Working Group 2
Data exchange

Lisbon, 2nd September 2025

Agenda

- Review and Approval of 2023 Minutes from Berlin
- Status of 2023 Action Items and summary of activities
  - FOLDS
  - RCM
  - Status of DAS Metadata proposal

Agenda

- Review and Approval of 2023 Minutes from Berlin
- Status of 2023 Action Items and summary of activities

Presentations

- DAS Metadata Proposal (J. Carter)
- Controlled Vocabularies for Seismology (A. Strollo/F. Haslinger)
- Data Formats for Large Datasets (L. Ermert)
- Recommendations/updates to StationXML from WG5 (W. Crawford)

- A.O.B.

Review and Approval of 2023 Minutes from Berlin

https://docs.google.com/document/d/17cYdIjYkz7vtMVNde0p0V009AjrAaFSIRqwgoJmUtII/edit?usp=sharing

Action Items from 2023

- **[Action Item – DAS Metadata]:** Foster integration and participation of WG2 (members) in the DAS RCN discussions. Keeping regular coordination between both and similar running projects would be desirable.
  - ☑ Proposal submitted and now under review!

- **[Action Item – miniseed3]:** Update all available resources in the FDSN site
  - ☑ Updated at https://docs.fdsn.org/projects/miniseed3/

Action Items from 2023

- **[Action Item – Development for Legacy data]:** Contact interested WG members who would like to form part of an advisory/review group to provide feedback during the development phase.
  ☑ Many interested data centres to test the development (T. Ahern).

- **[Action Item – Controlled vocabulary]:** Invite interested people to define a technical support for this, and specialists to populate the vocabularies. A link to the Action Item about Event types seems natural and should probably be considered.
  ☑ Approach presented during this Meeting.

Action Items from 2023

- **[Action Item - StationXML]:** Call for interested members to form a review team for the available guidelines and provide a recommendation to the WG2.
  👨💻 Meetings held with WG5 members. Next steps to be defined here.

- **[Action Item - QuakeML]:** ETH will open the discussion about QuakeML on the WG, including new functionality planned to be included in v2.0. (John Clinton).
  📌 ETH intends to develop in the next 1-2 years, but resources are scarce and it's not a top priority. Other teams would be welcome to proceed, with our collaboration, if they would like.

Action Items from 2023

- **[Action Item - Event types]:** Dmitry Storchak has agreed to lead an effort that will span both FDSN WG2 and COSOI regarding Event Types. He will inform the community about the progress on this discussion and joint effort. (Dmitry Storchak).
  📌 To be included in the Controlled Vocabularies approach?

- **[Action Item – Reproducibility]:** Call interested members to write a proposal for this, if enough support is found in the community.
  ✕ Not enough interest in the original proposal.

Action Items from 2023

- **[Action Item – GNSS Integration]:** GFZ will continue work to advance the integration of this type of data into our data centres. Also, to provide a guideline and/or a white paper on this topic.
  👨‍💻 Work in progress during EarthScope-ORFEUS Meetings. Session proposal for IASPEI 2027: Data Management for Large Datasets from seismology, GNSS and other disciplines (storage, formats, services).

- **[Action Item with WG3 – XML to JSON conversion]:** Form a group to define a JSON representation of StationXML and QuakeML.
  👨‍💻 Group formed, approach defined and first meeting set.

WG2 activities from 2023 to 2025

- DAS Metadata Schema - proposal preparation.
- Data formats EarthScope-ORFEUS collaboration
- Update of the Recommendations for DOIs
- Paper about the success story of 10 years of DOI adoption in Seismica (Thanks Wednesday group!)
- JpGU 2025 presentation to foster adoption of community seismological standard formats/services (jointly with WG3).
- Federation of Online Legacy Data in Seismology (FOLDS)
- Rapid Changing Metadata (RCM)
- Recommendations/updates to StationXML from WG5 (W. Crawford)

DAS Metadata Schema - Proposal

Toward a Metadata Standard for Distributed Acoustic Sensing (DAS) Data Collection

Voon Hui Lai[*1], Kathleen M. Hodgkinson[2], Robert W. Porritt[2], and Robert Mellors[3]

Cite this article as Hui Lai, V., K. M. Hodgkinson, R. W. Porritt, and R. Mellors (2024). Toward a Metadata Standard for Distributed Acoustic Sensing (DAS) Data Collection, *Seismol. Res. Lett.* **95**, 1986–1999, doi: 10.1785/0220230325.

Abstract

With increasing geophysical applications using distributed acoustic sensing (DAS) technology, there is a need to implement a metadata standard specifically for DAS to facilitate the integration of DAS measurements across experiments and increase reusability. We propose a metadata standard intended primarily for the DAS research community, which fully describes the five key components of a DAS experiment: (1) interrogator; (2) data acquisition; (3) channels; (4) cable; and (5) fiber. The proposed metadata

DAS Metadata Schema - Proposal

- DAS-RCN group (finished in May 2023)
- Paper published in SRL (Sept 2023-May 2024)
- Proposal preparation and technical implementation (Oct 2024-Mar 2025)
- Proposal submission (April 2025)
- Proposal accepted (May-Jun 2025)
- Contact Review Team for a full technical review (Sept 2025)
- Final result expected by the end of the year

EarthScope – ORFEUS collaboration

- 1st Meeting (Vienna): April 2024
  - DAS Metadata schema
  - DOI metadata curation
  - Improve/foster usage of FDSN Data Centre Registry
- 2nd Meeting (Corfu): September 2024
- 3rd Meeting (Vienna): May 2025
  - Test different approaches for Large-N and DAS experiments
  - Apache Iceberg, TileDB, Zarr
  - S3 storage as backend (AWS and on-premise)

Recommendations for DOIs (Dec 2023)

Summary of changes (compared to 1st version)

- Use `resourceTypeGeneral="dataset"`.
- Use ORCID and ROR where available.
- `<formats>` with two `<format>` elements.
- Updated citation examples in "Citation Format".
- Landing page suggestions.
- Updates to reflect that FDSN DOI services are operating and that StationXML has been updated.

REPORT
doi:10.26443/seismica.v4i1.1537
SEISMICA

All details in Rob Casey's presentation on Friday at 8:30!

DOI, licence and citation uptake for seismological waveform data after 10 years of implementation effort

Helle A. Pedersen * [1,2], **Jonathan Schaeffer** [2], **Florian Haslinger** [3], **Rob Casey** [4], **Javier Quinteros** [5], **Lesley Wyborn** [6,7], **Elisabetta D'Anastasio** [8], **Jonathan B. Hanson** [8], **Jerry Carter** [4]

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTerre, 38000 Grenoble, France, [2]Univ. Grenoble Alpes, CNRS, INRAE, IRD, METEO-FRANCE, OSUG, 38000 Grenoble, France, [3]Swiss Seismological Service at ETH Zürich, Sonneggstr. 5, 8092 Zurich, Switzerland, [4]EarthScope Consortium, 1200 New York Avenue NW, Suite 454, Washington, DC 20005, USA, [5]GFZ German Research Centre for Geosciences, 2.4 Seismology, Potsdam, Germany, [6]National Computational Infrastructure Australia, Canberra, Australia, 2601, [7]AuScope Ltd, Melbourne, Australia, 3053, [8]GNS Science Te Pū Ao, 1 Fairway Dr., Lower Hutt, WEL, 5011, New Zealand

Pedersen, et al. (2025). Seismica, 4(1). https://doi.org/10.26443/seismica.v4i1.1537

JpGU 2025 – WG2/WG3 presentation

- FDSN Standards for Data and Services. A success story and a vision for the Next Generation Data Services.

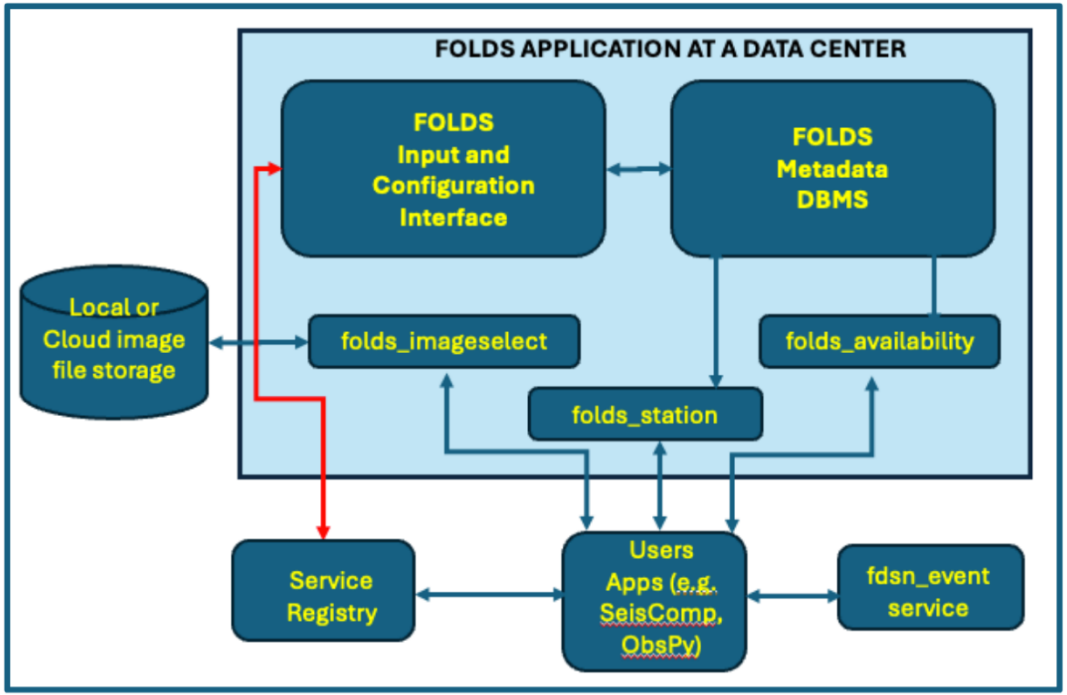
- We presented the current FDSN standards for data and services.
  - StationXML, miniseed(3), impact of DOI adoption

- Potential roadmap from WG2 and WG3.
  - DAS (meta)data formats
  - Manifest file (tomorrow in WG3 meeting)

Federation of Online Legacy Data in Seismology

- Metadata for legacy adopted during 2023.
- Funding to develop an application running at individual legacy datacenters forming a federated system.
- Work statement developed and will be sent to potential developers.
- 15 data centers with significant collections of analog data have agreed to be part of a beta test of the FOLDS application.
- A proposal is currently in preparation and will be submitted to the US National Science Foundation in December of this year.
- It is hoped that FOLDS will be funded and the beta test complete by the 2027 meetings in South Korea.

Federation of Online Legacy Data in Seismology



A Diagram Showing the Components of the FOLD Application

Rapid Changing Metadata (RCM)

- MERMAIDS have RCM due to movement in the water column and changes in latitude and longitude (10s of degrees).
- Roughly 5 years of RCM data have been captured in GeoCSV.
- Incorporation into StationXML. Recommendation is to support two options for including RCM data.
  1. Append RCM data at the end of StationXML volumes. This is suitable for use cases where the volume of RCM data is modest.
  2. Support of external reference at the station, or channel, level to GeoCSV files.

Rapid Changing Metadata (RCM)

1. Modify StationXML to allow RCM data to be included at the end of a StationXML file. This would likely be a simple modification.

2. Allow links to external files to be included at the station and/or channel portion of StationXML.

Similar case as other suggestions from WG5.

XML to JSON conversion

- **In collaboration with WG3**, we contacted interested people (seismologists+developers) to work on this topic.
- First meeting already organized for the group working on StationXML.
- List of interested persons to work on QuakeML. A Call still needs to be defined.

Presentations

- Proposal for a DAS Metadata Schema (J. Carter)
- Controlled Vocabulary for Seismology (A. Strollo)
- Data Formats for Large Datasets (L. Ermert)
- Recommendations/updates to StationXML from WG5 (W. Crawford)

Acknowledgements

- All reviewers that helped us during these last years. We need your active participation. A lot of topics waiting for interested WG members.

- Wednesdays group: Rob Casey, Jerry Carter, Elisabetta D'Anastasio, Jonathan B. Hanson, Florian Haslinger, Helle A. Pedersen, Jonathan Schaeffer, Angelo Strollo, Lesley Wyborn