Kobe 2017, FDSN Working Group III Minutes

2017 WGIII Attendance list: (25 in attendance)

Goran Ekstrom	ge21@columbia.edu	Columbia Univ
Jose Antonio Jara	joseantonio.jara@icgc.cat	ICGC
Bruce Beaudoin	bruce@passcal.nmt.edu	PASSCAL
John Clinton	j.clinton@sed.ethz.ch	ETHZ / SED
Fabien Engels	Fabien.engels@unistra.fr	RESIF
Marc Grunberg	marc.grunberg@unistra.fr	RESIF
Catherione Pequegnat	Catherine.pequegnat@univ-grenoble-alpes.fr	RESIF
Ruifeng Liu	liurf@seis.ac.cn	Institute of Geophysics, CEA
Alison Bird	alison.bird@canada.ca	Geological Survey of Canada
John Adams	john.adams@canada.ca	Geological Survey of Canada
David Jepsen	david.jepsen@ctbto.org	CTBTO
Wen-Tzong Liang	wtl@earth.sinica.edu.tw	IES Academia Sinica
Seiji Tsuboi	tsuboi@jamstec.go.jp	JAMSTEC, Japan
Florian Haslinger	haslinger@sed.ethz.edu	ETH Zurich
Mark Chadwick	m.chadwick@gns.cri.nz	GNS Science
Katrin Hafner	hafner@iris.edu	IRIS
Tormod Kværna	tormod@norsar.no	NORSAR
Reinould Sleeman	sleeman@knmi.nl	ORFEUS/KNMI
Xyoli Pérez Campos	xyolipc@sismologico.unam.mx	UNAM, Mexico
Chad Trabant	chad@iris.washington.edu	IRIS
Dmitry Storchak	dmitry@isc.ac.uk	ISC
Yinshuang Ai	ysai@mail.iggcas.ac.cn	IGGCAS, China
Adrienne Moseley	adrienne.moseley@ga.gov.au	Geoscience Australia
Wanye Crawford	crawford@ipgp.fr	IPGP
Tim Ahern	tim@iris.washington.edu	IRIS

Tim Ahern chairman
The meeting began at 12:03 Room 503.

Tim Ahern outlined the agenda with one added item related to allowing “https” for FDSN web-services. There was also a request by John Adams to have a discussion related to event types, it was indicated that there is a related item on the agenda where it can be discussed.

Tim Ahern requested comments or amendments to the minutes from the 2015 meeting in Prague. John Clinton raised the issue of “action items” in the minutes and what is the procedure for following up on the outcome of the noted actions. **[Note: the list of action items and their current status from the 2015 WGIII meeting are included at the end of these minutes]**

John Clinton moved that the minutes be adopted this was seconded by Göran Ekström and passed unanimously.

Recommendations for changes to current FDSN web services.
Chad Trabant outlined a change to the "fdsn-station" service and the use of the "matchtimeseries" parameter. This is used at IRIS to request metadata information for station time periods where data is likely to be available. The IRIS service uses broad "time extents" rather than fine grained data availability details.

Florian Haslinger asked whether this would require a "service" which had access to a data store. Chad Trabant replied "yes" and the implementation of this may need to be optional, or at least not required. Reinoud indicated that this was too strong, and Tim Ahern suggested "recommended". Florian noted that many features get promoted by being useful and suggested "optional".

Mark Chadwick asked whether this information also existed in StationXML, Chad Trabant noted that there was an extension to StationXML that IRIS uses for this purpose.

Tim Ahern suggested that it should be left as optional and that he would talk with GEMPA about adding the capability to SeisComp3.

Action Item 2017:1 IRIS to investigate with GEMPA about the possibility of adding this feature to SeisComp3. General consensus was to add matchtimeseries as an "optional" parameter in the station service.

Time format
The next topic was to allow "Z" as a suffix on all StationXML time values. There was no opposition but it was noted that it should not be a requirement.

Event Types
John Clinton led the conversation regarding the topic of adding an "event type" to the event service and that it should use QuakeML 1.2. The changes outlined were to the specification and to the text output format. He also asked the question as to what the default should be, either "earthquake" or "*".

Alison Bird outlined a document giving the requirements for discriminating event types as needed in Canada. Highlighted were two aspects not included, namely the language and the audience. In Canada there is a requirement for both French and English as well as the ability to add internal comments that are not intended to be published.

Alison Bird also described a requirement for a better event type hierarchy, such as subcategories of mining events. [Further details could be found in the document].

John Clinton suggested that the group not get into setting the event type definition. John Adams asked about the process for getting changes made in QuakeML. There was more discussion with Dmitry Storchak outlining the difficulty with any update path as all the English letters used to denote event types have been used.

Chad Trabant indicated that the change to the text output format would also require a new format version. Although it would be most practical if the event type could be added as a field at the end of each row.

Tim Ahern posed the question of where the Event Type classifications should be updated. Florian Haslinger mentioned that it came out of the work of NEIC-ISC-EMSC and as a result best describes events of interest of those who built the list, he also noted that they had run out of alphabet letters.

Dmitry Storchak noted that the system was designed so that it could work with other formats and not just QuakeML. John Adams raised the question of whether the ISC wants to move to QuakeML. Dmitry Storchak indicated multiple issues that require keeping the formats concurrent.

Tim Ahern noted that this was not for this Working Group and asked whether QuakeML could be moved to an FDSN standard.

John Clinton said the ETH was not actively developing the format and would like to move it to Working Group II.

There was no opposition from the room for Working Group II to handle QuakeML.

Action Item 2017:2 WG III recommends that QuakeML be made an FDSN standard and managed within WGII.

Waveform data availability service

Chad Trabant hoped that the working group would adopt the “waveform data availability” webservice as an official service.

Tim Ahern indicated that there may be a SeisComP3 solution. John Clinton suggested that it was more related to the EIDA data nodes than to SeisComP3 installations. He also mentioned that the metrics calculated should also be the same. [Note: the availability web service returns lists of start stop times of a timeseries. Metrics can be derived from this information (e.g. number of gaps, % available, etc) but it is not currently an agreed upon metric that is defined by any FDSN member.]

Florian Haslinger asked about what the service produces, Chad Trabant outlined that it was a raw list of time segments with start and end times at a fine grain level.

Reinoud Sleeman suggested that this was related to the QC metrics service, Tim Ahern thought that they were related but slightly different in the information being provided. Reinoud Sleeman also indicated that it would likely only be the EIDA nodes as it would not be so easy for base SeisComP3 installations. Ahern will investigate if the data availability service can be added to SeisComp3 to benefit a broader cross section of the FDSN community.

Waveform timeseries service
Chad Trabant outlined the IRIS timeseries service. It allowed processing on the fly and was popular from a user perspective.

John Clinton asked whether this could be achieved as an FDSN extension for a format change.

Waveform timeseriesplot service
Chad Trabant outlined the IRIS timeseriesplot service. The IRIS implementation allowed the user to build a URL and request a plot be generated. There was discussion on the output formats and the length of record plotted. John Clinton noted that this would be something that could be implemented in SeisComP3 relatively easily.

Waveform services
Tim Ahern called for a show of hands related to the three services. Results were more than half for the "availability" service (general consensus), about half for the timeseries service (general interest) and roughly half for the timeseriesplot service (some interest).

Action Item 2017:3 Ahern will see if these three capabilities (availability, timeseries, and timeseriesplot) can be added to SeisComp3 if funding can be found.

Network DOIs
Tim Ahern gave an overview of the FDSN DOI history and reported on the current status. More than 1/3 of permanent networks have a DOI minted and over ½ of the temporary networks. John Clinton asked whether these were only minted for PASSCAL deployments, Tim Ahern mentioned that DOIs could be minted for other deployments but the operator has to request them. PASSCAL has a standing policy that all temporary experiments have a DOI minted when the network code is assigned.

Reinoud Sleeman noted that the IRIS minted DOIs were not so rich in meta-data. Tim Ahern indicated that all fields may not have been filled in at the time of minting. John Clinton also noted that the form may not have all the information.

Action Item 2017:4 Tim Ahern to look at the DOI form and see if all required metadata be supported on the FDSN form from which DOIs are minted. Adjustments in the form may result.

It was suggested that networks should be tracked down and asked for DOIs. Tim Ahern noted the difficulty with maintaining the contacts databases and broader help from FDSN members is needed.

Federated Data Centres

Tim Ahern outlined the expansion of Federated Data Centres. The implementation of SeisComP3 allows federated data centres (those supporting at least the station and dataselect services) to be more easily established. IRIS is helping to establish SeisComP3 installations in Africa and Central Asia. JAMSTEC has a service but requires some investigation as to how to handle authenticated access to the data. Installations have also been ongoing in many Asian countries including Thailand, Nepal and Singapore. Tim Ahern is trying to contact more networks in Asia. Wen-Tzong Liang mentioned that his institute in Taiwan was not allowed to release data collected from the national agency (CWB) and asked whether there was any way to release only some data. The response was that SeisComP3 had a configuration flag for this situation.

Florian Haslinger asked what was meant by a "Federated Data Centre" to which Tim Ahern replied that it was a centre making FDSN webservices available and often implied an installation of SeisComP3.

Tim Ahern showed how the QC federated data centre availability metrics have been improving over time.

Florian Haslinger noted that he thought federation meant going to one data centre and having the ability to get back data from all other data centres.

Action Item 2017:5 Continue contacting organisations to become federated data centres. This would mean they still have control over their data, rather than sending it to another data centre where control could be lost.

Göran Ekström noted that there was no service to find where "this is the home datacentre for this network" rather than finding data but not knowing what is the authoritive datacentre.

Tim Ahern noted that there were two systems, IRIS and EIDA. Göran Ekström suggested that there should be a single federation (FDSN) service.

Tim Ahern described the IRIS system which used a cataloging mechanism, while Reinoud Sleeman described the routing service that EIDA provides. This is a higher-level service and that clients need to know how to utilise this information.

Tim Ahern then showed the data volumes currently being distributed. The amount of data in the federated cengtre is a bit more than twice that of IRIS alone and shows the power of the federated concept from a user's perspective. John Clinton noted that there

will be a change in Europe when the switch to webservices is made in six months time. Tim Ahern noted that SCEC noticed three times more data being distributed after their webservices went online.

Datacentre Registry
Chad Trabant spoke to the need for a data centre registry. He asked "should there be a central registry?" It would allow discovery of data and information about data centres including service endpoints.

Some options were presented, these included:

re3data.org (currently used by GEOFON, RESIF, IRIS).
roll you own perhaps under "fdsn.org".
or do something else.

It was suggested that this could form part of a working group. John Clinton suggested ETH could find people to work on it and Tim Ahern mentioned EarthCube has resources for this sort of effort and that they are pursuing a solution using re3data.org.

Action Item 2017:6 Explore building a working group to work on the issue of a registry for FDSN services.

Using https for web services
Reinoud Sleeman reported that there are European regulations that require services to be run over https but that currently running the services using port 443 violates the specification. There was discussion on why the port was in the standard in the first place and that there was a requirement that both be possible. Chad Trabant noted that moving to https would be tricky but that it would be easier if it was optional.

Action Item 2017:7:
Further discussion related to the use of https should be handled via email, but that a time limit should be imposed to allow for closure.

Quality Assurance
Reinoud Sleeman mentioned the "wfcatalog" service and coming to a consensus on what metrics are used. This needed co-ordination.

Tim Ahern outlined the Mustang system whose metrics are written in "R", it is a big and complex system and requires accessing the entire data archive when new metrics need to be computed. IRIS has also built a lighter system called ISPAQ. This is aimed at portable deployments or use by small networks and can calculate metrics in the field or at small data centers. It is publicly available and would be worth considering if other groups are planning to implement quality assurance algorithms.

Other Items
Wayne Crawford raised the question as to how, or where, to place extra information about processing steps that may be associated with the presented data, this is especially relevant to OBS collected data.

Tim Ahern said there were standard ways of doing it but that they can be non-trivial, and asked how should FDSN capture provenance information.

Chad Trabant suggested that a standard be adopted in Working Group II.

Action Item 2017:8 Develop a standard method of provenance within working group II.
Responsibility: WG II
Meeting finished at 13:30

Notes taken by Mark Chadwick.

Action Items and Status from 2015 FDSN WG III meeting in Prague

Action Item 2015:1 ETHZ will submit a proposal to include event type in fdsnws-event specification.
Status: Submitted by ETHZ and to be discussed at this meeting.
Considered closed with minor action required to update documentation.

Action Item 2015:2 Ahern to send email to request added capabilities to a processed data service.
Status: Being raised again at this meeting. No substantive progress since Prague.
Now considered closed as Action Item 2017:3 addresses this.

Action Item 2015:3 need to propose metadata versioning *task force* that can make recommendations to WGII (DMC?). Trabant offered to lead the effort.
Status: Unknown

Action Item 2015:4 Ahern: If you're interested in being on a FDSN Federated System task force- send Tim an email and it would be wonderful to get participation from other than US/Europe. The intent is to form a *task force* for this activity.
Status: The number of federated members has grown substantially. Currently there are 17 nodes that are participating with several other data centers with plans to participate. Specifically 5 data centers in Africa, 1 in Central Asia, 1 in Canada and 1 in Mexico are moving in this direction.
Considered closed as an action item but on-going.

Action Item 2015:5 Ahern to circulate an email to form a *task force* to discuss standardized quality assessment approach.

Status: This activity was focused on standardization of metric definitions and standardization of service invocations, Still a work in progress but as far as is known there is a harmonization of the US and European approaches.
Considered Closed.

Action Items from 2017 WG III meeting

Action Item 2017:1 IRIS to investigate with GEMPA about the possibility of adding this feature to SeisComp3. General consensus was to add matchtimeseries as an "optional" parameter in the station service.
Responsibility: Ahern

Action Item 2017:2 WG III recommends that QuakeML be made an FDSN standard and managed within WGII.
Responsibility: Clinton and Trabant as Chair and Vice Chair of WG II

Action Item 2017:3 Ahern will see if these three capabilities (availability, timeseries, and timeseriesplot) can be added to SeisComp3 if funding can be found.
Responsibility: Ahern

Action Item 2017:4 Tim Ahern to look at the DOI form and see if all required metadata be supported on the FDSN form from which DOIs are minted. Adjustments in the form may result.
Responsibility: Ahern

Action Item 2017:5 Continue contacting organisations to become federated data centres. This would mean they still have control over their data, rather than sending it to another data centre where control could be lost.
Responsibility: all FDSN members

Action Item 2017:6 Explore building a working group to work on the issue of a registry for FDSN services.
Responsibility: Clinton and Trabant/Stults (IRIS)

Action Item 2017:7:
Further discussion related to the use of https should be handled via email, but that a time limit should be imposed to allow for closure.
Responsibility: Sleeman