JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Data Amalgamation via the Heliophysics API (HAPI), or

'when time series datasets are accessible through a single API'

Alex Antunes, JHU / APL
Robert S. Weigel, George Mason University
Jon Vandegriff, JHU / APL

`https://hapi-server.org`

HAPI is a useful way to get data from many sources

ESAC HAPI Server
Cluster-3 Electric Wave Form Power Density

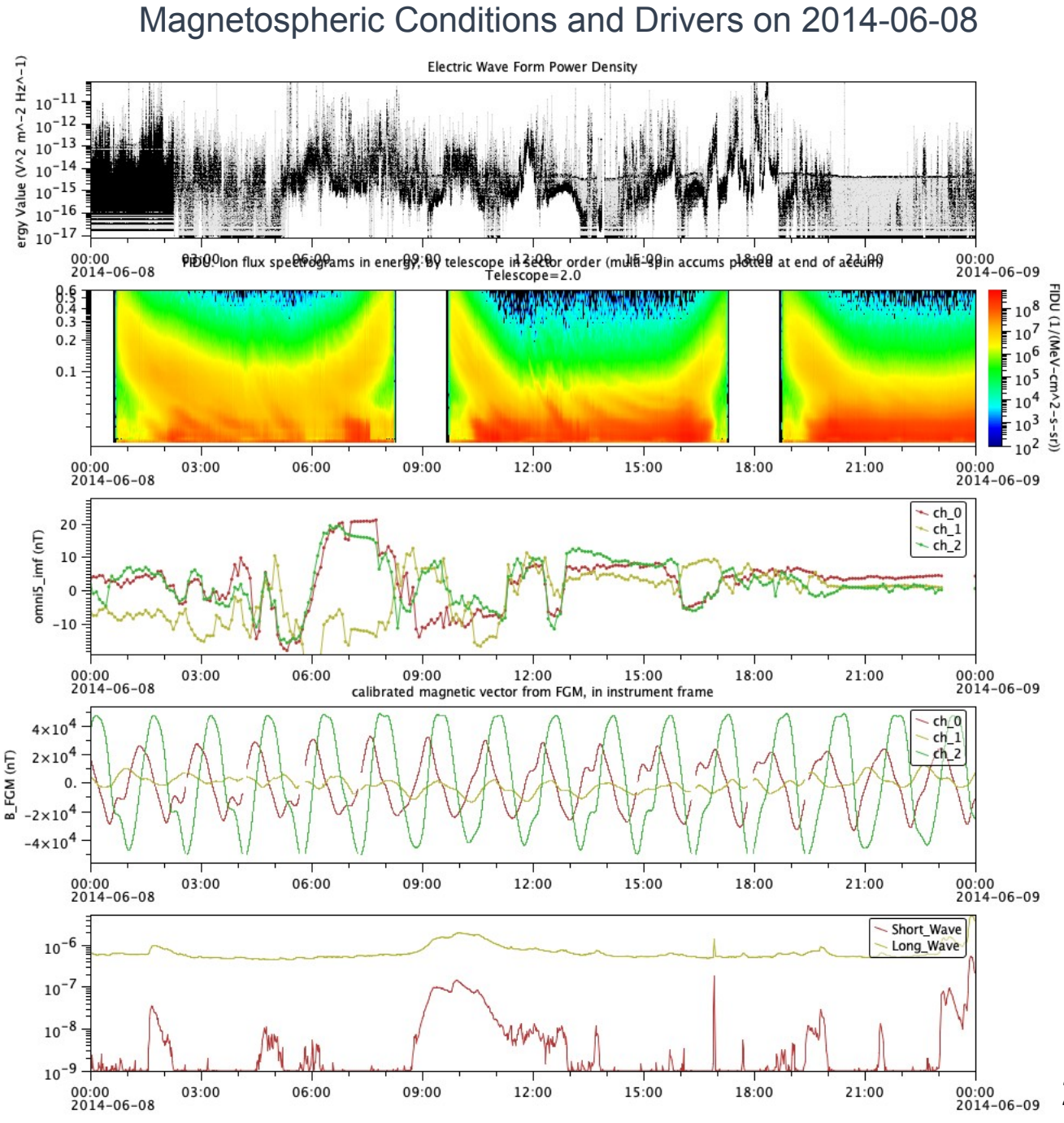
CDAWeb HAPI Server
Van Allen Probes, RBSPICE Ion Energy Spectrogram

AMDA HAPI Server
Omni MAG data

ViRES HAPI Server
GRACE-A MAG Data

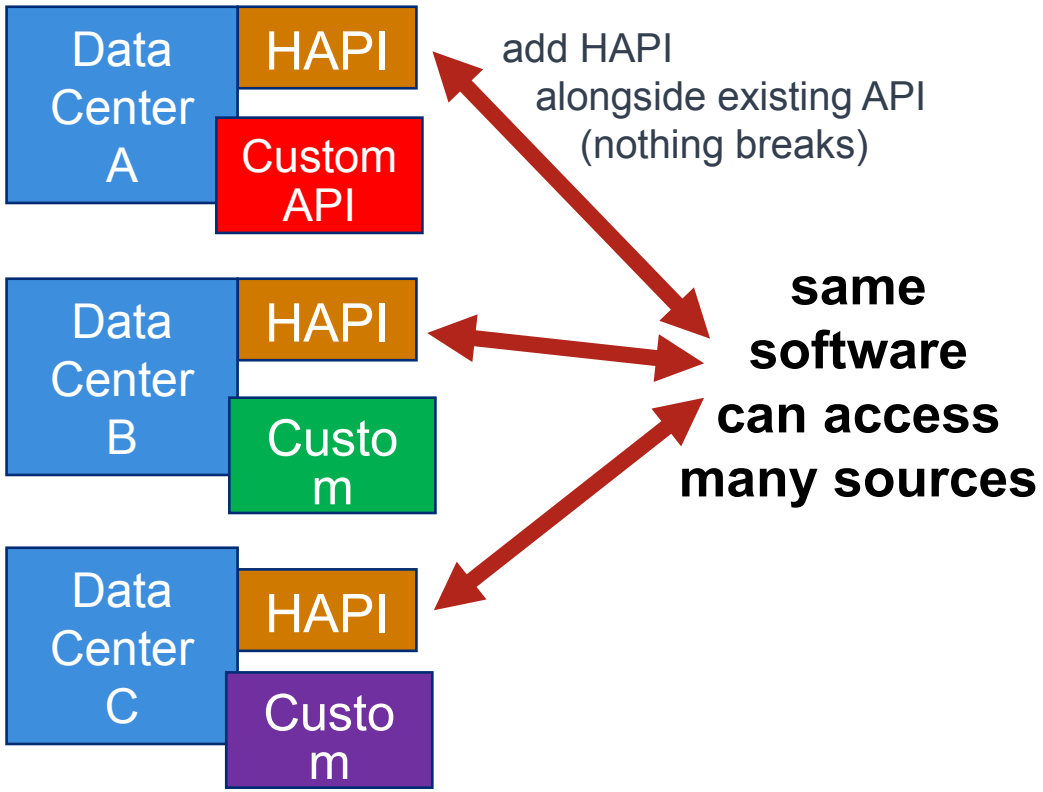
CCMC ISWAT HAPI Server
GOES X-ray flux

5 datasets
5 data providers
1 plot
(Autoplot - see autoplot.org)



Heliophysics Application Programmer's Interface

➔ no Heliophysics idioms: data model is just numbers vs. time



Reasons to Adopt

- **many data sources already available via HAPI**
  Helio, Planetary, in US, Europe, Canada, also NSF and NOAA are interested
- **clients already exist for access and display**
- **easy to implement server from the spec**
  (or use existing server: Python, Java, node.js)
- active development community (weekly meetings) and clients written outside the dev team
- recommended as official standard by COSPAR

What is the Heliophysics Application Programmer's Interface (HAPI)?

- Think of HAPI as "**http for data**" – a protocol for asking and getting data from a server
- RESTful – no state so that each request is independent
- URLs representing the requests can be thought of as (semi-)persistent identifiers
- Endpoints define the things you can ask of a HAPI server

Request:

```
http://server.org/hapi/data?dataset=ACE_MAG
                          &start=2004-183T00:00Z
                          &stop=2004-184T00:00Z
```

(line breaks for clarity)

Response:

```
2004-183T00:00:03.403Z,  1.0724e+02, -6.8993e+01, -5.1978e+02
2004-183T00:00:07.153Z,  1.0842e+02, -6.8956e+01, -5.1962e+02
2004-183T00:00:10.907Z,  1.0855e+02, -6.9063e+01, -5.2084e+02
2004-183T00:00:14.653Z,  1.0852e+02, -6.9049e+01, -5.2085e+02
2004-183T00:00:18.403Z,  1.0849e+02, -6.9035e+01, -5.2085e+02
2004-183T00:00:22.153Z,  1.0862e+02, -6.9142e+01, -5.2207e+02
2004-183T00:00:25.903Z,  1.0859e+02, -6.9128e+01, -5.2208e+02
```

(always the same format for all servers)

HAPI Offers many benefits for accessing, sharing data

For Scientists / Data Users

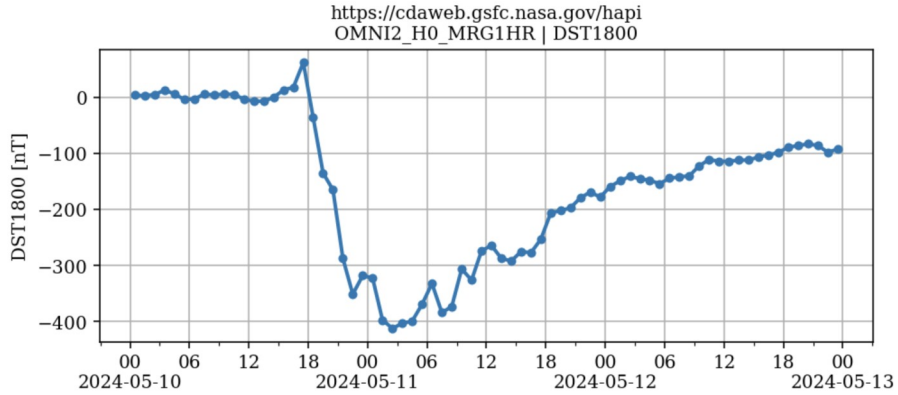
- **solves the "fill my array" problem**
- a standard access mechanism --> less data wrangling code
  - HAPI client library
    - functions like a HDF library or FITS reading library
    - gets you to data, not files
  - excellent metadata
  - HAPI common at multiple places
- can share HAPI URLs as a way to communicate data
- can share tools more easily

Since 2021, HAPI is a COSPAR-endorsed standard for Space Weather data access

HAPI Example (Python): INTERMAGNET and CDAWeb

```
from hapiclient import hapi
from hapiplot import hapiplot
start      = '2024-05-10T00:00:00Z' # Start and stop times
stop       = '2024-05-13T00:00:00Z'


                                                    CDAWeb
server     = 'https://cdaweb.gsfc.nasa.gov/hapi'
dataset    = 'OMNI2_H0_MRG1HR'
parameters = 'DST1800'
dataDST, metaDST = hapi(server, dataset, parameters, start, stop)
hapiplot(dataDST,metaDST)
```

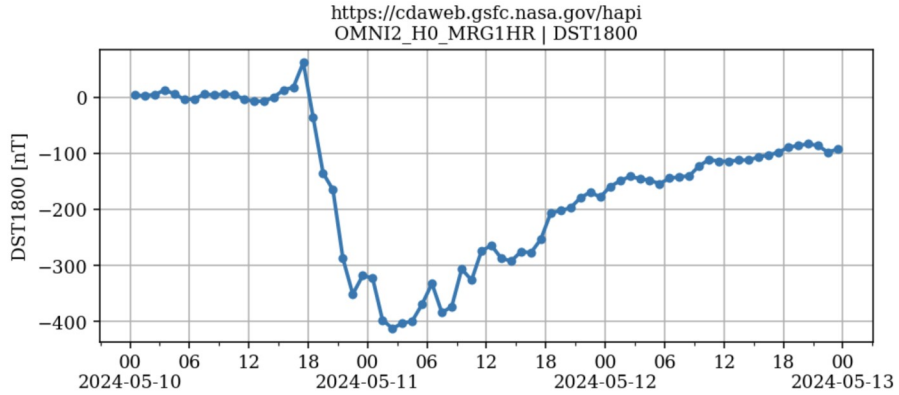


HAPI Example (Python): INTERMAGNET and CDAWeb

```
from hapiclient import hapi
from hapiplot import hapiplot
start      = '2024-05-10T00:00:00Z' # Start and stop times
stop       = '2024-05-13T00:00:00Z'
```

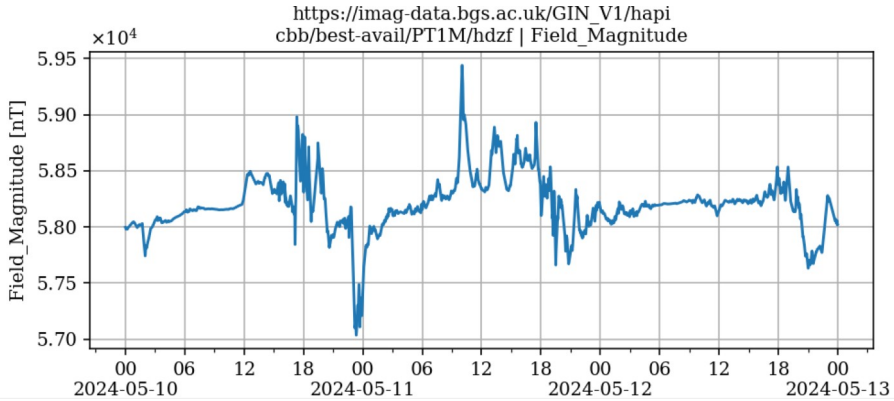
CDAWeb

```
server     = 'https://cdaweb.gsfc.nasa.gov/hapi'
dataset    = 'OMNI2_H0_MRG1HR'
parameters = 'DST1800'
dataDST, metaDST = hapi(server, dataset, parameters, start, stop)
hapiplot(dataDST,metaDST)
```



INTERMAGNET

```
server     = "https://imag-data.bgs.ac.uk/GIN_V1/hapi"
dataset    = "cki/best-avail/PT1M/hdzf"
parameters = "Field_Magnitude"
dataFM, metaFM = hapi(server, dataset, parameters, start, stop)
hapiplot(dataFM,metaFM)
```

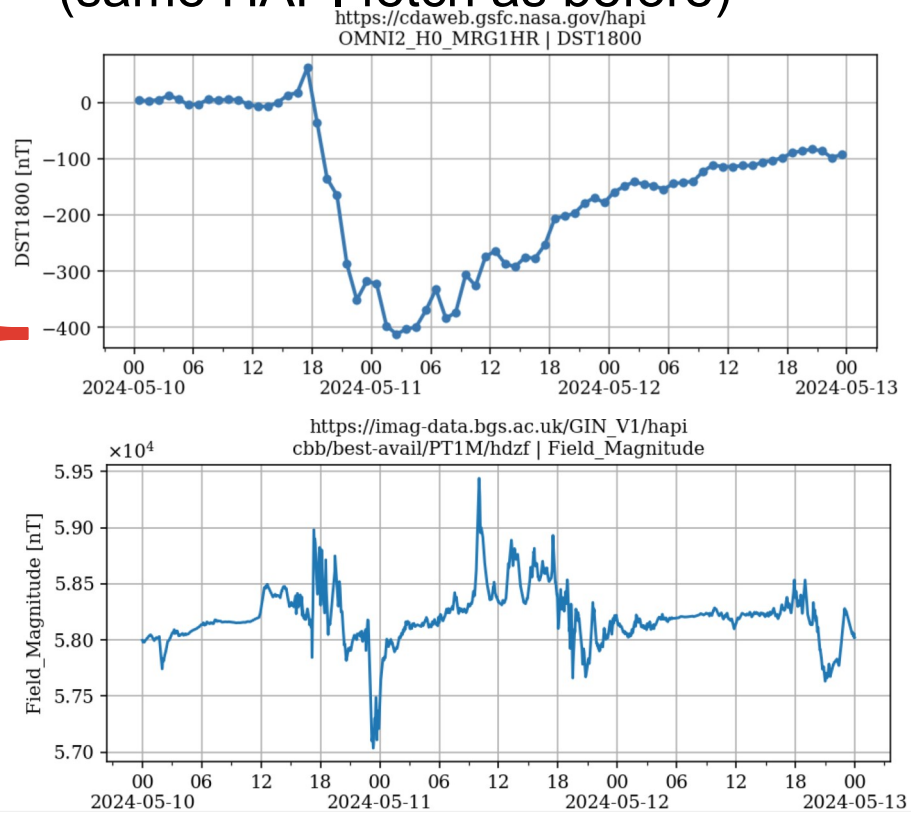


New Capabilities

CDAWeb & INTERMAGNET
(same HAPI fetch as before)

1) merge to unified time axis

```
from hapipytools import merge_hapi, hapi_to_df
dataMerged, metaMerged = merge_hapi(dataDST, metaDST, dataFM, metaFM)
hapiplot(dataMerged,metaMerged)
```



2) use in Pandas (DataFrame)

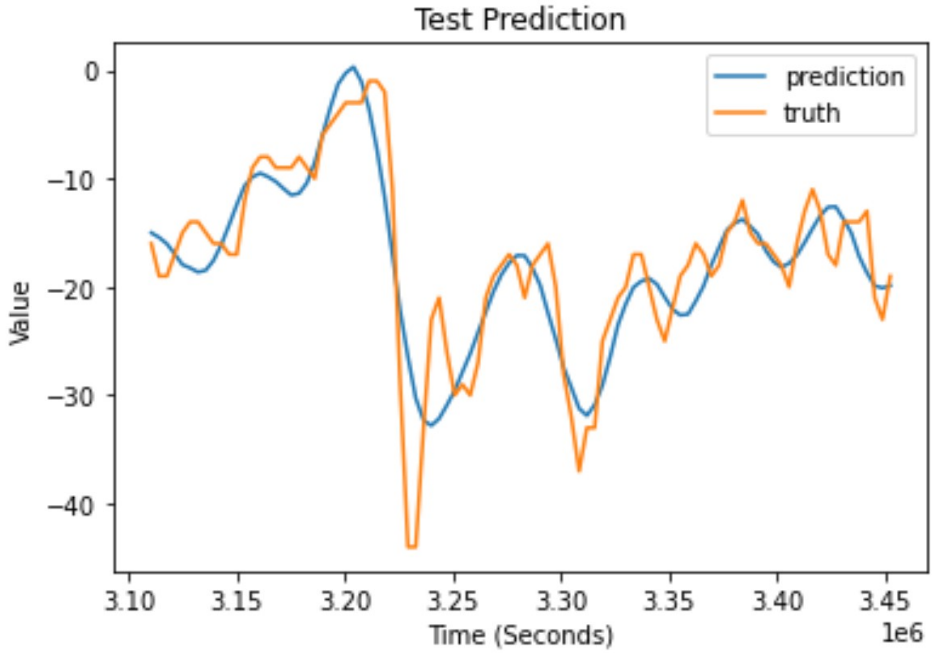
```
import pandas
merged_df = hapi_to_df(dataMerged)
```

More New

3) HAPI -> Machine learning (hapi-nn)
(PyTorch or Tensorflow)

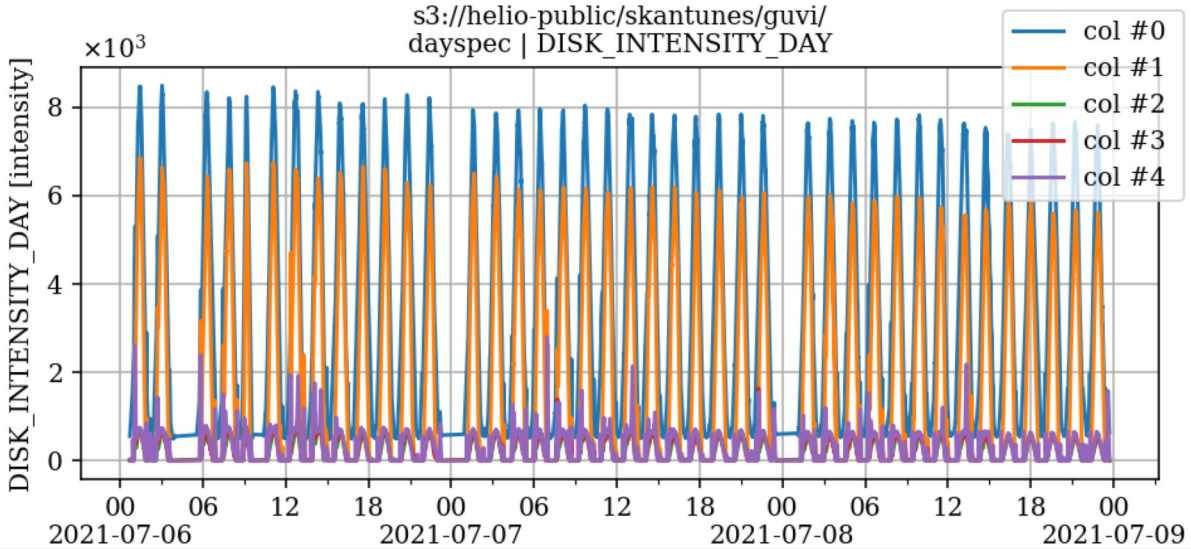
e.g. DST Clustering using CDAWeb OMNI2_H0_MGR1HR DST1800

```
import hapi-nn
#(fetch HAPI data)
trainer.set_hapidatas([data],
   xyparameters=[['DST1800'],['DST1800']])
trainer.prepare_data()
trainer.train(model, epochs, batch_size=batch_size)
predictions = tester.test(model)
```



4) local & cloud access, e.g. GUVI on AWS
(still in testing)

```
from hapiserver import hapi_localdisk
from hapiplot import hapiplot
start      = '2021-07-06T00:42Z'
stop       = '2021-07-09T00:00Z'
server     = 's3://helio-public/skantunes/guvi'
dataset    = 'dayspec'
params = 'DISK_INTENSITY_DAY'
dataGUVI, metaGUVI = hapi_localdisk(server, dataset, params, start, stop)
hapiplot(dataGUVI, metaGUVI)
```



What data is available now with HAPI?

Institution	Server	Type of Data	Num of Datasets
NASA	CDAWeb	Heliophysics	2800
	SSCWeb	Ephemeris	250
	SDAC	Solar Images (URLs)	50??
	CCMC	Space Weather Indices	250
IRAP Plasma Data Ctr	AMDA	Helio. & Planetary Data and Ephemeris	500
University of Iowa	Das2 Server	Helio. & Planetary	30
LASP Laboratory for Atmospheric and Space Physics	LISIRD	Solar Irradiance	40
SWARM Mission	ViRES Data Server	Space Mag Data	14
INTERMAGNET	INTERMAGNET	Ground-based Mag	~1000
Royal Netherlands Meteorological Institute	KNMI	Space Weather	~100
esa	ESAC / Cluster Mission Data	Helio. (magnetosphere)	

More HAPI servers coming soon

Institution	Server	Type of Data	Num. Datasets
	ESAC Solar Orbiter and others	Heliophysics	lots
JHU / APL	SuperMAG	global ground mag	~500
JHU / APL	TIMED / GUVI	ionospheric images	~10
	PDS PPI Node	Planetary Plasma, Particle, and Fields	~1000
NOAA / SWPC	SWPC	Space Weather	~50

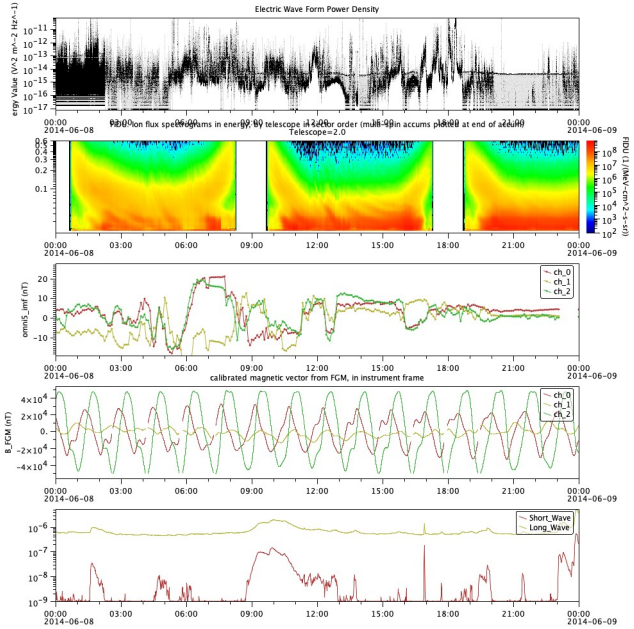
On the horizon			
CSA	Space Environment Canada (new initiative)	Ground-based ionospheric data	~1000
CEDAR / Madrigal	Open Madrigal	Space Weather	1000+ (??)
NCAR ? (in discussion)	*Multiple*	*Atmospheric*	*TBD*

Heliophysics Application Programmer's Interface

For Scientists: HAPI 'fills your arrays' for time series data
hapi-server.org lists servers, generates code clients for Python, Matlab, IDL, JS, Autoplot, wget, ...

For data providers: resources at *github.com/hapi-server*
we'll help you set up your server

HAPI is an established COSPAR-recommended standard with an evolving community



HAPI Specification and Libraries at Github,
https://github.com/hapi-server

All resources are linked off **https://hapi-server.org**

HAPI, Weigel et al, 2021