What is SPASE?

The Space Physics Archive Search and Extract (SPASE) effort is a heliophysics community based project with the goals of:

- Facilitating data search and retrieval across the Space and Solar Physics data environment with a common metadata language or interlingua
- Defining and maintaining a standard Data Model for Space and Solar Physics interoperability
- Demonstrating the Model’s viability
- Providing tools and services to assist SPASE users
- Working with other groups for other heliophysics data management and services coordination as needed

The Space Physics Archive Search and Extract (SPASE) effort is implemented by the SPASE Working Group composed of representatives of the heliophysics data community

The SPASE Working Group is the only international group supporting global data management for Solar and Space Physics
Heliophysics Virtual Observatories (VOs)

**NASA-Funded**
- VSO - Virtual Solar Observatory
- VSPO - Virtual Space Physics Observatory
- VMO - Virtual Magnetospheric Observatory
- VITMO - Virtual Ionosphere, Thermosphere, Mesosphere Observatory
- VHO - Virtual Heliophysics Observatory
- ViRBO - Virtual Radiation Belt Observatory
- VEPO - Virtual Energetic Particle Observatory
- VWO - Virtual Wave Observatory
- VMR - Virtual Model Repository

**Non-NASA-Funded**
- CAA - Cluster Active Archive
- CDPP - Centre de Données de la Physique des Plasmas
- CSSDP - Canadian Space Science Data Portal
- EGSO - European Grid of Solar Observations
- GAIA - Global Auroral Imaging Access
- VSTO - Virtual Solar Terrestrial Observatory
- ??
- ??
NSSDC Role in SPASE

• Leading the SPASE effort (biweekly teleconferences, periodic face-to-face meetings, presentations at appropriate conferences, etc.)

• Providing the website for the project (http://spase.gsfc.nasa.gov)

• Providing SPASE data set descriptions for key data sets

• Enabling the creation of the SPASE Registry of spacecraft, instrument, and personnel information
THE WEBSITE AND THE SPASE DATA MODEL

Version 1.4 of the Data Model is imminent and will be “frozen” for community usage.
6. Examples

As an example let us describe a person using SPASE metadata. This person is "John Smith" from Smith Foundation. While the SPASE data model is implementation neutral, XML representation is preferred. This example uses the SPASE XML form.

```xml
<spase>
  <person>
    <resourceID>spase://person/jsmith@smith.org</resourceID>
    <name>John Smith</name>
    <organization>Smith Foundation</organization>
    <address>1 Main St., Smithville, MA</address>
    <email>jsmith@smith.org</email>
    <phoneNumber>1-800-555-1212</phoneNumber>
  </person>
</spase>
```

4. The Data Model presented hierarchically

The taxonomy tree shows the inter-relationship of a "picture" view of the SPASE data model. This taxonomy element are contained in the data dictionary.

Notes: Occurrence specifications are enclosed in parentheses, e.g., +"..". For more, + = 1 or more
The SPASE Registry is a service for providing spacecraft, instrument, and personnel information for use in data set descriptions. The information comes from the NSSDC Master Catalog and Personnel database.

Present

- Next release will support:
  - Forms interface
  - Ability to display results in raw XML or in XHTML form
  - Specification of SPASE version
  - Choice of NSSDC or SMWG IDs for resource IDs
- Completed all but last item which is in progress
- Release expected to be late March 2009

Future

- Configuration of "git" repository on NSSDC computer
- Setting up "git" repository to synchronize with VMO repository
- Creation of software to update NIMS database from "git" repository contents for observatories, instruments, people
- Creation of software to update "git" repository contents based on updates to NIMS database
Space and Solar Physics
Data Environment

DATA ENVIRONMENT

VxQ's

VSO
VSPO
VMO
VITMO

Archive1
Archive2
Archive3
Archive4
Search Mechanism Across Virtual Observatories

• Many Virtual Observatories in Heliophysics, both new and old - often quite different from each other
• Many important Heliophysics Data Centers are not yet directly connected with a Virtual Observatory

To make data “findable” in the Heliophysics Data Environment a common metadata language and descriptions of data sets in the metadata language are important (SPASE)

• The Space Physics Archive Search and Extract (SPASE) project is an international collaboration begun through the NASA-funded Heliophysics Virtual Observatories initiative
SPASE-Based Searching

Searches across VOs based on the SPASE Data Model can be done by several methods:

- Any of the VOs can harvest and store all SPASE Data Descriptions and provide a cross-disciplinary search through their search interface.

- Registries can collect SPASE Resource Types and support searches
  - SPASE Resource Types: Catalog, Display Data, Numerical Data, Granule, Instrument, Observatory, Person, Registry, Repository, Service

- SPASE Query Language can be used to search generically across the VOs, assuming use of or mapping of SPASE metadata
Application Tools

Tools for working with SPASE metadata and the SPASE framework.

**Validator**
- Determines compliance with a version of the SPASE data model.
  - XML Validate

**Parser**
- Convert SPASE XML to internal structures
  - Parser

**Editor**
- Web-based Editors
  - Web Editor
- Standalone Editors
  - SPASE Assistant
- Editors with Database Storage
  - Web+DB Editor

**Generator**
- Create SPASE descriptions using external sources of information
  - Ruleset Description Generator

**Harvester**
- Extracts information from SPASE resource descriptions (or registries)
  - SPASE Registry Server
  - SPASE Database Query

**Wrapper**
- Converts or embeds SPASE metadata in other descriptions or forms (i.e., OAI)
  - Data Dictionary Lookup
  - SPASE-to-OAI mapping

**Correlator**
- Divide an XML document into individual resource descriptions into a well organized file system
  - Correlator

**Refcheck**
- Determine the validity of all references in a resource descriptions. Checks Resource IDs and URL
  - Refcheck

**There are additional tools in development:**
- SPASE Query Language
- Java-to-XML Binding Mechanism (JAXB)
- SPASE Guidelines Document
• How well does SPASE function as an *interlingua* among the Virtual Observatories and data archives?
• How effective is SPASE in describing data sets for data finding and usage?
• How much should SPASE be "inside" vs. "outside" the observatories, etc. to be effective?
• To what level of *detail* should data descriptions be created in order to fulfill the objectives of SPASE and the Space Physics data environment?
• What else is needed in SPASE for the non-NASA virtual observatories environment?
• Should SPASE data descriptions be centrally stored or distributed?
• What *interfaces* are most effective for SPASE searches?
• Should SPASE be expanded for usage in closely-related disciplines such as Planetary Science?