The Virtual Solar-Terrestrial Observatory*

Peter Fox (pfox@ucar.edu)
HAO/NCAR
(with Don Middleton, Stan Solomon, Deborah McGuinness, Jose Garcia, Patrick West, Luca Cinquini, James Benedict, Tony Darnell)

Work partially funded by NSF/CISE/SCI

NCAR

Outline

- Concept and user needs
- What's new about a VSTO?
- Phased implementation
- Current developments
  - User requirements
  - Use-case 1
  - Ontology development
  - Example encoding (OWL)
- Status

Concept and user needs

Goal - find the right balance of data/model holdings, portals and client software that a researchers can use without effort or interference as if all the materials were available on his/her local computer.

The Virtual Solar-Terrestrial Observatory (VSTO) is proposed to be:

- *a distributed, scalable education and research environment for searching, integrating, and analyzing observational, experimental and model databases in the fields of solar, solar-terrestrial and space physics*

VSTO would comprise:

- *a system which provides virtual access to specific data, model, tool and material archives containing items from a variety of space- and ground-based instruments and experiments, as well as individual and community modeling and software efforts bridging research and educational use*

What’s new in the VSTO?

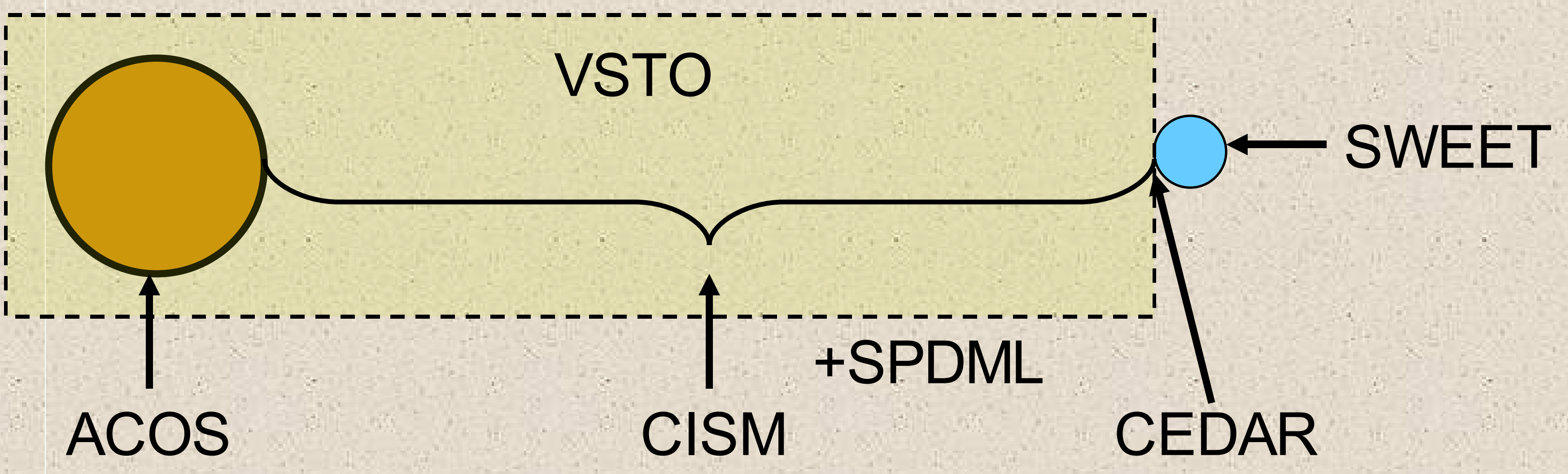
- Datasets alone are not sufficient to build a virtual observatory: VSTO will **integrate** tools, models, and data
- VSTO addresses the **interface** problem, effectively and scalably
- VSTO addresses the **interdisciplinary** metadata and **ontology** problem - bridging terminology and use of data across disciplines
- VSTO leverages the development of schema that adequately describe the **syntax** (name of a variable, its type, dimensions, etc. or the procedure name and argument list, etc.), **semantics** (what the variable physically is, its units, etc.) and **pragmatics** (or what the procedure does and returns, etc.) of the datasets and tools.
- VSTO provides a basis for a **framework** for building and distributing advanced data assimilation tools

Integration: catalog schema->ontologies

- Basic problem: schema are categorized rather than developed from an object model/class hierarchy -> significantly limits non-human use. However, they all form the basis to organize catalog interfaces for all types of data, images, etc.
- Directories, e.g. NASA GCMD, CEDAR catalog, FITS (flat) keyword/ value pairs, are being turned into ontologies (SWEET, GEON, VSTO)
- Markup languages, e.g. ESML, SPDML, ESG/ncML are excellent bases
- Evolve, recast, merge (where appropriate) using formal processes and tools, driven by a variety of *use-cases*
- Ontology - specification of a conceptualization - is the basis for interface specifications, allows reasoning, validation, etc.
- Allow for levels of semantic use of metadata (e.g. search, access, use)

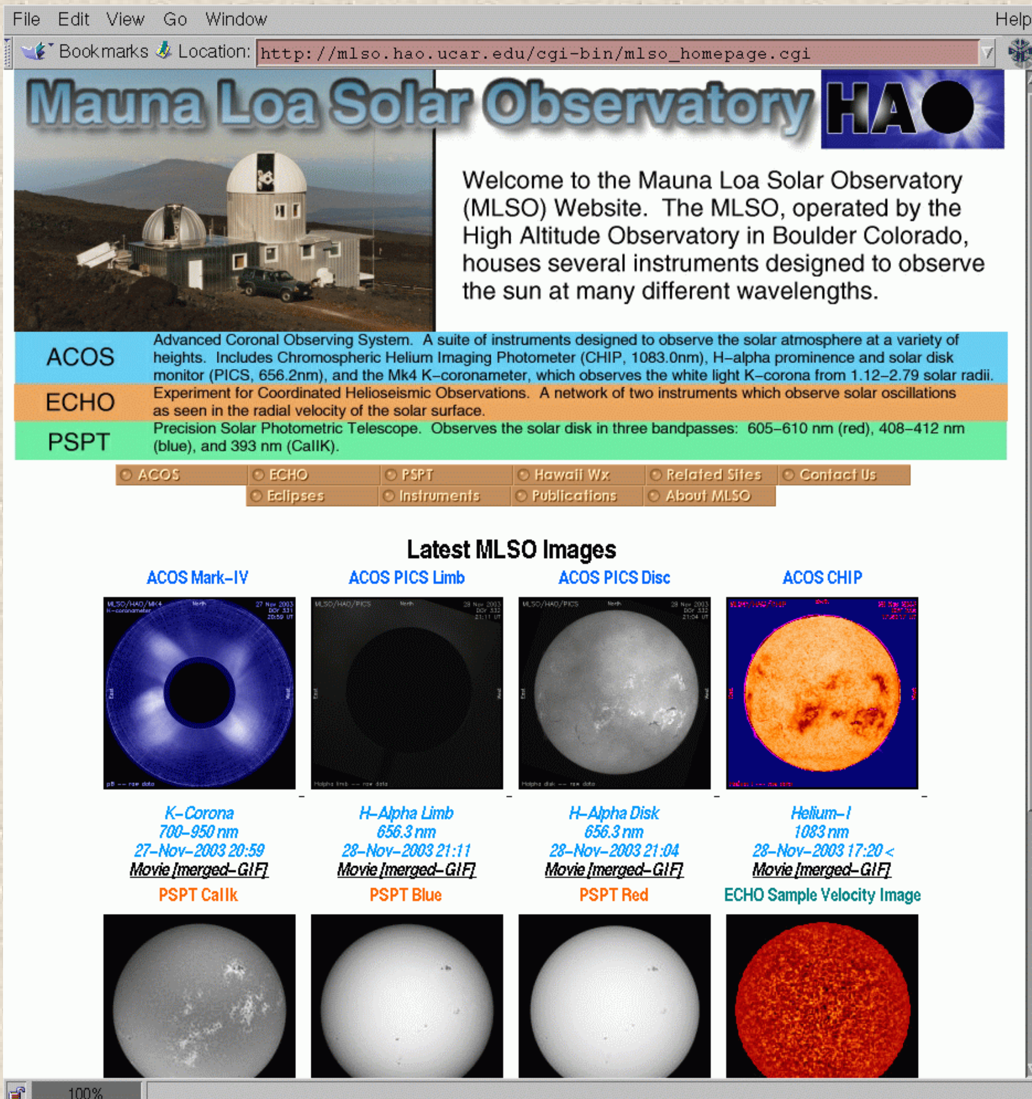
Phased implementation

- CEDAR, CISM, ACOS
- Realms (ontologies):
  - Covers middle atmosphere to the Sun + SPDML
  - Mesh with Earth Realm (SWEET)
  - Mesh with GEON



- Use-cases and user requirements

ACOS at the MLSO



Near real-time data from Hawaii from a variety of solar instruments, as a valuable source for space weather, solar variability and basic solar physics

CISM

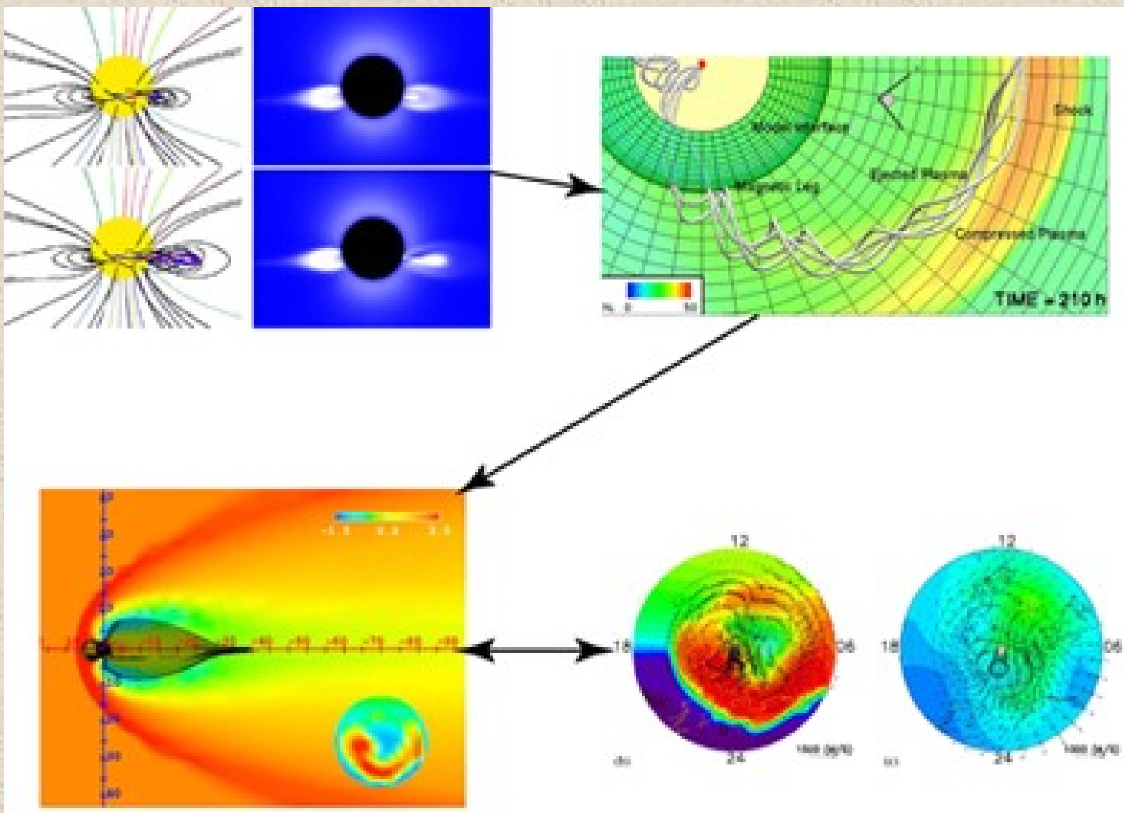
Goal: To create a physics-based numerical simulation model that describes the space environment from the Sun to the Earth.

THE USES OF SPACE WEATHER MODELING

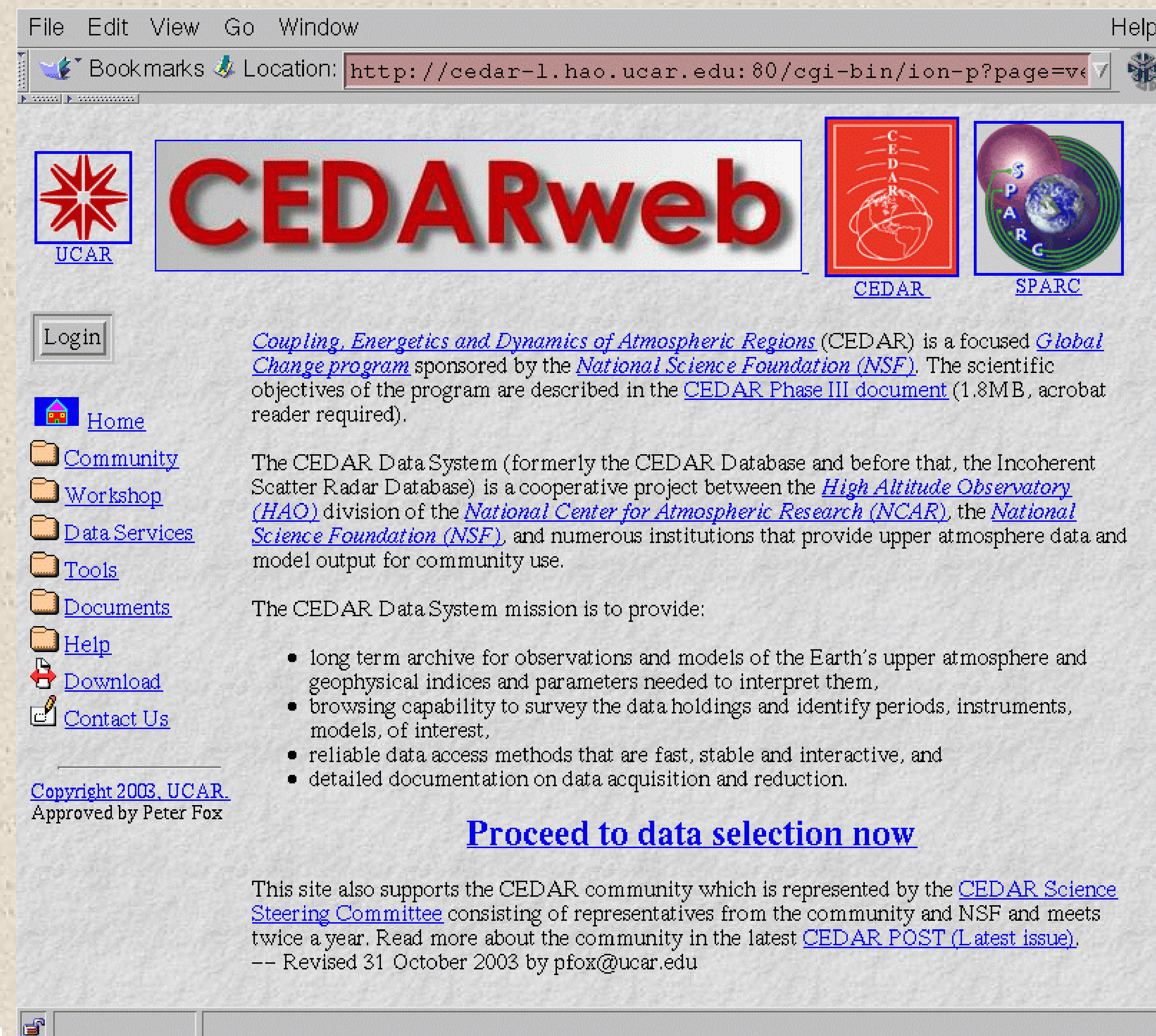
A scientific tool for increased understanding of the complex space environment.

A specification and forecast tool for space weather prediction.

An educational tool for teaching about the space environment.



CEDARWEB


File Edit View Go Window Help

Bookmarks Location: http://cedar-l.hao.ucar.edu:80/cgi-bin/ion-p?page=ve

UCAR

CEDARweb

CEDAR

SPARC

Login

Home
Community
Workshop
Data Services
Tools
Documents
Help
Download
Contact Us

Copyright 2003, UCAR.
Approved by Peter Fox

Coupling, Energetics and Dynamics of Atmospheric Regions (CEDAR) is a focused *Global Change program* sponsored by the *National Science Foundation (NSF)*. The scientific objectives of the program are described in the CEDAR Phase III document (1.8MB, acrobat reader required).

The CEDAR Data System (formerly the CEDAR Database and before that, the Incoherent Scatter Radar Database) is a cooperative project between the *High Altitude Observatory (HAO)* division of the *National Center for Atmospheric Research (NCAR)*, the *National Science Foundation (NSF)*, and numerous institutions that provide upper atmosphere data and model output for community use.

The CEDAR Data System mission is to provide:

- long term archive for observations and models of the Earth's upper atmosphere and geophysical indices and parameters needed to interpret them,
- browsing capability to survey the data holdings and identify periods, instruments, models, of interest,
- reliable data access methods that are fast, stable and interactive, and
- detailed documentation on data acquisition and reduction.

Proceed to data selection now

This site also supports the CEDAR community which is represented by the CEDAR Science Steering Committee consisting of representatives from the community and NSF and meets twice a year. Read more about the community in the latest CEDAR POST (Latest issue).
-- Revised 31 October 2003 by pfox@ucar.edu

Community data archive, documents, and support.

User requirements

- CEDAR
  - Search must return data (i.e. no null searches)
  - Search across instruments, models
  - Know about special time periods, campaigns, etc.
  - Allow selections based on (appropriate) geophysical conditions, e.g. Kp index
  - Usual format returned and in correct units
  - Must be able to easily re-create the search, access
  - Visual browsing
- MLSO
  - Same as CEDAR !!
  - + sampling interval choice, e.g. minutely, daily, average, best of the day, synoptic

CEDARWEB: Instruments/parameters now – flat and mixed

INSTRUMENT	CODE	ABBR	LAT	LONG	ALT
Jicamarca Peru I.S. Radar	10	JRO	-11.95	283.12	0.520
Arecibo P.R. I.S. Radar	20	ARO	18.35	293.25	0.0
Millstone Hill I.S. Radar	30	MLH	42.62	288.51	0.146
EISCAT I.S. Radar	70	EIS	69.58	19.22	0.087
Ground-Based Indices:					
Geophysical NGDC: AE	211	AEI	N/A	N/A	N/A
Geophysical NGDC: Dst	212	DST			
Model Outputs:					
NCAR TGCM/TIGCM Model	310	GCM	N/A	N/A	N/A
AMIE Model	311	ARE	N/A	N/A	N/A
HF Radars:					
Halley Antarctica HF Radar	820	HHF	-75.5	333.4	
MST Radars:					
Arecibo P.R. MST Radar	1040	ARM	18.35	293.25	0.0
Poker Flat Alaska MST Radar	1140	PKR	65.13	-147.46	
LF and MF Radars:					
Scott Base Antarctica MF Radar	1210	SBF	-77.85	166.75	
Poker Flat Alaska MF Radar	1375	RPK	65.13	-147.5	0.208
Passive Optical Instruments:					
Davis Antarctica Spectrometer	3010	DVS	-68.48	77.97	0.025
Arecibo P.R. Fabry-Perot	5160	AFP	18.345	293.25	0.0
LIDARs:					
USU ALO Rayleigh LIDAR	6330	USL	41.74	-111.81	1,466
USU Mesospheric Temp Mapper CCD Imager	7191	MTM	N/A		

Code	Full Name	Scale	Units	Name
	Time related parameters			
10	Year (Universal time)	1.0	yr	year
34	Time past 0000 UT	1.E-03	hour	uth
36	Time past 0000 UT	10.0	sec	uts
	Geographic parameters			
110	Altitude (height)	1.0	km	gdalt
120	Range	1.0	km	range
130	Mean azimuth angle (0=N, 90=E)	1.E-02	deg	azm
140	Elevation angle (0=horizontal, 90=vertical)	1.E-02	deg	elm
194	Declination angle (GEI coordinates)	1.E-02	deg	dec
	Basic ionospheric parameters			
510	Electron density	1.E+09	m^{-3}	ne
550	Ion temperature	1.0	K	ti
552	Ion temperature	0.1	K	tip1
560	Electron temperature	1.0	K	te
	Neutral atmosphere parameters			
800	Line of sight neutral velocity (away: >0)	1.0	m/s	vnlu
810	Neutral temperature	1.0	K	tn
	Vector field parameters			
1410	Direction 1 neutral wind (eastward)	1.0	m/s	vne
1420	Direction 2 neutral wind (northward)	1.0	m/s	vnn
1610	Direction 1 electric field (eastward)	1.E-05	V/m	ee
	Spectral parameters			
2400	Wavelength	0.1	nm	waveln

VSTO Use-case 1

UC1: Plot the observed/measured Neutral Temperature (Parameter) looking in the vertical direction for Millstone Hill Fabry-Perot interferometer (Instrument) from January 2000 to August 2000 (Temporal Domain) as a time series .

Precondition: portal application is authorized to access the backend data extraction and plotting service

5. User accesses the portal application
6. User goes through a series of views to select (in order) the desired observatory, instrument, record-type (kind of data), parameter, start and stop dates, and the plot type is inferred. At each step, the user selection determines the range of available options in the subsequent steps. NB, an alternate path is selection of start and stop dates, then instrument, etc.
3. The application validates the user request: verifying the logical correctness of the request, i.e. that Millstone Hill is an observatory that operates a type of instrument that measures neutral temperature (i.e. check that Millstone Hill <isA> observatory and check that the range of the measures property on the Millstone Hill Fabry Perot Interferometer subsumes neutral temperature). Also, the application must verify that no necessary information is missing from the request.
4. The application processes the user request to locate the physical storage of the data, returning for example a URL-like expression: find Millstone Hill FPI data of the correct type (operating mode; defined by CEDAR KINDAT since the instrument has two operating modes) in the given time range (Millstone Hill FPI <hasKindofData> 1701 <intersects> TemporalDomain [January 2000, August 2000])
5. The application plots the data in the specified plot type (a time series). This step involves extracting the data from records of one or more files, creating an aggregate array of data with independent variable time (of day or day+time depending on time range selected) and passing this to a procedure to create the resulting image.

CEDAR ontology - instrument classes

- Radar
  - Incoherent Scatter
  - Ionospheric Doppler(aka HF)
  - Middle Atmosphere (aka MLT)
    - MST
    - MF
    - LF
    - Meteor Wind
- Sounders
  - Ion
- Optical (hasBand, measuresTo, etc.)
  - Interferometers
    - Fabry-Perot (Use Case 1)
    - Michelson
      - IR
      - Doppler
  - Spectrometers
    - IR ([OH])
  - Airglow Imagers
  - All-Sky Cameras
  - Lidar
  - Spectrometers
    - Polarimeter
    - Heliograph
  - Photometers (hasChannel)
    - Single-Channel
    - Multi-Channel

Instrument Class Properties

hasName:
alsoKnownAs:
isOperatedBy:
hasOperator:
isType:
isPartofProgram:
isFundedBy:
hasCoverage:
hasSensor:
hasOperation:
hasOperatingMode:
hasOperatingFrequency:
hasDataInterval:
hasDataSet:
hasLocation: LocationClass

LocationClass:
    hasAddress:
    hasCityTown:
    hasPostalCode:
    hasCountry:
    hasContinent:
    hasLatitude: (geo, magnetic, etc.)
    hasLongitude:
    hasLocalTimeatZeroUT:

hasSensor or hasDetector

CEDAR ontology – parameters

- PhysicalQuantity
  - refersToRealm
  - Density
    - ElectronDensity
    - NeutralDensity
    - measuredBy: mass, number
  - Pressure
    - NeutralPressure
  - Temperature
    - IonTemperature
    - ElectronTemperature
    - NeutralTemperature
  - Field
    - hasMagnitude: xsd::number
    - FieldComponent
      - hasDirection
      - hasCoordinateSystem
    - MagneticFieldComponent
    - ElectricFieldComponent
    - VelocityFieldComponent

Wavelength: hasInterval
Wavenumber: hasInterval

- Index
  - hasSamplingRepresentation
  - Geophysical
  - Solar

- StatisticalMeasure
  - Covariance
  - ChiSquare
  - ReducedChiSquare
  - CrossCorrelation
  - Coherence
  - Curtosis
  - …

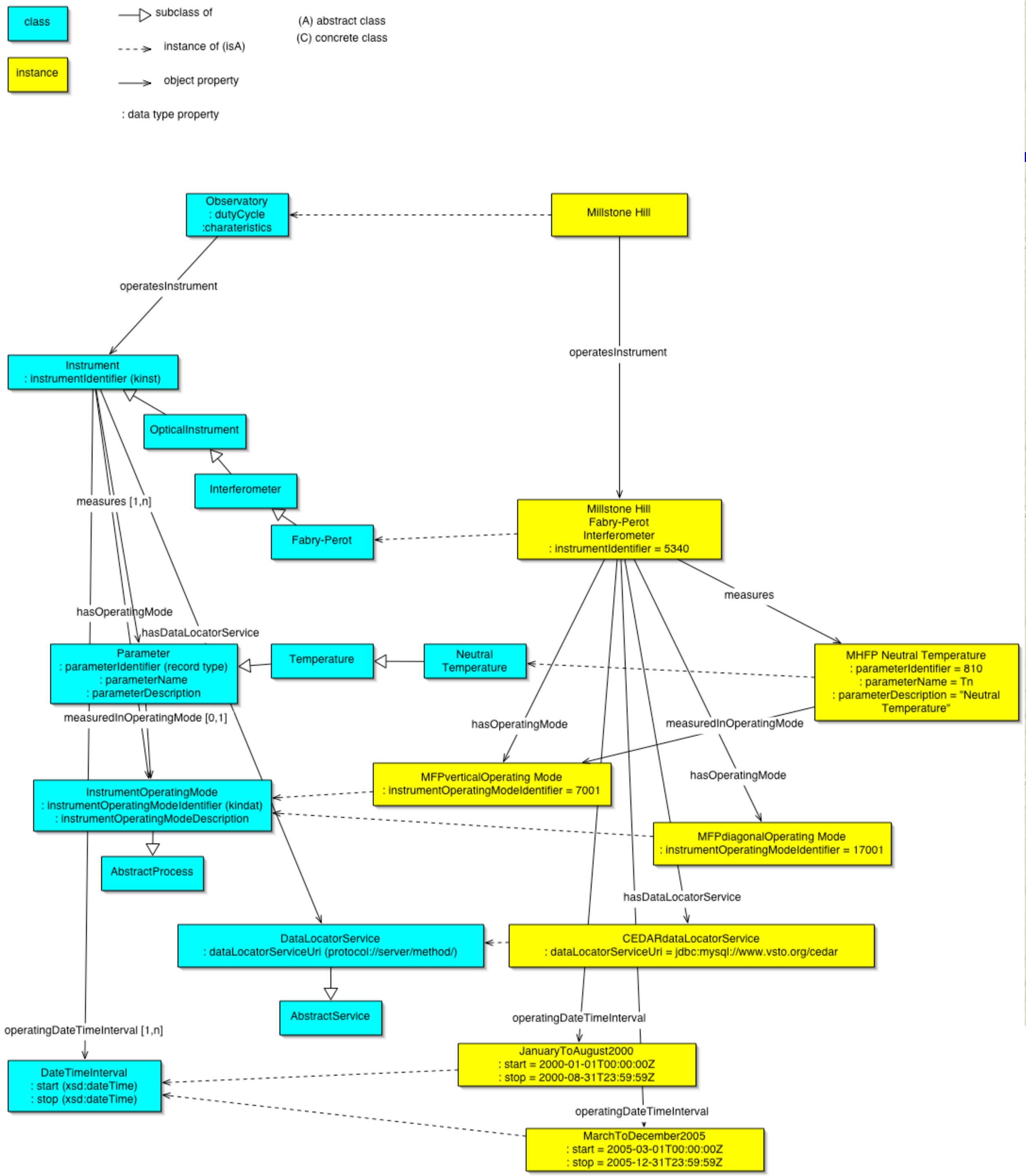
StatisticalOperation (SWEET)
numerics.owl

VSTO ontology

Classes – blue

properties

Instances – yellow



NCAR

Encoding the representation(s) – OWL-DL

```
<owl:Class rdf:ID="NeutralDensity">
<rdfs:subClassOf rdf:resource="#Density"/>
<owl:Restriction>
<owl:onProperty rdf:resource="http://sweet.jpl.nasa.gov/ontology/space.owl#isPartOf"/>
<owl:allValuesFrom rdf:resource="#Mesosphere"/>
</owl:Restriction>
</owl:Class>

<owl:Class rdf:ID="IonTemperature">
<rdfs:subClassOf rdf:resource="#Temperature"/>
<owl:Restriction>
<owl:onProperty rdf:resource="http://sweet.jpl.nasa.gov/ontology/space.owl#isPartOf"/>
<owl:allValuesFrom rdf:resource="#Ionosphere"/>
</owl:Restriction>
</owl:Class>

<owl:Class rdf:ID="Density">
<rdfs:subClassOf>
<owl:Restriction>
<owl:onProperty rdf:resource="http://sweet.jpl.nasa.gov/ontology/numerics.owl#hasDefaultUnit"/>
<owl:hasValue
        rdf:resource="http://sweet.jpl.nasa.gov/ontology/units.owl#kilogram_perMeterToPower3"/>
</owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf rdf:resource="#MassConcentrationRelatedQuantity"/>
</owl:Class>
```

EarthRealm and Upper atmosphere

EarthRealm excerpt:

```
<owl:Class rdf:ID="Thermosphere">
<rdfs:subClassOf>
<owl:Restriction>
<owl:onProperty rdf:resource="#hasAverageLowerBoundaryReferenceHeight"/>
<owl:hasValue
     rdf:datatype="http://www.w3.org/2001/XMLSchema#double">85000</owl:hasValue>
</owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf>
<owl:Restriction>
<owl:onProperty rdf:resource="#hasAverageUpperBoundaryReferenceHeight"/>
<owl:hasValue
     rdf:datatype="http://www.w3.org/2001/XMLSchema#double">500000</owl:hasValue
     >
</owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf rdf:resource="#AtmosphereLayer"/>
</owl:Class>
```

SunRealm and Active Regions

SunRealm excerpt:

```
<owl:Class rdf:ID="ActiveRegion">
<rdfs:subClassOf rdf:resource="#MagneticRegionType"/>
<rdfs:subClassOf>
<owl:Restriction>
<owl:onProperty
      rdf:resource="http://sweet.jpl.nasa.gov/ontology/space.owl#isPartOf"/>
<owl:allValuesFrom rdf:resource="#SunSurfaceLayer"/>
</owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf>
<owl:Restriction>
<owl:onProperty rdf:resource="#hasMagneticRegionType"/>
<owl:allValuesFrom rdf:resource="#SunMagneticRegion"/>
</owl:Restriction>
</rdfs:subClassOf>
</owl:Class>
```

Status

- VSTO ontology, version 0.1, (vsto.owl)
- Initial contributions to SWEET
- Initial mapping of VSTO ontology to SWEET
- VSTO specific instances (aka metadata)
- Detailed design and implementation for Use-case 1
- Use-case 2 (solar) and Use-case 3 (radar) are being formalized
- **http://vsto.hao.ucar.edu/** – project web site
- Collaborations: CISM, GEON, NMI, SWEET, SciFlo, CDP, ECHO, NSDL/DLESE/OAI, SPDML, OPeNDAP, SRB
- Please contact (pfox@ucar.edu, or vsto.org) for more information

NCAR

Additional material

CEDARWEB current architecture