HELIOPHYSICS

Heliophysics Virtual Observatories, Registries, Data Model, and Possible IVOA Role: A Largely NASA perspective

D. Aaron Roberts, NASA GSFC

IVOA Interoperability Meeting

17 May 2010

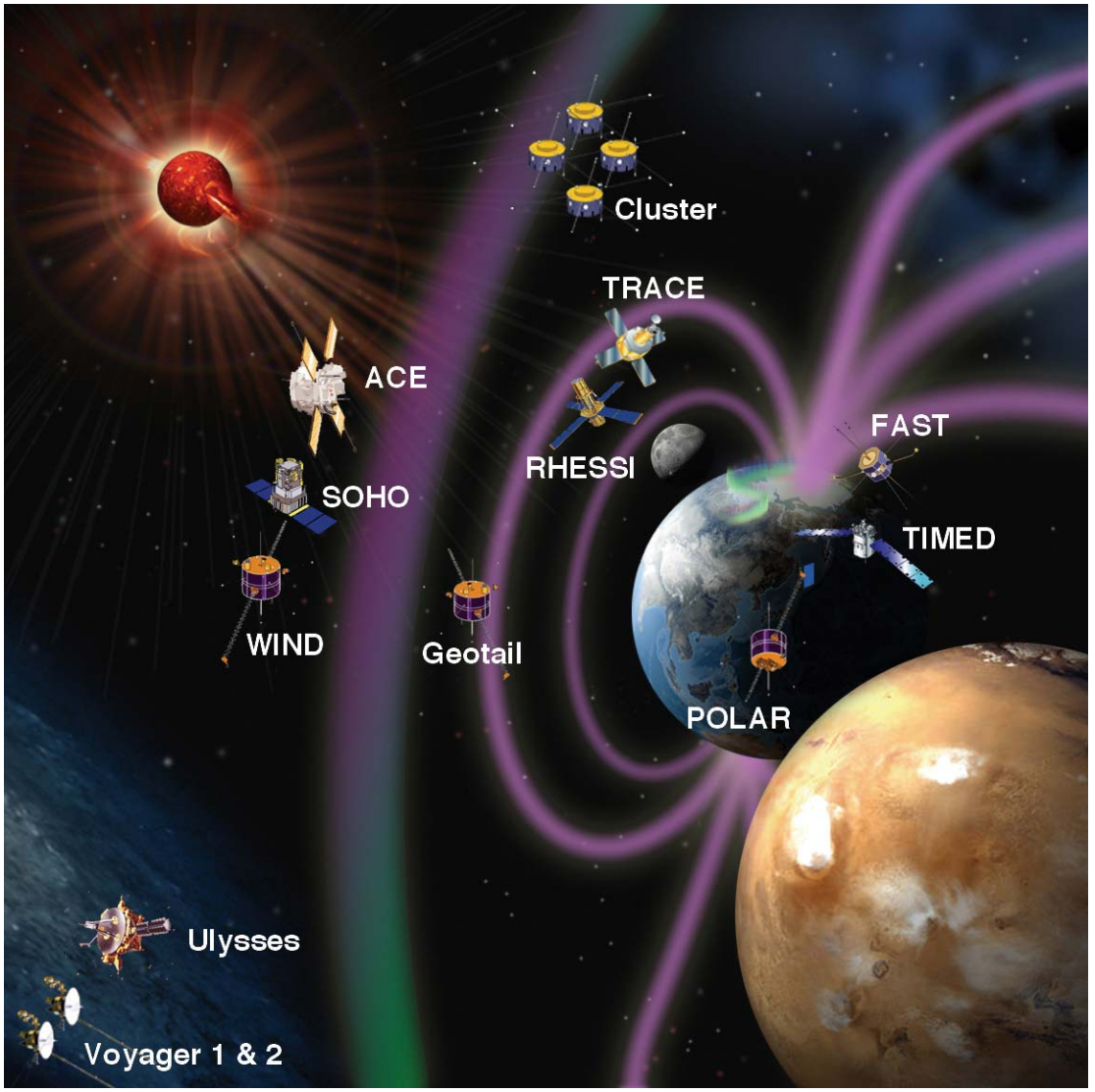
What is Heliophysics?

- **The study of the system composed of the Sun and its heliosphere and the the objects that interact with it, e.g. planetary atmospheres and magnetospheres, and the interstellar medium. ("In situ astrophysics")**
- **Combines the disciplines of space physics and solar physics.**
- **Largely concerned with matter in a plasma state.**
- **Closely tied to the study of space weather (how solar particles and fields affect human endeavors).**

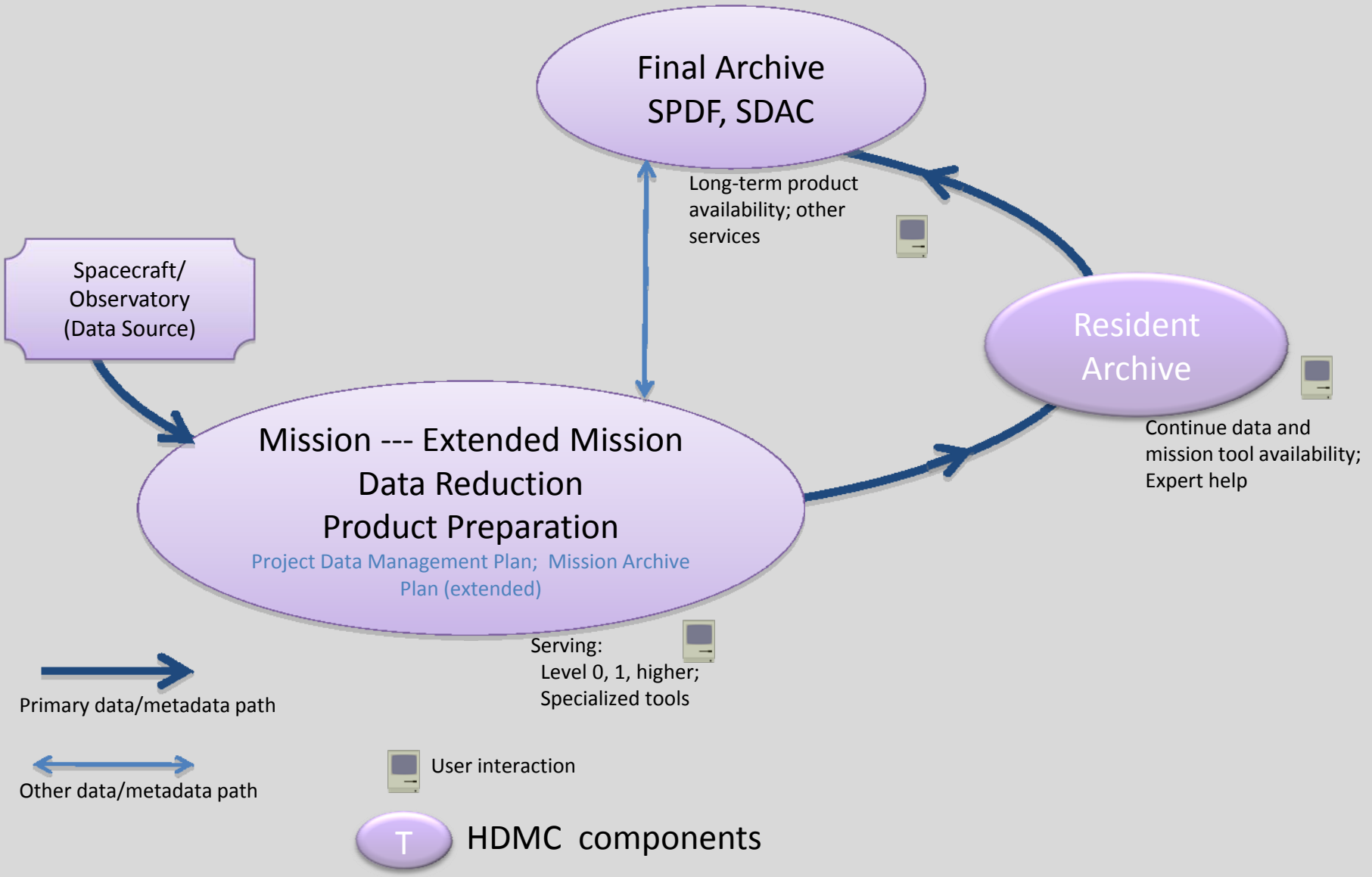
The Heliophysics System Observatory

E pluribus unum:

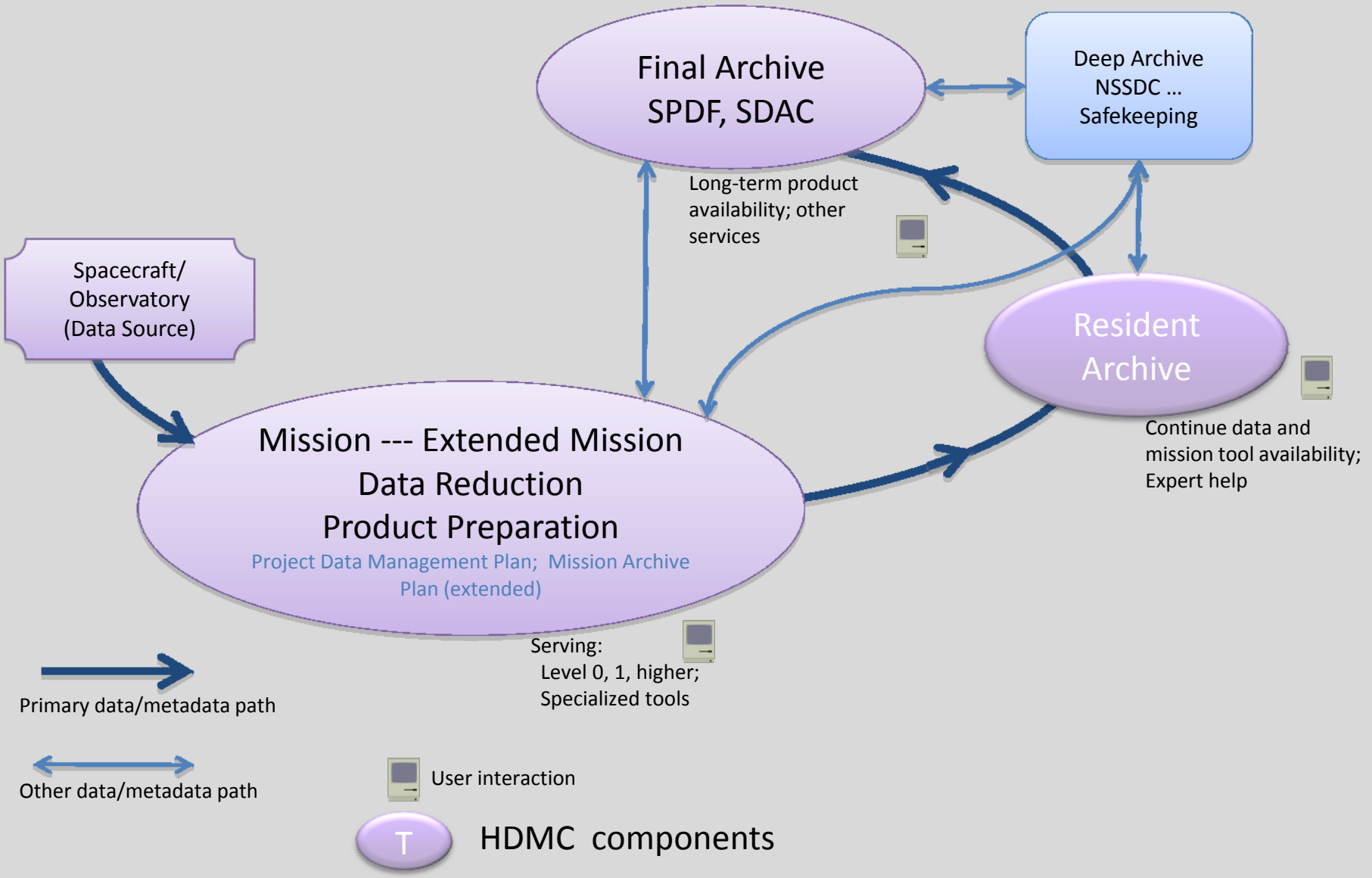
A Goal of the Our Data Environment: A Heliophysics Virtual Observatory



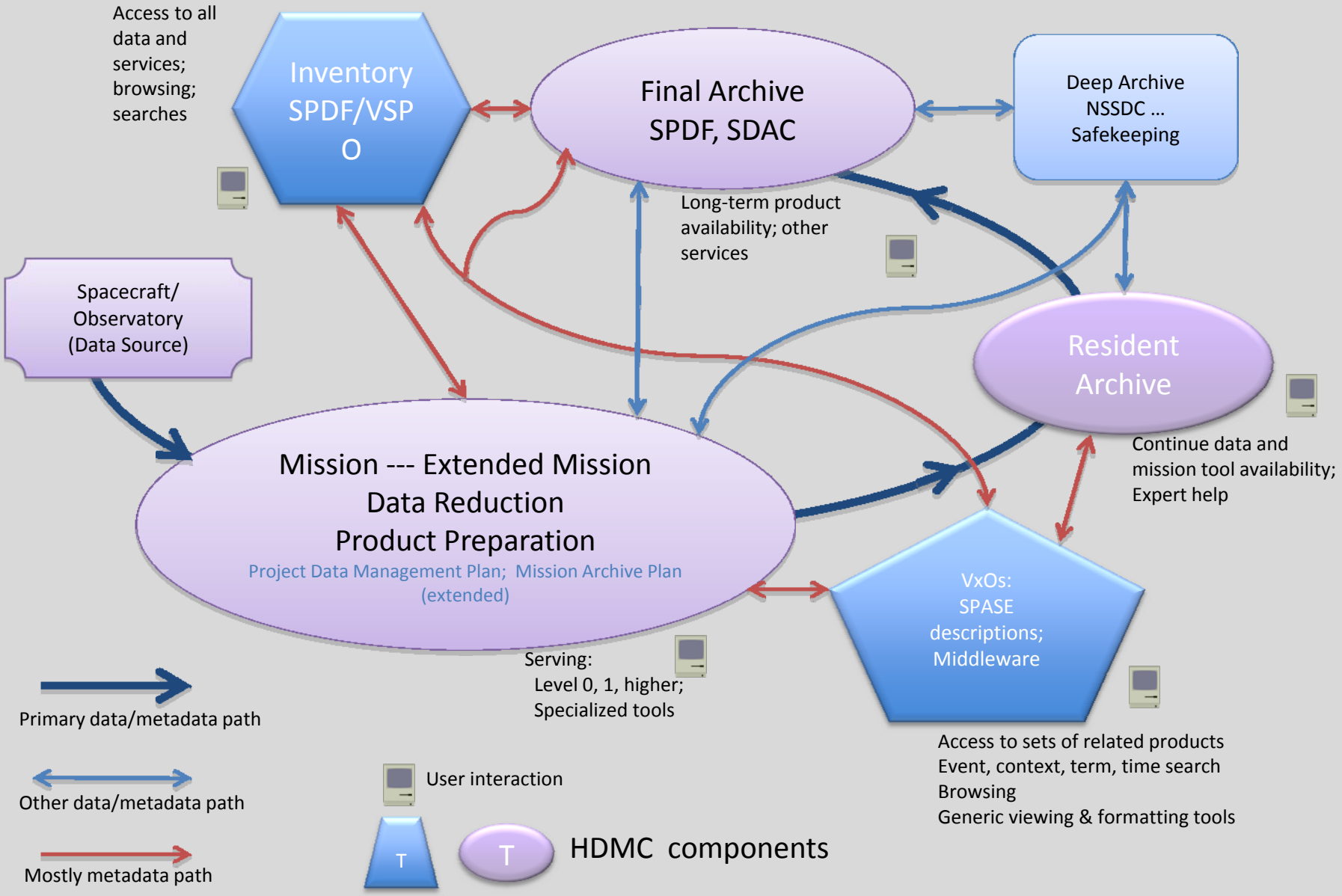
Heliophysics Data Environment – Main Data Path



Heliophysics Data Environment – Main Path + Backup



Heliophysics Data Environment – with Full Access



What is SPASE?

- Acronym for:
  Spase Physics Archive Search and Extract
- An organization to set community-based standards with the goals of:
  - Defining a **Data Model** (“ontology”) for Space Physics
  - Demonstrating its viability
  - Enabling interoperability in a federated environment

So that…

- Resources to be easily registered, found, accessed, and used

SPASE Today

- Data Model
  - Defined a standard data model for Heliophysics
    Current release version 2.1.0 (April 2010)
  - Vetted by research communities in many domains:
    - Magnetospheres
    - Waves
    - Ionosphere-Thermosphere-Mesosphere
    - Radiation Belts
    - Energetic Particles
    - Solar Physics
    - Models and Simulations
- Services
  - Initial work on metadata sharing (registries), including a general inventory and access to it ("VSPO").
  - Distributed queries (SPASE-QL)

SPASE Conceptual Model

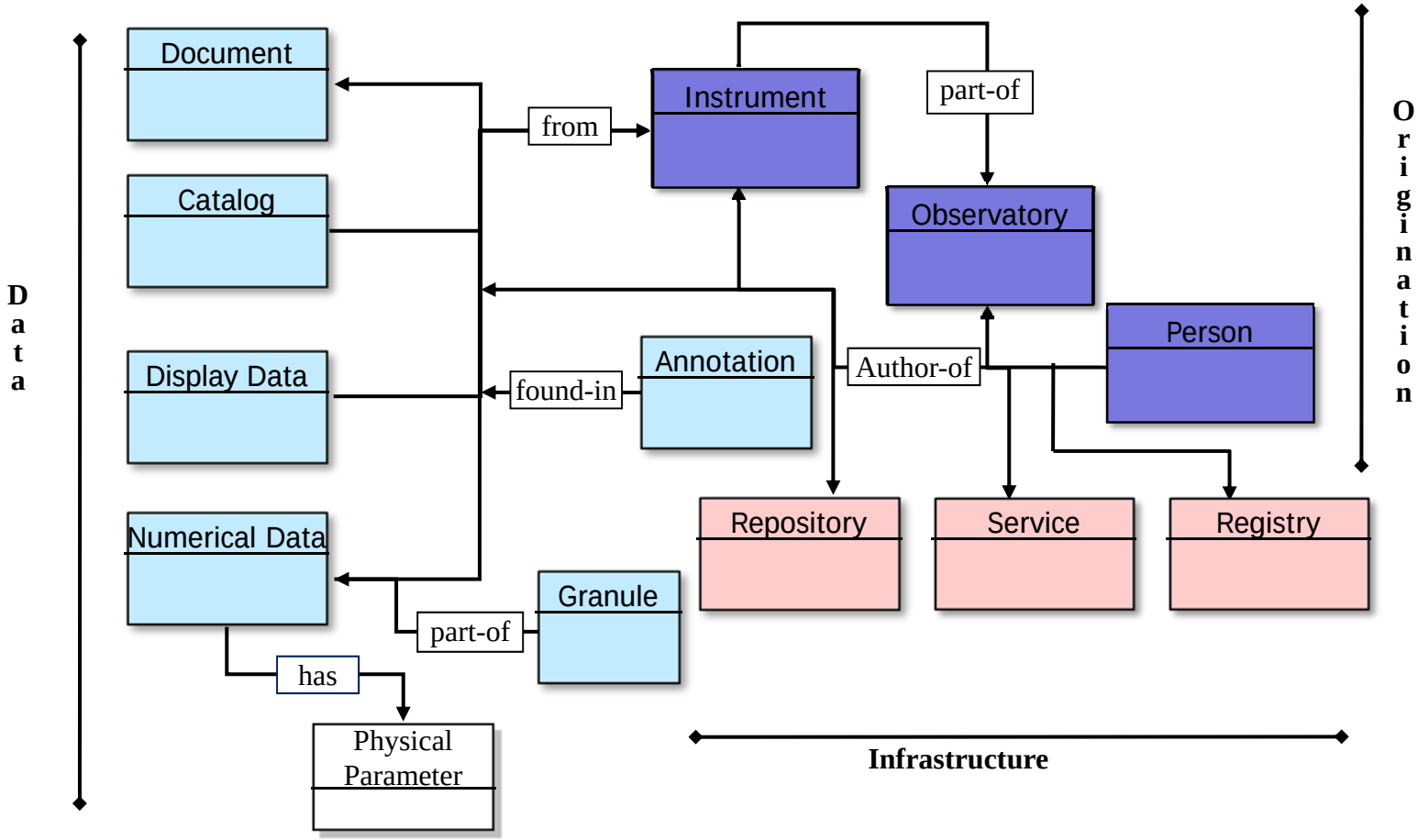


Figure 1. A simplified conceptual model of resources (classes) in the SPASE data model. Arrows point in the direction of association. Cardinality is not shown.

SPASE Speak

- The SPASE Data Model was developed by a team of scientists, IT specialists, data engineers and developers.
- Most documentation is oriented towards the user, so our nomenclature is less formal.

What we Say	In UML
Resource	Class
Element	Attribute
Container	Component Class
Association	Association

SPASE and ISO-11179

SPASE adheres to the principals of ISO-11179 (Metadata Registry) standard

- without formally adopting all aspects.

SPASE has a

- Data element concept (DEC)
- Representation
  - Value Domain
    - Enumerated
    - Non-enumerated
    - Value meaning
- Data Element Relationships
- Data Element Formulation Rules
- Naming and identification principles
- Maintains a simple metadata registry

Additional Details

- The SPASE Data Model is a Semantic Data Model.
  - Defines the meaning of data within the context of its interrelationships with other data.
  - Defines the scientific context of data.
- …with Ontology features
  - Typed associations between resources.
- The SPASE Data Model is implementation neutral.
- Chosen reference implementation is XML.
  - XML Schema
  - Numerous XML style sheets for converting metadata.
    - To HTML
    - To OAI
  - And lots of tools.

SPASE Tools

- To demonstrate the viability of the model and provide basic support for its adoption a set of tools have been developed to support the reference implementation (XML) :
  - **Validator** - Determines compliance with a version of the SPASE data model.
  - **Parser** – Convert SPASE XML to internal structures.
  - **Editor** – Create SPASE descriptions by hand.
  - **Generator** – Creates SPASE descriptions using external sources of information.
  - **Harvester** – Extracts information from SPASE resource descriptions (or registries).
  - **Wrapper** – Converts or embeds SPASE metadata in other descriptions or forms (i.e., OAI).
  - **Correlator** – Divide an XML document into individual resource descriptions into a well organized file system.
  - **Refcheck** – Determine the validity of all references in a resource descriptions. Checks Resource IDs and URL.

Example Person Resource (XML)

```
<?xml version="1.0" ?>
  <Spase xmlns="http://www.spase-group.org/data/schema">
  <Version>2.0.0</Version>
  <Person>
    <ResourceID>spase://SMWG/Person/Todd.King</ResourceID>
    <ReleaseDate>2007-06-07</ReleaseDate>
    <PersonName>Todd King</PersonName>
    <OrganizationName>UCLA/IGPP</OrganizationName>
    <Address>3846 Slichter Hall
             Los Angeles, CA
             90095-1567
    </Address>
  </Person>
  </Spase>
```

Clarity comes from usage

Current Users (to different degrees):

- United States
  - NASA's Heliophysics Data Environment (HPDE)
    - 9 Virtual Observatories (some more, some less)
    - Service providers (Autoplot, HELM, etc.)
  - National Science Foundation (NSF)
    - SuperMAG project
- Canada
  - Canadian Space Science Data Portal (CSSDP)
- European Union
  - Cluster Active Archive (CAA)
  - HELIO
- Beginnings of use in Japan (JAXA).

SPASE Members

Augsburg College http://www.augsburg.edu/home/physics/spacephysics.htm
Canadian Space Science Data Portal http://www.cssdp.ca/
CNES/CNRS Plasma Physics (CDPP) Data Archive http://www.cnes.fr/
European Grid of Solar Observations (EGSO) http://www.egso.org/
Inter-university Upper atmosphere Global Observation NETwork (IUGONET) http://www.iugonet.org/
Institute of Geophysics and Planetary Physics – UCLA http://www.igpp.ucla.edu
Institute of Space and Astronautical Science (ISAS/JAXA) http://www.isas.ac.jp/e/index.shtml
Jet Propulsion Laboratory http://www.jpl.nasa.gov/
John Hopkins University - Applied Physics Laboratory http://www.jhuapl.edu/
NASA/Goddard Space Flight Center http://hpde.gsfc.nasa.gov/
NOAA/National Geophysical Data Center http://www.ngdc.noaa.gov/
Planetary Data System- Plasma Physics Interactions Node http://ppi.pds.nasa.gov/
Rutherford Appleton Laboratory http://www.scitech.ac.uk/
Southwest Research Institute http://www.swri.org/

Table 1. Institutions with members who are participating in the SPASE Group

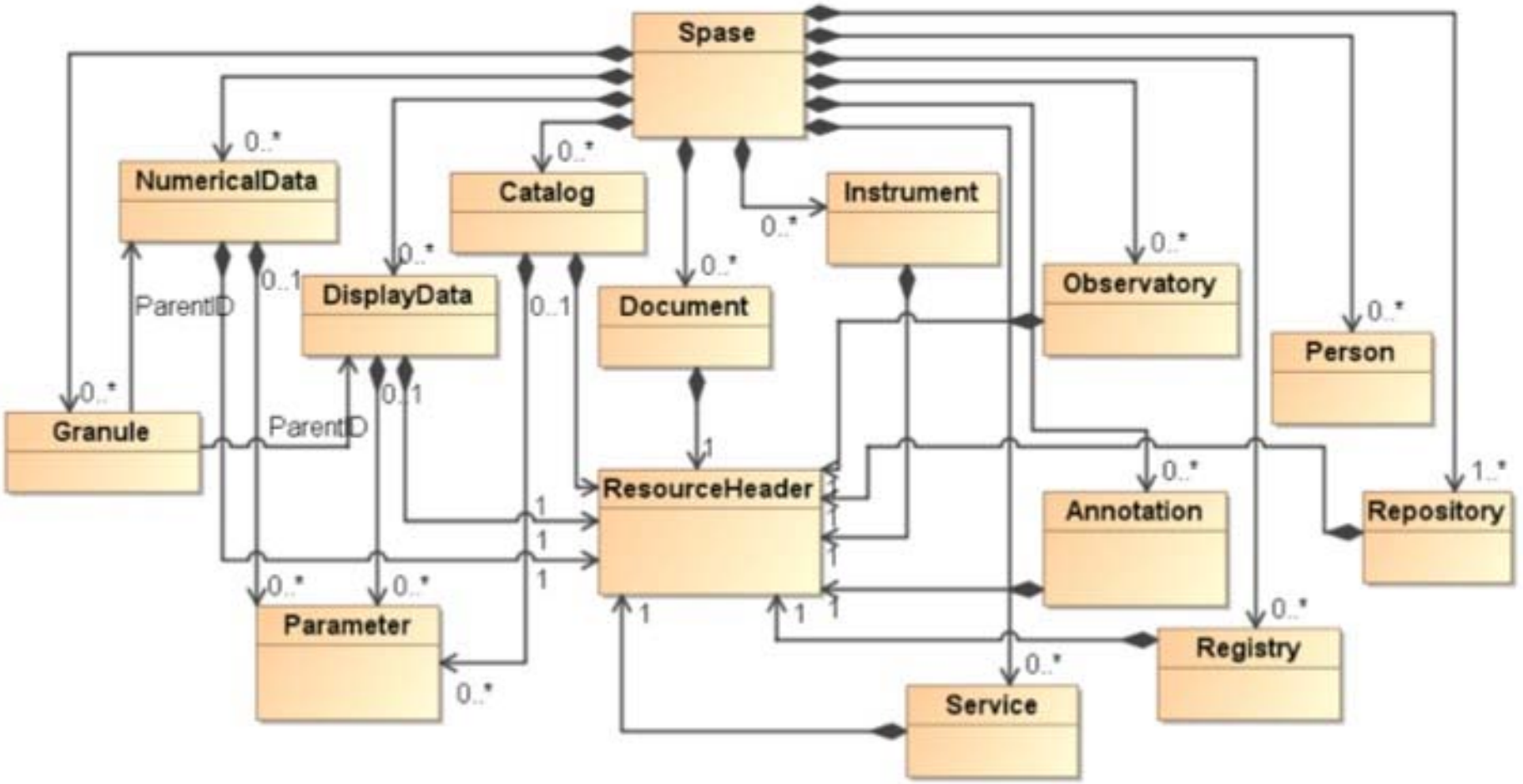
The Future – Data Model

- Provide improved support and services:
  - Data Model version migration
  - Improved editors
  - Improved style sheets
  - Registries
  - …
- Documentation:
  - Tutorials
  - Guides
  - References

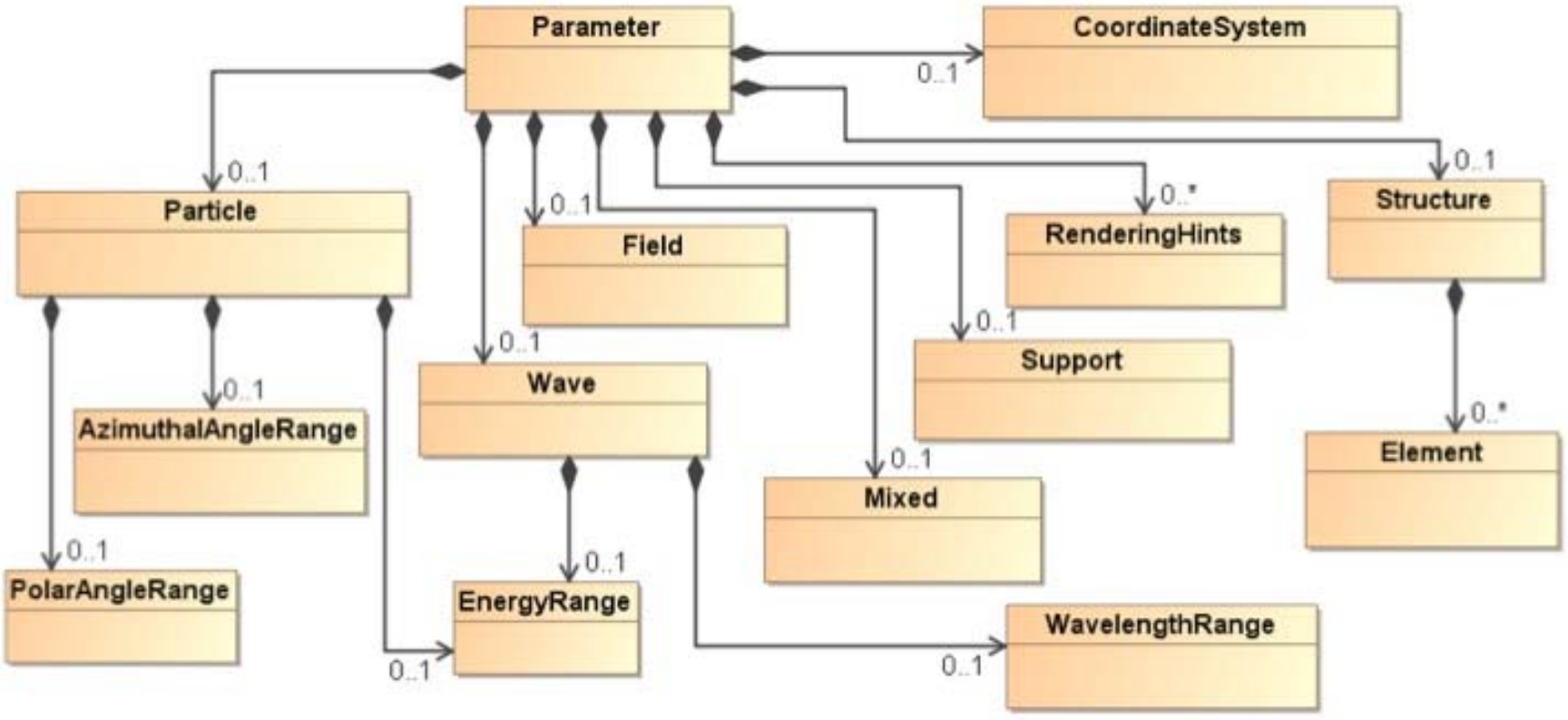
The Future – Interfaces

- Services
  - Resource Registry
  - Reporting
  - Visualization (“Autoplot”); how deeply based on SPASE?
  - Domain search engine; various exist at significantly different levels of “SPASE compliance”
- Query API:
  - SPASE-QL
  - REST; application-based access to resources, e.g., using IDL.

SPASE Data Model



SPASE Data Model: Parameter Class Diagram



Possible IVOA Connections

- Actual scientific overlap
  - Stellar/solar winds; sources, interaction with interstellar medium
  - Magnetospheres
  - Planets
  - Plasma physics problems
- Metadata standards
  - SPASE “header” information (people, observatories, etc.) “VOResource” (sort of)
  - VOTables
  - VOEvents
  - Coordinate systems
  - Other?
- Would it be better to just learn the other system when needed? What are the tradeoffs?

Some useful links

- Heliophysics Data Environment:
  - http://hpde.gsfc.nasa.gov
- SPASE
  - http://www.spase-group.org/
- Virtual Space Physics Observatory ("Heliophysics Resource Gateway")
  - http://vspo.gsfc.nasa.gov