



The SPASE Metadata Model: Standard Metadata for Space Science Data Description

*Todd King
James Thieman
D. Aaron Roberts
SPASE Consortium*

The Heliophysics Data Environment (HPDE)

... a place for scientific discovery.

Based on the principals of

- Open Source
- Open Systems
- Open Data

Because most projects that acquire data are publicly funded and research is increasingly multi-domain. Our goal is to ...

**... enable sharing and collaboration in
space physics research.**

HPDE Preferred Data Formats

Must be:

- Self Documented Data
- Established in the Heliophysics Data Policy
http://hpde.gsfc.nasa.gov/Heliophysics_Data_Policy_2009Apr12.html

Preferred (Commonly Used) Formats:

- HDF-5 (primarily in Earth Science; netCDF is now related to this),
- FITS (e.g., in Astronomy and Solar Physics),
- CDF (increasingly common in Space Physics),
- ASCII files with headers and/or independent documentation.

A Brief Note on CDF-A

- CDF-A is an archive acceptable form of the CDF data format.
 - Based on the ISTP CDF specification.
 - Include the SPASE resource identifier
 - CDF global attribute `spase_DatasetResourceID`
 - Note: Granules use `spase_GranuleResourceID`
 - Extensions to allow inclusions of SPASE description.
 - CDF global attribute `spase_DatasetResource`.

Sharing In Three Simple Steps

1. Create SPASE Resource Descriptions (metadata)
2. Submit description to metadata repository.
3. Use a Virtual Observatory (or registry services)

... and now for the explanation of what this means.

What is SPASE?

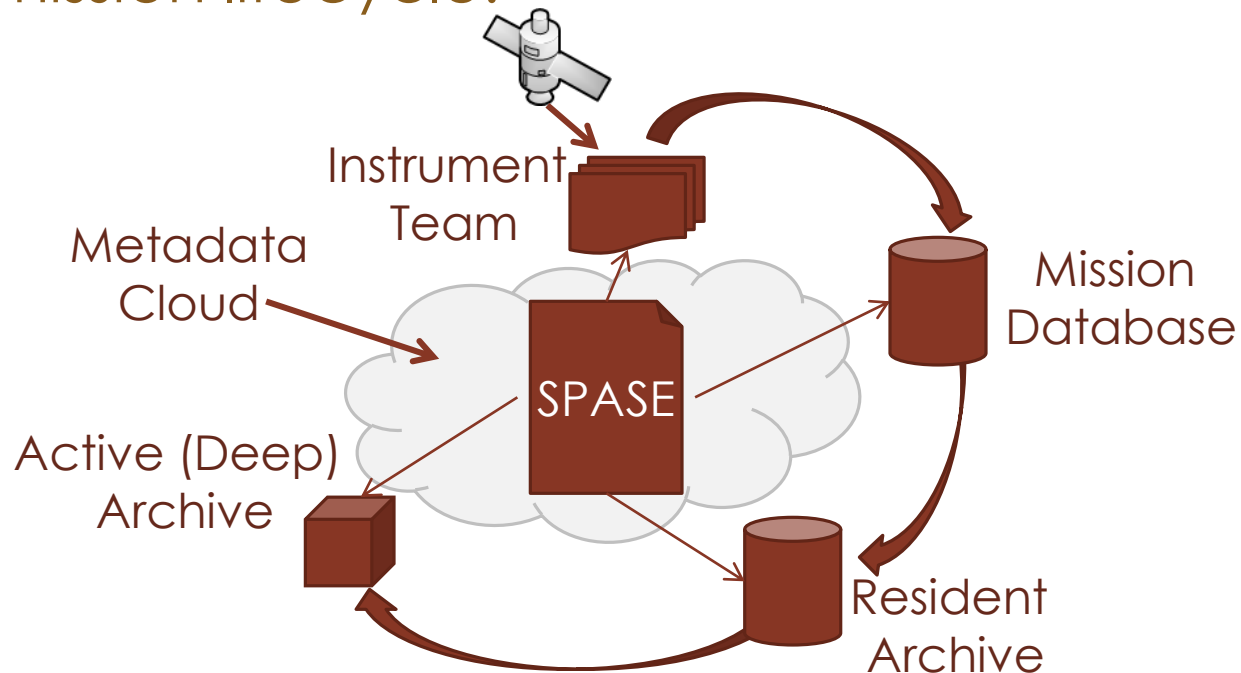
<http://www.spase-group.org/>

Space Physics Archive Search and Extract (SPASE)

- An International, community-based standards organization with the goals of:
 - Easing data search and retrieval across the Space and Solar Physics data environment
 - Defining and maintaining a Metadata Model for Space and Solar Physics interoperability
- Initiated in late 2002 as a discussion among space physics data archive representatives.
- Organized in 2003 as an international consortium with an open invitation for anyone in the community to participate
- U.S. participants funded by NASA in July 2005.
- Receive "permanent support" from NSSDC in 2007.

Metadata Cloud

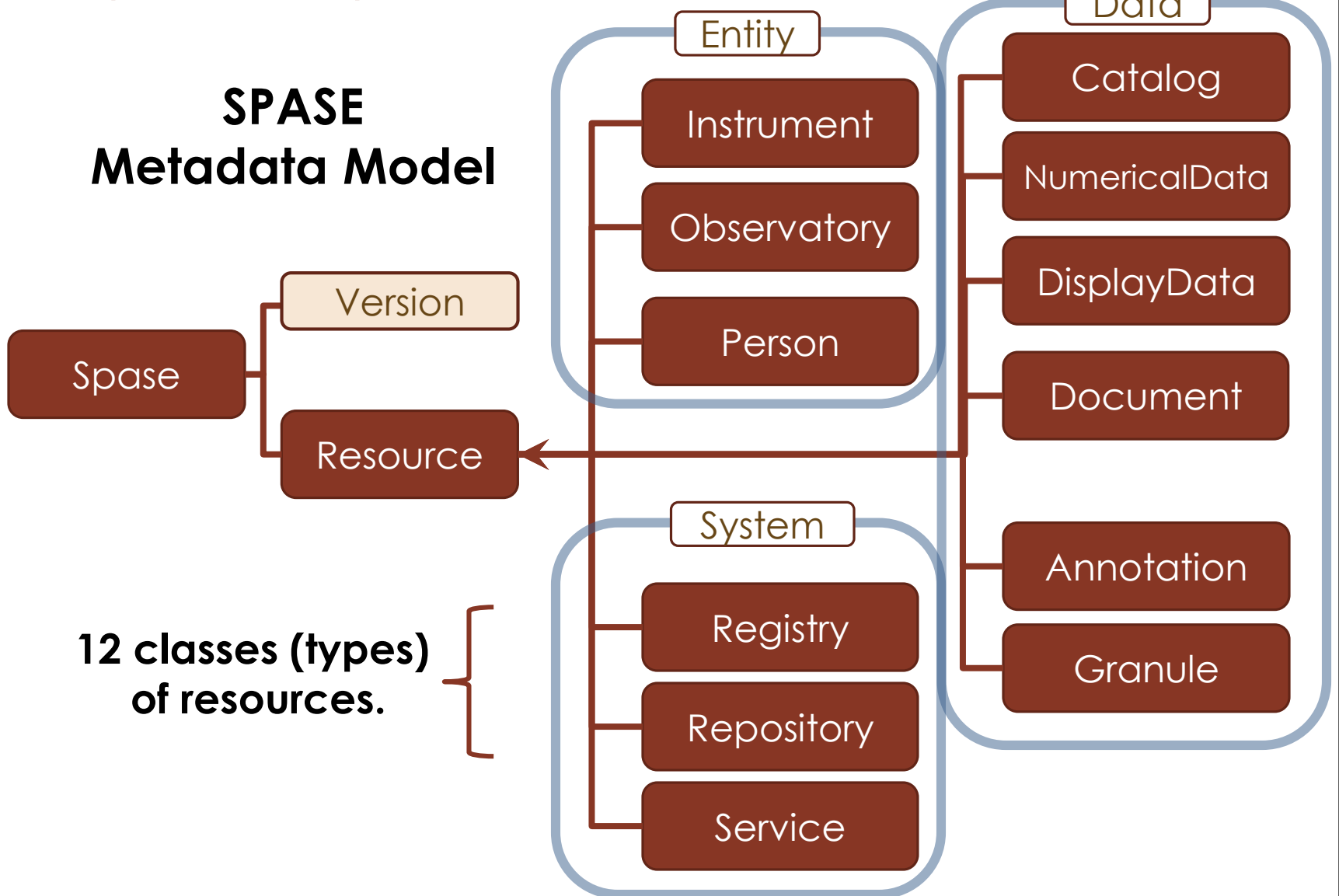
- SPASE expects data to move during the mission lifecycle.



- Data and metadata are managed separately.

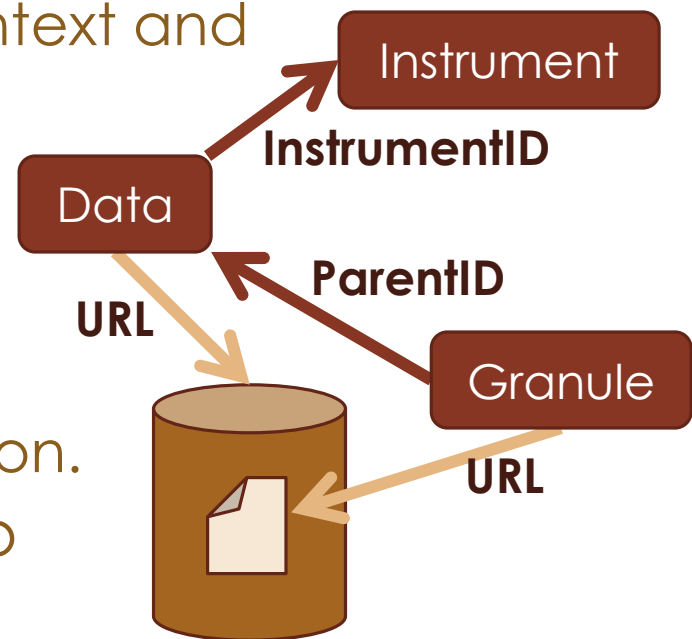
Space Physics Archive Search and Extract

SPASE Metadata Model



Associations

- Describe each unique resource once.
 - Assign a universally unique identifier
- Then use associations to provide
 - Provenance, scientific context and location of sampled data.
- Like kinds of data are described as collections of files.
 - A "Granule" describes a member (file) of a collection.
- Metadata is connected to data through URLs.



Example SPASE Instrument Description: Wind/WAVES

```
<?xml version="1.0" encoding="UTF-8"?>
<Spase xmlns:x0="http://www.spase-group.org/data/schema/spase-2_0_0.xsd">
<Version>1.0.0</Version>
<Instrument>
  <ResourceID>spase://SMWG/Instrument/Wind/WAVES</ResourceID>
  <ResourceHeader>
    <ResourceName>Wind Waves Thermal Noise Receiver</ResourceName>
    <ReleaseDate>2011-08-17T21:10:14Z</ReleaseDate>
    <Description>This investigation is designed to measure the intensity and arrival direction for both propagating and in situ waves originating in the solar wind near the earth. These waves depict the state of the solar wind impinging on the earth's magnetosphere. The instrument contains five subsystems within the main electronics box, plus the antenna subsystems which include a spin-axis and two spin-plane electric antennas (all spacecraft supplied) and a triaxial search coil (supplied by the plasma wave consortium). The five subsystems in the main electronics box are the radio frequency receivers, the comb filter receiver, the fast envelope sampler, the waveform analyzer, and the power distribution subsystem. The radio frequency receivers sweep over the band from about a fraction of a Hertz up to about 14 MHz for the electric field and to about 3kHz for the magnetic field. The comb filters have selectable bandwidths of 0.5, 1, or 2 Hz, with a total frequency range of 5 to 100 kHz. The fast envelope sampler is designed to capture transient events over four possible commandable decade ranges: 0.2-2, 0.6-6, 2-20, and 6-60 kHz. The waveform analyzer operates in the frequency regime below 16 kHz.</Description>
    <Acknowledgement/>
    <Contact>
      <PersonID>spase://SMWG/Person/Jean-Louis.Henry.Bougeret</PersonID>
      <Role>PrincipallInvestigator</Role>
    </Contact>
    <InformationURL>
      <Name>NSSDC's Master Catalog</Name>
      <URL>http://nssdc.gsfc.nasa.gov/database/MasterCatalog?sc=1994-071A&ex=5</URL>
      <Description>Information about the Plasma and Radio Waves (WAVES) experiment on the Wind mission.</Description>
    </InformationURL>
  </ResourceHeader>
  <InstrumentType>WaveformReceiver</InstrumentType>
  <InvestigationName>Thermal Noise Receiver</InvestigationName>
  <ObservatoryID>spase://observatory/example</ObservatoryID>
</Instrument>
</Spase>
```

Unique Identifier

An Association

Natural Order of Resource Creation

Based on potential associations.

1. Person
2. Observatory
3. Instrument
4. Data (NumericalData, DisplayData, Catalog)
5. Granule
6. Annotation

Naming Authorities

Naming Authorities allow for parallel effort without identifier collision.

CSSDP	Canadian Space Science Data Portal (CSSDP). Assigns resource IDs for data registered in the Canadian Space Science Data Portal (CSSDP). http://www.cssdp.ca/
SMWG	SPASE Metadata Working Group (SMWG). The community location for entities such as persons, observatories, instruments, repositories, registries and services. http://www.spase-group.org/registry
VEPO	Virtual Energetic Particle Observatory (VEPO). http://vepo.gsfc.nasa.gov/
ViRBO	Virtual Radiation Belt Observatory (ViRBO). http://vmo.nasa.gov
VITMO	Virtual Ionosphere, Thermosphere, Mesosphere Observatory (VITMO). http://vitmo.jhuapl.edu/
VMO	Virtual Magnetosphere Observatory (VMO). http://vmo.nasa.gov
VSPO	Virtual Space Physics Observatory (VSPO). http://vspo.gsfc.nasa.gov
VSO	Virtual Solar Observatory (VSO). http://sdac.virtualsolar.org/
VWO	Virtual Wave Observatory (VWO). http://vwo.nasa.gov

Available Resources

- SMWG Registry
 - Find Document, Instrument, Observatory, Person, Registry, Services
<http://www.spase-group.org/smwg/explorer/>
- HPDE Registry
 - All resources (data, granule and entities) in one place
<http://www.spase-group.org/registry/explorer/>
- Git Repository
 - Metadata management system (more on this later)
<http://vho.nasa.gov/scm/>

Available Tools

- **Resource Tools:** Collections of tools and applications for working with resource descriptions

[SPASE Resource Tools](#) A set of command-line applications to generate, validate, referentially check, use and organize resource descriptions written in SPASE XML. **Resource ID Maker** Create a SPASE Resource ID with a few form entries.

[Resource ID Maker](#) A simple web page to create a SPASE Resource ID with a few form entries.

- **Validator:** Determines compliance with SPASE data model.

[XML Validate](#) Test a web accessible SPASE description against a selected version of the data dictionary.

- **Parser:** Parse SPASE XML

[Parser](#) The SPASE XML parser is a collection of Java classes which can parse XML descriptions and load the information into a directly accessible form. Source code and class files: [parser.jar](#)

- **Registry:** Services to delivery SPASE metadata.

[SPASE Registry Server \(Try it!\)](#) The SPASE Registry Server is a java application which can harvest resource descriptions expressed in SPASE XML and provide a search service for this descriptions. It can also chain to other registry servers and aggregate all results, returning all matches in a self-organized network of registry servers. It can be run as either a servlet or bean. It uses the [SPASE XML Parser](#) package to harvest resource descriptions.

- **Editor Web-based Editors:**

[Web Editor \(Try it!\)](#) The web SPASE XML editor can be used to create or alter SPASE XML descriptions.

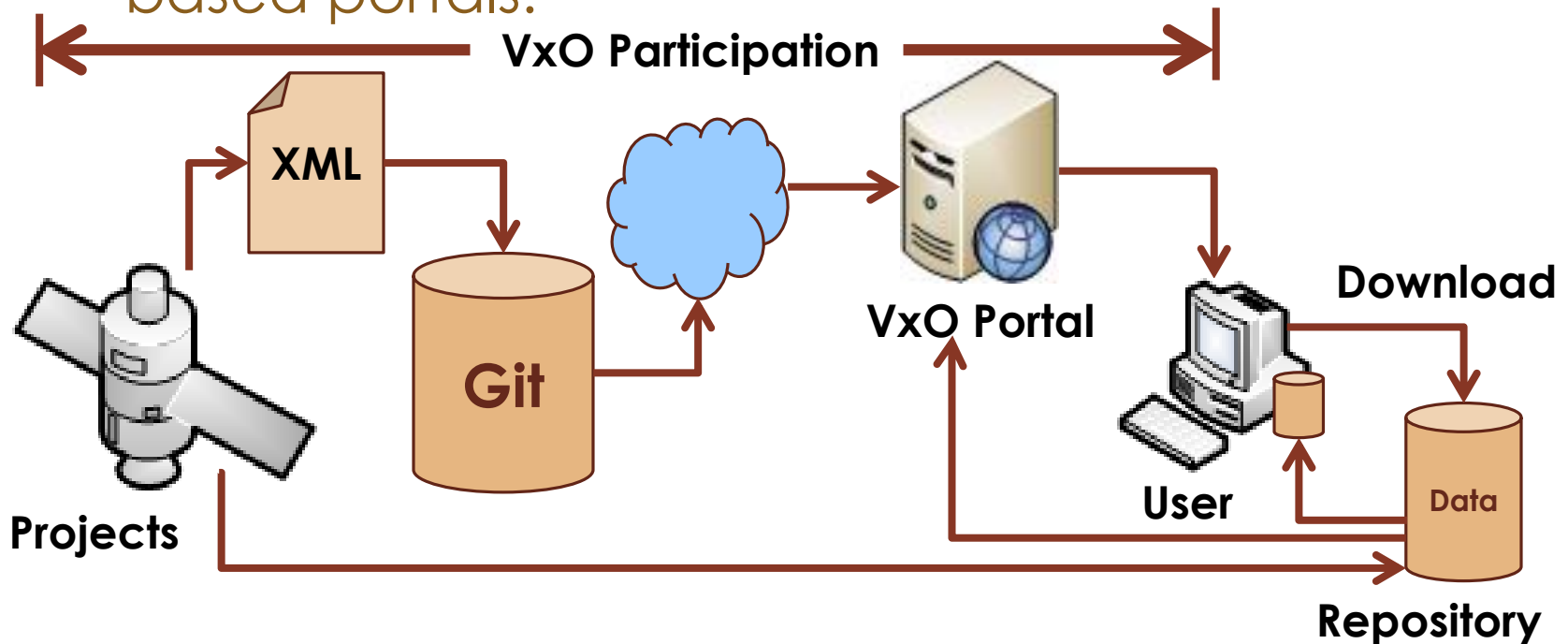
Managing Metadata

- Metadata is expressed in XML.
- Each resource (object) has a corresponding XML file.
- Git is used for managing the files and provides:
 - Revision control (versioning)
 - Distributed updates (clone/commit/push)
 - Open sharing (clone/pull)

**The HPDE has
registered
3,557 collections and
1,259,753 Granules.**

Virtual Observatories

- A source of expertise within a community
 - To aid missions in generating metadata.
 - To aid researchers in locating data.
- Use the metadata to provide community based portals.



Step 1: Initial search.

Enter Some Keywords

Enter Time Range

Home Data Services Documents Tools News Links Help

Find: Enter Keywords Here Time from 0000-00-00 00:00:00 to 2012-07-11 23:59:59 Search Per Page: 50

Current facet:

- Find Data
 - Spacecraft/Mission (33)
 - Ground Based Data (17)
 - Measurement Type (13)
 - Observed Region (4)
 - Instrument Type (24)
- Services
 - Autoplot
 - MANGO
 - SPASE
 - SPASE Data Model Explorer
 - SPLASH
- Quick Links
 - FAQ 7
 - Citation Policy
 - Help for Data Users
 - Help for Data Providers
 - Inventory

Recently Released Data Products

Data Product	Release Date
Spherical Elementary Current (SEC) Amplitudes derived using the Spherical Elementary Currents Systems (SECS) technique at 10 sec Resolution in Geographic Coordinates spase://VMO/NumericalData/SECS/Magnetometer/SECS/PT10S	2012-05-23
Equivalent Ionospheric Currents (EICs) derived using the Spherical Elementary Currents Systems (SECS) technique at 10 sec Resolution in Geographic Coordinates spase://VMO/NumericalData/SECS/Magnetometer/EICS/PT10S	2012-05-23
IBEX-Lo Release-3 Counts for flow longitude analysis spase://VEPO/NumericalData/IBEX/Lo/Release_3/Counts	2012-05-08
IBEX-Lo Release-3 Counts for flow latitude analysis spase://VEPO/NumericalData/IBEX/Lo/Release_3/Angular	2012-05-08
IMAGE FUV SI plots of O/N2 Ratios spase://VEPO/DisplayData/IMAGE/FUV/SI/O_N2_Ratios	2012-04-18

[More...](#)

Welcome to the VMO

Step 2: Refine

Firefox

VMO - Search

vmo.igpp.ucla.edu/search/

VIRTUAL MAGNETOSPHERIC OBSERVATORY SPASE inside

Home Data Services Documents Tools News Links Help Contact

Find: propagated solar wind Time from: 2009-02-27 00:00:00 to 2009-02-27 23:59:59 Search Per Page: 50

Current facet: (measurementtype:[" TO *"])

Find Data

- Spacecraft/Mission (33)
- Ground Based Data (17)
- Measurement Type (13)
 - Activity Index
 - Electric Field
 - Energetic Particles
 - Ephemeris
 - Image Intensity
 - Instrument Status
 - Ion Composition
 - Impedance
 - Magnetic Field**
 - Neutral Atom Images
 - Spectrum
 - Thermal Plasma
 - Waves (2)
- Observed Region (4)
- Instrument Type (24)

Search for Available Resources

To begin a search enter keywords in the "Find" box above, then click "Search"

Search results 42 matches in 0.002 seconds

Showing 1 - 42

1) [Ulysses/COSPIN/HET readout by readout counts and accumulation times for spin-averaged counting rates H1-H9.](#)

Resource ID:spase://VMO/NumericalData/Ulysses/COSPIN/HET/FullResolution/Rates/Omni1

Start: Ulysses Cadence: High Energy Telescope Resource: NumericalData

2) [Ulysses/HISCALE/PHA/G12](#)

Resource ID:spase://VMO/NumericalData/Ulysses/HISCALE/PHA/G12

Start: 1976-11-30 00:00:00 Observatory: Ulysses Cadence: 1 day
Stop: 2009-06-30 23:59:59 Instrument: Heliosphere Instrument for Spectra, Composition and Anisotropy at Low Energies Resource: NumericalData

3) [ACE Linearly Interpolated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSE Coordinates](#)

Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Processed/GSE/PT60S

Start: 1997-09-01 00:00:00 Observatory: ACE Cadence: 60 seconds
Stop: ACE Magnetic Field Instrument Resource: NumericalData

http://vmo.igpp.ucla.edu/search/search....facet=(measurementtype:(ActivityIndex))

Step 3: Inspect

Firefox

VMO - Search

vmo.igpp.ucla.edu/search/

VIRTUAL MAGNETOSPHERIC OBSERVATORY SPASE inside

Home Data Services Documents Tools News Links Help Contact

Find: Time from: to Per Page:

Current facet: (measurementtype:(MagneticField))

Find Data

- Spacecraft/Mission (33)
- Ground Based Data (17)
- Measurement Type (13)
 - Activity Index
 - Electric Field
 - Energetic Particles
 - Ephemeris
 - Image Intensity
 - Instrument Status
 - Ion Composition
 - Irradiance
 - Magnetic Field
 - Neutral Atom Images
 - Spectrum
 - Thermal Plasma
 - Waves (2)
- Observed Region (4)
- Instrument Type (24)

Services

Search for Available Resources
To begin a search enter keywords in the "Find" box above, then click "Search"

Search results 18 matches in 0.004 seconds
Showing 1 - 18

1) ACE Linearly Interpolated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSE Coordinates

Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Processed/GSE/PT60S

Start: 1997-09-01 00:00:00	Observatory: ACE	Cadence: 60 seconds
Stop: 2010-01-31 23:59:00	Instrument: ACE Magnetic Field Instrument	Resource: NumericalData

2) ACE Linearly Interpolated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSM Coordinates

Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Processed/GSM/PT60S

Start: 1997-09-01 00:00:00	Observatory: ACE	Cadence: 60 seconds
Stop: 2010-01-31 23:59:00	Instrument: ACE Magnetic Field Instrument	Resource: NumericalData

3) ACE Weimer Propagated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSE Coordinates

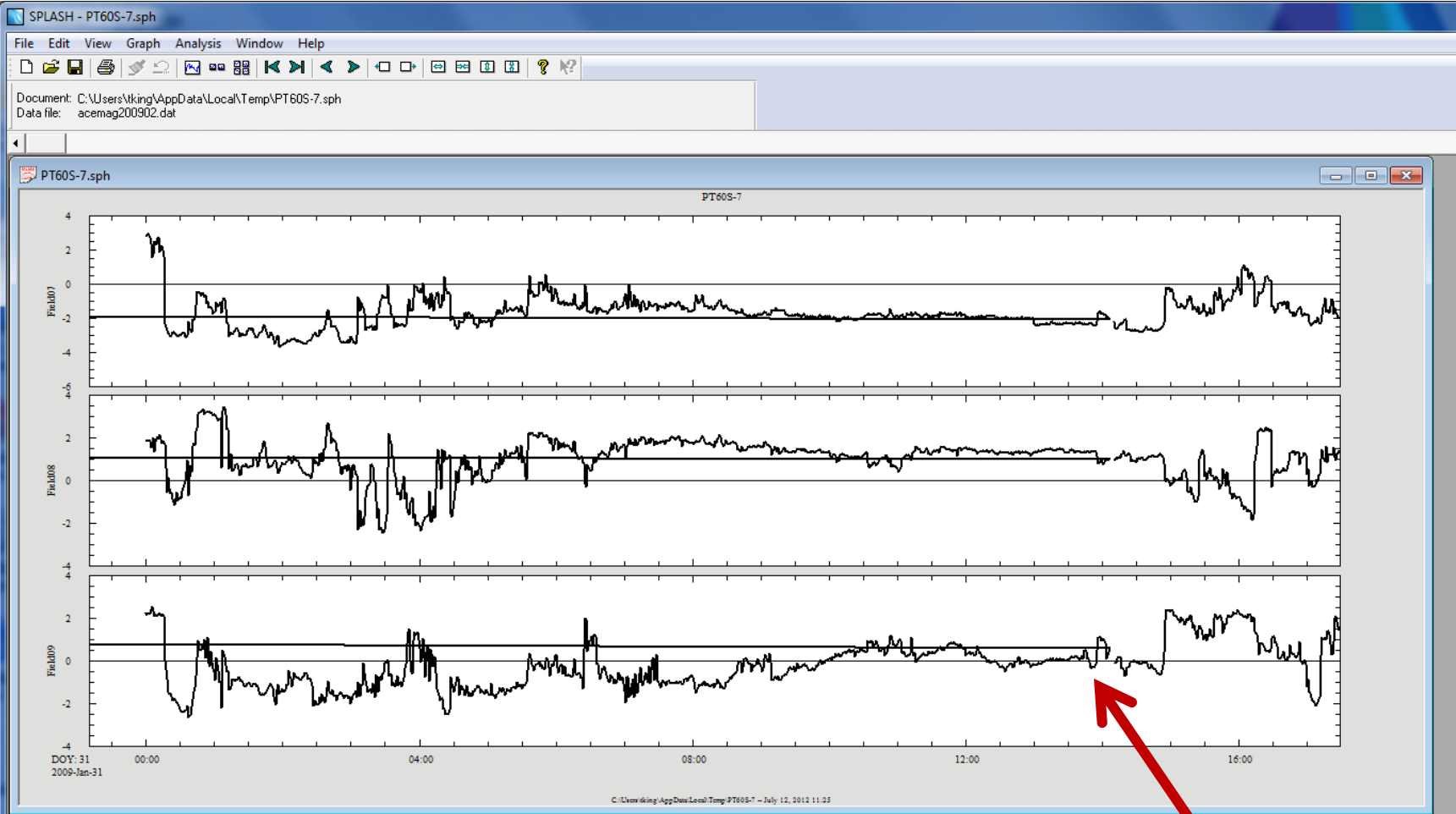
Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Propagated.SWEPAM/GSE/PT60S

Start: 1998-02-01 00:00:00	Observatory: ACE	Cadence: 60 seconds
Stop: 2010-01-31 23:59:00	Instrument: ACE Magnetic Field Instrument	Resource: NumericalData

4) ACE Weimer Propagated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSM Coordinates

Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Propagated.SWEPAM/GSM/PT60S

Start: 1998-02-01 00:00:00	Observatory: ACE	Cadence: 60 seconds
----------------------------	----------------------------------	---------------------



Looks promising!

Step 4: Download

Firefox

VMO - Search

vmo.igpp.ucla.edu/search/

NASA

VIRTUAL MAGNETOSPHERIC OBSERVATORY

SPASE inside

Home Data Services Documents Tools News Links Help Contact

Find: propagated solar wind Time from: 2009-02-27 00:00:00 to 2009-02-27 23:59:59 Search Per Page: 50

Current facet: (measurementtype:(MagneticField))

Find Data

- Spacecraft/Mission (33)
- Ground Based Data (17)
- Measurement Type (13)
 - Activity Index
 - Electric Field
 - Energetic Particles
 - Ephemeris
 - Image Intensity
 - Instrument Status
 - Ion Composition
 - Irradiance
 - Magnetic Field
 - Neutral Atom Images
 - Spectrum
 - Thermal Plasma
 - Waves (2)
- Observed Region (4)
- Instrument Type (24)

Services

Search for Available Resources

To begin a search enter keywords and click "Search"

Search results 18 matches in 0.004 seconds

Showing 1 - 18

1) ACE Linearly Interpolated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSE Coordinates

Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Processed/GSE/PT60S

Start: 1997-09-01 00:00:00	Observatory: ACE	Cadence: 60 seconds
Stop: 2010-01-31 23:59:00	Instrument: ACE Magnetic Field Instrument	Resource: NumericalData

2) ACE Linearly Interpolated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSM Coordinates

Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Processed/GSM/PT60S

Start: 1997-09-01 00:00:00	Observatory: ACE	Cadence: 60 seconds
Stop: 2010-01-31 23:59:00	Instrument: ACE Magnetic Field Instrument	Resource: NumericalData

3) ACE Weimer Propagated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSE Coordinates

Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Propagated.SWEPAM/GSE/PT60S

Start: 1998-02-01 00:00:00	Observatory: ACE	Cadence: 60 seconds
Stop: 2010-01-31 23:59:00	Instrument: ACE Magnetic Field Instrument	Resource: NumericalData

4) ACE Weimer Propagated 60 s Resolution Tri-axial Fluxgate Magnetometer in GSM Coordinates

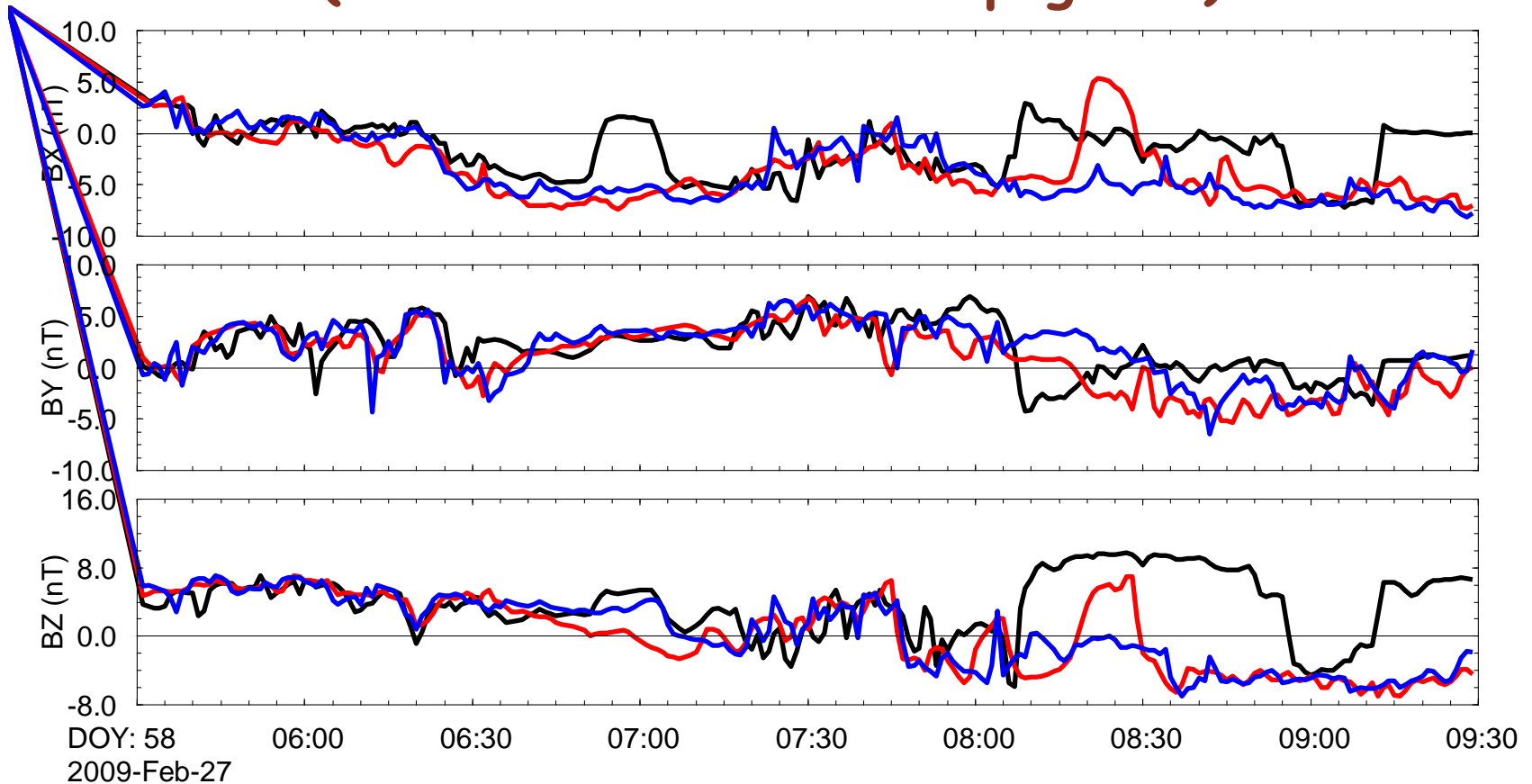
Resource ID:spase://VMO/NumericalData/Weygand/ACE/MFI/Propagated.SWEPAM/GSM/PT60S

Repeat and

... science

Step 5: Analyze and publish

IMF on February 27, 2009 (GSM Coordinates - Propagated)



C:\Research Projects\Causes\Feb 29 2009\ACE Wind Geotail IMF GSM Propagated -- July 05, 2011 21:40

ACE (GSM) 240.0, 29.71, 6.60

Geotail (GSM) 19.43, -14.7, -15.6

Wind (GSM) 197.5, -10.6, -30.3

Conclusions

SPASE metadata is:

- Designed to describe something once and migrate as needed.
- Can be used by a project internally and to openly share data.
- Enables connecting new mission data to historical through a standard common vocabulary.

And...

- Existing services and tools make the task of creating and using SPASE metadata easy.

תודה
Dankie Gracias
Спасибо شكراً
Merci Takk
Köszönjük Terima kasih
Grazie Dziękujemy Děkojame
Ďakujeme Vielen Dank Paldies
Kiitos Täname teid 谢谢
Thank You Tak
感謝您 Obrigado Teşekkür Ederiz
Σας Ευχαριστούμ 감사합니다
ບອບດຸນ
Bedankt Děkujeme vám
ありがとうございます
Tack