Extending the Space Science Data Environment through Semantics and Knowledge Representation

Tom Narock – UMBC, NASA/GSFC
Adam Szabo – NASA/GSFC
Jan Merka – UMBC, NASA/GSFC

What is a Virtual Observatory?

- The next phase of NASA's Heliophysics Data Environment

- A suite of software applications on a set of computers that allows users to uniformly find, access, and use resources (data, software, document, and image products and services using these) from a collection of distributed product repositories and service providers.

  - *D. Aaron Roberts*: presented October, 2004, Greenbelt, MD.

Current Situation – NASA VxOs

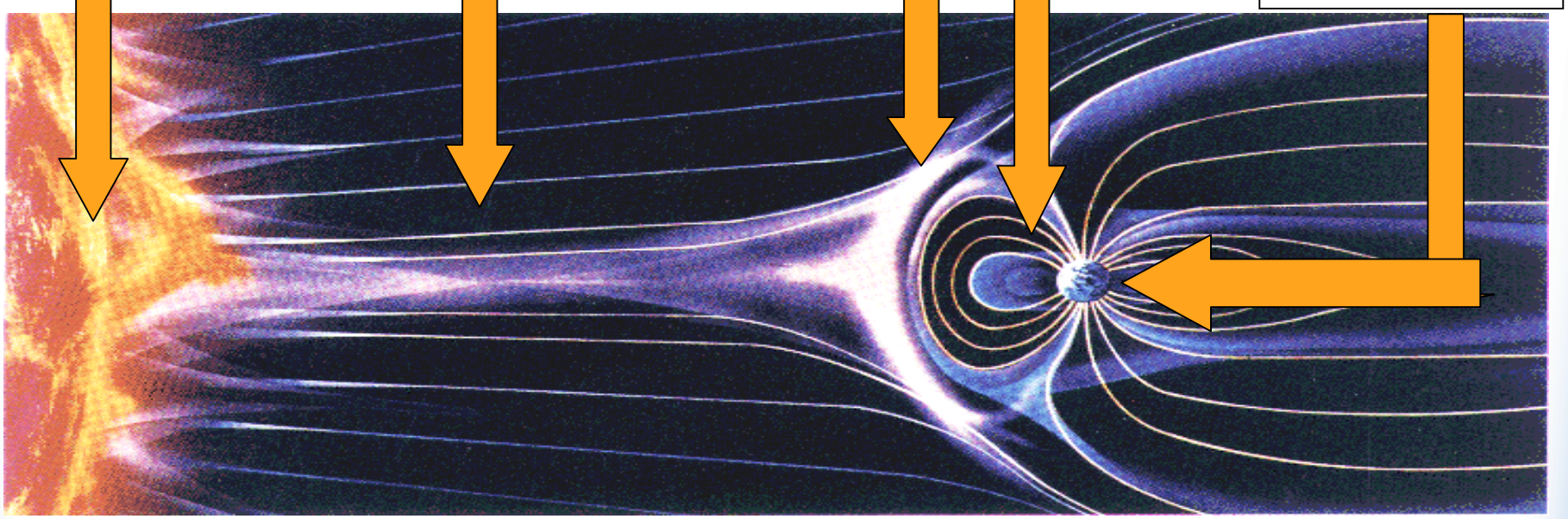
Virtual Solar Observatory (VSO)

Virtual Heliospheric Observatory (VHO)

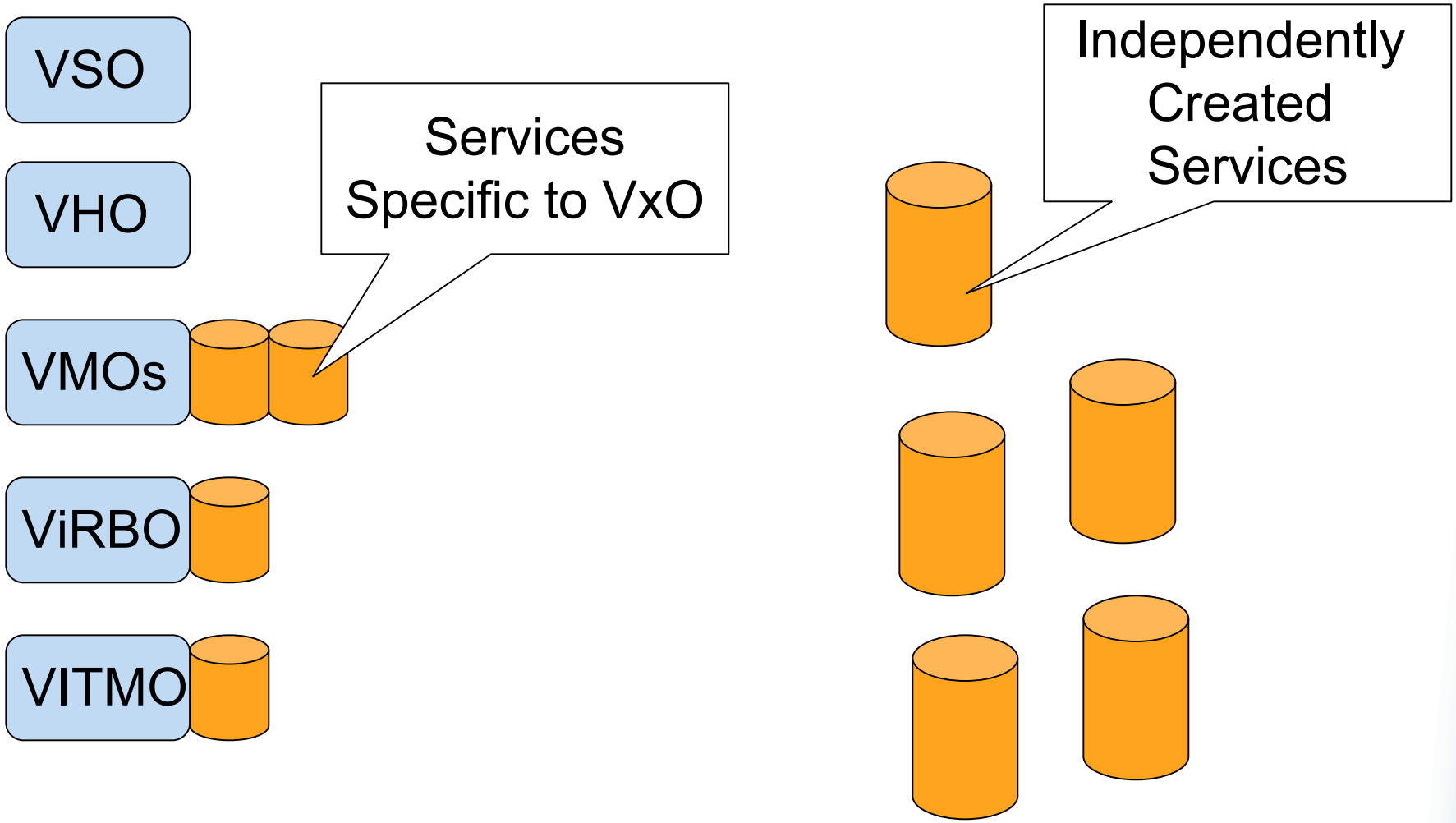
Virtual Magnetospheric Observatories (VMOs)

Virtual Radiation Belt Observatory (ViRBO)

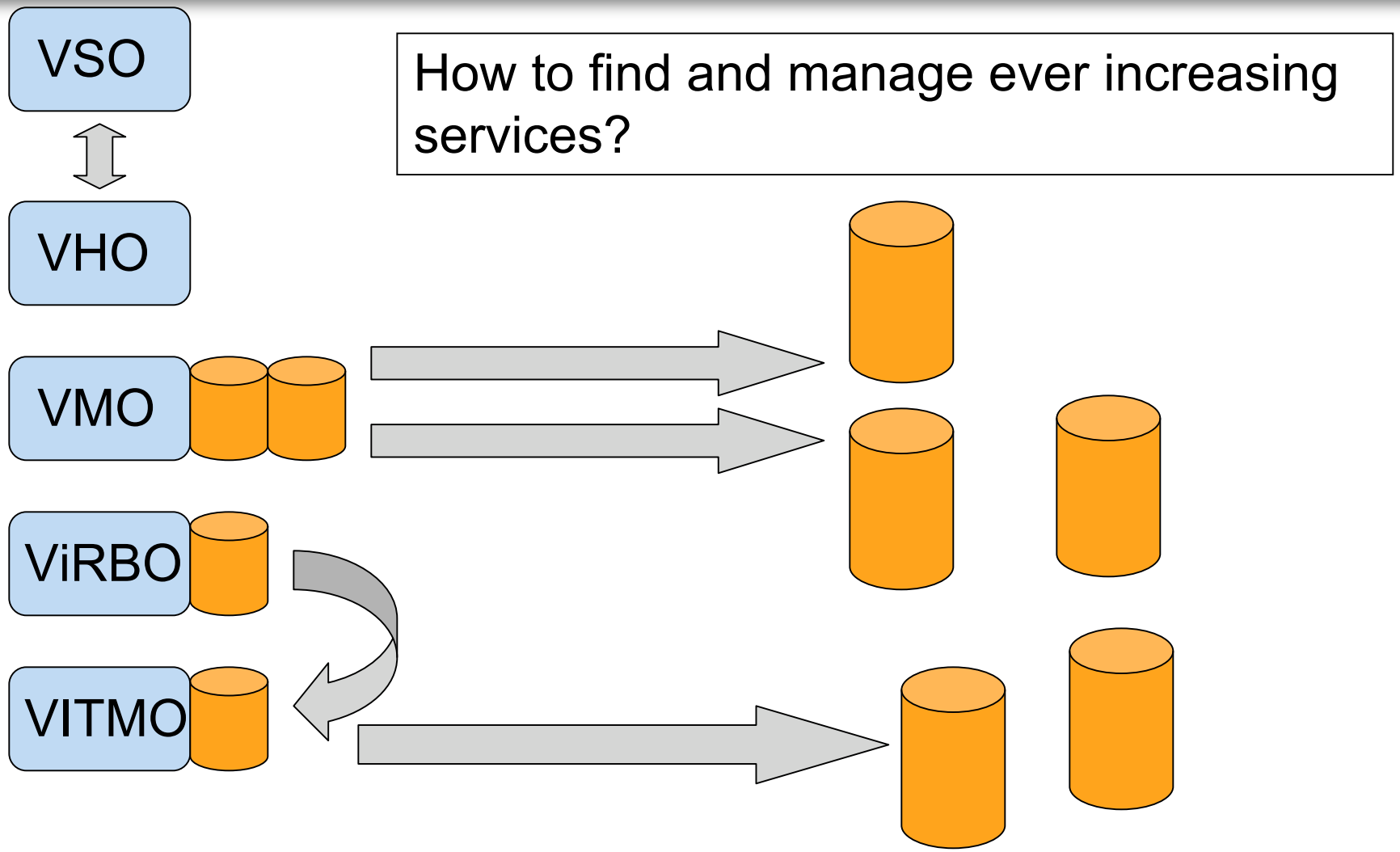
Virtual Ionospheric, Thermospheric, Mesospheric Observatory (VITMO)



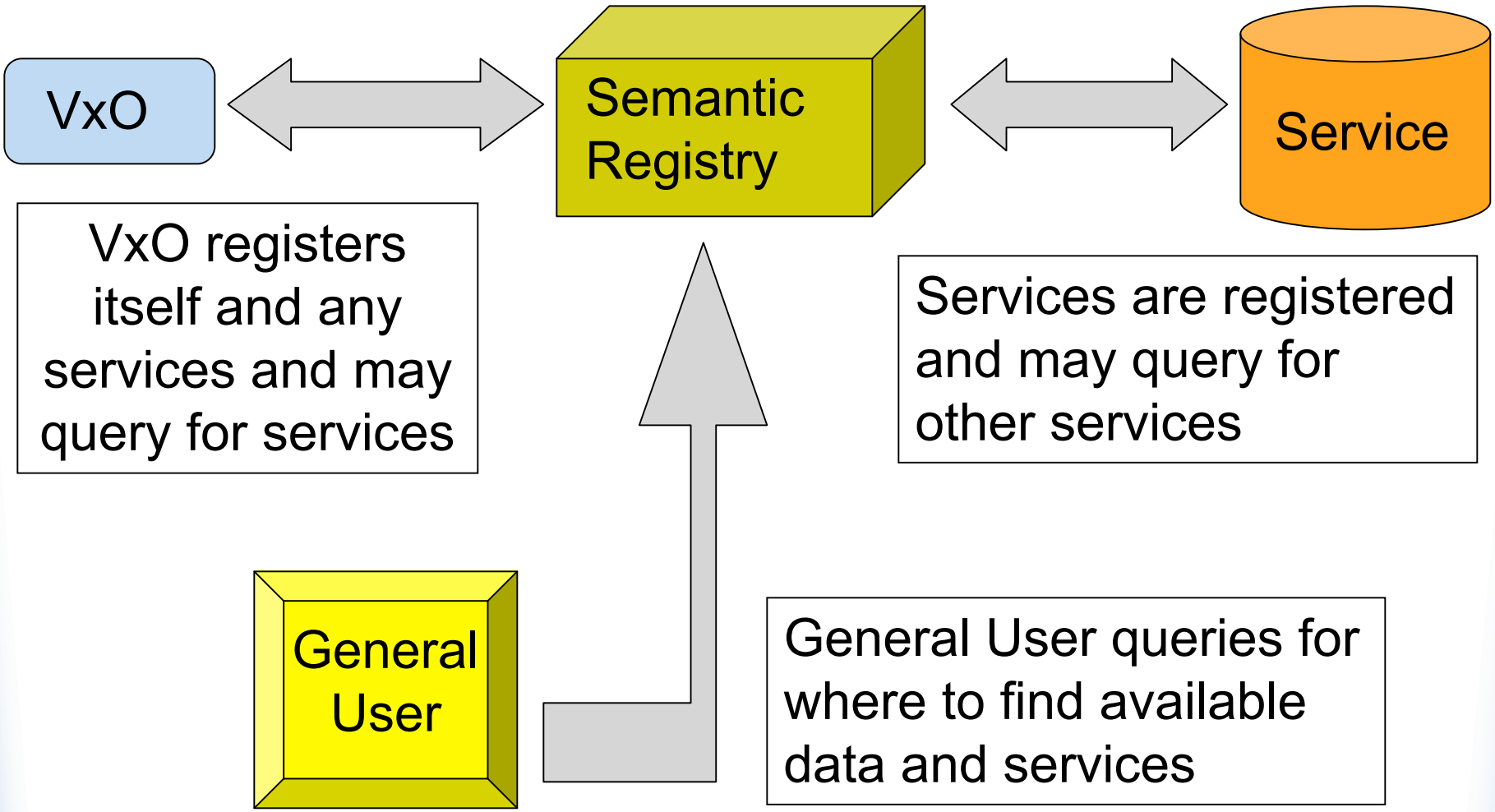
The Problem


VSO
VHO
VMOs
ViRBO
VITMO
Services Specific to VxO
Independently Created Services

The Problem


VSO
How to find and manage ever increasing services?
VHO
VMO
ViRBO
VITMO

Semantic Registry



Semantic Registry

- Conducts searches via an ontology
- Has advantages over traditional keyword or categorized searches

1. Allows for semantic description of inputs/outputs
2. Allows for more flexibility in searching than otherwise available

The Ontology

- SPASE – Space Physics Archive Search and Extract
- International group preparing a data model of space physics terms and descriptions
- Exists as a model for creating XML metadata descriptions of data, instruments, spacecraft and people.

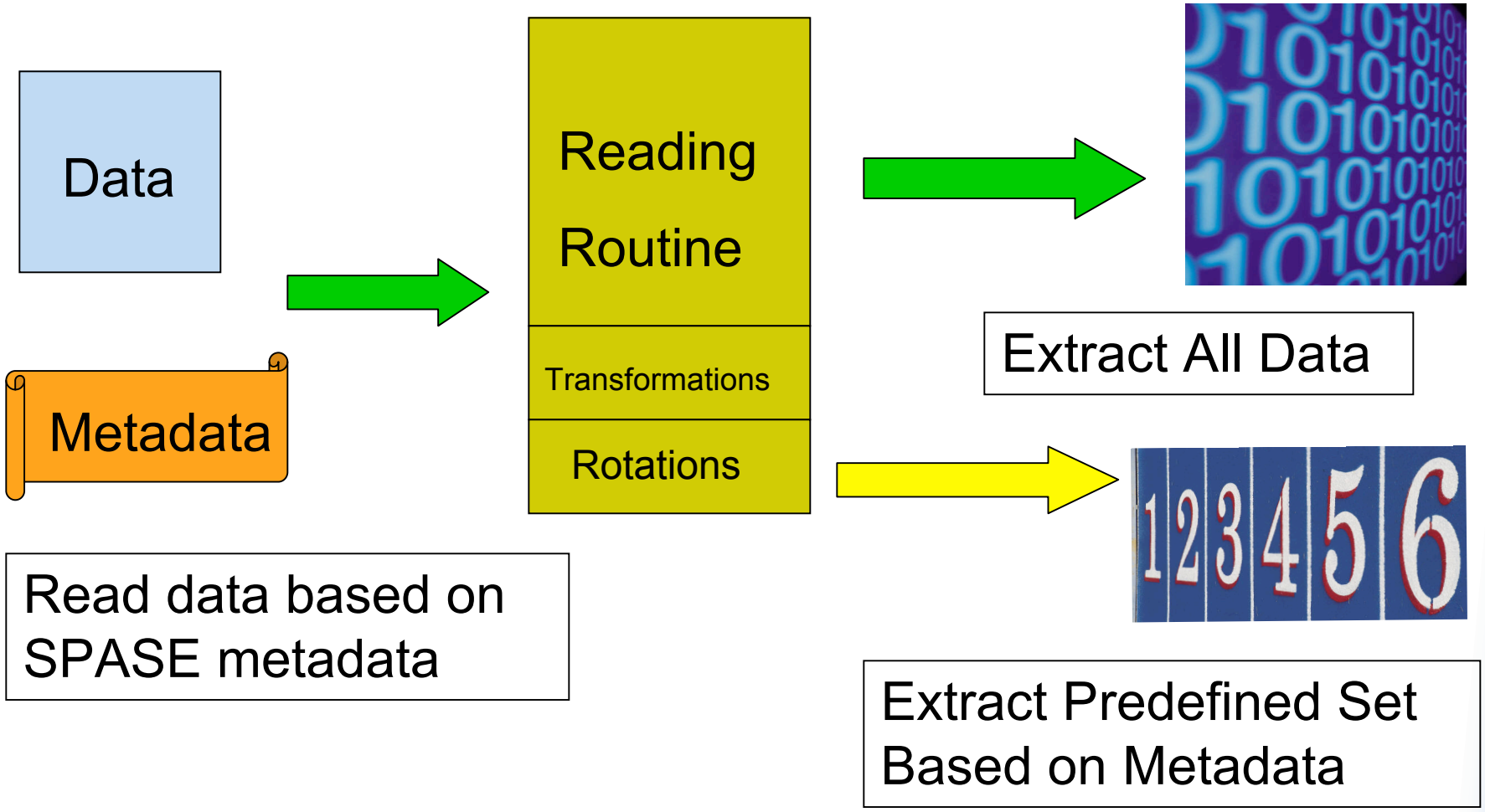
Semantic Registry Results

- Exact Match – Inputs/Outputs exactly match users request
- Direct Relation – request input is less specific or equal, service output is more specific and direct child
- More Specific – request input is less specific or equal, service output is more specific but not direct relation
- Less Specific – request input is less specific or equal, service output is less specific

Example Search

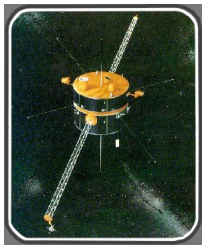
- Epsilon parameter is derived from magnetic field and plasma measurements
- Assume service that calculates Epsilon given Magnetic Field and Plasma measurements
- Registry Request: Input = Magnetic Field/Plasma data, Output = Derived Parameters
- Epsilon calculation service is a Direct Relation match

Tool for Service Providers



Tool Continued

- General Scientific Use: One reading routine for all data
- Data is immediately comparable and usable

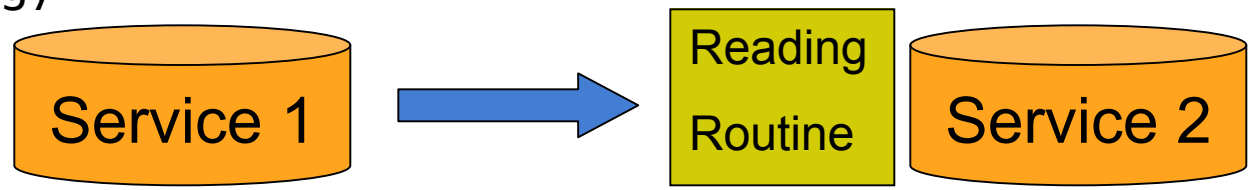


WIND
CDF
T (km/s)

Reading Routine

ACE
HDF
T (K)

- Works well for workflows and chaining of services
- Currently rules and transformations are hard-coded, move toward ontology



VxO Participation

- Working with Virtual Heliospheric Observatory (VHO) and Virtual Magnetospheric Observatory (VMO)
- Registering VxO and associated services
- Using registry from within VxO

Summary and Status

- Finalizing development and testing of semantic registry
- In near future will populate registry and make available to the public
- Developing reading tool and moving towards ontology
- Working with the SPASE community to extend descriptions to include how data is laid out in the file