

# Space Physics Data Facility (SPDF) Data Archives and Services

Robert Candey, Project Scientist  
Lan Jian, Deputy Project Scientist

## **Curation Scientists:**

Dieter Bilitza, Scott Boardsen, John Cooper (emeritus), Leonard Garcia,  
Andriy Koval, Guiping Liu, Sonya Lyatsky, Emily Mason, Natasha  
Papitashvili, Pertti Makela, Aaron Roberts, Jonathon Smith

## **Technical and Development Staff:**

Sarah Fooks, Bernie Harris, Rita Johnson, Tami Kovalick (lead), Howard  
Leckner, Mike Liu, Kyle Marshall, Uthra Rao, Jen Sun, Ron Yurow

# Introduction to SPDF

[spdf.gsfc.nasa.gov](http://spdf.gsfc.nasa.gov)

- ❑ SPDF is the **active and final archive of non-solar data** from NASA heliophysics missions and collaborative missions with other US and foreign agencies
  - Facilitate scientific analysis of multi-instrument and multi-mission datasets
  - Enhance the science return of the many missions, providing context with other missions
  - Facilitate open science and long-term archiving
  - Provide quality control and other support to the missions
  - Make data available via many access methods (HTTP, FTP, REST, HAPI)
- ❑ We also archive other data **relevant to NASA heliophysics science objectives**
  - Related data from *planetary missions*, such as MAVEN, New Horizons
  - Heliophysics data from some NOAA and DoD satellites, such as GOES, DSCOVR, LANL
  - Non-US missions such as Arase and Formosat
  - Ground-based magnetometers, aurora cameras, radars, etc., which are funded by NSF or other agencies
- ❑ We work closely with missions from early in their development on data issues and planning, particularly in implementing data standards

# SPDF Collaborations and Standards

- ❑ SPDF collaborates with the Solar Data Analysis Center (SDAC), Community Coordinated Modeling Center (CCMC), HelioAnalytics, and Heliophysics Data and Model Consortium (HDMC) to form the Heliophysics Digital Resources Library (HDRL) for fully open and coordinated access to all NASA Heliophysics data and software, and full integration of data and modeling resources
- ❑ Member of **International Heliophysics Data Environment Alliance** (IHDEA.net)
- ❑ SPDF also supports the SMD-wide data catalog and science standards efforts, instigated by the recent SMD Data Policy
- ❑ We also are collaborating with the Astrophysics Data System (ADS) digital library portal in adding heliophysics resources
- ❑ SPDF also builds critical infrastructures for the **Heliophysics Data Environment**:
  - Common Data Format (CDF) self-describing science file format ([cdf.gsfc.nasa.gov](http://cdf.gsfc.nasa.gov))
  - Heliophysics Data Portal discipline-wide data inventory and access service and its underlying SPASE metadata
  - ISTP/IACG/SPDF Metadata Guidelines of standardized internal metadata for understanding datasets

# Over 200 Missions/Projects Supported by SPDF

\* Only orbit data available

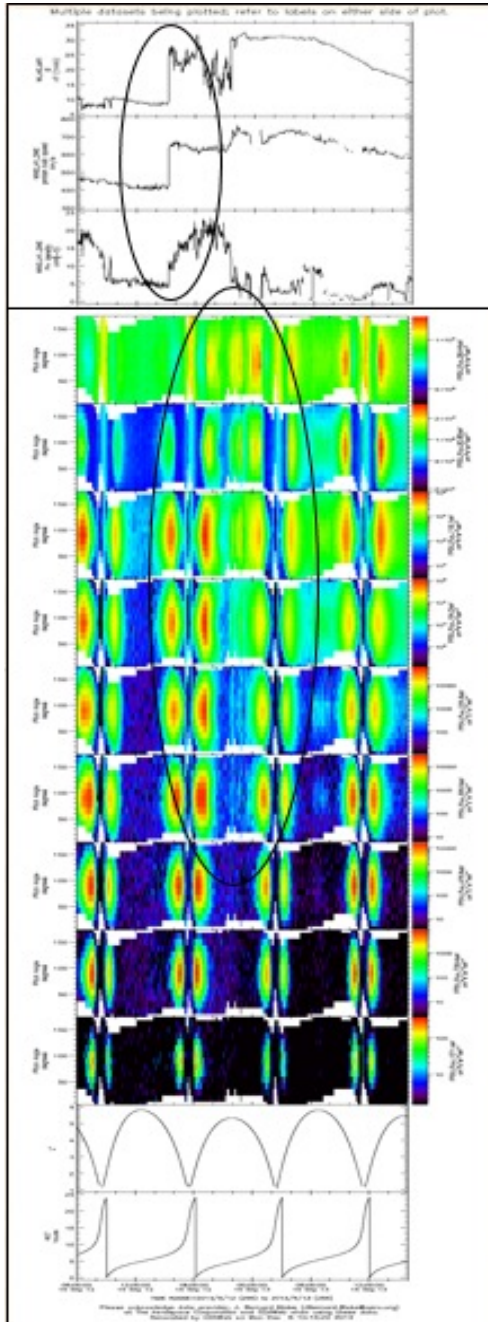
ACE		Cluster		GOES		Kepler		Parker Solar Probe		Spitzer	
Active*		Cosmos 900		GOLD		LANL		Phobos		Sputnik 1	
Aeros		C-NOFS		GPS		LRO		Pioneer		STEREO	
AIM		CRRES		GMS 3		LUNA		Pioneer 10		Suisei	
Akebono*		CSSWE		GRACE*		Magsat		Pioneer 11		Swarm	
Alouette1		Dawn*		Granat		MAP		Pioneer Venus		Tatiana	
Alouette2		DEMETER*		Hawkeye		Mariner 10		Polar		THEMIS	
AMPTE		DMSP		Helios		Mars		Prognoz		TIMED	
APEX-MAIN*		Double Star*		Hinode		MAVEN		Reimei		TRACE	
Apollo		DSCOVR		Hinotori		MESSENGER		Rosetta*		TWINS	
Aqua		DE		IBEX		Microlab 1		RHESSI		UARS*	
Ariel-4		Equator-S		ICON		Mir*		SAMPEX		Ulysses	
Arase (ERG)		Explorer		IMAGE		MMS		Sakigake*		Van Allen Probes	
ARCAD		FAST		IMP 7		MRO		San Marco		Vega	
ARTEMIS		FIREBIRD*		IMP 8		MSL		SCATHA*		Venera	
ASTRID II*		Formosat		IMP_early		MSX*		SDO		Viking	
AE		Freja*		Interball		Munin		SET-1/DSX		Voyager	
Aura		Galileo*		ISEE		New Horizons		SMILE		Voyager 1	
Aureol2		GCOM W1		ISEE 3-ICE		NOAA*		SNOE		Voyager 2	
BARREL		Genesis		ISIS		Oersted		SOHO		Wind	
BepiColombo		Geotail		ISS		OGO		Solar Orbiter		XMM-Newton	
CALIPSO		Geotail		Jason 2		Ohzora		SORCE		Yohkoh*	
Cassini*		Giotto*		Juno		PARASOL		Spartan-A		Zond	
Cassiope		GOCE									

Total: ~10,000 datasets, ~400 TB data

Recent average monthly data ingestion rate: ~0.6 million data files, ~13.7 TB data



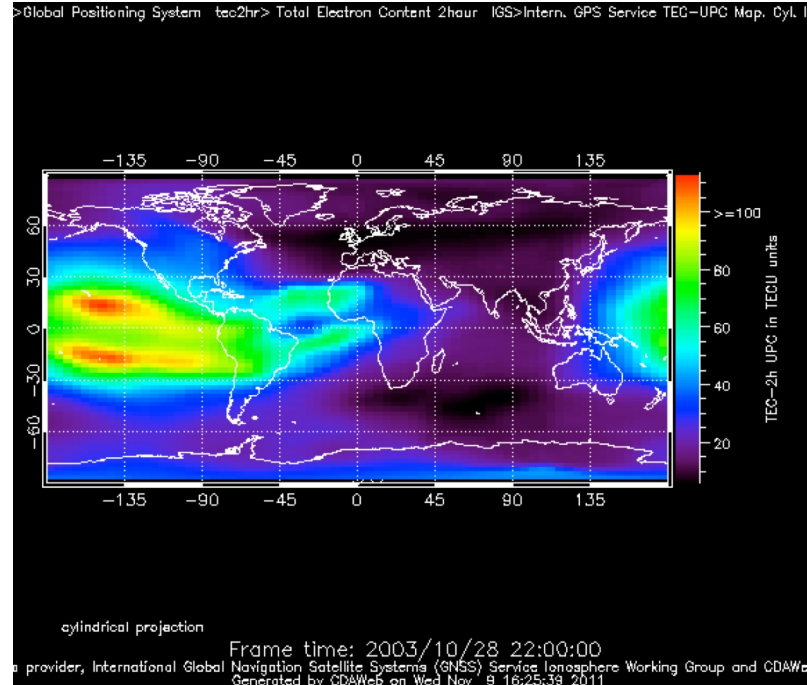
WIND MFI & SWE  
Van Allen Probe A ECT & MagEIS



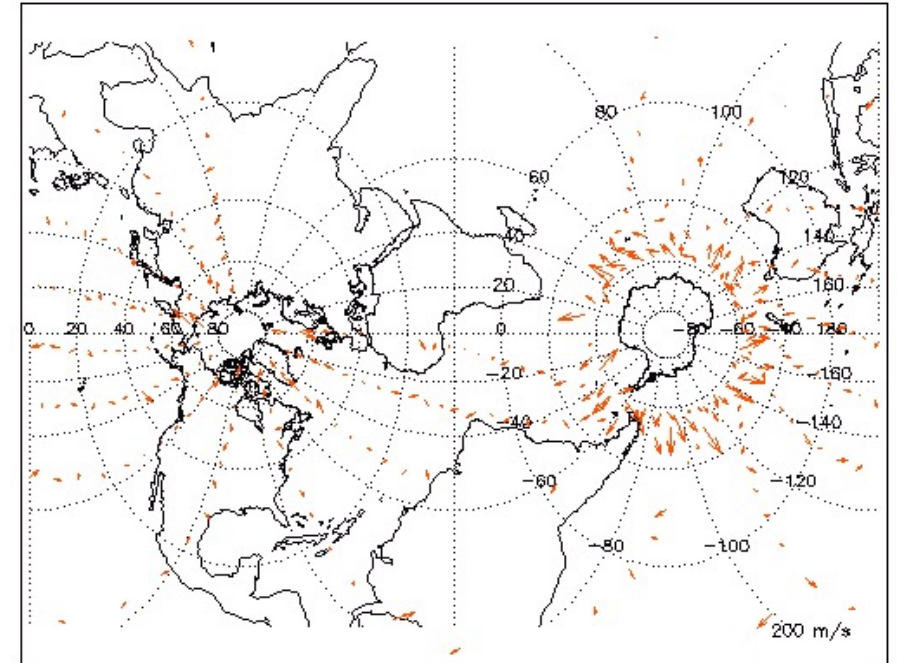
# Coordinated Data Analysis Web (CDAWeb)

Serve ~ 80 missions/projects

Example Parameter Displays



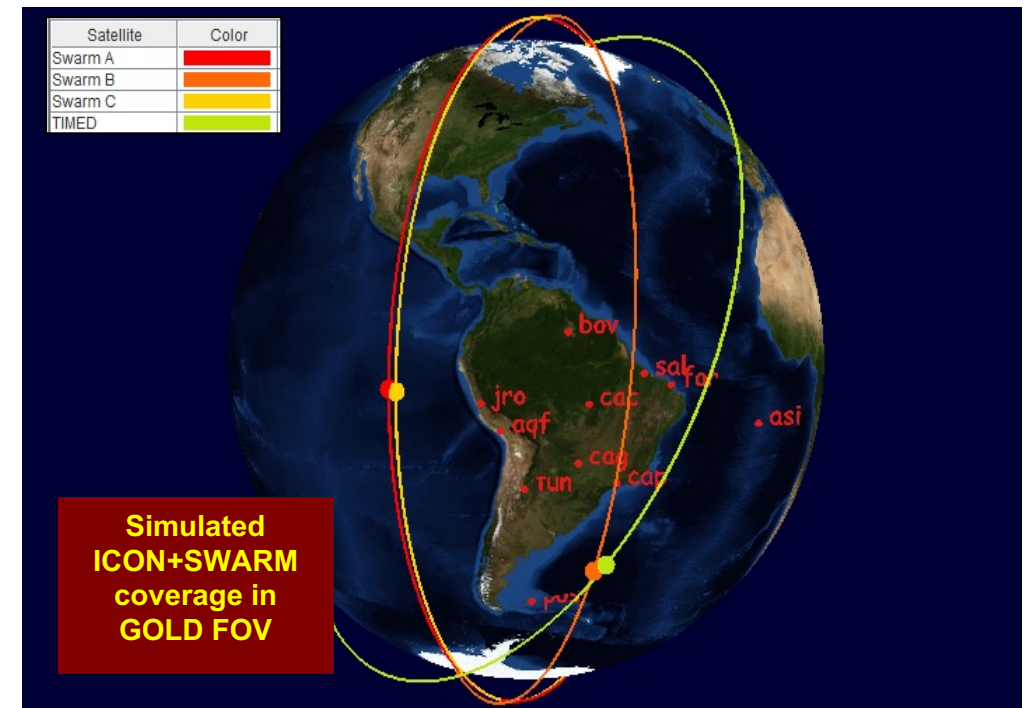
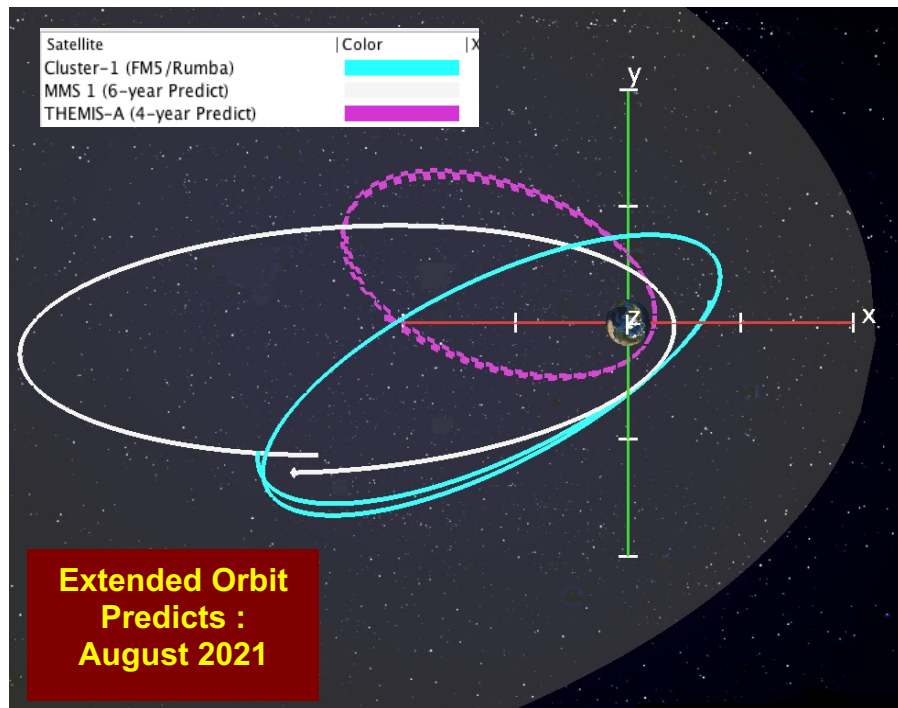
GPS International GNSS Service Total  
Electron Content



TIMED/TIDI Wind Vectors Movie  
Transverse Mercator Projection

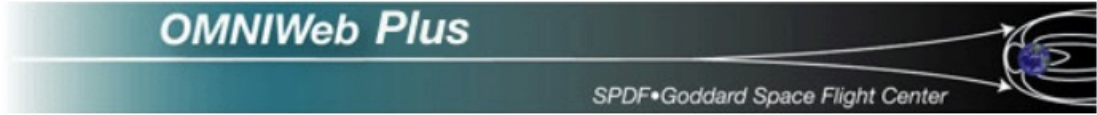
# Satellite Situation Center (SSCWeb)

- Include most heliophysics satellites and many ground stations
- Plot and list orbits of multiple spacecraft in a variety of coordinate systems
- **4D Orbit Viewer:** Interactive 4D animation of orbits
- Query for satellite-satellite and satellite-ground station conjunction





- OMNIWeb Plus, Home
+ ABOUT THE DATA
+ABOUT THE INTERFACE
+Data from command line
+ SPDF/FTP
+ Citing OMNI data usage
DATA via FTPBrowser
Energetic Particle fluxes
ATMOWeb main page
CGM transformation



Paths to Magnetic field, Plasma, Energetic particle data relevant to heliospheric studies and resident at Goddard's Space Physics Data Facility.

- OMNI data (spacecraft-interspersed, near-Earth solar wind data)
  - Low resolution OMNIWeb (1-hour, 1 and 27 days, 1963 - current)
  - High resolution OMNIWeb (1-min, 5-min, 1981 - current)
- Spacecraft-specific data sets (near 1 AU, including near-Earth)
  - + ACE
  - + Geotail
  - + IMP-8, IMP6&7
  - + Wind
  - + Explorer 33&35, Genesis, ISEE 3, Prognoz, SOHO, GOES
  - + Moon Related Spacecraft
  - + DSCOVR
- Deep space data
  - COHOWeb-formatted hourly solar wind field, plasma and proton fluxes
  - + Pioneer
  - + Ulysses
  - + Voyager
  - + Cassini, Helios, Mariner, STEREO
- Interfaces for comparing multi-source data
  - + Merged Magnetic field and Plasma 1-min
  - + Magnetic field
  - + Plasma
  - Energetic particle fluxes
    - Multi-source spectra of energetic particle fluxes (MSSP)
    - + IMP8/CPME, GOES and ACE/SIS proton fluxes, 1-hour

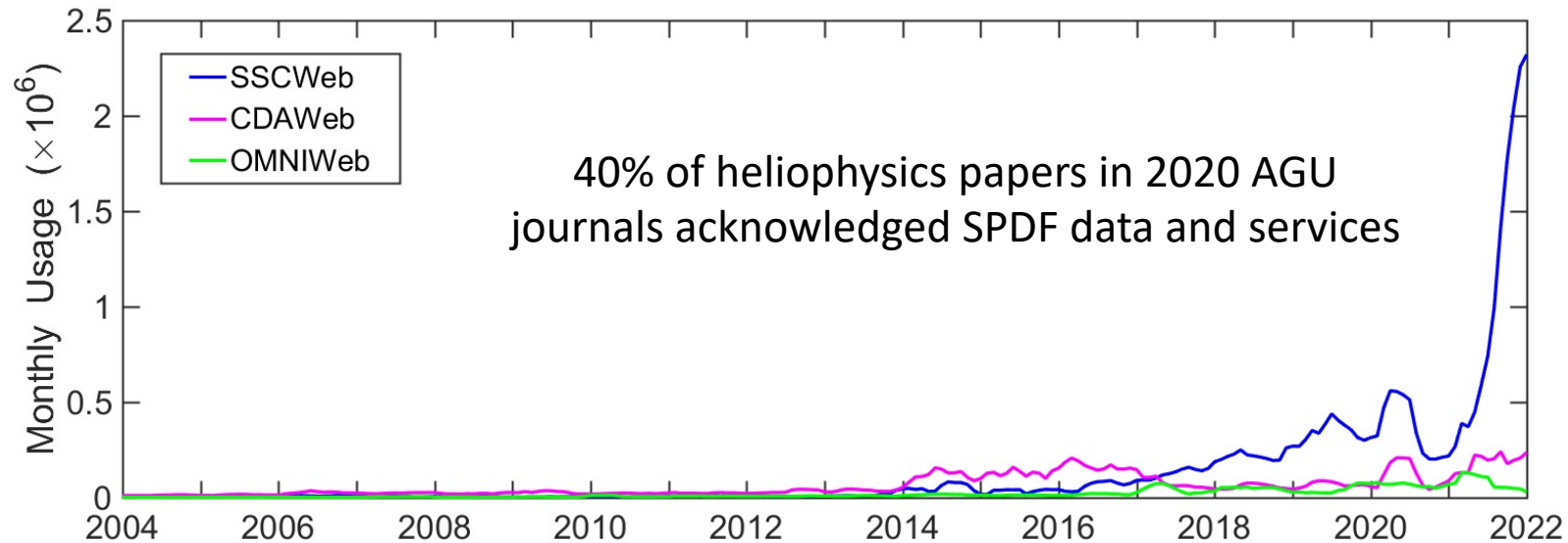
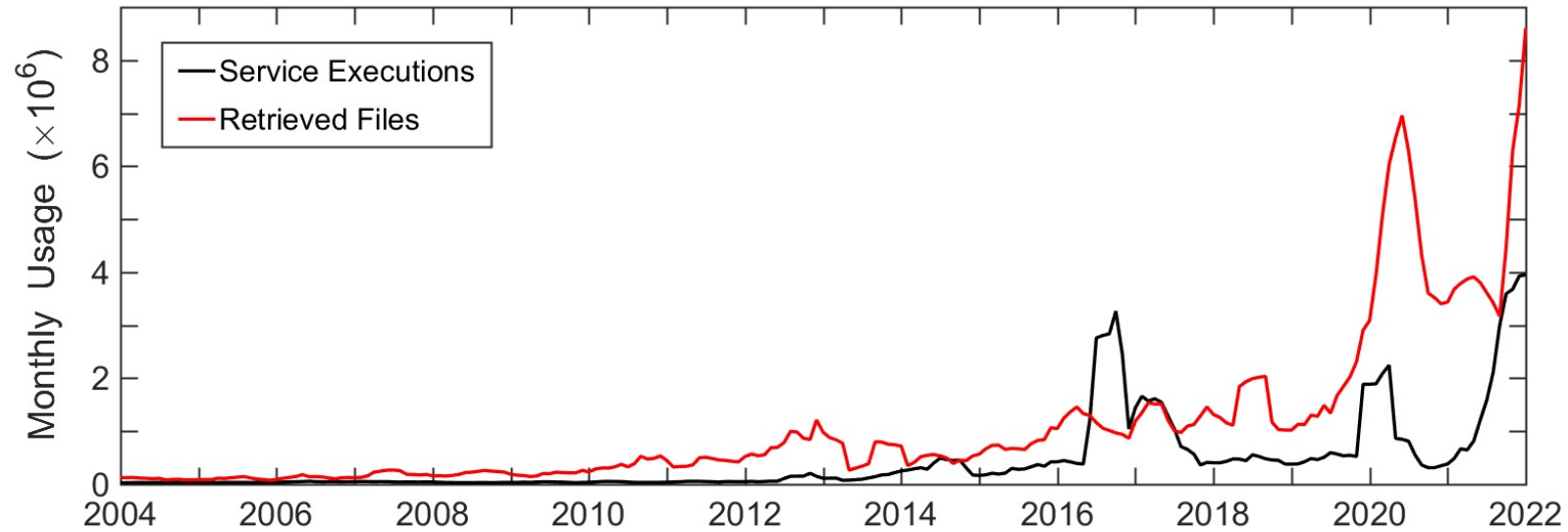
[Heliocentric Trajectories for Selected Spacecraft, Planets, and Comets](#)

# OMNIWeb Plus

- OMNI Data: Database of solar wind magnetic field and plasma parameters mapped to the nose of the Earth's bow shock
- Based on a large volume of quality-controlled satellite measurements (since Nov. 1963)
- **COHOWeb**: Solar wind field, plasma, and energetic particle fluxes in other locations of the heliosphere, especially useful for planetary studies and heliospheric model validation
- Interface for plotting, filtering, and downloading the data

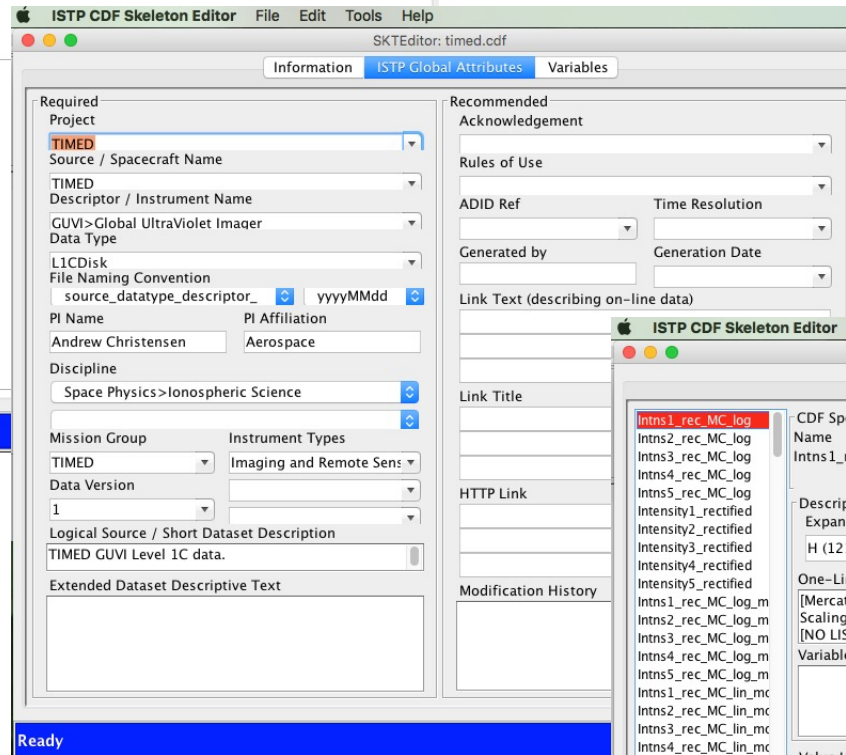
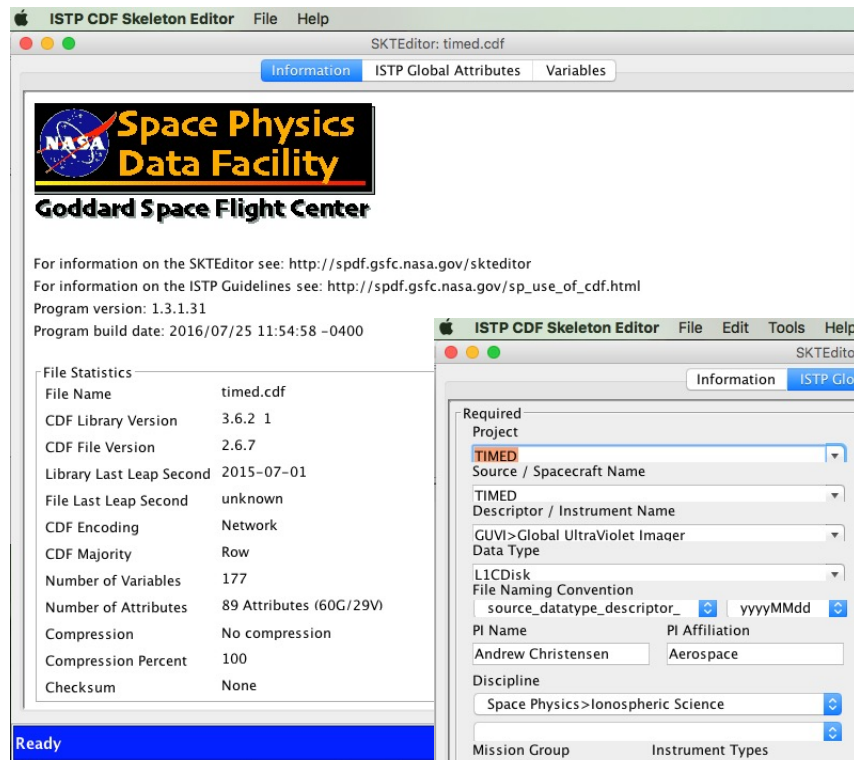
# SPDF Statistics

reports at [cdaweb.gsfc.nasa.gov/publiclogs/](https://cdaweb.gsfc.nasa.gov/publiclogs/)

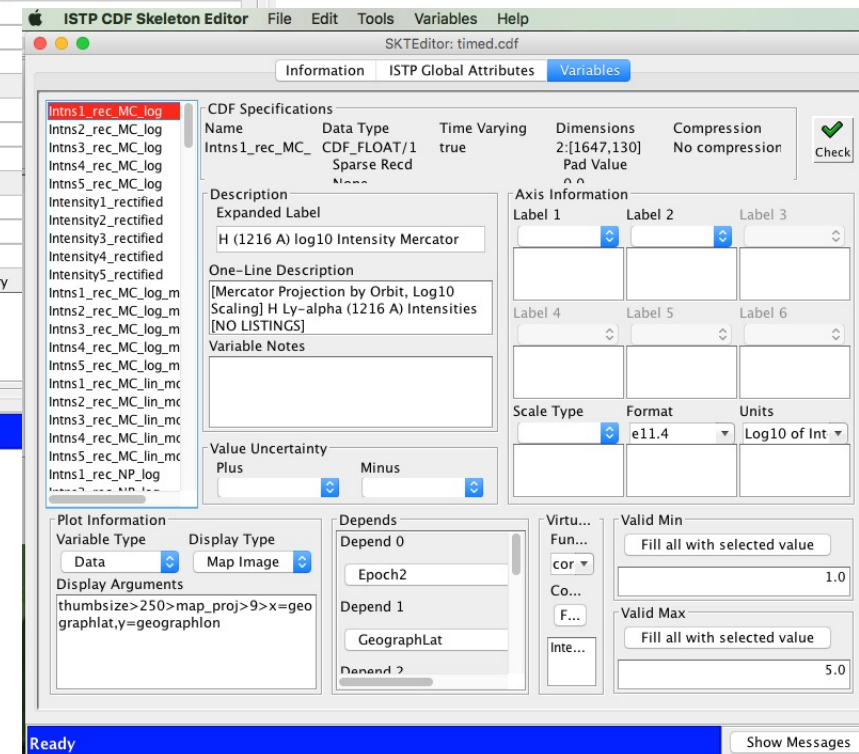




# ISTP Metadata Guidelines for CDFs and netCDFs SKTeditor tool



Update SKTeditor tool away from Java and include SPASE metadata creation at the same time



The following variables are not ISTP-compliant:

- Intns1\_rec\_MC\_log  
DISPLAY\_TYPE attribute value 'map\_image>THUMBSIZE>250>MAP\_PROJ>9>x=GeographLat,y=GeographLon' is not all lower case.  
DISPLAY\_TYPE attribute value changed to 'map\_image>thumbsize>250>map\_proj>9>x=geographlat,y=geographlon'.  
DISPLAY\_TYPE error: invalid keyword 'thumbsize'
- DEPEND\_2 is not 1 dimensional
- DEPEND\_2 is wrong size
- DEPEND\_1 is not 1 dimensional

# SPDF Plans

- ❖ HelioCloud support
- ❖ Improving the ISTP/IACG Metadata Guidelines with version control, etc.
- ❖ Replacing current Java-based 4D orbit viewer with browser-based viewer
- ❖ Updating CDAWeb displays with interactive plotting and improved interface
- ❖ Upcoming web redesign and adding user guides and how-to's for our services and tools
- ❖ Supporting CDFs in cloud object storage, perhaps with Zarr like netCDF is exploring
- ❖ CDF-JSON, added to converters for CDF, CDFML, netCDF, HDF-4, FITS, and to PDS-3

# **Space Physics Data Facility (SPDF)**

**[spdf.gsfc.nasa.gov](http://spdf.gsfc.nasa.gov)**

Backup



# SPDF provides multiple services and access methods

- Direct file downloads via FTPS and HTTPS <<https://spdf.gsfc.nasa.gov/pub/data/>>
- Orbit and ground track displays/queries via SSCWeb and 4D Orbit Viewer
- CDAWeb services:
  - Data files, plots and listings with supersets or subsets by time & selected variables, time-binning
  - Web service interfaces (REST, SOAP, IDL, Matlab, Java, Python) <<https://cdaweb.gsfc.nasa.gov/WebServices/>>
  - New HAPI (Heliophysics API) <<https://cdaweb.gsfc.nasa.gov/hapi>>
  - Autoplot [autoplot.org/help#CDAWeb](http://autoplot.org/help#CDAWeb)
  - Other methods such as IDL <[https://cdaweb.gsfc.nasa.gov/alternative\\_access\\_methods.html](https://cdaweb.gsfc.nasa.gov/alternative_access_methods.html)>
- SPDF complements the services of the mission and instrument teams
- SPDF auto-ingest scripts check all supported mission data sites daily to retrieve new data files, and CDF files are validated and ingested
- Master CDFs add or improve metadata for use in CDAWeb
- The **SPASE** (Space Physics Archive Search and Extract <<http://www.spase-group.org/>>) team use the master CDFs to generate SPASE IDs and descriptions for all datasets, to add entries to the **Heliophysics Data Portal** <<https://heliophysicsdata.gsfc.nasa.gov>> and mint DOIs for each dataset

# SPDF activities in past year

- Added many new datasets from ICON, GOLD, Parker Solar Probe (87), IBEX (40), Solar Orbiter (81), MMS (16), BARREL (219), FAST, Voyager PWS waveform, and many other spacecraft, rocket, balloon, and ground instruments
- Final data from Van Allen Probes (RBSP), with most datasets entirely reprocessed
- Automated ingest pipeline for > 75 missions out of over 200 missions for a total of ~4,000 datasets using ~400 TB (ingest and usage logs: <https://cdaweb.gsfc.nasa.gov/publiclogs/> )
- Creating CDFs from SOHO in-situ data and finishing making CDFs for IBEX data and Wind STICS
- Continue population of OMNI, COHO, SSC databases
- CDAWeb plot and display improvements, waveforms, inventory plots, time slices, audification
- Adding SPASE Resource IDs and DOIs to CDAWeb metadata and displays
- Working towards a grand vision as part of NASA's Heliophysics Digital Resource Library, including 6 new curation scientists (part-time)

# CDF Plans

- High-level functions to read variables or whole CDF into a map structure for IDL, Java, Perl and C#
- CDF-JSON converter
- Improve Windows installer, autoconf/make build/install, Mave/Ant/Gradle installs
- Improve documentation, beginner's guides, add to Wikipedia CDF entry
- Standardize ISTP/IACG Metadata Guidelines with version control, etc.
- New SKTeditor in Javascript or Python, perhaps also SPASE metadata creation
- Look into supporting CDFs in cloud object storage, perhaps Zarr like netCDF is exploring
- Define CDF MIME type and international standard
- Apache 2 license in place of current custom license
- Update CDFML and its corresponding JSON representation with cdf.xsd use more specific datatype (e.g., xs:dateTime, xs:integer, xs:float, etc.) instead of just xs:string

# Abstract

In order to improve access to the data and models of the Heliophysics System Observatory (HSO) and NASA-funded research projects, the NASA Heliophysics archive and modeling groups are collaborating to create a Heliophysics Digital Resources Library (HDRL) for improved cross-mission and observation-model comparison, machine learning and other large-scale and collaborative analysis, increased discoverability and usability of data and model results, software and services, and more complete metadata and provenance and quality control. Observational data are archived and served by the Solar Data Analysis Center (SDAC) and the Space Physics Data Facility (SPDF). The Community Coordinated Modeling Center (CCMC) provides empirical and first-principles simulations and analysis and display tools. A number of largely cross-cutting registry, access, and analysis standards and tools are provided by the Heliophysics Data and Model Consortium (HDMC).

As part of this effort, SPDF, as the active and final archive for non-solar NASA Heliophysics data, works with current operating missions and the Heliophysics community to ingest, preserve and serve a wide range of past and current public science-quality data from the mesosphere into the furthest reach of deep-space exploration. SPDF facilitates scientific analysis of multi-instrument and multi-mission datasets to enhance the science return of the many missions. SPDF develops and maintains the Common Data Format (CDF) and the associated ISTP/SPDF metadata guidelines. SPDF services include CDAWeb, which supports both survey and burst mode data with graphics, listings and data superset/subset functions. SPDF is currently receiving and serving data from missions including Parker Solar Probe, Solar Orbiter, MMS, Van Allen Probes, THEMIS/ARTEMIS, GOLD, ICON, ACE, Cluster, IBEX, Voyager, Geotail, Wind and many others, and >120 Ground-Based investigations. SPDF also operates the multi-mission orbit displays and query services of SSCWeb and 4D Orbit Viewer, as well as the Heliophysics Data Portal (HDP) discipline-wide data inventory and access service, and OMNIWeb and COHOWeb for near-Earth and deep-space solar wind plasma, magnetic field, and energetic particle database, respectively.

## Plain-Language Summary:

Working in cooperation with current operating missions and the Heliophysics community, Space Physics Data Facility (SPDF <https://spdf.gsfc.nasa.gov>), as one of the NASA Heliophysics active final archives, preserves and distributes in-situ data. SPDF ingests, preserves and serves a wide range of past and current public science-quality data from the mesosphere into the furthest reach of deep-space exploration from a wide variety of Heliophysics missions. SPDF is collaborating with the other NASA Heliophysics archive and modeling groups to create a Heliophysics Digital Resources Library (HDRL) for improved cross-mission and model comparison, machine learning and other large-scale and collaborative analysis, increased discoverability and usability of observation data and model results, software and services, and more complete metadata and provenance and quality control.



# HELIOPHYSICS SYSTEM OBSERVATORY

- 20 Operating Missions with 27 Spacecraft
- 12 Missions in Formulation or Development
- 6 Under Study

- FORMULATION
- IMPLEMENTATION
- PRIMARY OPS
- EXTENDED OPS

CubeSats		
In Development		On Orbit
AEPEX	CuSP	LLITED
AERO / VISTA	DAILI	MinXSS-3
CIRBE	Dione	petitSat
CODEX	GTOSat	REAL
CURIE	LAICE	SPORT

OPERATING & FUTURE



# Introduction to Heliophysics Digital Resources Library (HDRL)

- NASA Heliophysics archive and modeling groups collaborate to create a Heliophysics Digital Resources Library (HDRL) for:
  - Improving access to the data and models of the Heliophysics System Observatory (HSO) and NASA-funded research projects
  - Improving cross-mission and observation-model comparison, machine learning and other large-scale and collaborative analysis
  - Increasing discoverability and usability of data and model results, software and services, with more complete metadata and provenance and quality control.
- HDRL components
  - Solar Data Analysis Center (SDAC) archives observational solar data
  - Space Physics Data Facility (SPDF) archives observational non-solar data
  - Community Coordinated Modeling Center (CCMC) provides empirical and first-principles simulations and analysis and display tools
  - Heliophysics Data and Model Consortium (HDMC) provides largely cross-cutting registry, access, and analysis standards and tools

# Heliophysics Digital Resources Library (HDRL) Architecture



## Mission Operations

