

Standardizing Access to Heliophysics Data - HAPI Specification Updates and Some New Usages for Cloud and Model Data

Jon Vandegriff¹, Eric Winter², Alex Antunes¹, Brent Smith², Robert Weigel³, Jeremy Faden⁴, D Aaron Roberts⁵, Robert Candey⁶, Eric Grimes⁷, Scott Boardsen⁸, Bernard Harris⁹, and Todd King¹⁰

¹Johns Hopkins University Applied Physics Laboratory

²Applied Physics Laboratory Johns Hopkins

³George Mason University

⁴Self Employed

⁵NASA Goddard SFC

⁶NASA Goddard Space Flight Cntr

⁷Auburn University

⁸NASA GSFC

⁹NASA Goddard Space Flight Center

¹⁰UCLA/ESS

November 24, 2022

Abstract

The Heliophysics Application Programmer's Interface (HAPI) is a standard mechanism for accessing time series data. Because of its adoption at multiple Heliophysics (HP) and Space Weather (SW) data centers, it is now a useful way to reach many different resources within the community. It is also a COSPAR standard. The use of HAPI so far has been as a standard layer on top of a traditional mission or instrument archive, where HAPI lives alongside an existing, custom web-based computer interface. We will give highlights of recent additions to the specification which is now at version 3.0, with 3.1 around the corner. We also will present explorations into two new ways in which HAPI can be utilized. 1) HAPI as a way to access output from model runs, which can generate large volumes of data at various time cadences and spatial distributions. 2) HAPI as an interface over cloud-based data resources. In the cloud, HAPI can connect large volumes of data to scientist-friendly front-end analysis capabilities, such as a JupyterHub or potentially a Pangeo-like environment. The evolution of HAPI and its uses is expected to keep enhancing interoperability among Heliophysics and Space Weather resources.

Standardizing Access to Heliophysics Data - HAPI Specification Updates and Some New Usages for Cloud and Model Data

The poster is divided into four main sections:

- What is HAPI?** HAPI is a data access standard for solar system data. It provides a specification and a way to use data from multiple HAPI sources. A callout states: "Software needs 3 different interfaces to talk to 3 data centers." Why HAPI? Without HAPI, there is less interoperability for multiple needs.
- New use case: Access to Cloud Data** HAPI can operate in a cloud-based environment. This year, we are building a system for doing this with a large volume of Heliophysics data. A diagram shows Cloud Data being accessed via HAPI to Cloud Compute. A note mentions that more details on solar model access to storage instead of regular file systems are available.
- New use case: Access Model Data** HAPI can serve as an interface over model data. Plan model outputs are otherwise difficult to be downloaded and organized before they can be generally used. A diagram shows HAPI connecting to a model and then to a user. A specific example is given: Geophysics SWO Modeling CAMERA. Earth's field lines (sampled from one time step) are shown, and CAMERA output data is too big to download - so we need HAPI to access model data in the cloud.
- HAPI Adapters and Plans** HAPI provides a standard interface for data providers to use or adapt. There are HAPI adapters for data providers to use or adapt. There is a verification interface to test any server - only available for server developers. There are HAPI plans for analyzing HAPI accessible data. For programming access, clients in Python, C#, Matlab, R, Java. Analysis Applications: SPEDAS, pySPEDAS, Bstplot. Many institutions have or are adopting HAPI. Logos for NASA, Community Coordinated Modeling Center (CCMC), Coordinated Data Analysis Web (CDAWeb), and SuperMAG/CAMERA are shown.

This website uses cookies to ensure you get the best experience on our website. [Learn more](#)



Jon Vandegriff, Sandy Antunes, Eric Winter, Brent Smith, Robert Weigel, Jeremy Faden, Aaron Roberts, Robert Candey, Eric Grimes, Scott Boardsen, Bernard Harris, Todd King

ENTER NAMES OF AFFILIATED INSTITUTIONS



PRESENTED AT:

AGU FALL MEETING
New Orleans, LA & Online Everywhere
13-17 December 2021

Poster Gallery brought to you by **WILEY**

WHAT IS HAPI

HAPI: a data access standard for time series data

The specification:

github.com/hapi-server/data-specification (<https://github.com/hapi-server/data-specification>)

To see new features in 3.0, visit the link above or scroll to the bottom of this panel.

Try it now:

<http://hapi-server.org/servers/> (<http://hapi-server.org/servers/>)

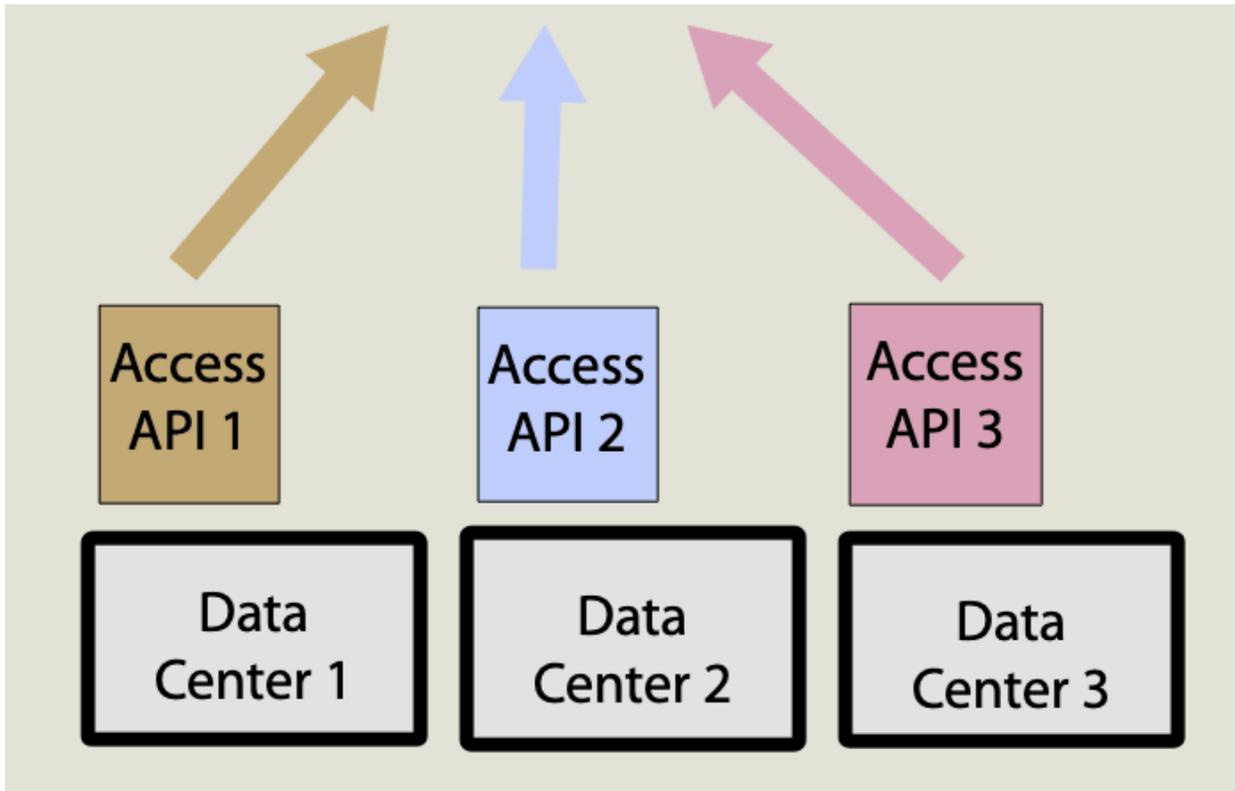
This lets you get and plot data from multiple HAPI servers, **and for every example, see functioning code showing how to programmatically get data yourself.**

Why HAPI?

Without HAPI, there is less interoperability for analysis tools.

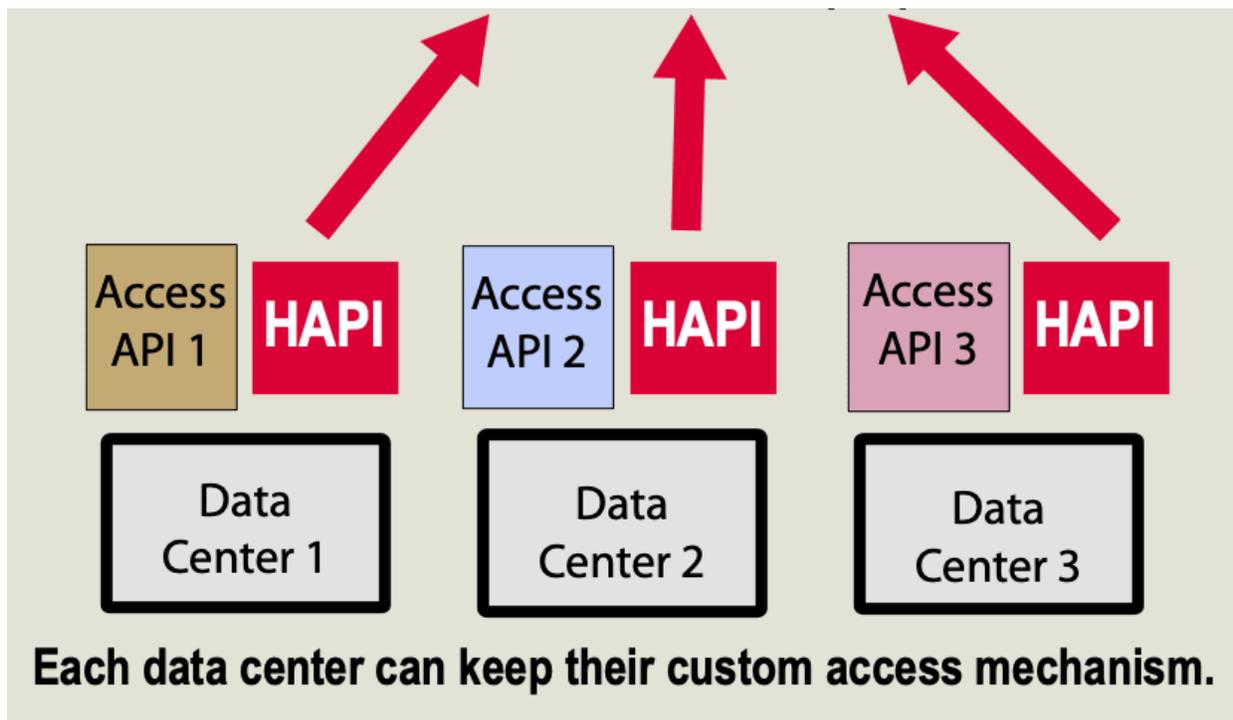


Software needs 3
different interfaces to
talk to 3 data centers.



With HAPI, data access is identical across data centers.

Software can use one
interface to talk to all
data centers.



A HAPI server offers simple end points that allow listing of available data and parameters, and then requests for streaming of data for an arbitrary time period. Endpoints:

`hapi/about`

`hapi/capabilities`

`hapi/catalog`

`hapi/info`

`hapi/data`

HAPI 3.0 Changes

Non-backward compatible changes to the request interface in HAPI 3.0:

1. The URL parameter `id` was replaced with `dataset`.
2. `time.min` and `time.max` were replaced with `start` and `stop`, respectively.
3. Addition of a new endpoint, "about", for server description metadata.

These three changes were discussed in issue #77. HAPI 3.0 servers must accept both the old and these new parameter names, but the HAPI 2 specification requires an error response if the new URL parameter names are used. In a future

version, the deprecated older names will no longer be valid.

HAPI 3.0 Schema Changes

1. Ability to specify time-varying bins (see issue #83)
2. Ability to use JSON references in info response (see issue #82)
3. Ability to indicate a units schema (if one is being used for units strings) (see issue #81)

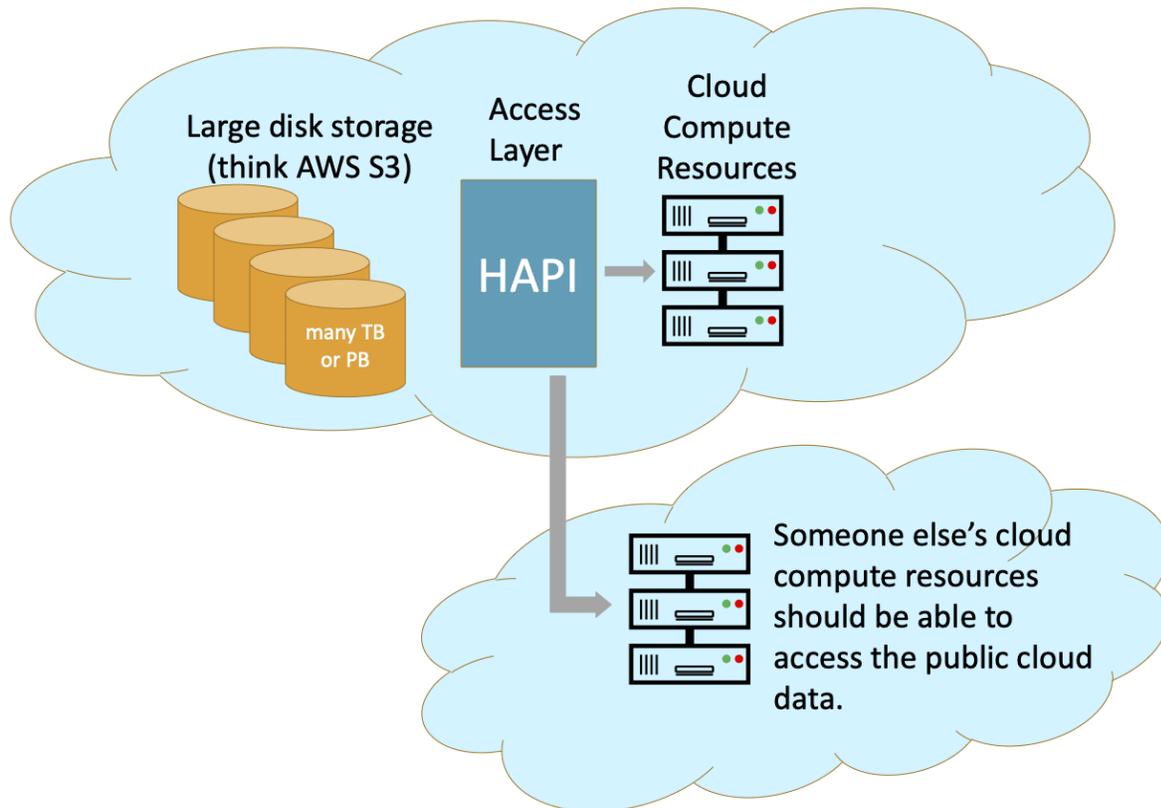
NEW USE CASE: ACCESS TO CLOUD DATA

HAPI can operate in a cloud-based environment.

This year, we are building a system for doing this with a large volume of Heliophysics data.



More details: code needd to access S# storage instead of regular file system data.



Most data formats are adding support for reading directly from S3 storage. For netCDF, it works!

Pseudo code to read netCDF from S3

(this is here to demonstrate that we've tried it)

```
import s3netcdf

floc['bucket'] = 'INSERT BUCKET HERE'
floc['key']     = 'INSERT KEY HERE'
floc['secret'] = 'INSERT SECRET HERE'

# AWS S3 file search
filenames = s3netcdf.s3_search("2019204')
                                                    # yyyydoy
... select one filename...

s3handle = s3netcdf.s3open(access)
# add of yyyydoy to filestem
doy = fname[28:35]
dirname = floc['bucket'] + '/' + doy + '/'

# open xarray via S3 bucket
science_data = s3netcdf.s3data(s3handle,dirname+fname)
```

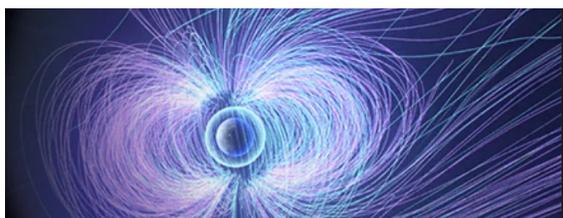
NEW USE CASE: ACCESS MODEL DATA

HAPI can serve as an interface over model data.

Raw model outputs are diverse and need to be transformed into regularized forms before they can be generically used.

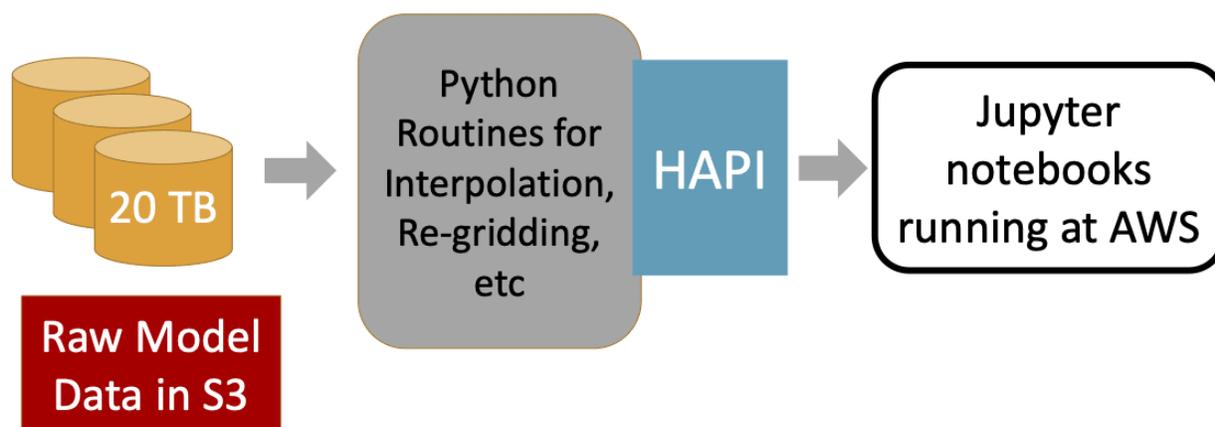


Geophysics MHD Modeling: GAMERA



Earth's field lines (snapshot from one time step).

GAMERA output data is too big to download - we are creating HAPI-based analysis tools in the cloud:



See also: Data-Model comparison poster by Micheal Wiltberger:

SuperMAG

<http://supermag.jhuapl.edu>



Data from SuperMAG sensors:

<http://supermag.jhuapl.edu/data/hapi>

Data from the same set of SuperMAG sensors formatted in exactly the same way as the HAPI data for the measurements, but coming from a simulated magnetic field (such as the GAMERA model - see below)/

<http://supermag.jhuapl.edu/model/hapi>

Data-model comparison can be made much easier.

Note that these model output mechanisms for HAPI are still in development!

HAPI ADOPTERS AND PLANS

Related poster: HAPI for SuperMAG and TIMED/GUVI ([https://agu2021fallmeeting-
agu.ipostersessions.com/Default.aspx?s=3E-29-0E-B3-35-50-69-B5-45-FB-0C-3F-44-C5-60-D5](https://agu2021fallmeeting-
agu.ipostersessions.com/Default.aspx?s=3E-29-0E-B3-35-50-69-B5-45-FB-0C-3F-44-C5-60-D5))

There are **HAPI servers** that data providers can use or adapt. See the list at <http://hapi-server.org> (<http://hapi-server.org>)

There is a **verification mechanism** to test any server -- very useful for server developers!

<https://github.com/hapi-server/verifier-nodejs> (<https://github.com/hapi-server/verifier-nodejs>)

There are **HAPI clients** for analyzing HAPI-accessible data.

For programmatic access, clients in: Python, IDL, Matlab, R, Java

Analysis Applications: SPEDAS, pySPEDAS, Autoplot

Many institutions have or are adopting HAPI:



Community Coordinated
Modeling Center (CCMC)
Coordinated Data Analysis
Web (CDAWeb)



SuperMAG,
GAMERA
in progress



Planetary Data System
(node for plasma, particles
and fields). *in progress*



Physics Department
(Autoplot)



Automated Multi-Dataset
Analysis (AMDA) at Plasma
Physics Data Centre (CDPP)



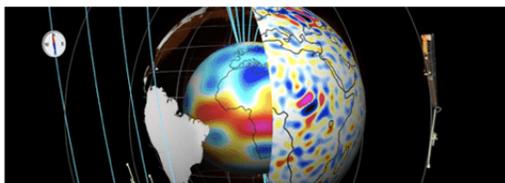
European Space Astronomy
Centre (ESAC). *in progress*



LASP Interactive Solar
Irradiance Data Center
(LISIRD)



Physics Department



ESA's SWARM Mission
(VirES toolkit)

DISCLOSURES

We thank NASA for supporting this work on the following grants:

Sustaining HAPI Through Promulgation and Development

NASA Grant 80NSSC22K0194

Enhancing Science Analysis with Cloud-Based Tools and Frameworks

NASA Grant 80NSSC22K0226

Creating Interoperable Data Views for Two Important Heliophysics Resources: SuperMAG and TIMED/GUVI

NASA Grant 80NSSC19K0874

Enhancing and Standardizing Python Readers for HAPI Data

NASA Grant 80NSSC20K0177

Enhancing Heliophysics Python Library Interoperability by Adapting Common Data Models

NASA Grant 80NSSC21K0304

AUTHOR INFORMATION

JHU/APL: Jon Vandegriff, Sandy Antunes, Eric Winter, Brent Smith

George Mason University: Robert Weigel

Cottage Systems, Inc.: Jeremy Faden

GSFC: Aaron Roberts, Robert Candey, Bernard Harris

UCLA: Eric Grimes, Todd King

UMBC / GSFC: Scott Boardsen

ABSTRACT

The Heliophysics Application Programmer's Interface (HAPI) is a standard mechanism for accessing time series data. Because of its adoption at multiple Heliophysics (HP) and Space Weather (SW) data centers, it is now a useful way to reach many different resources within the community. It is also a COSPAR standard. The use of HAPI so far has been as a standard layer on top of a traditional mission or instrument archive, where HAPI lives alongside an existing, custom web-based computer interface. We will give highlights of recent additions to the specification which is now at version 3.0, with 3.1 around the corner. We also will present explorations into two new ways in which HAPI can be utilized. 1) HAPI as a way to access output from model runs, which can generate large volumes of data at various time cadences and spatial distributions. 2) HAPI as an interface over cloud-based data resources. In the cloud, HAPI can connect large volumes of data to scientist-friendly front-end analysis capabilities, such as a JupyterHub or potentially a Pangeo-like environment. The evolution of HAPI and its uses is expected to keep enhancing interoperability among Heliophysics and Space Weather resources.

REFERENCES

HAPI - Heliophysics Application Programmer's Interface

<http://hapi-server.org>

<https://github.com/hapi-server/>

HAPI: An API Standard for Accessing Heliophysics Time Series Data

Robert Weigel, Jon Vandegriff, Jeremy Faden, Todd King, D Aaron Roberts, Bernard Harris, Robert Candey, Nand Lal, Scott Boardsen, Chris Lindholm, Doug Lindholm, Thomas Baltzer, Larry Brown, Eric Grimes, Baptiste Cecconi, Vincent Génot, Benjamin Renard, Arnaud Masson, Beatriz Martinez

JGR Space Physics, 18 November 2021, <https://doi.org/10.1029/2021JA029534>

SuperMAG - a collection of regularied, global magnetometer data at 1-minute and (newly) 1-second cadence

<http://superMAG.jhuapl.edu>

Gjerloev, J. W. (2009), A Global Ground-Based Magnetometer Initiative, EOS, 90, 230-231

Gjerloev, J. W. (2012), The SuperMAG data processing technique, J. Geophys. Res., 117, A09213, doi:10.1029/2012JA017683

GAMERA Model - Grid Agnostic MHD for Extended Research Applications

<https://cgs.jhuapl.edu/Models/gamera.php>

Zhang, B., Sorathia, K. A., Lyon, J. G., Merkin, V. G., Garretson, J. S., & Wiltberger, M. (2019). GAMERA: A Three-dimensional Finite-volume MHD Solver for Non-orthogonal Curvilinear Geometries. The Astrophysical Journal Supplement Series, 244(1), 20, <https://doi.org/10.3847/2F1538-4365/2Fab3a4c>.