

How to use the Madrigal database for atmospheric science

Bill Rideout

MIT Haystack Observatory

brideout@mit.edu

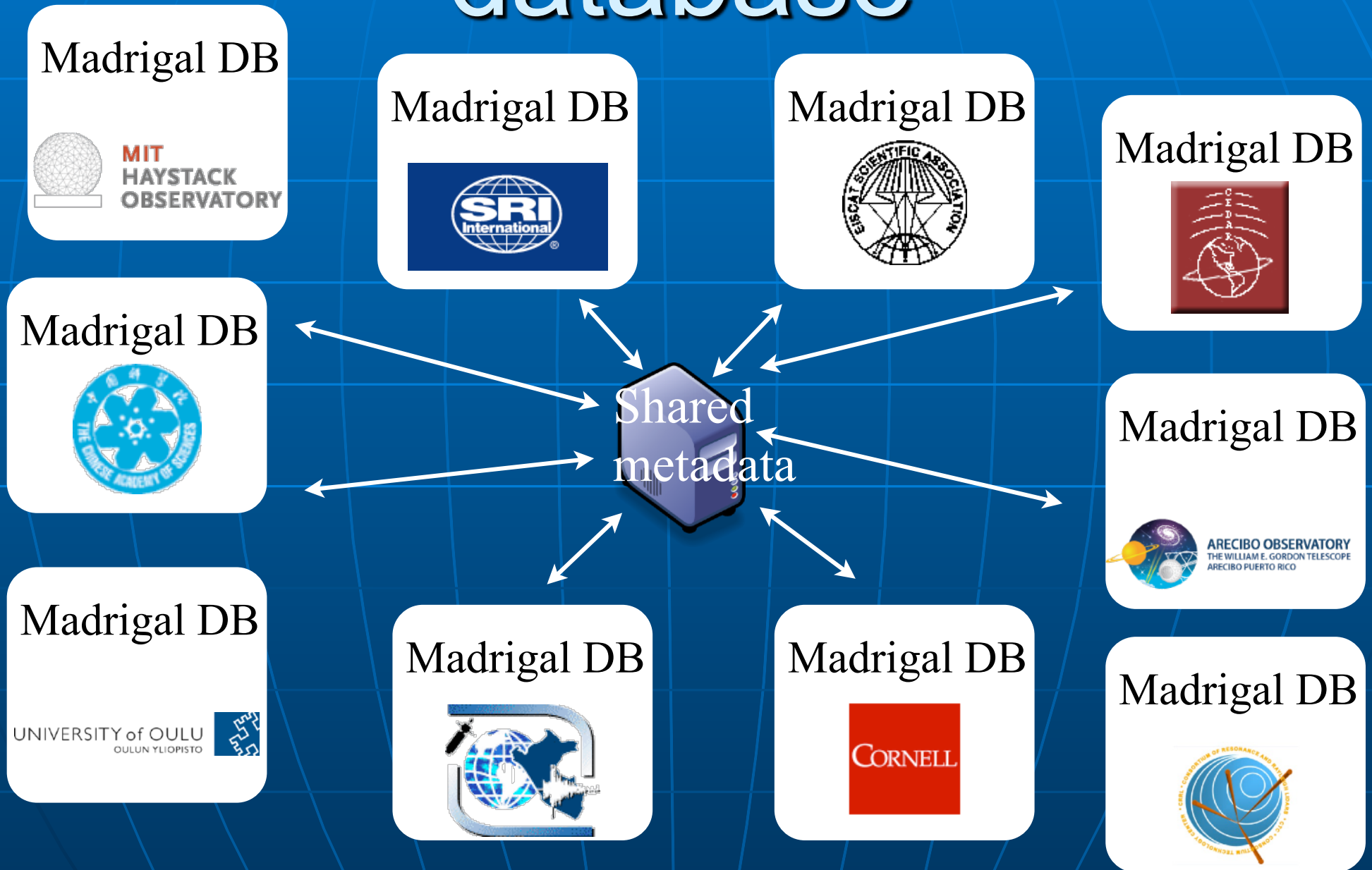
2023 ISR School
Fairbanks, Alaska

Outline

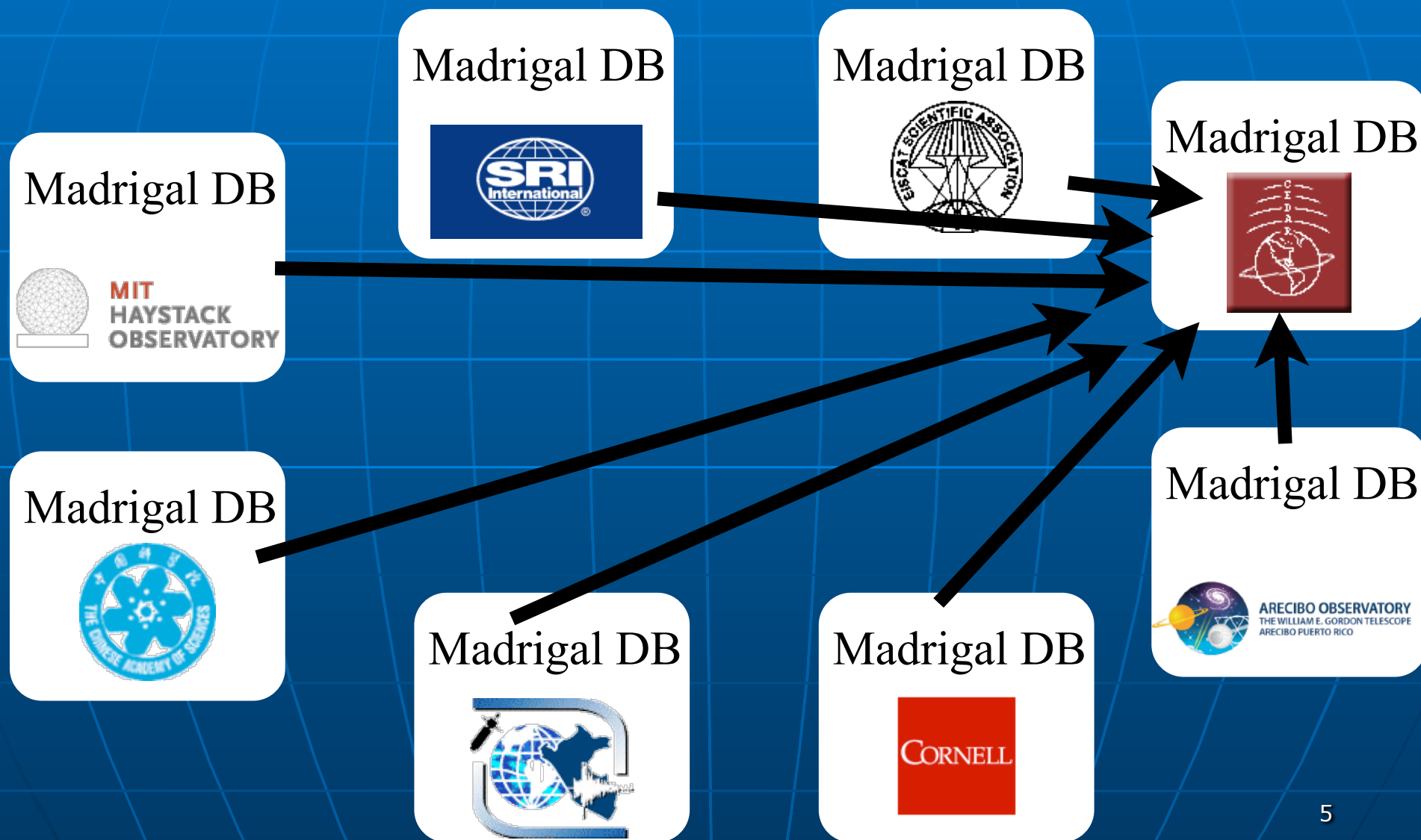
- What is Madrigal?
- What is the CEDAR database format?
 - Exercise - open a file with Hdfview
- In what formats can I get Madrigal data?
- How do I use Madrigal?
 - Background
 - The website
 - Exercise - use the website
 - Script data access
 - Exercise - write a script

What is Madrigal?

Madrigal is a distributed database



Cedar Madrigal archive imports all data weekly



The Madrigal database stores data from a wide variety of upper atmosphere research instruments

Incoherent Scatter Radar



TEC via GPS



MF Radar



Examples of number of instruments in Madrigal:

- Incoherent scatter radars: 22
- MST radars: 3
- MF radars: 16
- Meteor radars: 11
- FPI: 32
- Michelson Interferometers: 6
- Lidars: 9
- Photometers: 7

Other examples:

- GPS TEC
- DMSP

Madrigal is open-source

[CEDAR Home](#) [Access data](#) [Access metadata](#) [Run models](#) [Documentation](#) [Other Madrigal sites](#) [OpenMadrigal](#)

Welcome to the Madrigal3 CEDAR Database

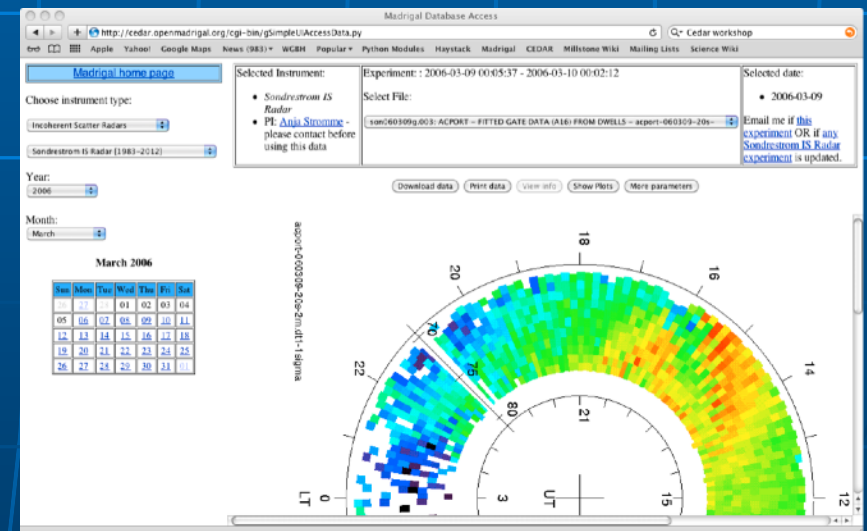
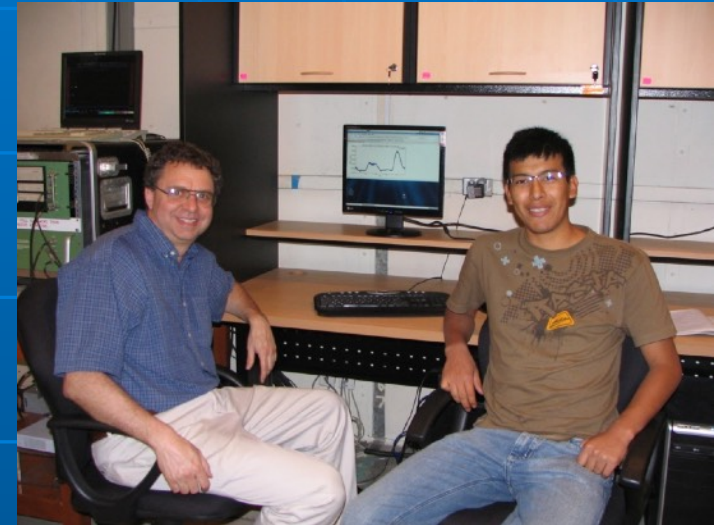
Madrigal is an upper atmospheric science database used by groups throughout the world. Madrigal is a robust, World Wide Web based system capable of managing and serving archival and real-time data, in a variety of formats, from a wide range of upper atmospheric science instruments. Data at each Madrigal 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.

To see a list of all Madrigal sites, use the *Other Madrigal sites* pull down menu. Data can also be accessed directly, using *APIs* which are available for several popular programming languages (Matlab, python, and IDL). A Subversion archive of all Madrigal software and documentation is available from the *Open Madrigal* Web site. The latest version of Madrigal and the remote APIs may also be downloaded from there.

Use of the Madrigal Database is generally subject to the CEDAR Rules-of-the-Road . Prior permission to access the data is not required. However, the user is required to establish early contact with any organization whose data are involved in the project to discuss the intended usage. Data are often subject to limitations which are not immediately evident to new users. Before they are formally submitted, draft copies of all reports and publications must be sent to the contact scientist at all data supplying organizations along with an offer of co-authorship to scientists who have provided data. This offer may be declined. The data supplied by the organizations must be clearly labeled as to the knowledge of the user and no publication or other use of this data is made available through another database. If you have any questions about appropriate use of these data, contact brideout@haystack.mit.edu

If you want to use the old Madrigal 2 version of the CEDAR Madrigal database, it is still temporarily available at <http://madrigal.haystack.mit.edu>. If you are using the old version because of a problem with Madrigal3, please contact brideout@haystack.mit.edu to describe the issue.

[Link to Subversion \(source code\)](#)

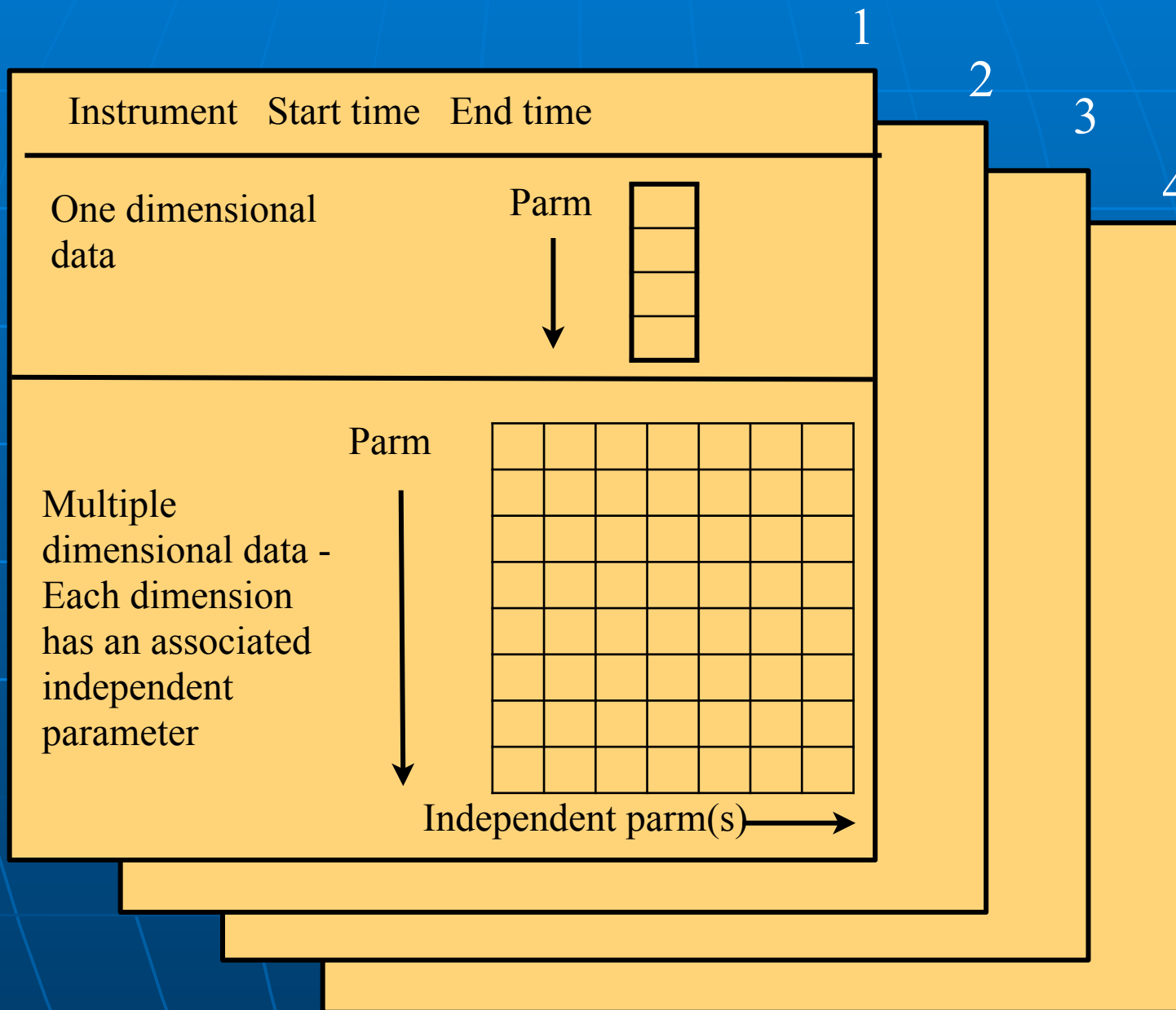


- What is the CEDAR database format?

Well defined parameters

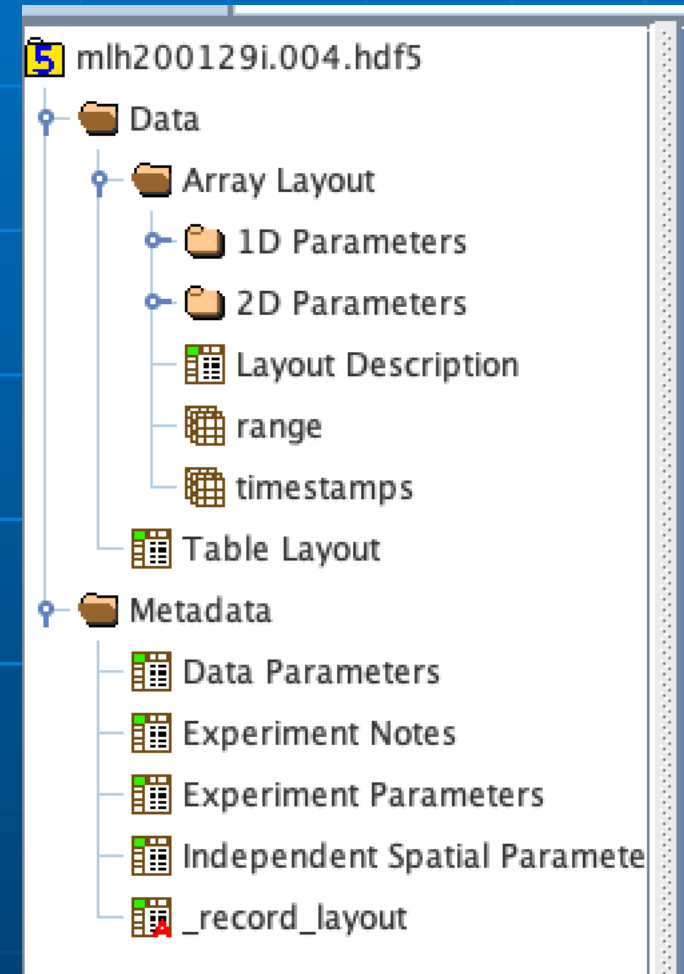
- Standard descriptions of all parameters
- Allows the existence of derivation engine
- Madrigal allows extended descriptions
- All parameters have corresponding error parameters
- Missing, Assumed

Cedar file data model



Cedar file format: Structured Hdf5

- Self-describing - all parameters defined, notes added
- Scientific standard
- Parms float, integers or strings
- Table Layout, Array Layout possible if repeating independent parms



Exercise: use hdfview to examine Madrigal files

- Go to <https://tinyurl.com/2023ISR>
- Go down to Monday - Madrigal exercise 1
- Install hdfview (<https://www.hdfgroup.org/downloads/hdfview/>) link also on web page
- Download example Poker Flat and Millstone ISR files with hdfview

In what formats can I get
Madrigal data?

Madrigal outputs

Underlying format	Hdf5
Output formats	Ascii, Hdf5, netCDF4
Formats with deriv parms	Ascii, Hdf5, netCDF4

- What do I need to understand about Madrigal to use it?

Madrigal Data Model

Madrigal site

(typically a facility with scientists and a Madrigal installation)



Instruments

(ground-based, typically with a set location)



Experiments

(typically of limited duration, with a single contact)



Experiment Files

(represents data from one analysis of the experiment)



Records

(measurement over one period of time)

Madrigal DB



Data shared among all Madrigal sites



Data unique to one Madrigal site

Madrigal Derivation Engine

- Derived parameters appear to be in file
- Engine determines all parameters that can be derived
- Easy to add new derived parameters using code written in C or Fortran



Classes of derived parameters

- Space, time
 - Examples: Local time, shadow height
- Geophysical
 - Examples: Kp, Dst, Imf, F10.7
- Magnetic
 - Examples: Bmag, Mag conjugate lat and long, Tsyganenko magnetic equatorial plane intercept
- Models
 - Examples: MSIS, IRI

- Using Madrigal

Rules of the road

- If you want to use data in Madrigal in a paper or talk, contact the PI
- The PI is listed on every web page

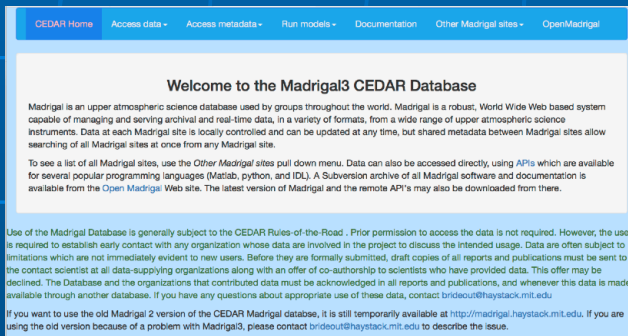


Use of data without informing PI may lead to seven years of bad luck in grant writing - don't let this happen to you!

How can the Madrigal database be accessed?



User



Web interface

Web services API

- From anywhere on internet
- Python API
- Matlab API
- IDL API

Typical use - data discovery

Create scripts using web interface: 99.9% of time no need to read documentation!

- Faster - download files
- More flexible - define parameters, filter

Live demo of Madrigal web page

- CEDAR Madrigal site
 - <https://data.amisr.com/madrigal/>

Exercise 2: using the Madrigal web interface

- Go to <https://tinyurl.com/2023ISR>
- Go down to Monday - Madrigal exercise 2
- Work through exercise

How can I put my instrument's data on Madrigal?



Send data to CEDAR Madrigal

Madrigal DB



- Send data to MIT Haystack in your own format
- Loading program written by MIT Haystack, verified by you
- Add new data in batch or via automated upload (eg, sftp, web access, etc)

Set up your own Madrigal site



- MIT Haystack will help with installation and writing needed loading programs
- You control when data uploaded
- Automated backup to central CEDAR Madrigal site²⁴

Remote Access to Madrigal Data

- Use the web interface to write scripts
- Built on web services
- Like the web, available from anywhere on any platform
- Read only API
- Complete Python, Matlab, and IDL APIs written (python 2 and 3)
- More APIs available on request or via contribution

Remote Access to Madrigal Data

Use the web interface to write these scripts

Simple

Most complex



I just want to download lots of files to my PC

globalDownload



I want to only download filtered data with derived parameters

globalIsprint



I want to write my own script

Work with API's

Live demo of Madrigal script creation page

- CEDAR Madrigal site
 - <http://cedar.openmadrigal.org>
 - Access Data -> Create a command to download multiple exes

Scripts generated by web site

- No need to read any documentation!

The image shows a screenshot of the CEDAR website's navigation menu. The menu is blue with white text. The main navigation bar includes 'CEDAR Home', 'Access data', 'Access metadata', 'Run models', and 'Docume'. The 'Access data' menu is expanded, showing options: 'List experiments', 'Select single experiment', 'Create a command to download multiple exps', and 'FTP-like access'. Below this, there are two buttons: 'Download a group of files as is' and 'Download data with selected parms/filters'. Two black arrows point from the 'Create a command to download multiple exps' option to the two buttons below. A white box with black text and a black border contains the text 'First choose Create a command'.

First choose
Create a command

Then decide to just download files as they are, or to select parms and/or filters.

Create a script to download files

[CEDAR Home](#)[Access data ▾](#)[Access metadata ▾](#)[Run models ▾](#)[Documentation](#)[Other Madrigal sites ▾](#)[OpenMadrigal](#)

Create a script command to download a series of existing Madrigal files in the format of your choice

Choose an instrument category if desired:

Incoherent Scatter Radars ▾

Start date

2017-01-01

Choose one instrument *(Year range shows data available):*

Jicamarca IS Radar [1966-2017] ▾

End date

2017-12-31

File format to download:

Hdf5 Space-delimited ascii netCDF4

Choose scripting language:

python Matlab IDL

Optional filters: kinds of data, experiment names, file status ☰

Choose one or more kinds of data:

All kinds of data
Ionospheric F-region data Electron and Ion temperature
VERTICAL VELOCITY COMPONENT ONLY
Faraday Doble pulse Version 1(7-lags)

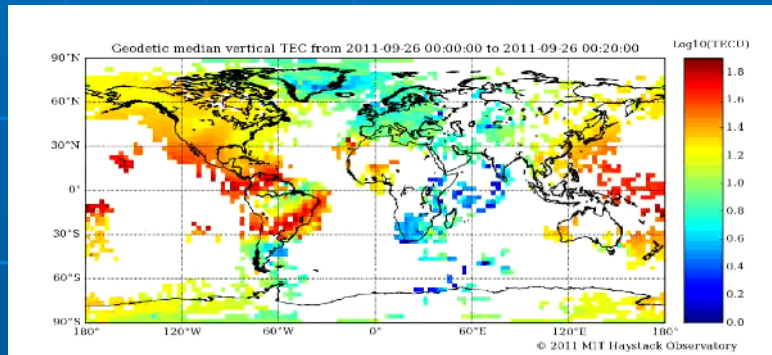
Filter experiments by name:

Filter files by description:

Generate command

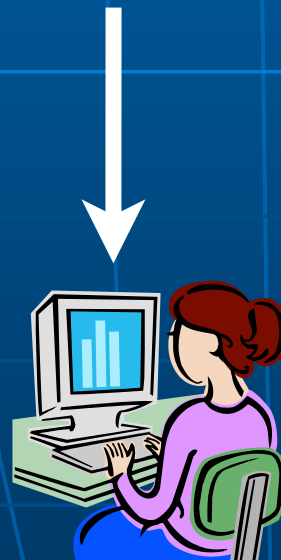
```
globalDownload.py --verbose --url=http://cedar.openmadrigal.org --outputDir=/tmp --user_fullname="Bill+Rideout" --user_email=brideout@mit.edu --user_affiliation="MIT" --format="hdf5" --startDate="01/01/2017" --endDate="12/31/2017" --ins t=10 --kindat=1040
```

More advanced globalprint scripts



<http://madrigal.iggeas.ac.cn/madrigal/>

TEC data: instrument id = 8000



Show me only data
where $TEC > 100TEC_u$
when $K_p > 7$ in 2003

Create a script to select parms/filters - globalsprint

[CEDAR Home](#)

[Access data](#) ▾

[Access metadata](#) ▾

[Run models](#) ▾

[Documentation](#)

[Other Madrigal sites](#) ▾

[OpenMadrigal](#)

Create a script command to download Madrigal data with selected parameters and filters

Choose an instrument category if desired:

Incoherent Scatter Radars ▾

Choose one instrument *(Year range shows data available)*:

Jicamarca IS Radar [1966-2017] ▾

Start date

2017-01-01

End date

2017-12-31

File format to download:

Hdf5 Space-delimited ascii netCDF4

Choose scripting language:

python Matlab IDL

Select parameters (required) ≡

Optional filters: kinds of data, filter by parm, experiment names, etc ≡

Generate command

```
globalIsprint.py --verbose --url=http://cedar.openmadrigal.org --parms=YEAR,MONTH,DAY,SEC,VIPE,DVIPE --output=/tmp --u
ser_fullname="Bill+Rideout" --user_email=brideout@mit.edu --user_affiliation="MIT" --startDate="01/01/2017" --endDate=
"12/31/2017" --inst=10 --kindat=30003
```

But if you want to read the docs

- write your own script

- Use python, Matlab, IDL
- Methods
 - getInstrumentsWeb
 - getExperimentsWeb
 - getExperimentFilesWeb
 - getParametersWeb
 - isprintWeb
 - madDownloadFile
 - madCalculatorWeb
- Methods match Madrigal model

Exercise 3: using the Madrigal script creation interface

- Go to <https://tinyurl.com/2023ISR>
- Go down to Monday - Madrigal exercise 3
- Work through exercise

Extending/contributing to Madrigal

- Madrigal is completely open source
- See www.openmadrigal.org for source
- All new code is Python or C. Imported derivation methods sometimes in Fortran.
- We appreciate all contributions
 - Suggestions and ideas
 - Finding bugs