Ontologies in Practice - a Solar Terrestrial Example

Peter Fox (pfox@ucar.edu)
HAO/ESSL/NCAR
NSF and NASA funded
What’s different?

• VSTO addresses the **interdisciplinary** metadata and **ontology** problem - bridging terminology and use of data across disciplines
• VSTO leverages the development of schema describing
  • **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, its uses, etc.) of the datasets and tools.
Why we were led to semantics

- When we integrate, we integrate concepts, terms
- In the past we would ask, guess, research a lot, or give up
- It’s pretty much about **meaning**
- Semantics can really help find, access, integrate, use, explain, trust…

➤ What if you…
- could not only use your data and tools but remote colleague’s data and tools?
- understood their assumptions, constraints, etc and could evaluate applicability?
- knew whose research currently (or in the future) would benefit from your results?
- knew whose results were consistent (or inconsistent) with yours?...
- Funders: What if you ...
- could identify how one research effort would support other efforts?
- (and your fundees) could reuse previous results?
- (and your fundees) could really interoperate?
VSTO USE CASE #2 (MLSO) ONTOLOGY WORKFLOW

1. Choose instrument & dataset
2-4. Choose year, month, day
5. Choose data product

Send data request

Receive data response
VSTO USE CASE #1 (CEDAR) ONTOLOGY WORKFLOW

1. Choose observatory
   - Observatory

2. Choose instrument
   - Instrument

3-5. Choose year, month, day, number of days
   - DateTimeInterval

6. Choose instrument operating mode
   - InstrumentOperatingMode

7. Choose parameter
   - Parameter

8. Infer possible plot types
   - TimeSeriesPlot

Send data request
   - DataRequest

Receive data response
   - DataResponse
OpticalInstrument
  Interferometer
    Fabry–PerotInterferometer
    MichelsonInterferometer
    IRMichelsonInterferometer
    DopplerMichelsonInterferometer
AirglowImager
AllSkyCamera
Lidar
Spectrometer
  Polarimeter
  Heliograph
Photometer
  SingleChannelPhotometer
  MultiChannelPhotometer

Taxonomy of instruments covering content areas. Currently expanding and evaluating.

Approach:
  • identify instruments & parameters
  • organize hierarchically
  • compare/extend SWEET (realms, properties, space, …)
  • scientific expert review
  • ontology expert review
  • related scientific review
  • populate instances (including meta-data)
  • use-case driven
hasInstrumentOperatingMode +

hasMeasuredParameter +

hasContainedParameter +

Instrument OperatingMode

NeutralTemperature

hasCoordinate

Time

Parameter

hasCoordinate +

TimeDependentParameter
### CEDAR ontology – parameters

<table>
<thead>
<tr>
<th>PhysicalQuantity</th>
<th>Wavelength: hasInterval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wavenumber: hasInterval</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Density</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>ElectronDensity</td>
<td>hasSamplingRepresentation</td>
</tr>
<tr>
<td>NeutralDensity</td>
<td>Geophysical</td>
</tr>
<tr>
<td></td>
<td>Solar</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>IonTemperature</td>
</tr>
<tr>
<td>ElectronTemperature</td>
</tr>
<tr>
<td>NeutralTemperature</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Field</th>
<th>StatisticalMeasure</th>
</tr>
</thead>
<tbody>
<tr>
<td>hasMagnitude: xsd::number</td>
<td>Covariance</td>
</tr>
<tr>
<td>FieldComponent</td>
<td>ChiSquare</td>
</tr>
<tr>
<td></td>
<td>ReducedChiSquare</td>
</tr>
<tr>
<td>hasDirection</td>
<td>CrossCorrelation</td>
</tr>
<tr>
<td>hasCoordinateSystem</td>
<td>Coherence</td>
</tr>
<tr>
<td>MagneticFieldComponent</td>
<td>Curtosis</td>
</tr>
<tr>
<td>ElectricFieldComponent</td>
<td></td>
</tr>
<tr>
<td>VelocityFieldComponent</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>...</th>
</tr>
</thead>
</table>

| StatisticalOperation (SWEET) |
| numerics.owl |
APL, May 26 2006 - Ontology Workshop - Solar-Terrestrial as example of ontology use
VSTO SOFTWARE ARCHITECTURE

OWL ONTOLOGIES

- cedar_instances.owl
  - import
  - cedar.owl
  - import
  - vsto_core.owl
  - import
  - mlso.owl
  - import
  - vsto.owl
  - import
  - mlso_instances.owl

Automatic generation

packages ncarm.sto.auto, ncarm.sto.auto.impl

Java interfaces
  - implement
  - extend
  - my Java interfaces
  - implement
  - create
  - VSTOFactory

Java classes
  - extend
  - my Java classes

Objects representing OWL classes + stub extensions for inserting custom functionality
APL, May 26 2006 - Ontology Workshop - Solar-Terrestrial as example of ontology use

Java Protege-OWL API + supporting jars

VSTOfactory - objects representing OWL classes + stub extensions for inserting custom functionality

PELLET Reasoning Engine

VSTOService

CEDARservice

MLSOservice

CEDAR DB

MLSO DB

CEDAR specific service extensions

MLSO specific service extensions

USE CASES WORKFLOW SIMULATION PROGRAMS

JUNIT tests

VSTO WEB PORTAL

SPRING controllers and data beans

JSP

USER INTERFACE CONTROL COMPONENTS

USER INTERFACE VIEWS
**CEDAR Use Case #1**

**Step 2 of 8: Choose Instrument**

- **Instrument:** Millstone Hill Fabry-Perot

A device that measures a physical phenomenon or parameter. At a minimum, it possesses a detector which produces a signal from which the desired quantity is calculated or inferred. The detector signal possesses the information needed to either obtain the value of the phenomenon (e.g., temperature: a voltage is converted into a temperature unit) or infer its value using further processing and computation (magnetic fields: detector intensities at a different wavelengths and...
KINST = 5340
KINDAT = 7001
PARAMS = 810
NPARAMS = 1
YEAR = 2000
MONTH = 7
DAY = 10
NDAYS = 10
NPARAMS = 2
PLOTSIZE = 600

Instrument: 5340 — Millstone Hill Fabry—Perot
Record Type: 820/7001 — Vı
Parameters:
  810 — tn — Neutral temperature
Starting Date: July 10, 2000
Ending Date: July 20, 2000

These plots are produced for visual browsing of the data and should not be used in publications without citing the data provider and CEDARWEB.
Integrative use-case:

Find data which represents the state of the neutral atmosphere anywhere above 100km and toward the arctic circle (above 45N) at any time of high geomagnetic activity.

How to translate this into a complete query for data?

What information do we have to extract from the use-case?

What information we can infer (and integrate)?

Most of all – what will be returned? Data from instruments, indices and models!!!
Translating the Use-Case – non-monotonic?

Input
Physical properties of neutral atmosphere
- Above 100km
- Toward arctic circle (above 45N)
- High geomagnetic activity
Action: Return Data

GeoMagneticActivity has ProxyRepresentation
GeophysicalIndex is a ProxyRepresentation (in Realm of Neutral Atmosphere)
Kp is a GeophysicalIndex
hasTemporalDomain: “daily”
hasHighThreshold: xsd_number = 8
Date/time when KP => 8

Specification needed for query to CEDARWEB
Instrument
Parameter(s)
Operating Mode
Observatory
Date/time
Return-type: data

APL, May 26 2006 - Ontology Workshop - Solar-Terrestrial as example of ontology use
NeutralAtmosphere is a subRealm of TerrestrialAtmosphere
hasPhysicalProperties: NeutralTemperature, Neutral Wind, etc.
hasSpatialDomain: [0,360],[0,180],[100,150]
hasTemporalDomain:
NeutralTemperature is a Temperature (which) is a Parameter

GeoMagneticActivity has ProxyRepresentation
GeophysicalIndex is a ProxyRepresentation (in Realm of Neutral Atmosphere)
Kp is a GeophysicalIndex
hasTemporalDomain: “daily”
hasHighThreshold: xsd_number = 8
Date/time when KP => 8

FabryPerotInterferometer is a Interferometer, (which) is a Optical Instrument (which) is a Instrument
hasFilterCentralWavelength: Wavelength
hasLowerBoundFormatioHeight: Height

ArticCircle is a GeographicRegion
hasLatitudeBoundary:
hasLatitudeUpperBoundary:

GeophysicalIndex has ProxyRepresentation (in Realm of Neutral Atmosphere)
Integrative use-case:

Find data which represents the state of the charged terrestrial atmosphere in response to a recent solar activity particle ejection event.

How to translate this into a complete query?

Translate the terms from the use-case and bridge the terminologies?

A lot of information can be inferred and thus integrated

What will be returned? Data from the Sun at relevant times from relevant events and data from the upper atmosphere all from instruments, indices and models!
Education – semantic integration

Sun–Earth Connections

A 9th grade teacher is preparing a lesson plan aimed at getting students to learn more about the ‘northern lights’, addressing NSES content standards in earth science. The teacher wants the students to learn the scientific terminology, where the phenomena occurs and retrieve some data or graphics for a recent occurrence. The goal of the lesson plan is to engage students, using authentic data from the aurora, as part of an inquiry-based program.

The aim of this use-case is to translate the teacher’s requirements into a search for, and to find a specific dataset appropriate to the education task from a Virtual Observatory. Data would need to be provided in a usable format, as well as in a graphically accessible way.

At present, in order for a teacher to access VO data, they would need to know the appropriate scientific vocabulary, the types of data available, spatial locations and directional operating modes of instruments (also models and indices) to be able to locate, retrieve and use the data from them. This use-case will demonstrate how ontologies, and semantically enabled interfaces can bridge the current gap in terminology and significantly reduce the level of detail that a person has to know about the data.
Vocabulary and beyond

Teacher (hasgradelevel=9) - primary actor
Lesson plan - intermediate product which connects actors
Students - secondary actor
Learn - process outcome
‘Northern lights’ - identified phenomenon in non-science vocabulary
NSES content standards (controlled vocabulary), ISO std. and ontology
Earth science (application domain)
Scientific terminology (controlled vocabulary)
Where phenomena occurs (spatial, coordinates)
Retrieve data (outcome, service)
Retrieve graphics (outcome, service)
Recent occurrence (time)
Authentic data (validation, verification, domain and range)
Aurora (phenomenon)
Inquiry-based program
-> develop ontology, i.e. classes and subclass relations, properties, inherit existing ontologies, rapid prototype
Ontology Workshop - Solar-Terrestrial as example of ontology use
APL, May 26 2006 - Ontology Workshop - Solar-Terrestrial as example of ontology use
VSTO ONTOLOGIES HIERARCHY

vsto_core.owl
vsto:
core classes and properties

cedar.owl
cedar:
CEDAR specific classes

mlso.owl
mlso:
MLSO specific classes

cedar_instances.owl
cedar:
CEDAR specific instances

mlso_instances.owl
mlso:
MLSO specific instances

vsto.owl
container for all ontologies

import >
Phased implementation

- Leverage programs – CEDAR, CISM, ACOS
- Realms (ontologies):
  - Covers middle atmosphere to the Sun
  - Mesh/map with Earth Realm
  - Mesh/map with Solid Earth/Interior

Use-cases and user requirements
MLSO Use Case #2

Images Request
1. Dataset: PICS H-Alpha Disk
2. Year: 1997
3. Month: December
4. Day: 2
5. Product: Best PICS Disk Image

Data Request Completion

Images: 19971202.194133.dpm.asl.fits.gz
        19971202.194133.dpm.asl.jpg

< Back  Cancel  Finish

MLSO DPM
HALPHA
1997-12-02
19:41:33 UT
DOY: 336
Community aspects

- www.geospaceontology.org
- Publish your
  - experiences,
  - results,
  - lessons learned, AND
  - your ontology!!! AND
  - discuss!!!
In progress

- VSTO ontology version 0.4, (vsto.owl) for use-case 3
- Production UI for August release, include security, etc.
- API/Web services encapsulation of semantic interfaces being documented via use-cases from APL and Madrigal
- Use-case 5 – the Sun–Earth one
- More use-cases to drive the completion of the ontologies – filling out the instrument ontology esp.
- Educational ontology
- Test configuring a VxOs in 2007
What is an Ontology?

Catalog/ID

Terms/glossary

Thesauri
“narrower term” relation

Informal
is-a

Formal
is-a

Frames
(properties)

General
Logical
constraints

Formal
instance

Value
Restrs.

Disjointness, Inverse, part-of...

*based on AAAI ’99 Ontologies panel – McGuinness, Welty, Ushold, Gruninger, Lehmann