# Incorporated Research Institutions for Seismology

# Request for Proposal

# IRIS MT Software Development

# September 22, 2020

**RFP**

IRIS MT Software Development

## Table of Contents



Appendices:
Appendix A – Sample Contractual Agreement
Appendix B – EMTF Resources and Documentation (https://drive.google.com/drive/folders/1lBcMkusNokZmcMcSADEPmyPiazcoSe7d?usp=sharing)

# Request for Proposals
# IRIS MT Software Development

## I. Background

### A. Information about IRIS

The Incorporated Research Institutions for Seismology (IRIS) is a consortium of over one hundred U.S. universities dedicated to exploring the Earth's interior through the collection and distribution of heterogeneous geophysical data.

IRIS programs contribute to scholarly research, education, earthquake hazard mitigation, and the verification of the Comprehensive Test Ban Treaty, in addition to the long-term stewardship of collected data.

Support for IRIS comes from the National Science Foundation (NSF), other federal agencies, universities, and private foundations.

IRIS is a 501 (c) (3) nonprofit organization incorporated in the state of Delaware with its primary headquarters office located in Washington, DC.

### B. Information about the IRIS MT Program

Under its current cooperative agreement "Seismological Facility for the Advancement of Geoscience" (SAGE) with the National Science Foundation, IRIS is establishing a new pool of portable magnetotelluric (MT) systems at the PASSCAL Instrument Center (PIC) in Socorro, New Mexico. This effort aims to make crust and lithosphere-scale MT capabilities widely available to principal investigators (PIs) by providing modern instrumentation with user training and data handling support at no extra cost. Instruments from the PIC are used worldwide and often form the backbone of international, collaborative geophysical research. By utilizing the resources of the PIC, IRIS is providing a foundation to broaden the participation of PIs in MT research by increasing opportunities for experiments by non-specialists and encouraging combined seismic-magnetotelluric deployments.

## II. Project Description

## III. Tasks and Deliverables

### A. Description of Need

We invite proposals to develop software resources which will facilitate the exchange and research utilization of data collected by PIs using SAGE MT instruments from the PIC and also be broadly useful to the international community of MT practitioners:

1. Tools to write and read MTH5 (an HDF5 based exchangeable MT format) from common archival and downloadable data and metadata formats used by the SAGE facility.
2. A data processing workflow for multi-station MT surveys to process raw time series of electric and magnetic fields into transfer functions, the starting point for inversions of MT-based observations into models of Earth structure.

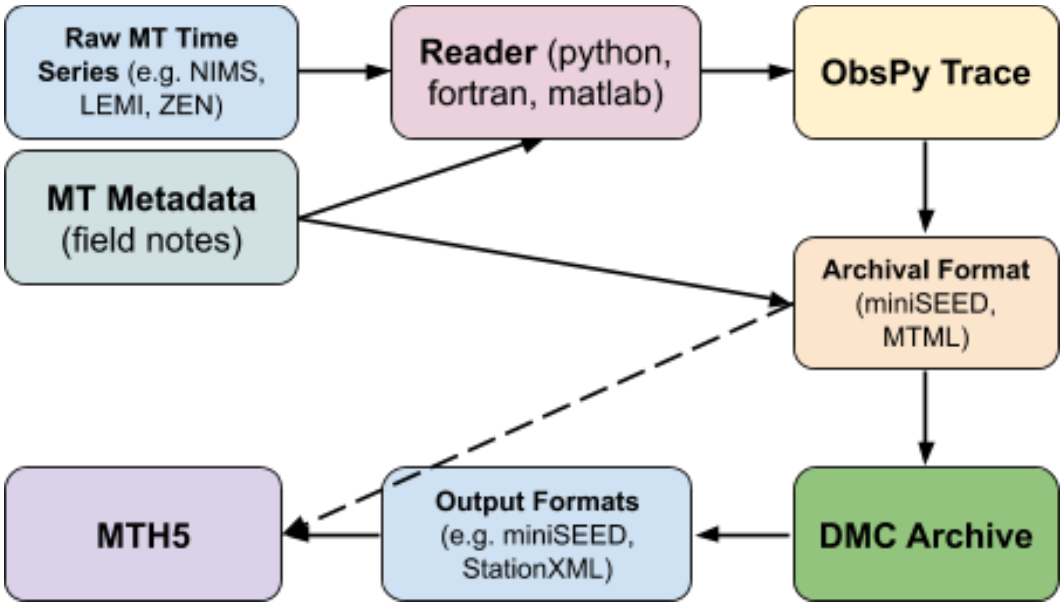
Contingent on the selected proposal, we may also utilize one development consultant with domain expertise in MT data processing and data storage formats to provide additional guidance and oversight of the project. Initial software resources, which these tools will either improve or replace, and opportunities for feedback on the development process and products will be available to the developer and project consultant throughout the project. We will describe these needs and provide detailed supporting documentation as part of this Request for Proposals (RFP).

## B. Requirement Details

**Element 1 - Tools to service a standardized MT data format (MTH5)**

Archived MT datasets use standards and formats that vary widely. The SAGE MT program loans instruments to PIs, who are in turn responsible for publicly archiving all collected data in the SAGE Data Management Center (DMC) within two years. The PIC supports software tools for PIs to ensure that their data adhere to established data and metadata formats, and initially for MT will use miniSEED for time series data and a variant of StationXML that includes MT fields (MTML) for metadata. The USGS also must publicly archive its MT datasets, and in collaboration with the MT community IRIS and the USGS have drafted MTH5 (https://mth5.readthedocs.io/en/tables/index.html), a provisional exchangeable MT data format with accompanying standardized metadata for storing MT time series and all supporting data and metadata.

We illustrate the stages of MT data flow from the instrument and field notes to the SAGE DMC and subsequent output to an end user. Because the archival and downloaded formats are not innately suited as an exchangeable research-ready format for MT, we require software tools to support MTH5.




The development of MTH5 should:

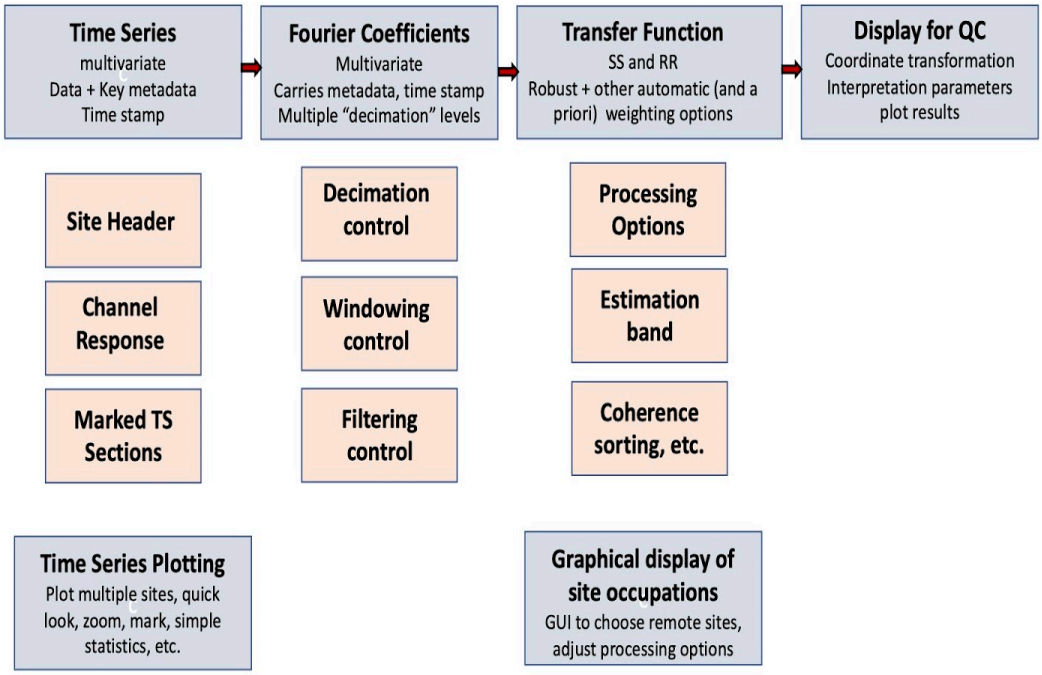
- Refine and complete, in collaboration with Jared Peacock (USGS), the generalized library for MTH5 (https://github.com/kujaku11/MTarchive).
- Provide tools to author, validate, scan, and query MTH5 files.
- Include the capability for MTH5 to store instrument responses in a look-up table for ease of reference (relevant to element 2) and include a routine to export responses as poles and zeros if needed.
- Include a module to author MTH5 from continuous time series and metadata archived with IRIS and downloaded using ROVER (https://iris-edu.github.io/rover/).
- Allow for parallel utilization of MTH5 (desired).
- Utilize architecture that is plug-in ready, such that specially authored readers for MT instruments can be modules for converting to MTH5.
    - Specifically, integrate the LEMI-424 reader being developed by the PIC as the first example of this module capability.
- Provide an example case using a group of to be determined stations taken from the EarthScope MT-TA dataset (http://ds.iris.edu/gmap/#network=_US-MT-TA&planet=earth)

**Element 2: MT Data Processing Software**

Gary Egbert (Oregon State University) authored MT data processing software (EMTF) that is well-recognized and widely used throughout the international MT community. This element will encompass a rewrite of the core EMTF workflow in Python and potentially a rewrite or reutilization of other supporting tools in this workflow. The original EMTF codebase and related components that were authored by Gary and others are provided. Our goal is to place these data processing tools into a consistent, modern codebase and leverage existing Python packages like xarray, dask, scipy.stats which support potential multi-processing. This software will serve a growing community of

new users, we would like this software to provide a simple, well-documented, user-friendly interface for MT data processing that is open source and designed in a way that encourages community contributions and targeted developments in the future. The workflow should be modular, such that alternative data processing schema can be substituted for different steps of the workflow.

This element encompasses all parts of the proposed data processing workflow (pictured below), which entails several different components:




- MTH5 will serve as the default input data format. For users with their own preferred format, this should be designed in a way such that they can substitute their own input module.
- These blocks relate to existing Fortran programs (or editable input/control files) in EMTF, and/or to matlab functions, scripts, or classes.
- The upper boxes (time series, Fourier coefficients, etc.) represent the main "objects" that need to be implemented and the main steps in routine processing. The object structure may be more granular--that is probably a programming decision.
- Support objects (site header, decimation control, etc.) are organized under the upper level object that they are first associated with--but in most cases they are (or at some point may be) used in other main objects.
- The lower boxes represent additional optional modules/objects which will be useful in practice. For example, for a quick look at data quality, a user may run through the upper level, with some default setting of options, ending with a quick initial plot of results. If those look good, that could be it (especially for in-field processing). If there are quality issues, the first step would be to plot the time-series, to get some idea of what the issues are. This may lead to various actions--

marking time series to omit some sections, modifying processing options, doing remote-reference processing with several candidate sites, etc. Interactive tools (which would be embedded in the two boxes in the lower row) would facilitate this.
- Transfer functions must be output in XML format (https://library.seg.org/doi/10.1190/geo2018-0679.1) that is ready to archive with the SAGE DMC (http://ds.iris.edu/spud/emtf).

**Element 3 (Development Consultant): To be utilized depending on the outcome of the solicitation**

The development consultant (DC) is envisioned as an optional component of this project, depending on the responses to elements 1 and 2. The DC will collaborate with the developer to ensure that elements 1 and 2 are faithful to the requirements of this project. We envision this task requiring up to 8 hours/week through the duration of the project. The consultant will:

- Have demonstrated experience in MT/EM data science.
- Serve as a front-line resource for consultation by the developer.
- Work collaboratively with developer and project team to validate all software.
- Participate in software design reviews.

## C. Other Criteria

- Anticipated funding for this project is **$65,000**. We are targeting $53,000 for elements 1 and 2 and $12,000 for element 3.
- Elements 1 and 2 should be bid together. Elements 1, 2, and 3 may be bid together by a developer with extensive subject matter expertise in MT and experience with MT data formatting and processing.
- Some resources (e.g. LEMI-424 data reader, schema for MTML formatting—similar to StationXML) for element 1 are under development and will be made available upon request of the developer.
- Element 2 will be supported by a suite of established MT data processing codes, pseudo code, commented algorithms, functional descriptions, and a video tutorial provided as part of this solicitation (Appendix B).
    - The main modules and supporting objects are required.
    - The optional modules are desired.
    - Implementation for high performance computing systems or multiple processors is desired.
- Responses to this RFP should include:
    - CV or professional resume
    - Examples of projects using Python
    - Examples of fully developed software tools, preferably open source
    - Examples of projects involving multi-processing and/or high-performance computing
    - Demonstrated familiarity with EMTF or knowledgeable commentary on documentation provided for the data processing workflow

    - Detailed description of proposed design approach to the software
    - Proposed development process and review milestones
- The IRIS project team will advise this project and participate in the review and acceptance of the deliverables. This team includes Gary Egbert (Oregon State University), Anna Kelbert (USGS), David Goldak (PIC), Jared Peacock (USGS), Maeva Pourpoint (PIC), and a TBD member of IRIS Data Services.

### D. Project Deliverables

As part of furnishing the data formatting and processing software resources requested above:

- Source code should be object oriented, using Python 3.6 or higher.
- Source code should be adequately commented.
- Software should be developed in a way to be extensible for future developments (i.e. substitution of data input or specific processing modules).
- Software will be fully open source and must rely on established open source libraries.
- Software must run on recent versions of common operating systems (e.g. Mac OS X and Linux, Windows).
- Documentation is sufficient to install and operate the software, authored in Sphinx, and portable to readthedocs.org with examples.
- Software must produce MTH5 files and expected transfer functions using an example dataset provided by IRIS.
- IRIS will review and approve all delivered products before paying an invoice.

## IV. Contractual Conditions

This Contractual Agreement will be issued as a Fixed Price Contract. Funding for this activity is provided to IRIS through a Cooperative Agreement from the National Science Foundation. Contractor(s) must comply with all applicable federal, state, local laws and regulations and all applicable orders and regulations of the executive and other departments, agencies, and instrumentalities of the United States as stipulated in the Agreement (Appendix A).

## V. Respondent Instructions

### A. Intention to Bid

If your company intends to respond to this RFP, please send an email acknowledgement by **October 6, 2020,** to the contact below:

<andy.frassetto@iris.edu>

## B. Proposal Evaluation Criteria

Proposals will be evaluated on the basis of both the written proposal and any written responses to questions IRIS may receive. However, IRIS may use information other than that provided by the Respondent in its evaluation.

An evaluation panel consisting of members of the IRIS Instrumentation and Data Services program staff and MT program governance groups will rank proposals on the following criteria (not listed in order of priority):

- Quality and suitability of proposal for element 1;
- Quality and suitability of proposal for element 2;
- Quality and suitability of proposal for element 3;
- Cost;
- Availability;
- Previous experience; and
- Staff credentials.

Proposals must be sure to address all criteria specified in this RFP.

## C. References

Proposals must include three (3) client references. The minimum information that must be provided about each reference is:

- Name of individual or company
- Address of individual or company
- Name and phone number of contact person
- Type of services or materials provided to reference and date provided

## D. Proposal Length

The proposal should be limited to 8 pages (excluding cover page, figures, and other supporting materials) in length. Margins shall be 1-inch all around with a 12-point font. The 8-page proposal must stand on its own with respect to the evaluation.

## E. Transmittal Information

The proposal should include a cover page with the following information:

- Name and Address of the Responding Organization
- Project Title
- Project Period of Performance
- Name, Title and Contact Information of the Key Technical Personnel,
- Name, Title and Contact Information of the Key Administrative Personnel, and
- Name, Title, Contact Information and Signature of the person authorized to submit the proposal.

## F. Supplemental Information

Any supplemental information furnished by a Respondent after the due date for receipt of proposals will not be considered, unless IRIS formally requested such information. IRIS may request a Respondent to furnish such supplementary information as is required, in the opinion of IRIS, to assure that the Respondent is sufficiently competent and financially sound to successfully perform the contemplated work.

## G. Rights of Retention

Following submission of proposals and final evaluation, IRIS will have the right to retain the proposals, maintaining them in confidence. All documents submitted in response to this RFP shall become the property of IRIS.

## H. Clarification of RFP

Any questions, requests for clarification or requests for data in connection with this RFP shall be made no later than **Tuesday October 6, 2020, at 8 PM EST** via email to:

<andy.frassetto@iris.edu>

Questions and all responses will be publicly posted at
**https://www.iris.edu/hq/rfp/mtdev/q&a**.

## I. Cost of Proposal

IRIS will not reimburse the Respondent for their cost of preparation and submission of a proposal.

## J. Confidentiality

IRIS shall treat responses to this RFP as proprietary and confidential property (hereinafter "Proprietary and Confidential Information"). News or other information releases pertaining to this RFP shall not be made without prior written approval from IRIS.

## K. Proposal Submission

Proposal submission deadline is **Tuesday, October 20, 2020, at 8 PM EST.** Submit one electronic copy (PDF or MS-Word) to:

<andy.frassetto@iris.edu>
The proposal must be received by the proposal submission deadline to be considered for this award.

### L. Other

By submitting a proposal, Respondent agrees to all applicable provisions, terms and conditions associated with this Request for Proposal.

## VI. Selection

IRIS reserves the right to make its selection based solely on the information provided, to reject any or all proposals, to accept any proposals, or to affect any combination of proposals. IRIS reserves the right to conduct discussion or request proposal revisions, if deemed necessary.

The selection will be made on the basis of IRIS evaluation and determination of which proposal will provide the greatest benefit to IRIS, not necessarily on the basis of lowest price. IRIS has no obligation to reveal how proposals were assessed. Therefore, proposals should contain your best terms within the proposed functional and technical approach.

IRIS reserves the right to reject any or all proposals that are deemed to be non-responsive, late in submission or unsatisfactory in any way. IRIS shall have no obligation to award a contract for work, goods and/or services as a result of this RFP.

Contracts will be contingent upon the availability of funds and IRIS will solely be responsible for determining which, if any, proposals will be funded.

## VII. Preliminary Schedule

The following schedule may be changed or modified by IRIS:

1. RFP issued: 9/22/20
2. Respondents' requests for clarification of RFP due: 10/6/20
3. IRIS response to request for clarification of RFP due: 10/13/20
4. Proposals due: 10/20/20
5. Selection(s) made: 11/6/20
6. Respondents notified of determination: 11/11/20
7. Contract issued: 11/25/20
8. Delivery of software: 3/31/21