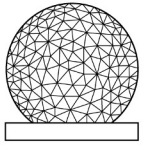
Using the Madrigal Database for Atmospheric Science

Katherine Cariglia
MIT Haystack Observatory
cariglia@mit.edu

MIT HAYSTACK OBSERVATORY

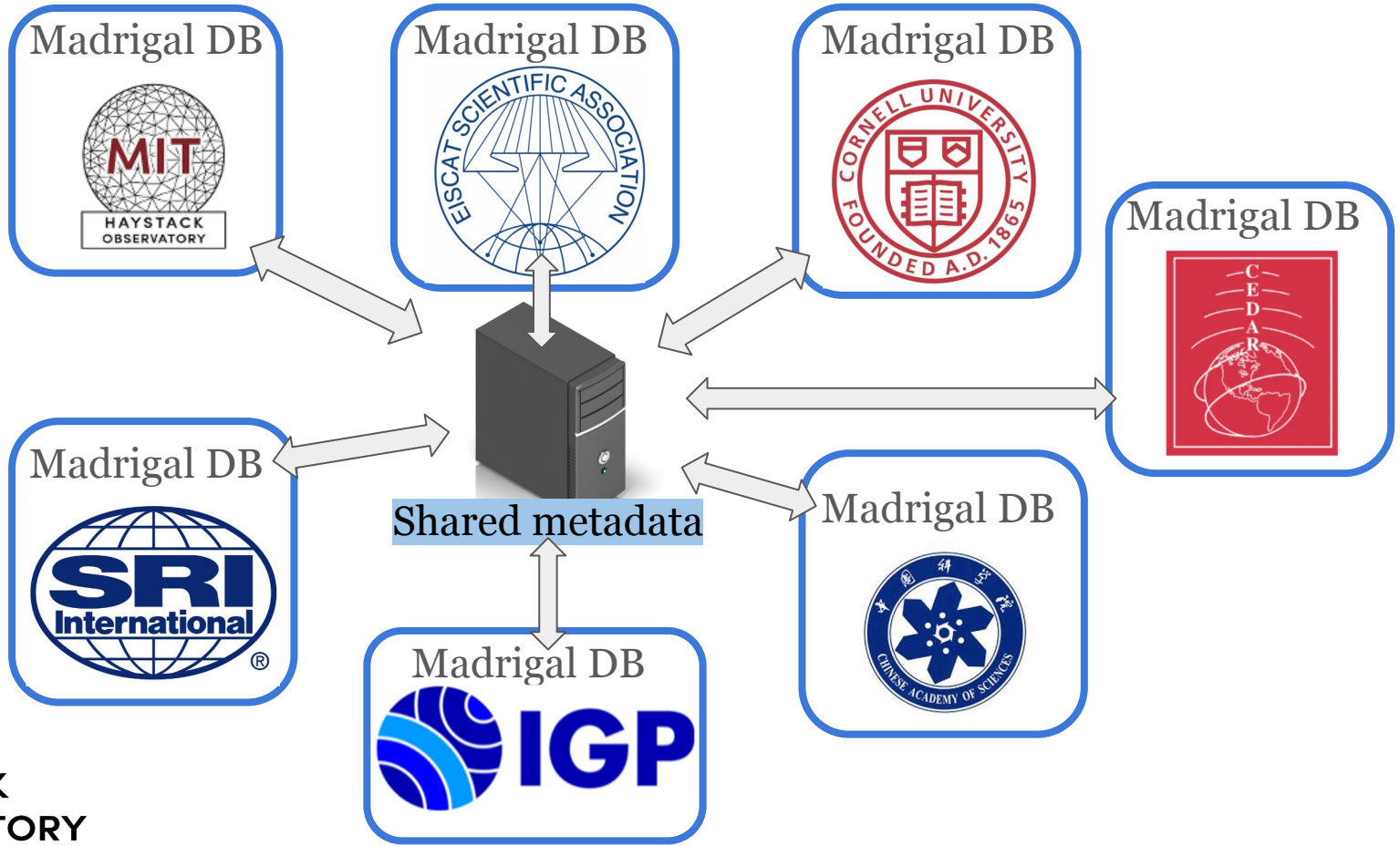
Introduction: the CEDAR Madrigal Database

- CEDAR is a research community investigating the near-Earth space environment
- CEDAR Madrigal database is a community resource for upper atmospheric science data from various ground and space-based instruments from around the world
- MIT Haystack Observatory Atmospheric and Geospace Science Group maintains the CEDAR Madrigal database
- Madrigal is part of Millstone Hill Geospace Facility - supported by NSF



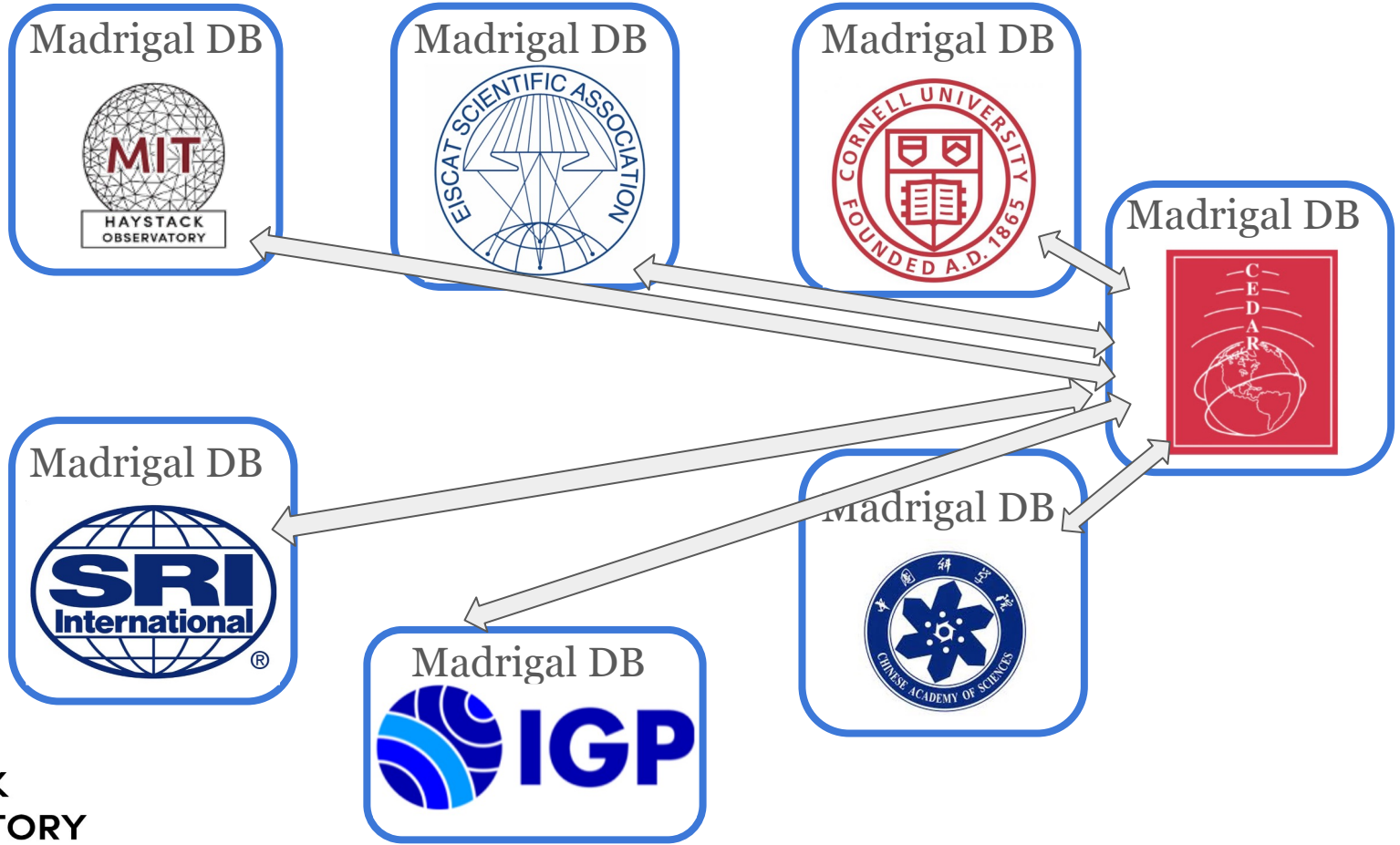
MIT
HAYSTACK
OBSERVATORY

Madrigal: A Distributed Database



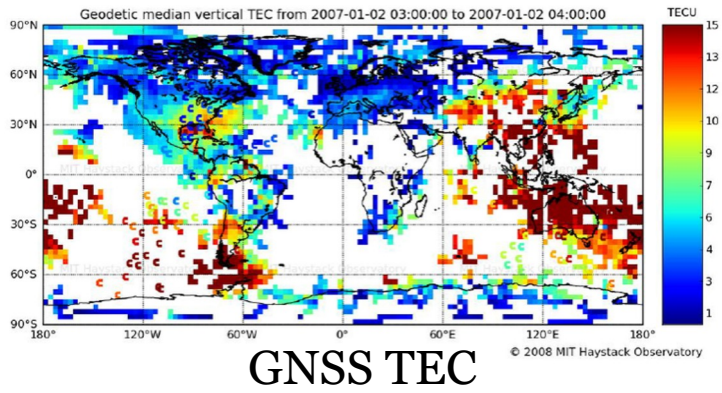
MIT
HAYSTACK
OBSERVATORY

CEDAR Madrigal Archive Imports All Data Weekly



MIT
HAYSTACK
OBSERVATORY

What Kind of Data is Stored in Madrigal?



GNSS TEC

200 Diverse Instruments in Madrigal

Incoherent scatter radars (ISR): 28

Lower/middle atmosphere radars: 21

Photometers: 7

Fabry Perot Interferometers: 38

Defense Meteorological Satellite Program (DMSP)

Michelson Interferometers: 6

Lidars: 10

Meteor radars: 18

GNSS Total Electron Content (TEC)

MIT HAYSTACK OBSERVATORY

Madrigal is Open Source

- Download/update Madrigal - includes Madrigal server and client APIs
  - Subversion Source Control
- OpenMadrigal Admin Mailing List
- OpenMadrigal Annocements Mailing List
- Administering OpenMadrigal

The Open Madrigal Initiative

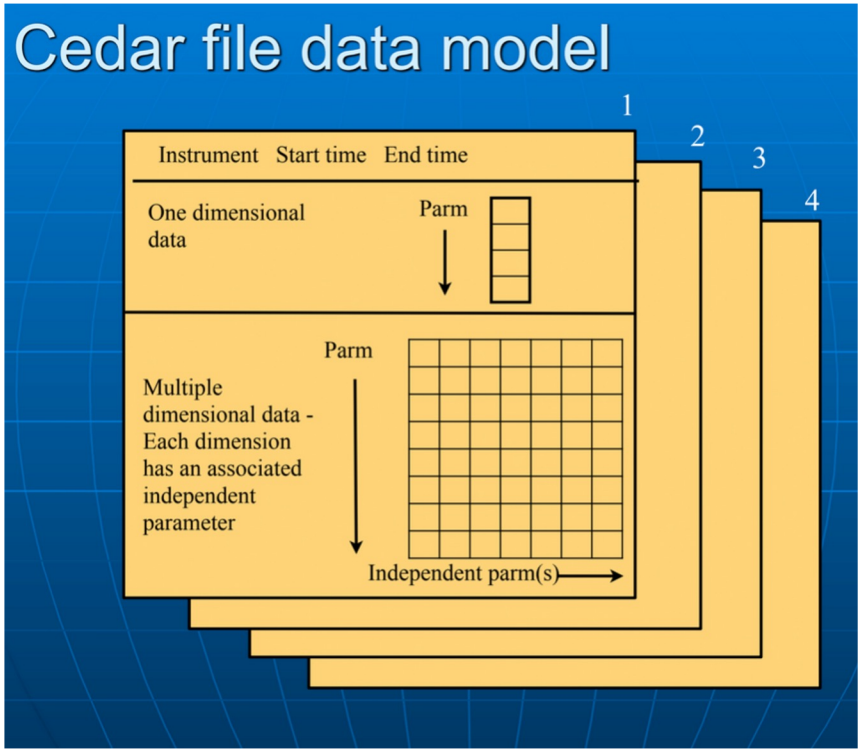
The OpenMadrigal project develops and supports the Madrigal database, which is a community resource for geospace data. The project has been led by MIT Haystack Observatory since 1980, and is now supported by the NSF and other community members. Madrigal is a robust, World Wide Web based system capable of managing and serving archival and real-time data, in a variety of formats, from a wide range of ground-based and some space-based instruments. Madrigal is installed at a number of sites around the world, and has a central archive site at the CEDAR Madrigal database. Data at each Madrigal site is locally controlled and can be updated at any time, but shared metadata between Madrigal sites allow searching of all Madrigal sites at once from any Madrigal site.

www.openmadrigal.org

MIT HAYSTACK OBSERVATORY

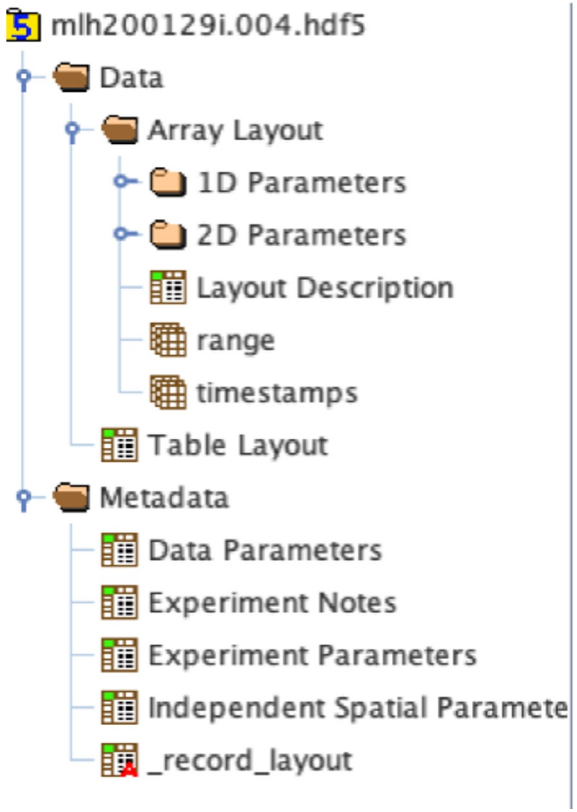
The CEDAR Madrigal Database Format

- Underlying data format is HDF5
- Well-defined parameters with standardized descriptions, units
- All parameters have corresponding uncertainty parameters
- Missing and assumed values included
- Well-defined parameter standards allow for the existence of a derivation engine, given a geographic location and datetime



CEDAR File Format: Structured HDF5

- Self-describing data: all parameters/units defined, notes/definitions included
- Table Layout always given
- Array Layout is default if independent parameters are given



Madrigal Metadata Model

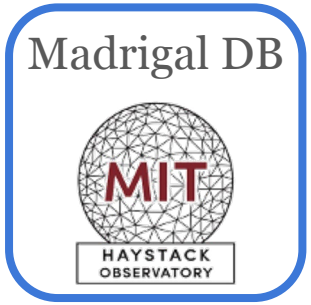
Madrigal site - a facility with scientists and a Madrigal installation

Instruments - ground based (set location) or satellites

Experiments - limited duration, single PI contact

Experiment files - data from one analysis of the experiment

Records - measurement over a single period of time



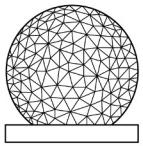
Data shared among all Madrigal sites

Data unique to one Madrigal site

MIT
HAYSTACK
OBSERVATORY

In What Formats Can Users Get Madrigal Data?

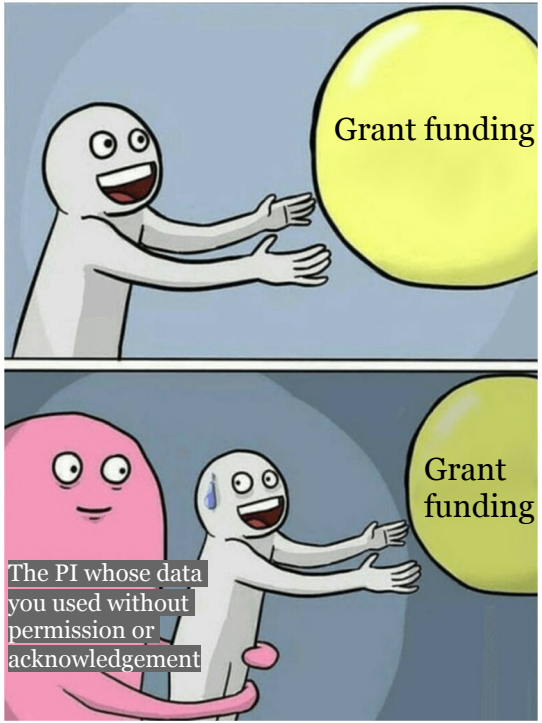
- Underlying format: HDF5
- Output formats: HDF5, netCDF4, ascii
- All output formats contain both observed and derived parameters



MIT
HAYSTACK
OBSERVATORY

Using Madrigal - Rules of the Road

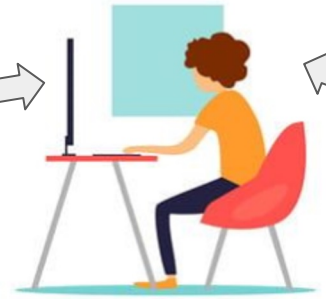
- If you wish to use Madrigal data in a paper or talk, **please contact the PI**
- PI for every experiment listed on data download page



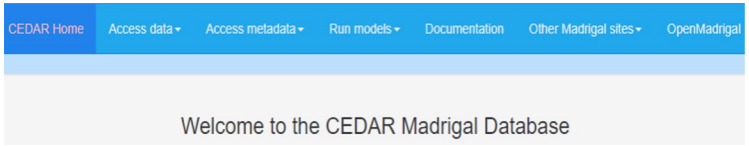
Use of data without informing the PI may lead to bad luck with grant writing-- don't let this happen to you!

Using Madrigal - Web or API Access

User



Web Interface



Web Services API

- Python
- Matlab
- IDL

Typical use case: data discovery

Create scripts using the MadrigalWeb API for all your data needs!

- Faster downloads compared to web interface
- Data flexibility - filter data by desired parameters/derived parameters

Installation of API(s)

MadrigalWeb available via pip, or you can download source

Download Python, Matlab, or IDL APIs to remotely access Madrigal

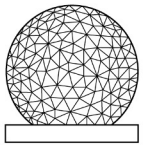
Madrigal is now easy to access from any platform remotely using web services. Since these API's have been tested on both Windows and Unix, they are available both as tar files and zip files.

To use any of these API's, there is usually no need to read any documentation. Just create the command to run using the web page at *http://cedar.openmadrigal.org/chooseScript/* .

New! - The python, Matlab, and IDL API's now all have a globalDownload method, which allows you to download any group of Madrigal files with one simple command. Also, they can now download Hdf5, netCDF4, and Ascii files from Madrigal 3 site such as http://cedar.openmadrigal.org.

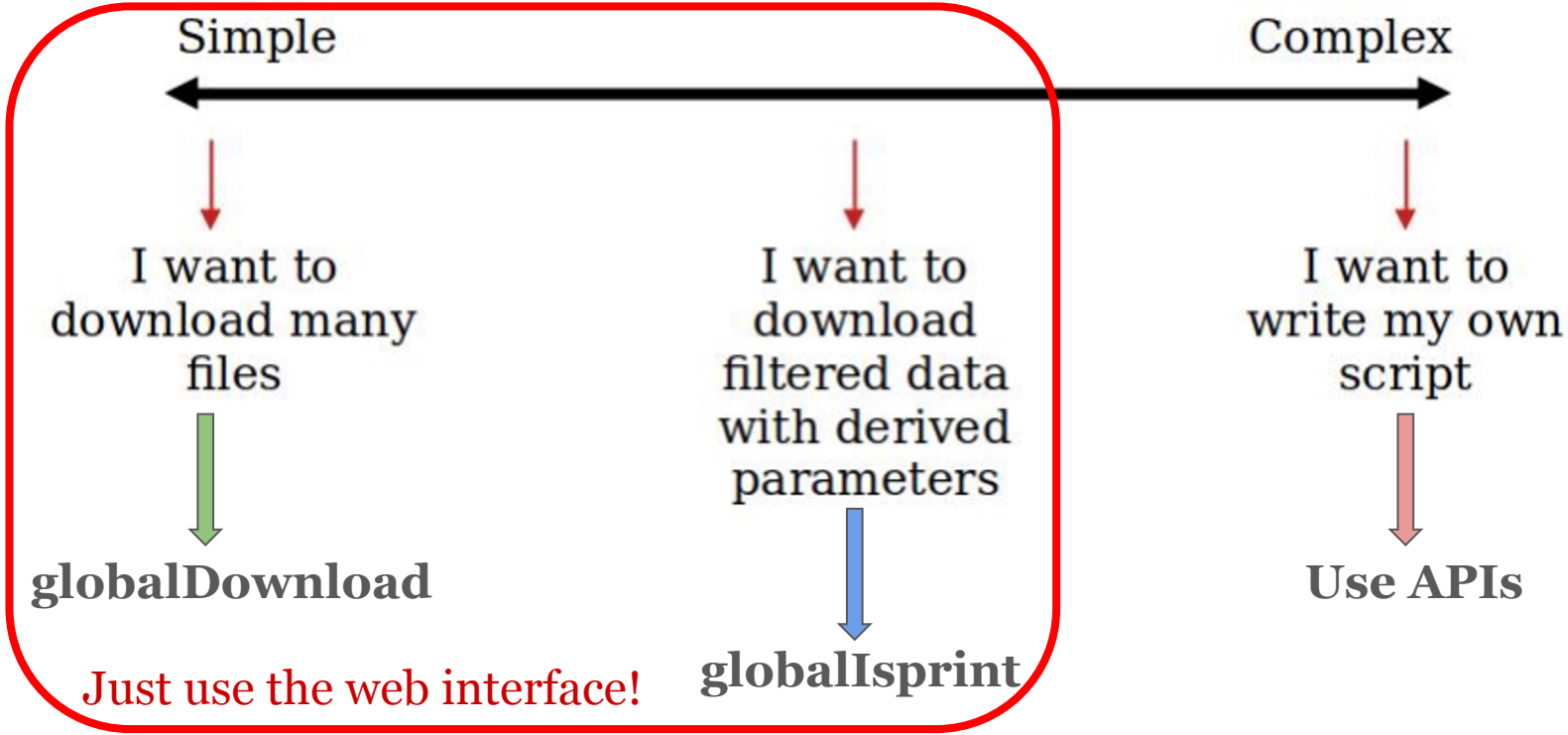
- **Download madrigalWeb python Madrigal API** - This 3.2.4 release works with both python 2 and python 3. It has a new feature that allows users to create a single citation that points to a list of CEDAR Madrigal files.
  - madrigalWeb-3.2.4.tar.gz (last updated Jun. 27, 2024)
  - madrigalWeb-3.2.4-py3-none-any.whl (last updated Jun. 27, 2024)
  - Can also be installed via pip: *pip install madrigalWeb*
  - Documentation - Scripts, Tutorial, Reference
- **Download remote Matlab Madrigal API** - This 2.4.8 release works fully with both Madrigal 3.0 and Madrigal 2.6. This release requires wget to download files or run isprint.
  - remoteMatlabAPI-2.4.8.tar.gz (last updated Nov. 15, 2022)
  - remoteMatlabAPI-2.4.8.zip (last updated Nov. 15, 2022)
  - Documentation - Scripts, Tutorial, Reference
- **Download remote IDL Madrigal API** - This 1.3.4 release works fully with both Madrigal 3.0 and Madrigal 2.6. This release requires wget be installed.
  - remoteIDL_API-1.3.4.tar.gz (last updated Jul. 29, 2019)
  - remoteIDL_API-1.3.4.zip (last updated Jul. 29, 2019)
  - Documentation - Scripts, Tutorials, Reference

http://cedar.openmadrigal.org/madrigalDownload



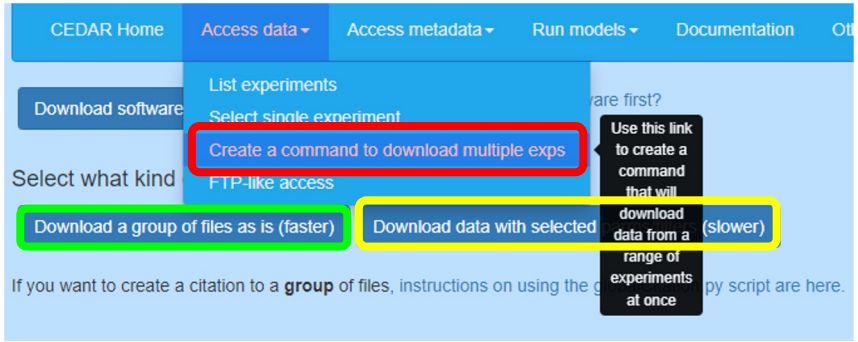
MIT
HAYSTACK
OBSERVATORY

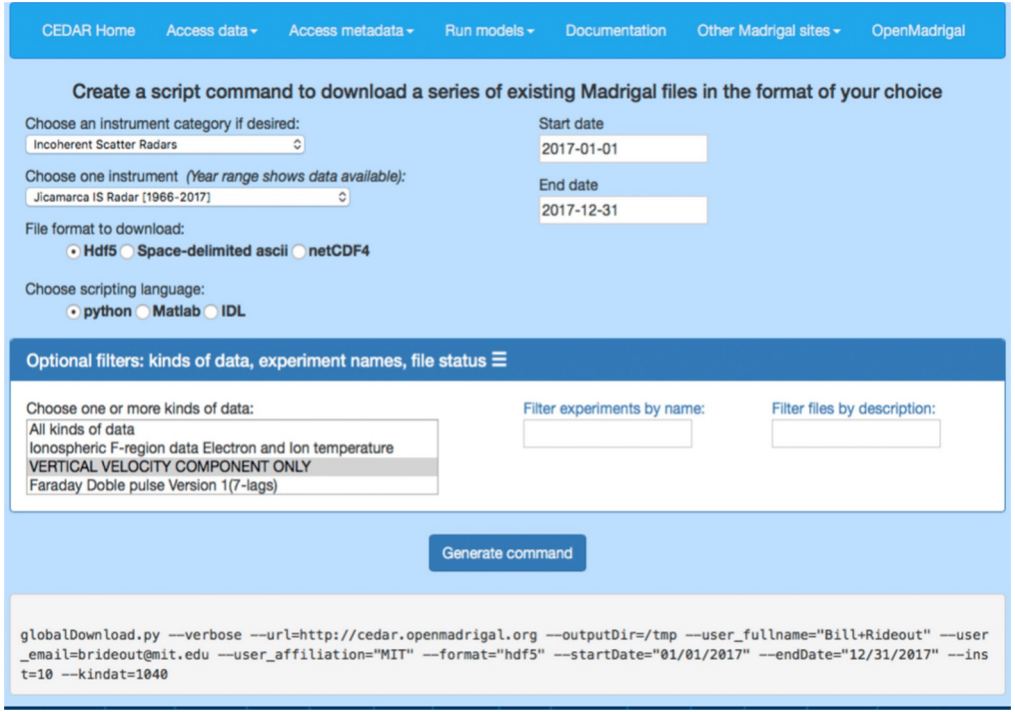
Use Cases for the API



Download Scripts Generated by Web

No need to read documentation!



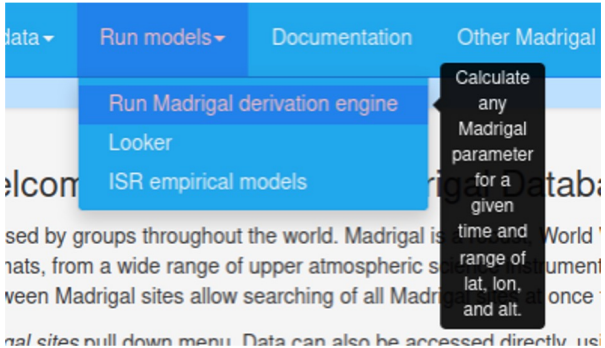


- First, create a command
- Next, download set of files as is or with selected parameters/filters

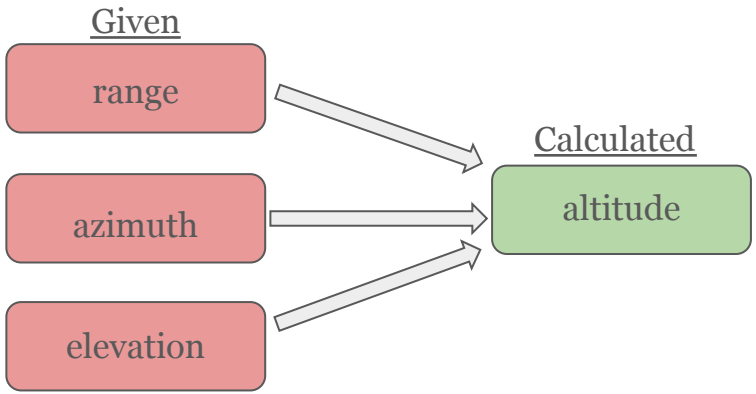
MIT
HAYSTACK
OBSERVATORY

Madrigal Derivation Engine

- Standardization of file parameters allows the existence of the derivation engine
- All parameters have corresponding uncertainty parameters
- Independent parameters built into standard
- Derived parameters appear in file
- Derivation engine determines which parameters can be derived
- New derived parameters easy to implement in C or Fortran



Derivation example for radar data:



MIT HAYSTACK OBSERVATORY

Types of Derived Parameters

Space and time

- Local time
- Shadow height

Geophysical

- Kp
- Dst
- Imf
- F10.7

Magnetic

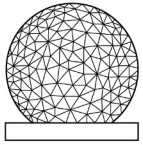
- Bmag, mag conjugate lat + lon, Tsyganenko magnetic equatorial plane intercept

Models

- MSIS - neutral atmosphere model
- IRI - International Reference Ionosphere model

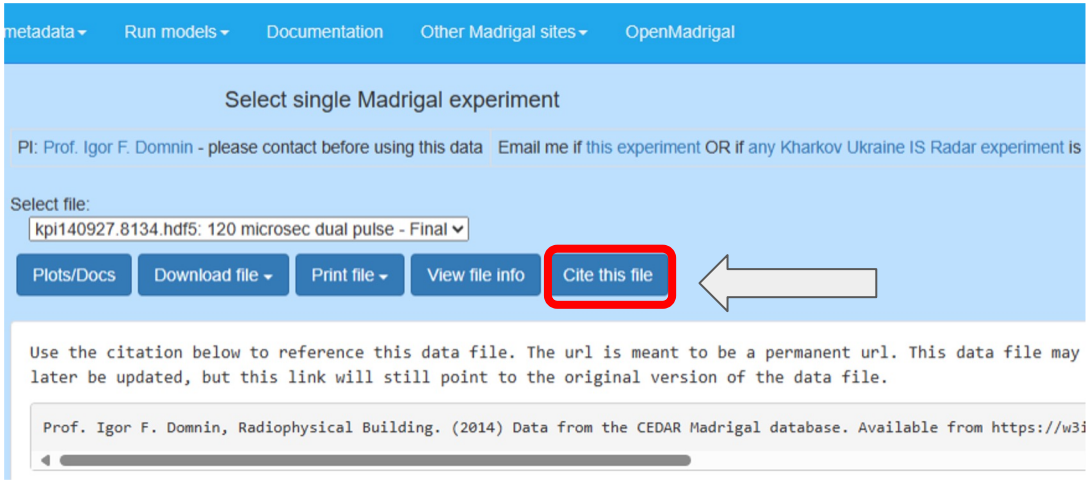
Other Models

- NATEC - North American Total Electron Content model
- Empirical ISR - developed at MIT Haystack



Data Reproducibility

- Madrigal files are never deleted - when new versions of experiment files are created, old versions marked as “history”
- Every file has a citable URL
- Python API allows creation of a single citation for multiple files



Citation generated:

Prof. Igor F. Domnin, Radiophysical Building. (2014) Data from the CEDAR Madrigal database. Available from https://w3id.org/cedar?experiment_list=experiments/2014/kpi/27sep14&file_list=kpi140927.8134.hdf5

MIT HAYSTACK OBSERVATORY

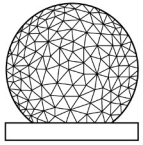
How to Get Your Instrument's Data into Madrigal

Method 1: Send data to CEDAR Madrigal

- Send data, documentation, and summary plots to MIT Haystack
- Loading program written by MIT Haystack, verified by you
- Add new data in batches or via automated upload

Method 2: Set up your own Madrigal site

- MIT Haystack will help with installation and loading programs
- You control when the data is uploaded
- Automated backup to CEDAR Madrigal site



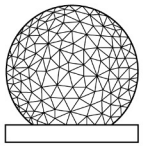
MIT
HAYSTACK
OBSERVATORY

Thank you for listening!

Many thanks to the organizers of this workshop for supporting early career researchers (such as myself!)

Any questions? Please contact

cariglia@mit.edu or **brideout@mit.edu**



MIT
HAYSTACK
OBSERVATORY

BOSTON
UNIVERSITY