WGII Meeting Minutes

Berlin, Germany

July 15th 2023

Chair: Javier Quinteros
Vice-Chair: Rob Casey
Minutes:

Attendees

Javier Quinteros - GEOFON/GFZ	javier@gfz-potsdam.de
Chad Trabant - EarthScope	
Jerry Carter - EarthScope	
Tim Ahern - EarthScope	
Carlo Cauzzi - ORFEUS	
Diego Franceschi - INGV	
Alfonso Mandiello - INGV	
Peter Danecek - INGV	
Florian Haslinger - ETH	
Martin Vallee - IPGP	
Wen-Tzong Liang - IESAS	
Yinshuang Ai - CAS	
Jan Wiszniowski - IGPAS	
Ludek Vecsey - IGCAS	
Nick Ackerley - NRCan	
Peter Voss - GEUS	
Pierre Sakic - IPGP	
Wayne Crawford - IPGP	
Gaby Laske - SIO, UCSD	
Nikolaus Horn - Geosphere Australia	
Milena Moretti - INGV	
Lucia Margheriti - INGV	

John Clinton - ETH	
Klaus Stammler - BGR	
Xyoli Pérez Campo - CTBTO	
John Ebel - Boston College	
Josep Batlló - ICGC	
Seiji Tsuboi - JAMSTEC	
Michelle Guy - USGS/NEIC	
Renate Hartog - PNSN, U. Washington	
Jerome Salichon - GNS, New Zealand	
Jeff McGuire - USGS	
Thomas Lee - Harvard	
Angelo Strollo - GFZ	
Will Yeck - USGS	
Tom Carth - ISC	
Jonathan Hanson - GNS, NZ	
Ryan Gallacher - ISC	
Mark Chadwick - GNS, NZ	
Dmitry Storchak - ISC	

Being more than 5 representatives (at least 8) from this WG-II and the minimum attendance is considered to be satisfied.

Presentation

Presentation slides made to WG can be found at: https://www.fdsn.org/wg/wgII/

Review of Proposed Agenda

Proposed agenda was presented by Javier Quinteros. With no comments or additions, the agenda was adopted.

Review and Approval of Minutes from the previous Meeting in Hyderabad 2021

Javier Quinteros presented the minutes from the 2021 meeting. They were approved without objections.

Status of 2021 Action Items

J. Quinteros reviewed the action items from the last meeting.

The Action Items about "**Metadata for Legacy Data**", "**miniSEED3**", and "**StationXML**" have been accomplished. All proposals were through all needed steps and after the final voting phase they were accepted and adopted as standards (if applicable).

There were 2 Action Items about "**GNSS Integration**". Both of them are on-going activities. A dedicated session (JS05[1]) was organized at IUGG2023 and is being held on the same day of this meeting with these items in mind. All discussions have been streamlined there, and the activity will continue until the Items are completed.

The Action Item about "**QuakeML evolution**" is paused for the moment because resources are scarce and it is not a top priority for ETH. However, other teams are welcome if they would like to start with this, and of course with the collaboration from ETH if they like.

The Action Item about creating an "**Event Types Nomenclature**" is slowly advancing. A slide from Dmitry Storchak reporting on this was presented and it's included in the presentation.

Summary of Action Items: 3 Complete / 3 in progress / 1 Paused.

Recommendation about Metadata for legacy data

Tim Ahern presented a set of slides about the state of the art of his proposal about "Recommendations about Metadata for Legacy Data" (which was approved and is now a standard). Also about the plans for future funding possibilities to create an application to help capture this information, a complete specification for a data provisioning system, which mimics somehow the web services for normal waveform data and its related metadata. Tim would like to create a Review Team to check and provide feedback about the evolution of this project during its expected lifetime, if it gets funded (around 2-3 years). The idea is that the evolution of this new proposal can be followed from the beginning by the seismological community and keep FDSN involved in it.

---

[1] JS05 Real-Time GNSS Data and Products Usage: Interoperability and Management Challenges (IASPEI, IAG, IAVCEI, IAPSO). Convener(s): Angelo Strollo (Germany, IASPEI). Co-Convener(s): Antonio Avallone (Italy, IAG), Yuhe Tony Song (USA, IAPSO), Clinton John (Switzerland, IASPEI), Giuseppe Puglisi (Italy, IAVCEI)

[Action Item]: Contact interested WG members who would like to form part of this kind of advisory/review group.

Future of miniSEED3

Chad Trabant presented details about the status of miniseed3, which had been approved as a standard some months ago. The new version of libmseed will natively support both versions (2 and 3), but will internally work with version 3. Convertors are also provided to make the transition as simple as possible for users.
A special remark was raised regarding the new source identifiers. All data centres and software providers will need to adapt and check that they will be able to support these new codes in the future. For instance, network/station names are up to 8 characters long and more advanced semantics could be included by means of underscores.
All documentation, links and related resources, which have to be updated, will be made available in the next few days on the FDSN site.
There are no plans to deprecate miniseed2.

[Action Item]: Chad to update (with the help of FDSN web masters) all available resources in the FDSN site.

DAS Metadata: Work done in the DAS RCN group

Jerry Carter presented a summary of the state-of-the-art regarding the definition and development of a metadata format for new DAS datasets. This has been coordinated within the DAS RCN group and the result is currently available on Github. They are working in a publication providing details about this proposal.
Jerry mentioned also that for the time being there will be "derived products" (downsampled versions) of the raw data in order to make them available/integrated in our normal systems. Full resolution data could go to Glacier storage and be provided on demand. The discussion on potential data formats to adopt is only starting.

[Action Item]: Foster integration and participation of WG2 (members) in the DAS RCN discussions. Keeping regular coordination between both and similar running projects would be desirable.

Controlled vocabulary for seismology

Angelo Strollo/Florian Haslinger called for collaboration by members who could be interested in defining and populating a controlled vocabulary to be used in the seismological context. Currently, our community does not have a standard vocabulary and we are using vocabularies from other disciplines that help us to standardize the usage of some information in a controlled way.

Wayne asked if this could also help in the case of units inside our StationXML metadata. This is actually true and one of the simple use cases we could have. We could also facilitate data discovery by tagging datasets in a way that researchers from other disciplines can easily understand what is it about without the need to know complex seismological metadata structure and semantics.

[Action Item]: 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 proposed.

Guidelines and checks for StationXML documentation

J. Quinteros mentioned that a review team will be formed to evaluate available guidelines and best practices about how metadata should be included in StationXML for different very particular cases. For instance, OBS data, Structural Monitoring, validation and coherency of StationXML metadata, and some others.
The idea will be to critically review these documents and add them to the official StationXML documentation.
As there are also some automated validation tools used in the community, this could also be a topic to include in this improvement of the documentation.

[Action Item]: Call for interested members to form a review team for the available guidelines and provide a recommendation to the WG2.

Reproducibility: how to define a dataset?

J. Quinteros presented a suggestion related to the reproducibility aspect of results based on our datasets. It is becoming normal that Journal Editors ask authors to provide a link to the dataset used for the paper. This is actually impossible in most of the cases, because of the storage needed to host it, and because it would need a DOI to be minted for this. Quite impractical. However, Datacite suggests a solution that seems feasible for our community. Namely, to issue a persistent identifier (DOI) to a query, which could be re-executed. Therefore, one could think of a plain text file with a detailed list of streams and time windows in a POST format (as specified for FDSN web services), which could be identified by a DOI. Even simple services like Zenodo could provide a DOI for this.
The main benefit is that the reader of a paper will know exactly which data needs to have to reproduce the results. FDSN (or someone else) could easily provide a client to regenerate the dataset.

[Action Item]: Call interested members to write a proposal for this, if enough support is found in the community.

Summary of retained and on-going action items

[Action Item - QuakeML]: ETH will open the discussion about QuakeML on the WG, including new functionality planned to be included in v2.0.
Responsible: John Clinton

[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.
Responsible: Dmitry Storchak

[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.

A.O.B.

Adjourn

Additional Resources

http://docs.fdsn.org/projects/stationxml/en/latest/