O4.3-165

Rapidly Changing Metadata (RCM): A Project to Add FAIR Metadata to Existing Standards

Tim Ahern[1] & Joel Simon[2]

[1]Director of Data Services Emeritus, EarthScope
[2]Data Manager, Bathymetrix, LLC / Princeton University

Presentation Date: 10 September 2025

Tim Ahern & Joel Simon

O4.3-165

Rapidly Changing Metadata (RCM): A Project to Add FAIR Metadata to Existing Standards

- The FDSN SEED format has been a major format used for data distribution since the late 1980s.
  - Earlier recorded digital data from many data centres have converted from their native format to SEED
  - Conversion tools exist to allow SEED and CTBTO data to be used seamlessly
  - It is a format that assumes static or slowly-changing metadata--it was not designed for stations that move

- This talk describes a straightforward method where Rapidly Changing Metadata (RCM) can be captured ensuring the data are Findable, Accessible, Interoperable, and Reproduceable (FAIR) using the GeoCSV format

- Comma Separated Values or CSV
  - Many disciplines use comma separated values (CSV) styles to capture data
  - GeoCSV came from the NSF EarthCube GeoWS project which identified a structure that made earth science data with additional metadata able to use GeoCSV for its data.
  - EarthScope/IRIS supports GeoCSV output in some of its web services

Rapidly Changing Metadata: (RCM): A Project to Add FAIR Metadata to Existing Standards

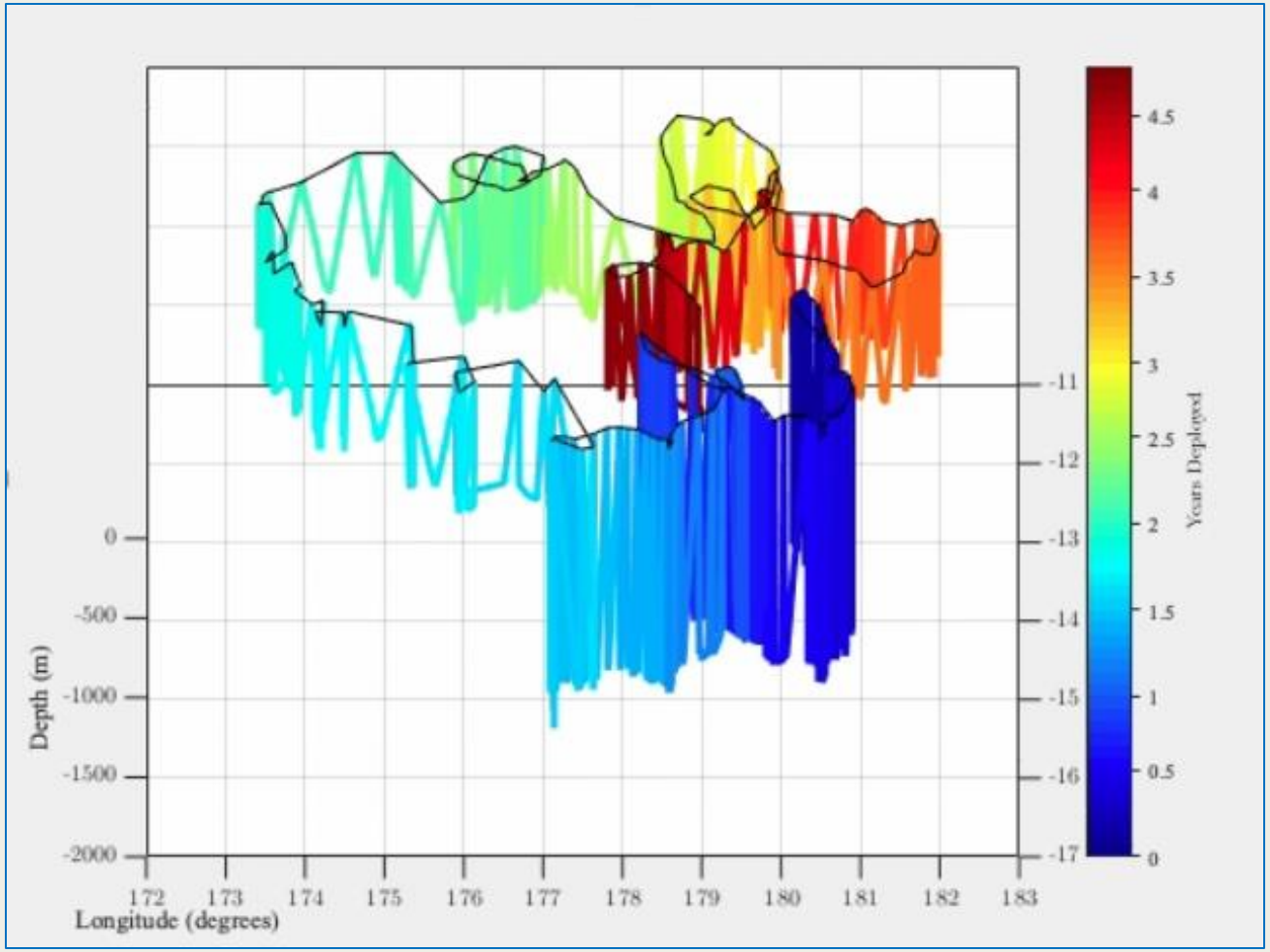
Tim Ahern & Joel Simon

MERMAIDS Mobile Earthquake Recording in Marine Areas by Independent Divers

O4.3-165

- MERMAIDS are autonomous drifting floats
- Equipped with a hydrophone
- Operate at 1500m depth
- When a P wave is detected,
  - MERMAIDS rise to the surface
  - determines location via GPS,
  - synchronizes clock,
  - transmits data window
- The location at the time of recording is interpolated using data from multiple dives

GeoCSV was able to capture the metadata to allow others to reproduce their results



- Figure showing the 3D location of one MERMAID over the past 4.5 years

Tim Ahern & Joel Simon

Introduction to GeoCSV files and RCM in Seismology

GeoCSV Headers

GeoCSV Column Headers and Fields

	Time and NSLC Element	Position Element	Sensor and Gain Element	Timing Element	Orientation Element	Additional Element
GeoCSV Fields	Start Time End Time Network Station Location Channel	Latitude Longitude Elevation Depth	Sensor Desc. Scale Factor Scale Freq. Scale Units	Sample Rate Time Delay Time Correction Applied	Dip Dip +/- Azimuth Az +/-.	Parm 1 Parm 2 Parm 3 …. Parm N

Tim Ahern & Joel Simon

GeoCSV Header

```
#dataset: GeoCSV
#created: 2025-06-09T22:15:37.918Z
#description: Metadata for drifting Mobile Earthquake Recording in Marine Areas by Independent Divers (MERMAID) hydrophones, www.EarthScopeOceans.org

#attribution: automaid v4.3.2 (https://github.com/earthscopeoceans/automaid [doi: 10.5281/zenodo.5057096])

#matlab_reader: https://github.com/joelsimon/GeoCSV/blob/master/readGeoCSV.m

#waterpressure2depth: 100 mbar is approximately equal to the pressure of 1 meter of water

#frequency_response: http://ds.iris.edu/data/reports/MH/MH.Mermaids.Response.V3.pdf

#lineterminator: '\n'
#delimiter: ','
```

Tim Ahern & Joel Simon

Column Identifiers- example for Geographic Position Fields

#field_unit: unitless	degrees_north	degrees_east	meters	meters
#Field_type: string	float	float	float	float
#MethodIdentifier	latitude	longitude	elevation	depth

Tim Ahern & Joel Simon

A Sample MERMAID GeoCSV file

#dataset: GeoCSV
#created: 2025-06-09T22:15:37.918Z
#description: Metadata for drifting Mobile Earthquake Recording in Marine Areas by Independent Divers (MERMAID) hydrophones, www.EarthScopeOceans.org
#attribution: automaid v4.3.2 (https://github.com/earthscopeoceans/automaid [doi: 10.5281/zenodo.5057096])
#matlab_reader: https://github.com/joelsimon/GeoCSV/blob/master/readGeoCSV.m
#waterpressure2depth: 100 mbar is approximately equal to the pressure of 1 meter of water
#frequency_response: http://ds.iris.edu/data/reports/MH/MH.Mermaids.Response.V3.pdf
#lineterminator: '\n'
#delimiter: ','

#field_unit: unitless	iso8601	unitless	unitless	unitless	unitless	unitless	degrees_north	degrees_east	meters	mbar	unitless	hertz	unitless	seconds	seconds
#field_type: string	datetime	string	string	string	string	string	float	float	float	float	string	float	integer	float	float
MethodIdentifier	StartTime	Network	Station	Location	Channel	DataQuality	Latitude	Longitude	Elevation	WaterPressure	InstrumentDescription	SampleRate	SampleCount	TimeDelay	TimeCorrection
Measurement:GPS:u-blox_NEO-M8N	2018-06-27T19:16:42.000Z	MH	P0006	nan	nan	nan	-14.453383	-179.485031	nan	nan	MERMAIDHydrophone(452.020)	nan	nan	0	nan
Measurement:GPS:u-blox_NEO-M8N	2018-06-27T19:16:53.000Z	MH	P0006	nan	nan	nan	-14.4534	-179.485031	nan	nan	MERMAIDHydrophone(452.020)	nan	nan	0	nan
Measurement:GPS:u-blox_NEO-M8N	2018-06-27T19:21:53.000Z	MH	P0006	nan	nan	nan	-14.45365	-179.485153	nan	nan	MERMAIDHydrophone(452.020)	nan	nan	-0.000031	nan
Measurement:GPS:u-blox_NEO-M8N	2018-06-27T19:22:04.000Z	MH	P0006	nan	nan	nan	-14.453667	-179.485153	nan	nan	MERMAIDHydrophone(452.020)	nan	nan	-0.000031	nan
Measurement:Pressure:KELLER_Series_6	2018-06-27T19:28:03.000Z	MH	P0006	nan	nan	nan	nan	nan	nan	730	MERMAIDHydrophone(452.020)	nan	nan	nan	nan
Algorithm(thermocline):automaid:v4.3.2	2018-06-27T19:35:41.808Z	MH	P0006	nan	nan	nan	-14.454387	-179.485443	nan	5000	MERMAIDHydrophone(452.020)	nan	nan	nan	nan
Measurement:Pressure:KELLER_Series_6	2018-06-27T19:43:55.000Z	MH	P0006	nan	nan	nan	nan	nan	nan	9590	MERMAIDHydrophone(452.020)	nan	nan	nan	nan
Measurement:Pressure:KELLER_Series_6	2018-06-27T20:00:10.000Z	MH	P0006	nan	nan	nan	nan	nan	nan	16050	MERMAIDHydrophone(452.020)	nan	nan	nan	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-02T11:40:17.000Z	MH	P0006	nan	nan	nan	-14.131583	-179.773087	nan	nan	MERMAIDHydrophone(452.020)	nan	nan	0	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-02T11:40:28.000Z	MH	P0006	nan	nan	nan	-14.131567	-179.773087	nan	nan	MERMAIDHydrophone(452.020)	nan	nan	0	nan
Measurement:Pressure:KELLER_Series_6	2018-08-02T11:47:53.000Z	MH	P0006	nan	nan	nan	nan	nan	nan	560	MERMAIDHydrophone(452.020)	nan	nan	nan	nan
Algorithm(thermocline):automaid:v4.3.2	2018-08-02T11:58:36.069Z	MH	P0006	nan	nan	nan	-14.130936	-179.773331	nan	5000	MERMAIDHydrophone(452.020)	nan	nan	nan	nan
Measurement:Pressure:KELLER_Series_6	2018-08-02T12:09:09.000Z	MH	P0006	nan	nan	nan	nan	nan	nan	9370	MERMAIDHydrophone(452.020)	nan	nan	nan	nan

Tim Ahern & Joel Simon

Ross Ice Shelf Example of Time & NSCL, Position

#dataset: GeoCSV 2.0									
#created: 2023-06-17T12:25:20Z									
#Reference url: http://ds.iris.edu/data/reports/XH_2014_2017/									
#delimiter: ','									
#lineterminator: '\n'									
#field_unit	ISO8601	unitless	unitless	unitless	unitless	degrees_north	degrees_east	meters	meters
#field_type	datetime	string	string	string	string	float	float	float	float
Method/Identifier	StartTime	Network	Station	Location	Channel	Latitude	Longitude	Elevation	Depth
GPS Q330 GPS Clock	2014-12-31T23:00:40Z	XH	DR01	*	*	-77.77508	178.34172	30	0
GPS Q330 GPS Clock	2015-12-31T03:10:28Z	XH	DR01	*	*	-77.76594	178.34611	19	0
GPS Q330 GPS Clock	2016-01-20T01:08:44Z	XH	DR01	*	*	-77.75806	178.34989	16	0
GPS Q330 GPS Clock	2014-12-31T23:30:38Z	XH	DR05	*	*	-78.64047	-179.09994	19	0
GPS Q330 GPS Clock	2015-12-31T22:50:24Z	XH	DR05	*	*	-78.63164	-179.09239	9	0
GPS Q330 GPS Clock	2016-11-16T17:53:01Z	XH	DR05	*	*	-80.86433	178.43481	8	0
GPS Q330 GPS Clock	2014-12-31T23:43:19Z	XH	RS01	*	*	-78.18889	169.96239	14.3	0
GPS Q330 GPS Clock	2015-12-31T17:58:39Z	XH	RS01	*	*	-78.18333	169.965	1.4m	0

Tim Ahern & Joel Simon

Orientation Element Onshore/Offshore OBS Example

#dataset: GeoCSV 2.0								
#created: 2023-06-19T18:52:20Z								
#Reference doi: 10.1785/0120160165								
#delimiter: ','								
#lineterminator: '\n'								
#field_unit	ISO8601	unitless	unitless	unitless	unitless	SEED Convention	Seed Convention	azimuth uncertinty
#field_type	datetime	string	string	string	string	float	float	float
Method/Identifier	StartTime	Network	Station	Location	Channel	dip	azimuth	azimuthal uncertainty
DLOpy	2006-04-22T00:00:00Z	YS	PL38	0	BH1	-90	179.37	2.20
STACH	2006-04-22T00:00:00Z	YS	PL38	0	BH1	-90	179.58	7.10
Laske et al	2006-04-22T00:00:00Z	YS	PL38	0	BH1	-90	178.60	0.73
DLOpy	2006-04-22T00:00:00Z	YS	PL40	0	BH1	-90	168.32	1.00
STACH	2006-04-22T00:00:00Z	YS	PL40	0	BH1	-90	278.38	3.61
Laske et al	2006-04-22T00:00:00Z	YS	PL40	0	BH1	-90	277.14	0.87
DLOpy	2006-04-22T00:00:00Z	YS	PL47	0	BH1	-90	277.13	1.14
STACH	2006-04-22T00:00:00Z	YS	PL47	0	BH1	-90	278.38	3.61

Orientation Element

DISC
Ut er
ea co

Tim Ahern & Joel Simon

Include only the RCM elements you need

- For instance, consider a seismometer on a glacier that is changing position rapidly and whose orientation is also changing due to changes in the direction of the surface of the glacier

Time and NSLC Element	Position Element	Orientation Element
Start Time End Time Network Station Location	Latitude Longitude Elevation Depth	Dip Dip +/- Azimuth Az +/-.

Tim Ahern & Joel Simon

Next Steps

- Including GeoCSV into SEED StationXML
  - Option 1: append needed GeoCSV information at the end of a StationXML file
  - Option 2: merge elements of GeoCSV interspersed in the StationXML to reduce the size of the StationXML
  - Both options would make the resulting StationXML very large in many cases
  - Option 3 (preferred): (preferred) embed a single addition to StationXML at the station/channel level pointing to an external file containing the GeoCSV when one exists

Tim Ahern & Joel Simon

Summary

- GeoCSV files have enabled projects to capture RCM in a straightforward and efficient manner
- The format enables FAIRness by identifying what the fields contain and what variable types they are
- GeoCSV is flexible as new elements can be added when needed
- GeoCSV can document how derived metadata was obtained to even the level of what algorithms were used
- GeoCSV is an easy way to capture objects that make the data FAIR

Tim Ahern & Joel Simon

Thank you for your attention