Un modèle de données pour les sciences planétaires

B. Cecconi et beaucoup d'autres personnes du projet Europlanet-IDIS

12 Mars 2013. Réunion ASOV, Observatoire De Paris

Contexte

* Europlanet/IDIS: construire un prototype d'OV planéto européen

* Plusieurs thématiques scientifiques: *atmosphères, surfaces, interieurs, petits corps, paramètres orbitaux, exploration in situ, plasma (ondes, particule et champs continus), radioastronomie...*

* Grande variété de type de données: *images, spectres, séries temporelles, films, spectres dynamiques, profils, cartes...*

* Encore plus grande variété de paramètres physique !

* Cela inclut: *mesures à distance ou in-situ, modèles, expériences de labo, analogues terrestres...*

Architecture
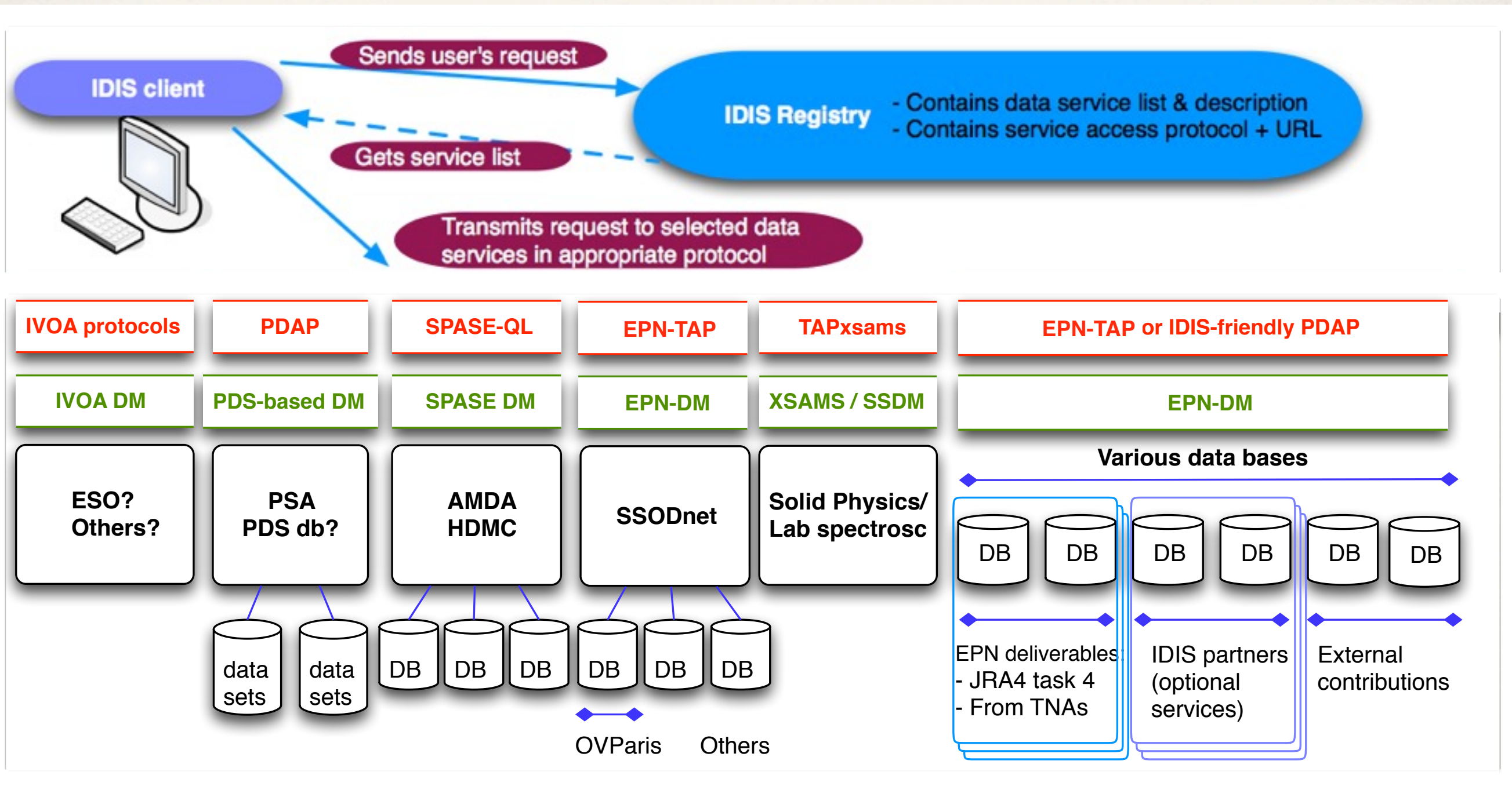
Qu'est qui existait pour les sciences planétaires ?

✤ IPDA (International Planetary Data Alliance):
- proche de PDS4 (nouvelle version du NASA Planetary Data System, en XML et interopérable), mais pas encore prêt.
- PDAP (Planetary Data Access Protocol)
=> Format et protocoles orientés archive.

✤ IVOA (International Virtual Observatory Alliance):
- Plein de modèles (*Characterisation, VOResource, VODataService, Space Time Coordinates...*)
- TAP (Table Access Protocol): *protocole générique pour accéder à des tables*
- Il existe un Registry

- SAMP (Simple Application Messaging Protocol): *voir les autres présentations.*
=> Modèles génériques (attention: coordonnées célestes pas adaptées)

✤ VAMDC (Virtual Atomic and Molecular Data Centre)
- XSAMS: Data Model pour les données spectroscopie atomique et moléculaire

✤ HDMC (Heliophysics Data and Model Consortium):
- SPASE (Space Physics Archive Search and Extract): Data model et langage de requête

✤ OGC-GIS (Open Geospatial Consortium):
- pas encore complètement testé...

Data Model
Métadonnées requises

✤ On veut décrire le contenu des Datasets et des Granules (i.e., les produits, ou fichiers, ou le niveau de granularité le plus petit que fournit un service). On ne décrit pas le format des fichiers.

   ✤ Identification de la Ressource

   ✤ Cible d'observation

   ✤ Instrument (incluant l'hôte de l'instrument)

   ✤ Les Axes (limites, unités, résolution...)

   ✤ Paramètres physiques (Nom, Unités, UCD...)

Data Model
Héritage

* Le Data Model IDIS-DM v1 a été construit à partir de zéro. Simple mais restreint et non-compatible IVOA.

* IDIS-DM v2, appelé Europlanet (**EPN**)-**DM**, est basé sur **VORessource**, **VODataService** and **VODataCollection**. Il utilise d'autres modèles IVOA, comme **STC**, **Utypes**, **UCDs, Characterization**...

* Unités: **EPN-Unit** spécifique (inspiré de Osuna&Salgado 2008)

* On voulait utiliser **ObsCoreDM**, mais ce n'est encore fait (*Je n'ai pas trouvé de schéma XML pour la validation des descripteurs*).

* Version actuelle de la documentation disponible ici:
  http://voparis-europlanet.obspm.fr/docs/PlanetaryScienceResource-DM-latest.pdf

Détails du DM
Data product type

The data product type describes the high level scientific organization of the data product being considered. The list of product values is:

* **Image**: associated scalar fields with two spatial axes, e.g. image with multiple color planes, from multichannel cameras for example.

* **Spectrum**: data product for which the spectral coverage is the primary attribute, e.g. a set of spectra

* **DynamicSpectrum**: consecutive spectral measurements through time, organized as a time series. voir baptiste 1D temp, 1D Spectral

* **SpectralCube**: set of spectral measurements with 1D or 2D spatial coverage, e. g. imaging spectroscopy. The choice between Image and Spectral_cube is related to the characteristics of the instrument

* **Profile**: scalar or vectorized measurements along one spatial dimension, e.g. atmospheric profiles, atmospheric paths, sub-surface profiles, etc.

* **Volume**: any measurement with three spatial dimensions

* **Movie**: set of chronological 2D spatial measurements

* **Cube**: multidimensional data with three or more axes, e.g. all that is not described by other 3D data types such as spectral cubes

* **TimeSeries**: measurements organized primarily as a function of time (with exception of dynamical spectra). A light curve is a typical example of a time series dataset.

* **Catalogue**: it can be a list of events, a catalog of object parameters, a list of feature, ..., e.g. list of asteroid properties

* **SpatialVector**: list of summit coordinates defining a vector, e.g. vector information from GIS, spatial footprints, ...

Comparaison PDAP et EPN-TAP
Mots-clés de requête

PDAP-core Query Keywords (V1.0 2011-Nov-09)	EPN-TAP core Query Keywords
RESOURCE_CLASS -	- RESOURCE_TYPE
	- *DATAPRODUCT_TYPE* (from list)
TARGET_NAME -	- TARGET_NAME
TARGET_TYPE -	- TARGET_CLASS (from list)
START_TIME/STOP_TIME -	- TIME (MIN/MAX)
TIME_SERIES EXTENSION ?	- *SAMPLING STEP* (MIN/MAX)
	- *EXPOSURE TIME* (MIN/MAX)
[WAVELENGTH] -	- SPECTRAL RANGE (MIN/MAX)
SPECTRUM EXTENSION ?	- *SPECTRAL SAMPLING* (MIN/MAX)
	- *SPECTRAL RESOLUTION* (MIN/MAX)
[SPACECRAFT_ALTITUDE] -	- SPATIAL RANGE (C1,C2,C3) (MIN/MAX)
	- *SPATIAL RESOLUTION* (MIN/MAX)
FLY-BY EXTENSION ?	- *SPATIAL FRAME TYPE* (from list)
	- *INCIDENCE ANGLE* (MIN/MAX)
	- *EMERGENCE ANGLE* (MIN/MAX)
INSTRUMENT_NAME -	- INSTRUMENT HOST NAME
INSTRUMENT_HOST_NAME -	- INSTRUMENT NAME
INSTRUMENT_TYPE -	- MEASUREMENT TYPE
RETURN_TYPE -	*(always VOTable)*

EPNcore DataModel

�֍ Restriction/Adaptation de EPN-DM pour TAP: «EPNcore data model»

✤ EPN-TAP = TAP avec des mots-clés spécifiques à la planéto

✤ **EPNCore** : 19 mots-clés obligatoires:
- *resource_type*: 'dataset' ou 'granule'
- *dataproduct_type*: décrit juste avant
- Cible d'observation: *target_name, target_class*
- Temporel: *time_min, time_min, time_exp_min, time_exp_max, time_sampling_step_min, time_sampling_step_max*)
- Spectral: *spectral_range_min, spectral_range_max, spectral_resolution_min, spectral_resolution_max, spectral_sampling_step_min, spectral_sampling_step_max*
- Spatial: *spatial_frame_type ('celestial', 'body', 'cartesian', 'cylindrical', 'spherical'), c1min, c2min, c3min, c1max, c2max, c3max, c1_resol_min, c2_resol_min, c3_resol_min, c1_resol_max, c2_resol_max, c3_resol_max*
- Géométrie d'observation: *incidence_angle, emergence_angle, phase_angle*
- Instrument: *instrument_host_name, instrument_name*
- Paramètre physique: *measurement_type (UCDs)*

✤ **EPNResponse** : description de la réponse d'un service EPN-TAP.

✤ Documentation technique:
`http://voparis-europlanet.obspm.fr/xml/TAPCore/doc/html/`

Centre de données EPN-TAP

✤ **VOParis** Data center, Paris-Meudon
- APIS (Auroral Planetary Imaging and Spectroscopy)
- BDIP (old planetary images database)
- Titan Atmospheric Profiles (from Cassini VIMS data)
- Exoplanet.eu database

✤ **Nançay Radioastronomy** Station
- DAM (Jovian Decametric Radio Emission Routine Observations)

✤ **CDPP**, Toulouse
- AMDA database

✤ **Graz**, Austria
- VEx/MAG data

http://vo.ari.uni-heidelberg.de/soft/dachs

Discussion

- Fin de la phase de test. Pas de gros changement depuis 1 an.

- Utilisation de EPN-TAP à VOParis et CDPP.
  Comparaison avec PDAP.

- Utilisation possible pour l'héliophysique.

- Futur version plus proche de ObsCoreDM ?

Unified Content Descriptors (*UCD*) pour les sciences planétaires et l'héliophysique

B. Cecconi et beaucoup d'autres personnes des projets Europlanet-IDIS et HELIO

12 Mars 2013. Réunion ASOV, Observatoire De Paris

Suggestions and questions Comets and Samples (1)

✤ Input from Italian EPN group (description of planetary samples)

✤ Suggested set of new UCDs:
  - **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 Heliphysics

* 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, just begun.

Remarques conclusives

* Travail en cours !

* De nouveaux cas d'études apparaissent à chaque nouvelle base qui implémente EPN-TAP

* Chaque suggestion sera étudiée avec les outils disponibles. On cherchera toujours à utiliser des UCDs existant avant d'en proposer de nouveaux.

* Pour l'instant: pas de besoin d'une nouvelle «main category».