>179.58</td>
<td>7.10</td>
</tr>
<tr>
<td>Laske et al</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL38</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>178.60</td>
<td>0.73</td>
</tr>
<tr>
<td>DLOpy</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL40</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>168.32</td>
<td>1.00</td>
</tr>
<tr>
<td>STACH</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL40</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>278.38</td>
<td>3.61</td>
</tr>
<tr>
<td>Laske et al</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL40</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>277.14</td>
<td>0.87</td>
</tr>
<tr>
<td>DLOpy</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL47</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>277.13</td>
<td>1.14</td>
</tr>
<tr>
<td>STACH</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL47</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>278.38</td>
<td>3.61</td>
</tr>
<tr>
<td>Laske et al</td>
<td>2006-04-22T00:00:00Z</td>
<td>YS</td>
<td>PL47</td>
<td>0</td>
<td>BH1</td>
<td>-90</td>
<td>277.14</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Experiment YS 2006-2007

**Times and SNCL Element**

- **ISO8601 Unitless**: 2006-04-22T00:00:00Z
- **String Station**: PL38
- **String Location**: 0
- **String Channel**: BH1

**Orientation Element**

<table>
<thead>
<tr>
<th>SEED Convention</th>
<th>Seed Convention</th>
<th>Azimuthal Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Float</td>
<td>Float</td>
<td>Float</td>
</tr>
<tr>
<td>-90</td>
<td>179.37</td>
<td>2.20</td>
</tr>
<tr>
<td>-90</td>
<td>179.58</td>
<td>7.10</td>
</tr>
<tr>
<td>-90</td>
<td>178.60</td>
<td>0.73</td>
</tr>
<tr>
<td>-90</td>
<td>168.32</td>
<td>1.00</td>
</tr>
<tr>
<td>-90</td>
<td>278.38</td>
<td>3.61</td>
</tr>
<tr>
<td>-90</td>
<td>277.14</td>
<td>0.87</td>
</tr>
<tr>
<td>-90</td>
<td>277.13</td>
<td>1.14</td>
</tr>
<tr>
<td>-90</td>
<td>278.38</td>
<td>3.61</td>
</tr>
<tr>
<td>-90</td>
<td>277.14</td>
<td>0.87</td>
</tr>
</tbody>
</table>
NEXT STEPS – GEOCSV EXAMPLES

- Continue capturing metadata in GeoCSV format for projects that have RCM and managing the data at existing federated data centers
  - MERMAID data is available now
  - Moving stations on glaciers and ice sheets
  - OBSIP and other outreach to OBS community

- Form a Technical Advisory Group from WG membership to
  - recommend how RCM should be incorporated into StationXML

Other Technical Details
- Wildcarding conventions
- Representation of unknown values
PROPOSAL DEVELOPMENT

• Pursue Funding Support for Technical Completion Phase of incorporating RCM into StationXML
  • Develop a work specification with the Technical Advisory Group
  • Select a consulting firm to do the development work
  • Submission of the proposal to a funding agency

• Timeline
  • Beta version 6 months after funding
  • Release version 18 months after funding

• Adoption by FDSN at 2025 IASPEI meeting

• FDSN develops plans to make modifications to relevant tools that use StationXML
THANK YOU
COMMITTEE RECOMMENDATION

GeoCSV Proposal Review

On February 9, 2022 the review team (Chin-Jen Lin, Adam Ringler, and Joachim Wassermann) discussed the GeoCSV proposal over Zoom. The meeting was also attended by Tim Ahern, where he clarified a number of questions the review team had.

The review team found the GeoCSV to be a good approach to temporary metadata that had a temporal dependance to it. One of the few shortcomings in the Standard for the Exchange of Earthquake Data (SEED) format is that it is unable to deal with a large number of metadata changes. GeoCSV seems like a good compromise to deal with cases where the metadata will have many changes, making epochal boundaries in SEED difficult to deal with.

The example spreadsheet provided something that seems easy to populate. We also think this being an interim step for the more complete solution of being included in StationXML is valuable. However, we propose that the WGII team should evaluate the pros and cons of other formats differing from CSV for easier and possibly automatic integration into stationXML (e.g., JSON or YAML) by keeping the format simple and human readable.

Given our findings the review team approves the GeoCSV proposal and sees no reason to not go ahead on the implementation.

Chin-Jen Lin, Adam Ringler, Joachim Wassermann