IRIS PASSCAL Data Policy

November 1, 2017

This policy provides the archiving requirements for data collected with IRIS PASSCAL equipment through the PASSCAL Instrument Center (PIC). IRIS’s policy is that delivery of data and associated metadata to the IRIS Data Management Center (DMC) is an obligation of the Principal Investigator (PI). It is important to IRIS that the PI acknowledges this obligation and meets it within the required time frame. Failure to complete this requirement not only deprives the community of a valuable data resource, but also may jeopardize future requests to borrow IRIS equipment. IRIS expects data delivery while the experiment is in the field for long-term deployments (via telemetry or service runs), or immediately at the conclusion of the field experiment. The DMC will make the data available only to the PI and their designated representative(s) for a period of up to two years after the completion of the experiment. The PI can waive this period of exclusive access. After the prescribed moratorium, the data will be made freely and openly available to the public. All metadata, including QC metrics, and any additional experiment reports will be immediately open and available to ensure data sharing after the moratorium period is facilitated.

The equipment in the PASSCAL facility represents a significant community resource. In addition, the data collected by the PI community using the PASSCAL equipment has tremendous value past the specific scientific goals of the original experiment. Abiding by this data policy ensures the continued contribution of the IRIS PASSCAL facility to seismological studies.

Digital Object Identifiers (DOI)

IRIS Data Services has worked with the FDSN to develop a procedure to create DOIs for data from temporary networks. DOIs promote the citation of a PI's work as well as the facility support for each experiment. All PASSCAL experiments with SEED network codes will have DOIs minted automatically when mobilization forms are submitted.

More information on DOIs can be found at: http://www.fdsn.org/services/doi/

Data

The actual format of the data and the amount of data depend upon the type of experiment. Most PASSCAL experiments are classified as passive source, controlled source, or more recently some combination of controlled and passive recording (mixed-mode). The current standards for data archive for passive and active source data are SEED and SEG-Y formats respectively. We are working on improved methods to archive data that will eventually include PH5. For current procedures for data archiving, please visit our site at:

https://www.passcal.nmt.edu/content/data-archiving

Passive Source Experiment Data

PIs conducting a passive source experiment will provide all of the data and appropriate metadata from their experiment to the IRIS DMC for archive in SEED format. It is expected that the PI will coordinate with the PIC to arrange delivery of the data on a continuing basis during the experiment (after service runs as appropriate, or in real time if available). The final data shall be delivered immediately after the experiment is over.

If PI’s send data directly to the DMC, the PIC must be notified in order to close out the experiment
requirements for data sharing.

**Controlled Source Experiment Data**

For conventional, shot-windowed data, data should be delivered in SEG-Y format. The data should include all of the necessary information on the geometry of the experiment (metadata) and SEG-Y data should be corrected for all known timing problems. Again, the formats are evolving, so refer to our data archiving site for the latest acceptable formats.

**Mixed-Mode Experiment Data**

More recently, experiments are being proposed and deployed that involve both controlled source and passive source experiment components. In these cases, we are working towards a single format to allow ingestion, reformatting and distribution of the particular components of the data in the appropriate formats as required by the data users (PH5). Refer to data archiving site for details.

**Experiment Report**

An Experiment Report provides information regarding the project's PIs, purpose of the experiment, description of its layout and execution, as well as details of any unusual situation encountered during project mobilization (changing locations, missing GPS timing, etc). All active source datasets are required to submit an Experiment Report. SEED format datasets are encouraged to include a report if the PI and PASSCAL staff deems it would be helpful for making the data more useful to future data users. For guidelines on what is required in an instrument report, please visit our site at: https://www.passcal.nmt.edu/content/what-needed-archive-your-active-source-data-set

**Non-Standard**

There will always be some experiments that do not fit directly into one of the above categories. In those cases the exact form of the data delivery will be negotiated between the PI, Data Services and PASSCAL.

**Proprietary Data**

Data of all types should be archived with the DMC, in the appropriate format with complete metadata, as soon as possible and well before the general release of the data. The DMC will only allow access to the waveforms to the PI and others designated by the PI during a moratorium period - not more than 2 years from the completion of the field experiment. If an experiment's field deployment is extended, it is recommended that the original data start to become accessible within 2 years of the originally planned demobilization date. This is to allow access to the early stages of datasets from experiments that have significant extensions (ie experiments that turn into temporary “networks”). Access will be controlled by the DMC with passwords given to the PI. The PI can share the password with anyone he/she wishes. The PI will be notified when anyone registers for access to a proprietary dataset.

Information about the experiment contained in the metadata and data quality characteristics will be made publicly available during the experiment; only waveform data will be limited in distribution during the proprietary period.

**Open Stations**

All passive experiments with five or more stations will designate at least one station as an “open station”. The data from the “open station(s)” will be made available to the public immediately upon being archived.

**Support Available from IRIS**
Every PI and their field staff need to attend training at the PIC prior to their experiment where they will receive training on instrumentation and installation techniques, and the handling of the data and development of the station metadata. Teams will be taught how to archive their data using PASSCAL field computers and software that is provided as a part of the experiment equipment set. Every field computer has the software necessary to accomplish the data delivery task and the PIC has personnel who can provide assistance to the PI during and after the experiment. We encourage PIs to interact with PIC staff before and after service runs and immediately after the experiment conclusion to allow refreshers on data handling and ensure immediate ingestion of the data to the DMC. Please note that many new tools for assessing data quality are now available for experiment datasets that can only be accessed once the data are uploaded.

The PI is encouraged to utilize these resources at all stages of the work. In all cases, however, the ultimate responsibility for delivery of the data rests with the Principal Investigator. The PI must ensure that adequate resources are budgeted to accomplish this task.

**Data Quality Assessment**

All data submitted to the IRIS DMC will be run through the data quality assessment tools (MUSTANG) and the data quality metrics will be made open without delay for data quality evaluation. MUSTANG data may be accessed through standard web services tools or through the use of the MUSTANG clients (like MUSTANGDataBrowser or LASSO). A list of clients can be found at [http://service.iris.edu/mustang/](http://service.iris.edu/mustang/)

**Experiment Completion and Delinquent Datasets**

A PASSCAL data submission is not considered complete until both the Principle Investigator and PASSCAL Data Group staff certify that data archived are complete, the metadata accurate, and any reports filed are sufficient to allow other members of the community to utilize the data.

Data are considered delinquent two (2) years after the demobilization of an experiment. PIC staff will remind PIs that data are due beginning 6 months prior to the data becoming delinquent and will offer guidance on how to ensure your data are properly uploaded to the DMC. If your dataset becomes delinquent, the IRIS PASSCAL Program Manager and PASSCAL Standing Committee chair will work with you to determine how best to make your data available to the community to ensure the program resources are being adequately allocated. We certainly hope that this escalation will not be required, but it is in the best interest of the entire IRIS community to ensure that when IRIS PASSCAL equipment is used, that the data are archived in a useable format as required.

**Exemptions for Strictly Educational Use Datasets**

If you are using equipment from the PASSCAL facility for strictly educational training on the instrumentation and are not using the equipment for observing a science target, PASSCAL management may grant an exemption from the data distribution policy if there is no useful information to be shared among the seismological community. This is at the discretion of the PASSCAL Program Manager. Of course, if you are willing to share the data, all datasets are welcome. All requests for exemption must be submitted to the PASSCAL Program Manager. Until the Program Manager grants an exemption the data are considered due to IRIS.

There may be other needs for exemptions, including complete failure of the experiment to capture data or no useable data were captured, and these will be granted on an as-needed basis.

This policy is effective as of November 1, 2017 and is subject to change and revision as needs dictate. For updated versions of the policy and additional information on data delivery see the PASSCAL and IRIS Data Services pages listed below. There are Data Services specific policies that could impact portable experiments – particularly those that create large data sets.