

UCD and Utypes in Planetary Sciences

B. Cecconi & VOParis Team & CDPP Team

IVOA Interop, May 2013, Heidelberg

Context

- ❖ Europlanet / IDIS: prototyping a Planetary Sciences VO
HELIO: Heliophysics Integrated Observatory
- ❖ Several of science thematics: *Sun, planets, atmospheres, surfaces, interiors, small bodies, orbital parameters, in situ exploration, plasma (waves, particle and fields), radio astronomy...*
- ❖ Large variety of data types: *images, spectra, times series, movies, dynamic spectra, profiles, maps...*
- ❖ Even larger variety of physical parameters
- ❖ Including: *remote data, in-situ data, models, lab experiments, field analogs*

Data Model

Required Metadata

- ❖ We describe Datasets and Granules (i.e., product, file, or the smallest granularity distributed by the service) content, not the access the the data.
 - ❖ Resource identification
 - ❖ Target
 - ❖ Instrument (including hosting facility)
 - ❖ Axis (including bounds, resolution, sampling, unit)
 - ❖ Physical parameter (including UCD, unit)

Data Model Heritage

- ❖ IDIS-DM v1 was build almost from scratch. Simple but restricted and not compatible with IVOA (especially registries).
- ❖ IDIS-DM v2, now called Planetary Science Resource **PSR-DM**, is based on **VOResource**, **VODataService** and **VODataCollection**. It makes use of several other IVOA data models, such as **STC**, **Utypes**, **UCDs**, **VOUnit**
- ❖ We planned to use **ObsCoreDM** at some point, but not done in curent implementation.
- ❖ **EPNCore** is a restriction of PSR-DM, and has same scope than ObsCore, with slight changes.
- ❖ Latest version (not final) of documentation is available here:
<http://voparis-europlanet.obspm.fr/docs/PlanetaryScienceResource-DM-latest.pdf>

UCDs

Unified Content Descriptors

- ❖ IVOA recommendation:
<http://www.ivoa.net/Documents/latest/UCD.html>
- ❖ From this recommendation:
 - UCD = «A Controlled Vocabulary for Astronomy»
 - *A UCD does not define the units nor the name of a quantity, but rather "what sort of quantity is this?"; for example phys.temperature represents a temperature, without implying a particular unit.*
- ❖ IVOA UCD List:
<http://www.ivoa.net/Documents/latest/UCDList.html>
- ❖ *A UCD is a string which contains textual tokens called 'words', separated by semicolons(;). A word is composed of 'atoms', separated by periods(.). The hierarchy is as follows:
atoms → words → composed words
UCD1+ are either single words, or a composition of several words.*

Comparison tool (beta) UCD/SPASE/PDS

- ❖ <http://typhon.obspm.fr/idis/ucd.php>

Unified Content Descriptor Lists

These lists have been compiled using the [UCD list provided by CDS](#) adding SPASE and PDS keywords.

ID	Name	Source	Definition	Possible Synonyms
245	phys.electron	cds	Electron	[spase]Particle.ParticleType.Electron [idis]phys.particle.electron
246	phys.electron.degen	cds	Electron degeneracy parameter	
293	phys.temperature.electron	cds	Electron temperature	
653	Particle.ParticleType.Electron	spase	An elementary particle consisting of a charge of negative electricity equal to about 1.602×10^{-19} Coulomb and having a mass when at rest of about 9.109534×10^{-28} gram.	[cds]phys.electron [idis]phys.particle.electron
915	phys.particle.electron	idis	Related to electron (An elementary particle consisting of a charge of negative electricity equal to about 1.602×10^{-19} Coulomb and having a mass when at rest of about 9.109534×10^{-28} gram).	

Number of record selected: 5

POWERED BY
Mac OS X Server

UCD for Planetary Sciences

- ❖ Main source of inputs: **Europlanet-RI/IDIS (EU-FP7) project.**
- ❖ The IDIS group developed a data model and metadata to describe datasets and granules (i.e. usually files), as well as physical parameters and axis.
- ❖ Quantities to be described (19 mandatory parameters in EPN-TAP):
 - *temporal, spectral and spatial axes, coordinate systems*
 - *physical parameters (remote sensing and in-situ; photons, waves, fields and particle;) and type of quantity*
 - *instruments and observatories*
 - *observation conditions*

UCD for Solar and Heliophysics

- ❖ Main source of inputs: **HELIO (EU-FP7) project**.
- ❖ The HELIO group developed data models and metadata to describe event catalogs, physical parameters, datasets, observatories...
UCD was one of the adopted standards.
- ❖ Quantities to be described:
 - *events, times, durations, rotation or revolution number (**time.epoch...**)*
 - *physical parameters (remote sensing and in-situ; photons, waves, fields and particle;) and type of quantity*
 - *location (**pos.bodyrc,...**)*
 - *measured, modeled, processed or derived parameters (as in **phys.modeled?**)*
- ❖ *NB: for space physics, the SPASE data model is comparable to UCDs (or may be closer to a SKOS description)*

Suggestions and questions

Comets and Samples (1)

- * Input from Italian EPN group (description of planetary samples)
- * Set of new UCDs under discussion:
 - **phys.color** (*color of an object, generally assigned at eye, NOT USABLE for spectral type of stars NOR for the color index, e.g.: 'black', 'reddish'*)
 - **phys.luster** (*luster of an object, e.g.: 'pearly', 'metallic', 'vitreous'... mineralogical property, generally assigned at eye*)
 - **phys.porosity** (*porosity percentage of the body*)
 - **phys.shape** (*shape of an object, e.g.: 'irregular', 'spherical'...*)
 - **phys.transparency** (*transparency of a solid, e.g.: 'opaque', 'translucent', 'transparent' mineralogical property, generally assigned at eye*)
- * Hierarchical propositions:
 - **phys.aspect** ? (*«phys.aspect.color», «phys.aspect.luster»...*)
 - **phys.sample** (*relative to sample: «phys.sample.aspect.color»*) or associated to the existing UCD **src.sample** ? (*«src.sample;phys.aspect.color»...*)

Suggestions and questions Comets and Samples (2)

- ❖ Input from Italian EPN group (description of planetary samples)
- ❖ Set of new UCDs proposed for evaluation :
 - **phys.mol.elecband** (*electronic band of the transition*)
 - **phys.mol.species** (*species*)
 - **phys.reflectance** (*reflectance of the body*)
 - **phys.sample.cluster** (*Eventually indicates the cluster to which the sample belongs*)
 - **phys.size.smedAxis** (*for 3d objects a third axis is necessary*) linked to *phys.size.smajAxis* and *phys.size.sminAxis*
 - **src.group** (*group, family or dynamical class of the object, e.g.: 'Halley type comet', 'AGNII', 'Themis family asteroid'*)
 - **src.orbital.TissJ** (*Tisserand parameter respect to Jupiter*)
 - **em.line.FeKalpha** (*Fe K alpha line at 6.4 keV*)
 - **em.molecline** (*Designation of molecular lines*)
 - **em.molecline.C2** (*number of C2 lines in the observed range*)
 - **em.molecline.C3** (*number of C3 lines in the observed range*)
 - **em.molecline.CH** (*number of CH lines in the observed range*)
 - **em.molecline.NH2** (*number of NH2 lines in the observed range*)
 - **em.molecline.CN** (*number of CN lines in the observed range*)

Suggestions and questions

Comets and Samples (3)

- ❖ Input from Italian EPN group (description of planetary samples)
- ❖ Suggested UCDs for which we found possible equivalent UCDs:
 - **phys.sample.magnetized** ('yes', 'no', 'partially'...)
=> **src.sample;phys.magField;meta.flag**
 - **phys.sample.mass** (*mass of the sample*)
=> **src.sample;phys.mass**
 - **phys.sample.parentbody** (*Parent body of the sample, it can be generic or specific, very hard to recognize for dust, e.g.: 'Itokawa', 'asteroid', 'Moon'...*)
=> **src.sample;meta.id.parent**
 - **phys.sample.retrloc** (*retrieval location of the sample, e.g.: 'Moon, Mare Serenitatis', 'Earth stratosphere, above Sahara desert', 'Interplanetary medium at 2 AU'...*)
=> **src.sample;pos**
 - **phys.sample.type** ('Cosmic dust', 'Artificial terrestrial contamination', 'Lunar basalt'...)
=> **src.sample;meta.note**
 - **src.id** (*Identifier of the object, e.g.: 'alpha CMa', 'Jupiter Sol-4', '2P/Encke', 'NGC 2683'*)
=> **src.sample;meta.id**
 - **src.orbital.smajAxis**
=> **src;phys.angSize.smajAxis** ? (Note: «angSize» implies sky observation, not 3D measurement)

Suggestions and questions

Space Physics

- ❖ Input from CDPP (Toulouse, France), LESIA (Meudon, France), IWF (Graz, Austria)
- ❖ Set of new UCDs proposed for evaluation :
 - **phys.count** (*same as phot.count, but for anything else than photons*)
 - **phys.particle**
 - **phys.particle.aerosol**
 - **phys.particle.alpha**
 - **phys.particle.atom**
 - **phys.particle.dust**
 - **phys.particle.electron**
 - **phys.particle.ion**
 - **em.pw** (*local plasma waves*)
 - **phys.energy.flux** (*instead of phot.energy.flux ?*)
 - **phys.flow** (*relative to flow of particles or matter*)
 - **phys.gyrofrequency**
 - **phys.plasmafrequency**
 - **phys.heatflux**
 - **phys.phaseSpaceDensity**
 - «**em.radio below 20 MHz ?**»
- ❖ Energy bands for particle (electrons, ions or neutral) measurement ? (in eV / keV / MeV) Same for Mass spectroscopy (in atomic mass unit) ? and for «Mass per charge» ?

Suggestions and questions Imaging and Spectroscopy

- ❖ Input from LESIA (Meudon, France)
- ❖ Set of new UCDs proposed for evaluation :
 - **em.UV.EUV** (*next to em.UV.FUV, but for anything else than photons*)
 - **em.band** (*similarly to em.line, but for molecular bands*)
 - **em.band.CH4**
 - **em.band.H2O**
 - **em.band.CO2**
 - **em.band....** (*many more possible, very long list !*)
 - **meta.id.CoPI** (*similarly to meta.id.coI and meta.id.PI*)
 - **meta.processed** (*obtained through a processing pipeline*)
 - **meta.derived** (*obtained from a combination of observation and models*)
 - **em.molecline.rotation**
 - **em.molecline.vibration**
 - **obs.calib.dark**
 - **phot.radiance**
 - **phot.reflectance**
 - **pos.occult => pos.limb;obs.occult**
 - **src.orbital.smajAxis** \neq **phys.angSize.smajAxis**
 - **src.orbital.sminAxis** \neq **phys.angSize.sminAxis**
 - **src.orbital.number** (*number of the current revolution*)
 - **time.period.number** (*number of the current rotation, e.g. day number on Earth*)

Suggestions and questions

Solar and Heliophysics

- ❖ Input from HELIO Project (Europe)
- ❖ Set of new UCDs proposed for evaluation :
 - **time.period.number** (*number of the current rotation, e.g. number of the Carrington rotation of the Sun*)
 - **instr.obsty.experiment**
 - **pos.heliographic** (*centered on the center of the sun as seen from observer*)
- ❖ Note: ongoing job.

UCD Concluding remarks

- ❖ Ongoing work !
- ❖ More new cases will come soon as we start up new database ingestion into EPN-TAP services (Europlanet Project)
- ❖ Each suggestion will be discussed using available tools and documents. We will always check if adapted existing UCD can be used.
- ❖ Often, the proposed new UCD are close to utype (as a reference to a DM element).
- ❖ Up to now, we found no need for new «main category».

Example of Utype in VOTable for Planetary Radio Astronomy

```
<GROUP ID="freq_table" utype="spec:Char.SpectralAxis">
  <PARAM value="Frequency" datatype="char" arraysize="" name="Spectral Axis Name"
    utype="spec:Char.SpectralAxis.Name"/>
  <PARAM datatype="int" name="Number of Frequencies" ucd="meta.number" value="400"
    utype="spec:Length"/>
  <PARAM value="10.0000" datatype="float" name="Minimum Frequency" unit="MHz"
    ucd="em.freq;stat.min" utype="spec:Char.SpectralAxis.Coverage.Bounds.Range.Min"/>
  <PARAM value="40.0000" datatype="float" name="Maximum Frequency" unit="MHz"
    ucd="em.freq;stat.max" utype="spec:Char.SpectralAxis.Coverage.Bounds.Range.Max"/>
  <PARAM datatype="float" name="Frequency" unit="MHz" ucd="em.freq" arraysize="400"
    utype="spec:Char.SpectralAxis.Coverage.Location.Value"
    value="10.0000 10.0750 10.1500 10.2250 10.3000 10.3750 10.4500 10.5250
    10.6000 10.6750 10.7500 10.8250 10.9000 10.9750 11.0500 11.1250 11.2000
    .....
    38.9500 39.0250 39.1000 39.1750 39.2500 39.3250 39.4000 39.4750 39.5500
    39.6250 39.7000 39.7750 39.8500 39.9250"/>
  <PARAM value="75.0000" datatype="float" name="Frequency Sampling Step" unit="kHz"
    ucd="em.freq" utype="spec:Char.SpectralAxis.Accuracy.BinSize"/>
  <PARAM value="30.0000" datatype="float" name="Integration Bandwidth" unit="kHz"
    ucd="em.freq" utype="spec:Char.SpectralAxis.Resolution"/>
</GROUP>
```

Spectral Axis group

Name + Length

difference with char:numbins ?

Bounds

List of frequency values

Sampling and filter size

not clear if those 2 last utypes are correct...

We propose here use uType attributes to identify parameters

Example of Utype in VOTable for Planetary Radio Astronomy

```
<GROUP ID="freq_table" utype="spec:Char.SpectralAxis">
  <PARAM value="Frequency" datatype="char" arraysize="*" name="Spectral Axis Name"
    utype="char:Axis.Name"/>
  <PARAM datatype="int" name="Number of Frequencies" ucd="meta.number" value="400"
    utype="spec:Length"/>
  <PARAM value="10.0000" datatype="float" name="Minimum Frequency" unit="MHz"
    ucd="em.freq;stat.min" utype="char:Axis.Coverage.Bounds.Range.Min"/>
  <PARAM value="40.0000" datatype="float" name="Maximum Frequency" unit="MHz"
    ucd="em.freq;stat.max" utype="char:Axis.Coverage.Bounds.Range.Max"/>
  <PARAM datatype="float" name="Frequency" unit="MHz" ucd="em.freq" arraysize="400"
    utype="char:Axis.Coverage.Location.Value"
    value="10.0000 10.0750 10.1500 10.2250 10.3000 10.3750 10.4500 10.5250
    10.6000 10.6750 10.7500 10.8250 10.9000 10.9750 11.0500 11.1250 11.2000
    .....
    38.9500 39.0250 39.1000 39.1750 39.2500 39.3250 39.4000 39.4750 39.5500
    39.6250 39.7000 39.7750 39.8500 39.9250"/>
  <PARAM value="75.0000" datatype="float" name="Frequency Sampling Step" unit="kHz"
    ucd="em.freq" utype="char:Axis.Accuracy.BinSize"/>
  <PARAM value="30.0000" datatype="float" name="Integration Bandwidth" unit="kHz"
    ucd="em.freq" utype="char:Axis.Resolution"/>
</GROUP>
```

Other representation probably closer what I understood from the utypes discussion

Testing Utypes in NcML

(NetCDF-XML representation)

We want to describe the same information in a NetCDF environment
NB: NcML schema has only 4 elements: dimension, group, attribute, variable.

```
<attribute name="namespace:spec">http://www.ivoa.net/xml/Spectrum/Spectrum-1.01.xsd</attribute>
<dimension name="dim_freq" length="400"/>
<dimension name="dummy1" length="1"/>
<group name="freq_table">
  <attribute name="ID">freq_table</attribute>
  <attribute name="utype">spec:Char.SpectralAxis</attribute>
  <variable name="Minimum Frequency" type="double" shape="dummy1">
    <attribute name="utype">spec:Char.SpectralAxis.Coverage.Bounds.Range.Min</attribute>
    <attribute name="unit">MHz</attribute>
    <attribute name="ucd">em.freq;stat.min</attribute>
    <values>10.0000</values>
  </variable>
  <variable name="Maximum Frequency" type="double" shape="dummy1">
    <attribute name="utype">spec:Char.SpectralAxis.Coverage.Bounds.Range.Max</attribute>
    <attribute name="unit">MHz</attribute>
    <attribute name="ucd">em.freq;stat.max</attribute>
    <values>40.0000</values>
  </variable>
  ...
```

Spectral Axis group

Bounds

Testing Utypes in NcML (NetCDF-XML representation)

...

```
<variable name="Frequency" shape="dim_freq" type="double">  
  <attribute name="utype">spec:Char.SpectralAxis.Coverage.Location.Value</attribute>  
  <attribute name="unit">MHz</attribute> List of frequency values  
  <attribute name="ucd">em.freq</attribute>  
  <values>10.0000 10.0750 10.1500 10.2250 10.3000 10.3750 10.4500 10.5250 10.6000 10.6750  
    10.7500 10.8250 10.9000 10.9750 11.0500 11.1250 11.2000 11.2750 11.3500 11.4250  
    .....  
    39.2500 39.3250 39.4000 39.4750 39.5500 39.6250 39.7000 39.7750 39.8500  
    39.9250</values>  
</variable>  
<variable name="Frequency Sampling Step" shape="dummy1" type="double">  
  <attribute name="utype">spec:Char.SpectralAxis.Accuracy.BinSize</attribute>  
  <attribute name="ucd">em.freq</attribute>  
  <attribute name="unit">kHz</attribute> Sampling and filter size  
  <values>75.0000</values>  
</variable>  
<variable name="Integration Bandwidth" shape="dummy1" type="double">  
  <attribute name="utype">spec:Char.SpectralAxis.Resolution</attribute>  
  <attribute name="ucd">em.freq</attribute>  
  <attribute name="unit">kHz</attribute>  
  <values>30.0000</values>  
</variable>  
</group>
```

Using utypes to link to DM outside IVOA

```
xmlns:spase="http://www.spase-group.org/data/schema"
```

...

```
<GROUP name="Time Range" utype="spase:Catalog.TimeSpan">  
  <PARAM arraysize="" datatype="char" name="Catalog Start Time" ucd="time.start"  
    utype="spase:Catalog.TimeSpan.StartDate" value="2008-09-24T18:00:00.000Z"  
    xtype="dateTime"/>  
  <PARAM arraysize="" datatype="char" name="Catalog Stop Time" ucd="time.stop"  
    utype="spase:Catalog.TimeSpan.StopDate" value="2008-09-24T18:59:99.999Z"  
    xtype="dateTime"/>  
</GROUP>  
<GROUP name="Contact" ucd="meta.id">  
  <PARAM arraysize="" datatype="char" name="Name" utype="spase:Person.PersonName"  
    value="Baptiste Cecconi"/>  
  <PARAM arraysize="" datatype="char" name="SPASE Person ID"  
    utype="spase:Person.ResourceID" value="spase://SMWG/Person/Baptiste.Cecconi"/>  
</GROUP>
```

No show-stopper found yet.

Missing feature: *being able to «point» to DMs from several community in order to be compatible with tools from each community.*

=> what about «multiple» utypes ?

Conclusions & Perspectives

- ❖ UCD are working with solar system (planets and sun), with minor additions
- ❖ UCD definition are sometimes not adapted to in-situ measurement (*e.g.: measured electron temperature vs. derived electron temperature*)
- ❖ Tests on utypes (old fashioned) are promising when trying to use Spectrum DM v1.01 (tests to be done with v2.0)
- ❖ Ongoing serialization tests of IVOA DM in other data format, such as NetCDF.