Proposed architecture for EuroPlaNet-IDIS Virtual Observatory

Draft 0.15b

EPN/JRA4-IDIS/Task 2

Latest version:
Previous version(s): draft 0.15

Editors:
Stéphane Erard (LESIA / VO-Paris)
Contributors:
Stéphane Erard, Baptiste Cecconi, Pierre Le Sidaner, Jérôme Berthier, Chiara Marmo, Michel Gangloff, Nataliya Bourrel, Christian Jacquey, Etienne Pallier, Nicolas André, Nicolas Lormant, B. Schmitt, Florian Topf

---

Abstract

This document outlines the architecture of the Virtual Observatory system for Planetary Science initiated during the Europlanet contract as part of the IDIS (Integrated and Distributed Information Service) activity. The system makes intensive use of the EPN Data Model and EPN-TAP protocol described elsewhere.

Status of this document

This document has been produced from documents approved by the IDIS partners and discussions during IDIS meetings. Later developments of the EPN-TAP protocol, EPN-DM, and EPN-TAP based services have been integrated.

Acknowledgements

This work has been conducted in the frame of Europlanet-RI JRA4 work package.
The EuroPlaNet-RI project is funded by the European Commission under the 7th Framework Program, grant #228319 "Capacities Specific Programme".

Document change record

Issue	Date	Description of the modification	contributions
Draft 0.1	08/10/2010	First draft from VO-Paris	S. Erard, LESIA/Obs.Paris (VO-Paris) P. Le Sidaner, DIO/Obs.Paris (VO-Paris) J. Berthier, IMCCE/CNRS (VO-Paris)
Draft 0.2	12/10/2010	Internal iterations before submission to other partners	
Draft 0.3	14/10/2010	internal iterations	
Draft 0.4	29/10/2010	First inputs from CDPP & Grenoble	Same+ B. Cecconi, LESIA/Obs.Paris (CDPP) J. Normand, DIO/Obs. Paris (VO-Paris) B. Schmitt (IPAG)
Draft 0.5	30/10/2010	Adjusted/completed	
Draft 0.6	30/10/2010	Internal iterations	Same+ M. Gangloff, CESR/CNRS (CDPP)
Draft 0.7	10/01/2011	After VO-Paris / CDPP meeting in Toulouse (22/11)	Same+ CDPP personal
Draft 0.8	19/01/2011	After Atelier VO-planéto, Paris (18/1/2011)	Input from potential users and data providers involved in the French VO project + A. Sarkissian (LATMOS)
Draft 0.9	21/01/2011	internal interactions	
Draft 0.10	24/01/2011	internal interactions	
Draft 0.11	24/01/2011	internal interactions	(version from CDPP) N. André, M. Gangloff, E. Pallier
Draft 0.12	4/02/2011	Extra input from Atelier VO-planéto	Inputs from Chiara Marmo (IDES) Finalization for IDIS approval
Draft 0.13	13/02/2011	Inclusion of name resolver demonstrator studies	Inputs from J. Berthier (IMCCE/VO-Paris)
Draft 0.13b	13/02/2011	Finalization for IDIS approval	SE
Draft 0.14	13/01/2013 18/07/2013	Reworked after finalization of EPN-TAP Updated appendices / VO tools	SE SE
Draft 0.15	23/01/2013	Global update. Solutions adopted for registry & EPN-TAP + reference to IVOA standards Reference list	SE

Reference documents

[RD1] Planetary data access protocol (PDAP). IPDA draft 1.0 (16 April 2013)
https://planetarydata.org/projects/active-projects-2012-2013/PDAP%20Core%20Specification%20

[RD2] Virtual Observatory Support Interface (VOSI)
http://ivoa.net/Documents/VOSI/20101206/index.html

[RD3] GAVO / DaCHS implementation:
http://vo.ari.uni-heidelberg.de/docs/DaCHS /

[RD4] EPN data model version 1.18a (last version to date) can be found here:
http://www.europlanet-idis.fi/documents/public_documents/Data_Model_v1.18a.pdf

[RD5] TAP protocol
http://ivoa.net/Documents/TAP/

[RD6] UCD + UType concept
http://ivoa.net/Documents/cover/UCDlist-20070402.html

[RD7] Name resolver returning body official names and astronomical coordinates at a specific time:
http://vo.imcce.fr/webservices/ssodnet/?resolver

[RD8] Report of the IAU Working Group on Cartographic Coordinates and Rotational Elements: 2009 B. A. Archinal et al, Celest Mech Dyn Astr (2011) 109:101–135.
+ Erratum to: Reports of the IAU Working Group on Cartographic Coordinates and Rotational Elements: 2006 & 2009. B. A. Archinal et al, Celest Mech Dyn Astr (2011) 110:401–403.

[RD9] Space time and coordinate in IVOA
http://ivoa.net/Documents/latest/STC.html

[RD10] IVOA Registry interface
http://ivoa.net/Documents/RegistryInterface/

[RD11] Unit in the IVOA (current draft)
http://ivoa.net/Documents/VOUnits/

[RD12] EPN-RI Coordinate systems, V0.2

[RD13] EPN-TAP documentation:
http://voparis-europlanet.obspm.fr/xml/TAPCore/

[RD14] ObsTAP and ObsCore
http://ivoa.net/Documents/ObsCore/

Document change record 2
Reference documents 3
Acronym and definition list 6
1 Introduction 7
1.1 Functions of a Planetary Science VO 7
2 Proposed architecture for a Planetary Science VO 8
2.1 Data scope 8
2.2 Data Access Protocols and Data Models 9
2.3 Registry, client 11
2.4 Answers to queries 12
2.5 Data format, reading and writing 12
2.6 Visualization tools 13
2.7 More elaborate functions 14
2.8 Mapping, mosaicking 15
2.9 Interfaces 15
3 Practical implementation 17
3.1 Inclusion of new resources 17
3.2 Queries handling 17
Appendix A – Demonstrators 19
A1 Queries of planetary data on web portals 19
A2 PDS on-line access 19
A3 Service validator 20
A4 Name resolver 20
A5 XML descriptor service 20
A6 Demonstrators in progress 20
Appendix B - IVOA tools potentially useful for EPN-IDIS 22
B1 Visualization tools 22
B2 Data exchange: protocols and exchange formats 23
B3 Registries management 23
B4 Frameworks 24
Appendix C - Other tools of interest 25
C1 Geographic Information Systems 25
C2 Data handling and visualization 25
C3 Dynamical visualization 26

C4 Computational environments 26

Acronym and definition list

EPN, Europlanet-RI	Europlanet Research Infrastructure (EC's FP7)
Europlanet-IDIS	Europlanet Integrated and Distributed Information Service
JRAx	Join Research Activity #x
VO	Virtual Observatory
PDS	Planetary Data System
PSA	Planetary Science Archive
DARTS	Data ARchive and Transmission System
IVOA	International Virtual Observatory Alliance
IPDA	International Planetary Data Alliance
PDAP	Planetary Data Access Protocol
PSR-DM	Planetary Science Resource Data Model
CDPP	Centre de Données de la Physique des Plasmas
SSDM	Solid Spectroscopy Data Model
VAMDC	Virtual Atomic and Molecular Data Centre
SSODNET	Solar System Objects Data Network
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
ESA	European Space Agency
FITS	Flexible Image Transport System
IDL	Interactive Data Language
GDL	GNU Data Language
ISIS	Integrated Software for Imagers and Spectrometers
USGS	U.S. Geological Survey
NAOA	National Oceanic and Atmospheric Administration
SAMP	Simple Application Messaging Protocol
KML	Keyhole Markup Language
GML	Geography Markup Language
OGC	Open Geospatial Consortium
SPASE	Space Physics Archive Search and Extract
CDS	Centre de Données Astronomiques de Strasbourg
XPA	X Public Access
CSV	Coma Separated Values

1 Introduction

This document is intended as a short, non-technical, description of the Virtual Observatory system initiated during the EuroPlaNet-RI program (JRA4 work package).

A general scheme is described, which identifies the main mechanisms, and where choices have been made. The retained solutions are mentioned, on the basis of preliminary studies conducted at VO-Paris and CDPP, and hints for future developments are identified.

The detailed solutions at each step are detailed in other documents, in particular concerning the EPN-TAP protocol and the EPN Data Model.

Action items are listed at the end of each section.
⌘ Practical recommendations are outlined in yellow and introduced by ⌘

1.1 Functions of a Planetary Science VO

The Virtual Observatory in Astronomy is maintained by a consortium named IVOA. The objective of the Virtual Observatory is to improve access to the data archives, and to provide researchers with standard tools to handle and process these data, allowing them to use a much larger fraction of the observational data produced. A Virtual Observatory in Planetary Science is expected to provide similar services, with a particular gain of time foreseen at the level of data search and quick-look functions, and possibly also at the level of automated data processing. Another expected potential benefit is to enable data producers to distribute small, elaborated, data in an interoperable context sets with minor efforts, so as to maximize their visibility. The multiplication of such small data services will rapidly enlarge the quantity of derived data available in all fields of Solar System research.

A general scheme is described in Fig. 1, which illustrates the sequence of steps in a typical working session. The user works on a specific scientific problem, and is looking for related data. The user is working at his computer, sends queries to data bases related to his problem, and gets answers. Then the data must be loaded in memory, can be plotted in various forms (images, spectra…), and are possibly sent to more elaborated tools performing specialized functions. These steps are commented below, and solutions are discussed.

At each step, light and economical solutions have been searched, so as to minimize the developments. Existing solutions and tools have been recycled as much as possible, also to remain compatible with other fields, in particular in Astronomy or Heliophysics.

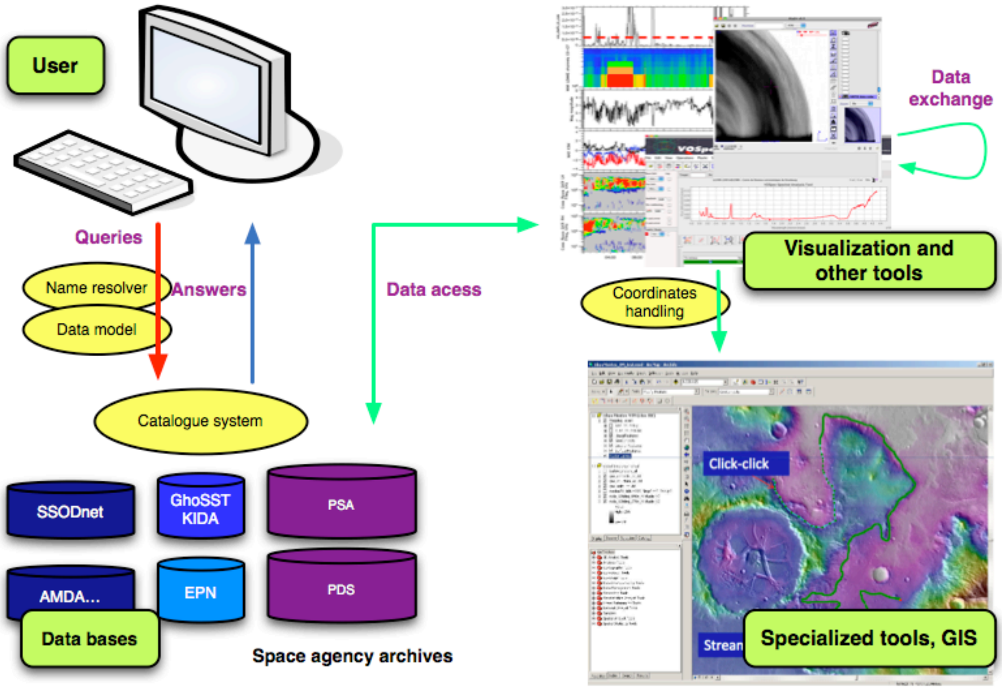

Fig.1: overall scheme of the EuroPlanet VO

2 Proposed architecture for a Planetary Science VO

2.1 Data scope

The perimeter of data to be accessed by the Europlanet VO derives from the objectives stated in the program proposal. It includes (Fig. 2):

- Data bases produced by various work packages during the program (including JRA4/task4, TNA3…)
- A selection of space borne data from planetary missions. This includes data from European space missions, i.e. access to the PSA.
- Data of particular interest selected by SA-IDIS participants to initiate the system.
- Data sets directly provided in a compliant form by data providers, typically as end product of a research activity, after scientific publication.
- Big data repositories including planetary science data and predating Europlanet are also natural targets to expand this system.

A database with standard software access allowing the user to search its contents and returning selected data is called a data service.

The Europlanet VO is intended to be sufficiently open to allow external data providers to include their data services with minimum efforts (for instance, data bases in the first category above). This includes observational data derived from space missions or ground-based telescopes, but also reference data acquired in the laboratory and simulations.

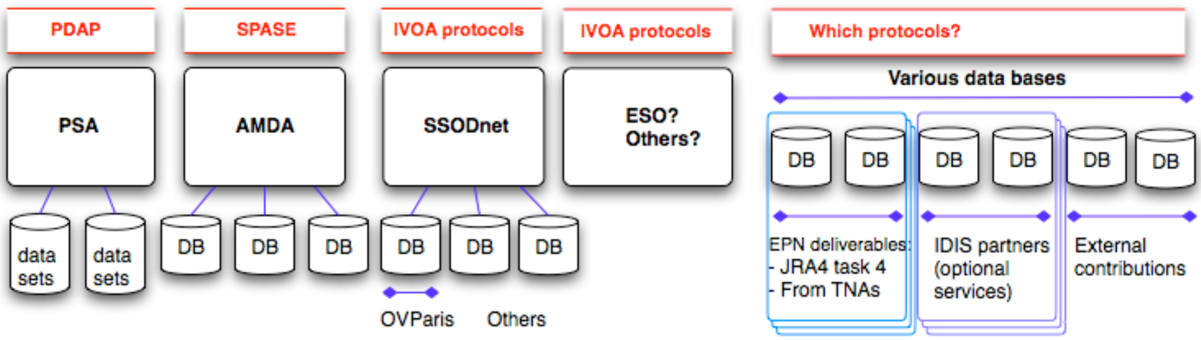

Fig.2: Data expected to be accessible by the EPN VO, and related access protocols

The minimum services to be provided include catalogue search, physical access to the data (i.e. reading the files), and basic visualization functions. It is stressed that there is a high level of heterogeneity in these data, e.g. with respect to astronomical data.

- Formats: although most data formats are readily handled, PDS3-formatted data from space missions are difficult to read due to the lack of standard/versatile software, and require a special input process. Many specific formats are also in use, e.g. CDF in plasma physics.
- Coordinate frames: sky images exist, but in contrast with astronomy the targets are moving and can be identified from their coordinates only at a given time. Most data are located on planetary surfaces, atmospheres, or interiors, and are described using specific coordinate frames.
- Nature: 1D/2D/3D/4D, bitmap vs vectorial, local vs event-related, observational vs laboratory vs modeling, celestial vs HR orbital images, electro-magnetic/light vs particles and other signals... Some of those call for particular display modes.
- Variability: many data experience temporal variations at different time scales (secular, seasonal, local time).

Studies at VO-Paris were based on telescopic and space-borne images and spectral cubes, plus laboratory spectra and simulations or computations (such as ephemeris), and time series (electro-magnetic parameters). Studies at CDPP were based on space plasma data of various nature (time series...) and format (CDPP, CDF ...). This perimeter is intended to gradually enlarge to encompass a broader field in Planetary Science.

2.2 Data Access Protocols and Data Models

Data search is performed by sending queries to data services. Such queries must be translated in a standard form and used to search a catalogue of available data. This procedure is called a Data Access Protocol in the IVOA framework.

Queries could be sent from a web form, from specialized software, or directly from visualization tools, as it is currently done in the IVOA system (e. g., from Aladin or TopCat). However, this requires the protocols to be implemented in the visualization tools, and therefore involves heavy developments.

Several protocols are used in different contexts, in particular:

- IVOA protocols, which allow searching the data according to various criteria (i.e., Cone Search to locate objects near a position in the sky, TAP for tabular data [RD5]...).
- The PDAP protocol, currently being defined in the IPDA framework [1]. A typical use of PDAP is to address the entire contents of the PSA by querying their database globally, as a single data service.
- The Spase-QL protocol used in plasma physics, e.g., to access the AMDA service at CDPP.

These solutions were studied to figure out whether they could answer the needs of the Europlanet VO. Most IVOA protocols appear to be related to objects located on the celestial sphere, and no solution currently exists to address data located e.g. on a planetary surface or in planetary atmospheres. Usability assessments of PDAP were performed at VO-Paris and CDPP. The protocol proved able to address the data of interest, provided topical extensions and perhaps some adjustments; however, PDAP is closely related to an implicit Data Model associated to the PDS3 format, therefore adapted to space borne data only and providing little support for telescopic observations or laboratory data. Another drawback is that PDAP is not implemented in the IVOA visualization tools, as IVOA protocols are.

The outcome of this study clearly identified the need to develop a simple protocol based on existing ones, but including specific parameters permitting to handle Planetary Science data in general. The selected solution was to define a restriction of the TAP protocol called EPN-TAP. The TAP mechanism is associated to a simple, specific Data Model (EPNcore, [RD4]). This Data Model is in practice a set of mandatory parameters to describe a data set, which is both implemented in all data services and used as query parameters by the protocol [RD13]. Most of these parameters are related to the description of the data axes (time, coordinates, spectral), target, measurement type, and origin of the data. The protocol also states that quantitative parameters have to be accessible in standard units in the catalogue, and that non-quantitative parameters (such as instrument names) are associated with identified reference lists. Most of these reference lists originate from the IAU or other independent sources such as NSSDC. This standardization makes it possible to send uniform queries to all EPN-TAP services.

A more general Data Model has also been defined to describe the data of interest in a uniform way (Planetary Science Resource Data Model, or PSR-DM). This Data Model could be used in the future with a more general protocol, or to describe data services in a registry. Such applications should use common metadata defined in the data model and the associated dictionary.

Specific domains may not be included in a single data model though, and can be handled with different systems. In particular, other data models are being defined in relation with laboratory databases of solid spectroscopy (SSDM for the GhoSST data service) and atomic and molecular spectroscopy of gases (XSAMS), both in collaboration with the VAMDC consortium, which provides access to Atomic and Molecular Data.

It may sometimes be difficult to match a user's query, mostly because of inadequate values provided either in the request or in the services. A name resolver is therefore required to process user's queries correctly, in particular for target names. Such a name resolver has been devised at VO-Paris as part of the SSODnet project, and is available on-line (see Appendix).

Actions:

• Propose extensions to EPN data model(s) from various fields. Identify necessary keywords common to SSDM and EPN data model to ensure EPN-SSDM interoperability.

• Develop/extend target name resolver (perhaps as a PDAP resource class).

• Consider interfaces with IVOA, HELIO, and other related programs such as VAMDC. This concerns e.g. coordinate systems [RD8, RD9, RD12] and UCDs [RD6].

Other activities:

• The EPN JRA4 partners also got involved in the Technical Expert Group of IPDA, proposing extensions to the PDAP core system to handle spectroscopy and time series.

2.3 Registry, client

A global EPN client has been developed at VO-Paris to send queries to both EPN-TAP and PDAP services. It includes a user interface, queries/answers handling, and transmission to standard visualization tools (Fig.3).

Queries should be sent to a global catalogue containing a description of all accessible data services and their capabilities. This description is provided according to the above data model, and allows the system to make a first order selection of services matching the user's query. In the IVOA, this is done using a system of mirrored registries where data providers can register their data. Declaring a new data service in the registry system is the normal way to make it available in the VO, and to publicize it [RD10]. Although the PDAP 1.0 document mentions a possible registry system, the level of detail involved is still unclear. IVOA-like registries exist at VO-Paris and ESAC. The VO-Paris registry is currently used to access both EPN-TAP and VAMDC resources.

We call EPN Service Data Model the data model used to describe the content of the EPN registry itself, i.e. all data services accessible through the EPN registry. The solution adopted for the Europlanet VO follows the IVOA scheme, which is easy to maintain.

The EPN registry only contains a short description of the available data services, including their address, and the protocols they support. The detailed description of the services is stored and maintained locally by the data providers. The queries are therefore processed at two successive levels (Fig.3): services of interest are identified at the registry level, the query is passed to them according to the protocol indicated, and is processed locally. In this context, a data service is a series of data sets located in a given place, accessible through the same protocol and described in a unique, local, catalogue.

The registry system needs to be maintained far beyond the program lifetime. Compatibility issues with other information systems are best anticipated by adopting the OAI-PMH standard, which is the widely used standard for metadata exchange (used in particular by the IVOA – this is a strong recommendation from the astronomy community). This standard includes the use of the "Dublin core" metadata.

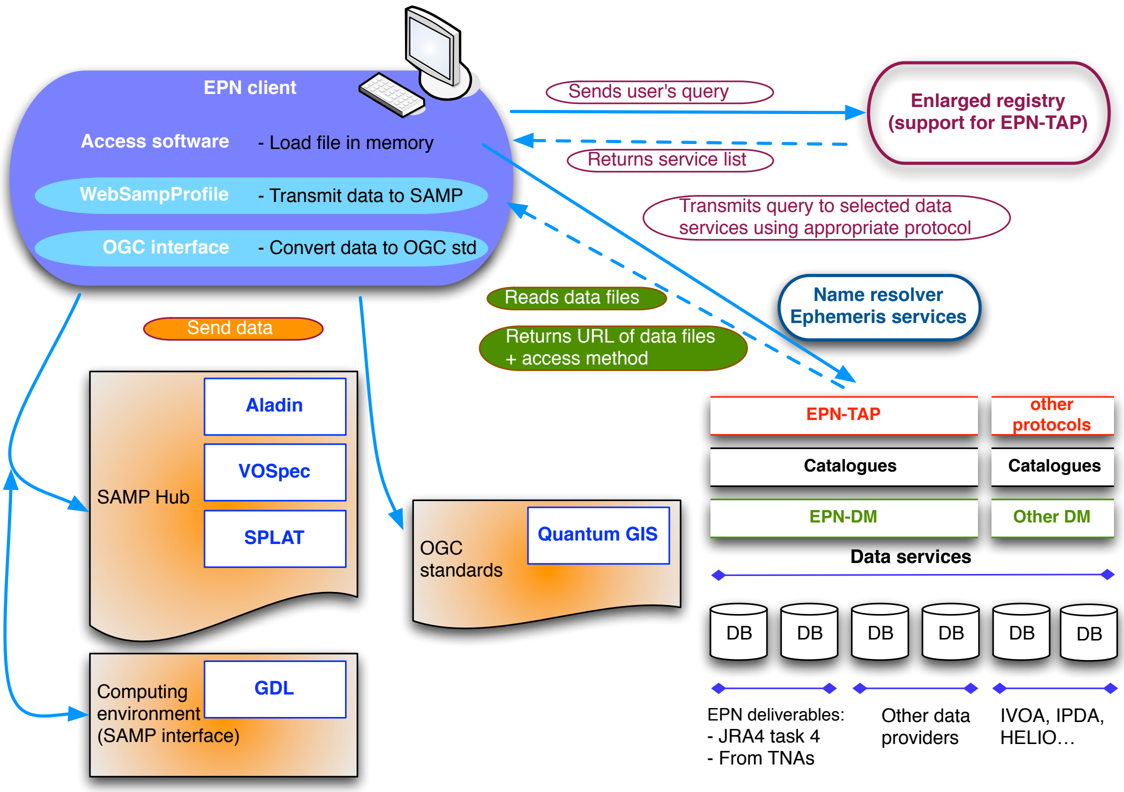

Fig.3: Data access architecture

2.4 Answers to queries

The answer to a query is sent back by the service to the user. It typically provides a description, a link to the data, and information about data interface, but in general not the data themselves (only metadata). To be directly usable, this answer must also be formatted according to a known standard. This formatting can be either related to a data model (one of the solutions currently adopted by the VAMDC program), or remain at a general level (e.g. VOTable).

The EPN-TAP protocol states that the answer is returned in a VOTable, like PDAP and most IVOA protocols. In this case, a URL to the selected data is usually embedded in the VOTable. In some cases when the data are directly integrated in the database view, they can be included in the result VOtable and have to be parsed.

2.5 Data format, reading and writing

1) Reading

Reading the data can be practically difficult, given the wide variety of formats used in Planetary Science.

A special problem arises with PDS3 data (used in all current space data archives), for which no versatile reader is available. Although PDS4 is expected to greatly simplify this issue in the future, availability of “historical” space data in PDS4 is an open question.

Several solutions may be adopted depending on the data:

- FITS data are readily accessible, including from IVOA compatible visualization tools.
- CDF data can be read in TopCat.
- Generic PDS3 data can only be read in the IDL (or GDL) environment. A demonstrator has been developed at VO-Paris to read the data in a local shell, write them back in FITS, then transmit them through the SAMP protocol to IVOA tools (see Appendix A).
- More specific formats need to be addressed case by case.

2) Writing

In terms of output, using the native format of the data is the baseline. In case a format can be arbitrarily selected, the general recommendation is to use only standard data formats commonly used in the field. Two specific recommendations have been identified:

⌘ - Due to the lack of official I/O software and to the fuzzy definition of PDS-compliance, PDS3 is strongly discouraged as an output format of data services outside the space agencies. FITS embedded in PDS3 is manageable, though.

⌘ - Light data files (such as vectors or small tables) can be transferred in ascii form whenever relevant. In this case, VOTable formatting is convenient because it is self-described and can be parsed directly by software, and first of all by visualization tools. However, this has proven to be a practical difficulty for direct data handling, and a standard ascii format (such as CSV) must *always* be available for direct use outside a VO context.

Actions:

- Develop PDS access to address other data types. Chain with PDAP search function on a use case (See Appendix A).

2.6 Visualization tools

Many visualization tools for basic data types were developed in the IVOA framework. A selection of tools of interest for EPN is given in Appendix B. Because development is a heavy activity, the baseline for EPN is to use only existing tools, maintained by external teams, even if this means developing interface layers on the EPN side or discussing evolutions with the developers.

- Image plotting is mainly performed in X/Y coordinates, but some tools can handle sky coordinates, which can also be used as a first order approximation of coordinate systems on planetary surfaces.

- Orthorectified images and maps need to be handled by specialized tools. OGS standards may apply in this case (see Appendix C).
- Specific data types may require special visualization tools, e.g. spectral laboratory data.

IVOA visualization tools implement a data exchange protocol called SAMP. Once the data are loaded by one of these tools, they are readily available to other tools through this protocol.

⌘ New tools developed in the EPN framework should implement the SAMP protocol (XML-RPC based). Web services can use the WebSampProfile interface to share data via the SAMP Hub.

2.7 More elaborate functions

In addition to data services, resources can also include computational services. Those may include standard computation/inversion algorithms. The question arises of the environment used to perform such operations. Several possibilities need to be studied:

- The Aladin plug-in system allows developers to implement basic operations (such as computing the average spectrum of a region of interest). These plug-ins are written in Java.
- IDL can exchange data with Aladin in some operating systems (not all, though). Compatibility with IDL or GDL would give access to a very wide range of applications available in public IDL libraries. This would also provide a link with ENVI, which is widely used for surface studies and imaging spectrometry (see Appendix C).
- IRAF includes VO functions since v. 2.16 (support for VOTable, registry, SAMP...).
- ISIS is a classical environment dedicated to planetary cartography, maintained by the USGS.
- The Orfeo Tool Box is a very large open source library of image processing algorithms developed by CNES for its Earth observation program. The French project Vahiné aims at providing a graphical interface to this library for Planetary Science, and to include data access capabilities.
- Other environments have been developed by space agencies for convenient data representation, such as 3Dview (CNES) or MATISSE (ASI), and are more or less accessible to the general user.
- The IVOA has developed several systems to handle automated workflows, including UWS.

Actions:

- Study compatibility with IDL, either to develop VO workflows in an IDL shell, or simply to allow the user to easily retrieve data from VO applications into IDL/ENVI to proceed with their treatments.
- Study possible link with Orfeo / Vahiné.

2.8 Mapping, mosaicking

Many planetary data need to be projected in a particular coordinate system. Different situations may occur:

- Sky coordinates, used e.g. in telescopic images of a target on background sky. This situation is expected to be handled by Aladin and similar tools. The IMCCE Skybot service is available from Aladin to identify moving targets in a Field of View at a given time.
- Planetary coordinates, e.g. for orbital images of planetary surfaces. This is similar to geographical coordinates on Earth. Apart from the geometric computation (which is expected to lie on the provider's side), plotting in such frames may also be an issue. High resolution imaging in particular requires a detailed description of planetary coordinates frames, including control point networks, which currently does not seem to be available. Converters between coordinate systems may also need to be developed.

Among the functions of interest to EPN, accurate registering of imaging data has a special importance. A large fraction of our community works with GIS to handle orthorectified data, either bitmap or vectorial. The GIS community has developed tools using standards elaborated in the framework of the Open Geospatial Consortium (OGC). These standards are discussed below (see Appendix C).

An open question is how imaging spectroscopy data can be handled in this context — such data sets often use extended pixels, in particular for comparison with high resolution images. More generally, no standard currently exist to provide coordinates together with the data.

Another question concerns the ability of GIS to handle images with arbitrary projections (in general, GIS tools expect images using conventional projections depending on map size and location, according to Earth standards). A potential issue is the general use of tiled, pre-projected, mosaics in common GIS software.

Actions:

• Define coordinate frame description in general, identify possible values for each target body, and include it in PDAP and other protocols.

• Study GIS standards, in particular Web Map Service (WMS) and Keyhole Markup Language (KML).

• Study VO-GIS interoperability. A minimum goal should be to allow sending VO data into a GIS when relevant. This implies studying VOTable to GML/KML translation.

• Similar situations requiring a link between a VO and other information systems may occur in specific fields, yet to be identified.

2.9 Interfaces

The EPN VO system currently relies mostly on adapted IVOA standards (Fig. 4).

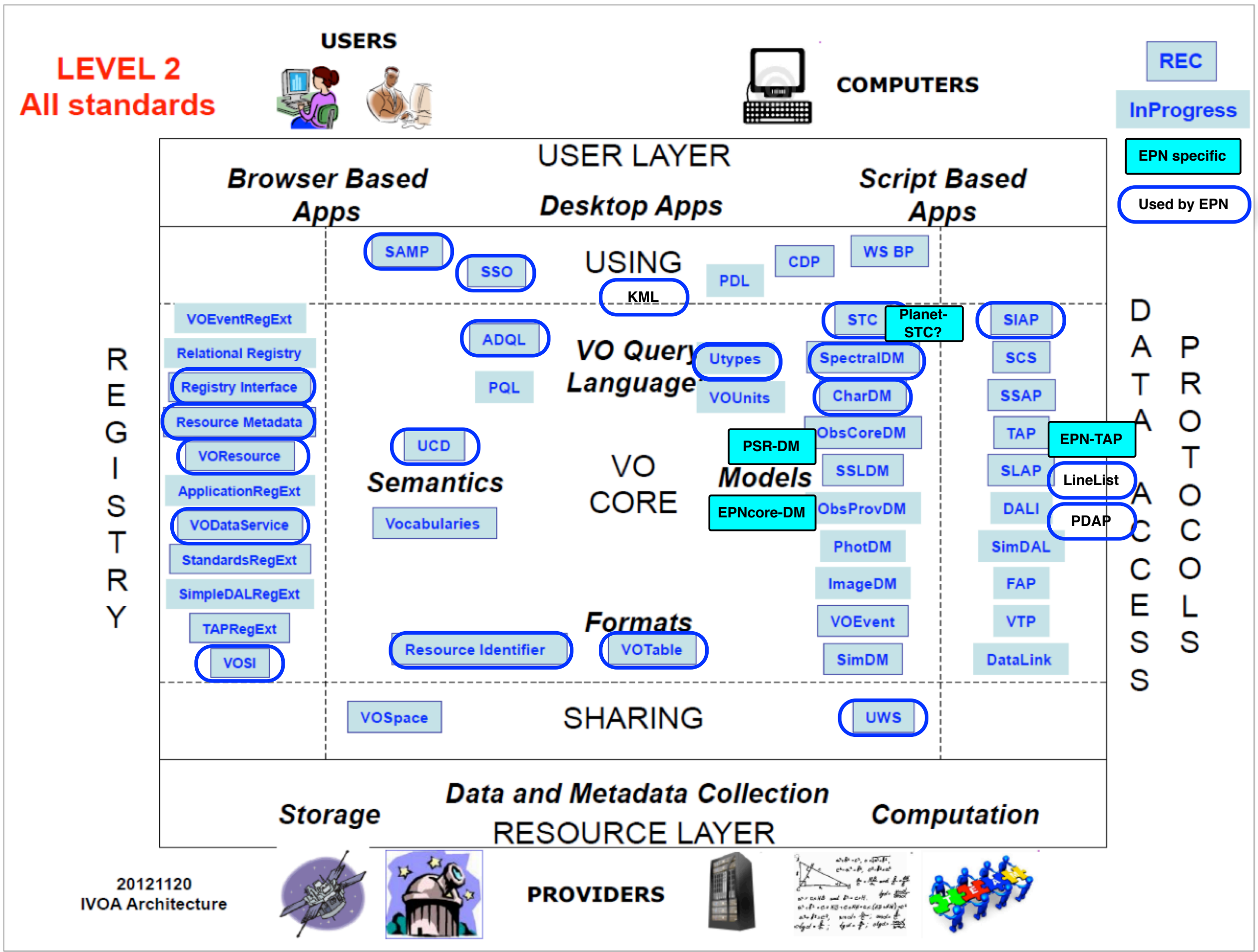

Fig.4: IVOA standards used in EPN VO

Additions consist in:
- Two Data Models for Planetary Science data in general: EPNCore-DM which is a minimal, descriptive DM, and PSR-DM which is a more comprehensive DM to describe datasets.
- The EPN-TAP protocol, which is mainly the TAP mechanism with a set of constraints based on EPNCore-DM. This is similar to the IVOA ObsTAP protocol based on ObsCore-DM [RD14].
- A specific description of coordinate frames in use for the Solar System. Although this could eventually be proposed as an extension to IVOA's STC, this description may actually be difficult to merge with this more general standard.
- The IVOA registries are used with extension to support extra namespaces related to Planetary Science, as well as VAMDC services.

Other standards come from other communities:
- The IPDA PDAP protocol is supported in its current version [RD1], although it may not be implemented on the current test services.
- The LineList protocol from VAMDC is accessible via the SpecView tool.
- OGC standards are also considered, starting with KML to provide footprints on planetary surfaces.

3 Practical implementation

3.1 Inclusion of new resources

New resources can be included in EPN in a VO-interoperable manner by any data provider. The process is the following:

- Set up an SQL database + install a framework supporting the TAP protocol (DaCHS and VO-Dance have been used successfully). A tutorial to install such a system is available on VO-Paris EPN site.
- Ingest your data and create a view named EPN_core describing all the “granules”. Your database does not need to be changed, but this view must contain all the mandatory parameters of EPN-TAP in the standard units.
- The data themselves can be either linked through the “access_url” parameter, or included in the view.
- Set up the framework to provide answers as a VOTable with compliant content.
- Fill up an XML file describing the data service and put it in the VO-Paris registry (see demonstrators in Appendix). Your service will be available from any EPN client, in particular the one at VO-Paris. It will also be accessible from TopCat directly (see doc on the VO-Paris EPN site).

A single DaCHS installation can accommodate several services, so grouping services may be a solution for small teams. Support can be provided at all these steps, e.g. efficient tools are available to write the XML descriptors (e.g., CDPP demonstrator, see Appendix A), to check service validity, and to register new data services.

3.2 Queries handling

User queries will be handled in several steps:

- The query is first interpreted at the EPN registry level. It will return a list of data services matching the query, plus basic information on these services — including a mention of access URL and supported protocols. The user may have the possibility to select a subset of services before the next step.
- The query is then passed to the candidate/selected services, following the protocol indicated. This query can therefore be expressed using any documented protocol (in a first step, this will be restricted to PDAP in addition to EPN-TAP).
- The query may involve two different granularity levels. When using the "DATA_SET" (PDAP) or “dataset” (EPN-TAP) resource class/type, it identifies data sets of interest inside this data service. When using the "PRODUCT" or “granule” resource class, a list of matching files is returned. In all cases, the result is a VOTable containing the description and access URL of every dataset or product matching the search criteria.

A typical way to use either EPN-TAP or PDAP is therefore to first query the EPN registry for data sets matching general criteria (e.g. Mars, orbital imaging), then to query directly the selected services for products in these data sets matching more restrictive criteria (e.g. images of a given region, with a minimum resolution).

The answer to a dataset query is essentially intended to retrieve the name of the matching data sets, so that a query at the granule level can be issued. The client may also provide an option to retrieve the whole data set on line; however, considering the size of modern data sets, this sort of request should be processed only after confirmation. At any rate, this must be authorized on the data provider side.

Appendix A – Demonstrators

Several demonstrators of the above mechanisms are available on-line.

A1 Queries of planetary data on web portals

• The EPN client at VO-Paris sends queries to all EPN-TAP declared in the VO-Paris registry, and to PDAP services.

http://voparis-europlanet-new.obspm.fr

The "Custom resource" tab makes it possible to query a resource not yet declared, by providing the URL to its VO schema (and should allow to query any TAP service). The "Advanced query form" tab in the service area of the result page allows the user to query all parameters from this service.

• Access to local planetary databases is implemented at VO-Paris development portal:

- From the first tab the user can send queries to VO-Paris data bases (astronomy + planetary science) using IVOA protocols. TAP has been implemented for planetary science:

http://voparis-srv.obspm.fr/portal/

- From the "Planetary Data" tab the user can send queries to Planetary Science services using the PDAP protocol. The query is sent to VO-Paris data service and also (as a first attempt) to the PSA and DARTS:

http://voparis-srv.obspm.fr/portal/ipda.php

• Access to planetary plasma data is implemented in the AMDA tool developed by the CDPP. Queries to distant databases (VEX-MAG in collaboration with IWF Graz, Cassini-MAPSKP…) are currently done using a SPASE-based connector:

http://cdpp-amda.cesr.fr

Access to external databases in AMDA has also been implemented using webservices (e.g., CDAweb). PDS also proposes their own webservices and we will study in the future how useful and practical they are in order to implement similar connections with AMDA.

A2 PDS on-line access

Access to PDS data has been demonstrated using VIRTIS/VEx spectral cubes. A cube is read in a local IDL session, stored in FITS and the reference is sent to Aladin and plotted in X/Y coordinates. In Aladin, spectra are extracted on mouse click and forwarded to VOSpec or (preferably) SPLAT for plotting:

http://voplus.obspm.fr/samp/SAMPWebProfile+FITS/demo.php

A3 Service validator

A service to test and validate VOTables and data services has been developed at VOParis. It includes a PDAP validation tool:
http://voparis-validator.obspm.fr/

A4 Name resolver

A name resolver has been developed as part of the SSODnet service at VO-Paris [RD7]:
https://vo.imcce.fr/webservices/ssodnet/?forms

The name resolver is implemented as a service using a PDAP-like syntax, with a pair of new PDAP "resource classes":

metadata.jsp?TARGET_NAME=CERES&RESOURCE_CLASS=TARGET

returns the standard name corresponding to the entry. Standard values are taken from the IAU Minor Planet Center list (which is probably different from the PDS').

metadata.jsp?TARGET_NAME=CERES&START_TIME=****&RESOURCE_CLASS=RESOLVER

returns the standard name with the RA/DEC location at the given time (following IVOA usage of the "RESOLVER" keyword). The answer can then be used to query data servers for this target.

A5 XML descriptor service

A service to support writing of XML data file descriptors is being developed by CDPP. It produces the XML files from a user-friendly web interface and can be bypassed for pipe-line processing:
http://oberoi.cesr.fr:8080/jaxfront/JAXFrontServlet?app=jaxfront&action=loadResource&resource=jumpStart/jumpStart.html

A6 Demonstrators in progress

Other demonstrators are being developed and expected to be ready in 2011, among which:

- Develop previous examples: chain PDAP search for VIRTIS data cubes (from VO-Paris portal) with reading and visualization in Aladin/VOSpec. Extend to a variety of data products with different dimensionalities.

- Chain VO-Paris and CDPP demonstrators: search with the AMDA tool at CDPP, identify a VIMS cube, read it, and plot it in Aladin/SPLAT

- Link between AMDA and the HST archive: display Cassini data at Saturn; search HST archive for Saturn auroral images within the corresponding time interval; visualization of the selected images with Aladin. A first version is available, based on the WebSampConnector from VO-Paris:
http://manunja.cesr.fr/~richard/amda/DDHTML/HTML/Hst2Aladin.html

- Search a planet in HST archive using name resolver, plot result in Aladin.

- Chain Name resolver with search functions, plot result in Aladin.

Appendix B - IVOA tools potentially useful for EPN-IDIS

Several types of tools have been developed in the frame of the astronomical VO in the past years. Using them would not only save time in developing the EPN VO, but would also insure to pick up optimal and standard solutions, which maturated through interactions within a large community.

Selected tools of interest for EPN are briefly mentioned below, grouped in functional categories. A more complete list is available at:
http://www.ivoa.net/cgi-bin/twiki/bin/view/IVOA/IvoaApplications

B1 Visualization tools

Most of these tools can exchange data through various protocols, and therefore complement each other the same way as plug-ins in a main application. The most common input data formats are FITS and VOTable.

• Aladin is a visualization tool working with stacks of registered images, which can directly access astronomical catalogues or databases. It is available both as a web-based applet and as a standalone application (at CDS). Aladin provides handy visualization of images and data cubes, allowing frame-by-frame display.
Aladin is interfaced with other VO tools through standard protocols. In particular, Aladin easily returns a catalogue of astronomical objects in the FOV, and the Skybot service is available to identify solar system objects in the FOV at the current date.
Aladin provides support for WCS (astronomical coordinates), which could be used to mimic planetary coordinates at the surface of solar system objects. In particular, it includes a footprint visualization service that may be used for orbital images. This would require the implementation of geographical coordinate systems on planetary bodies (including support for various control point networks).
Aladin is able to project images on the celestial sphere, and in the future also on a sphere (seen from the outside). It can handle multiresolution images with a maximum resolution of 4 mas, sufficient e.g. to process the most resolved orbital images of Mars.

• SAOImage / DS9 is a standard astronomical imaging and data visualization application. DS9 supports FITS images and binary tables, multiple frame buffers, region manipulation... It provides for easy communication with external analysis tasks and is highly configurable and extensible via XPA and SAMP.

• SPLAT is a Spectral Analysis Tool (UK VO, then JAC). It can handle vectors in general, but it is mainly oriented towards analysis of ISM/stellar spectra.

• TOPCAT is an interactive graphical viewer and editor for tabular data, including VOTable (AstroGrid). It can perform more general manipulation of vector data and volume data. It is also able to cross-correlate different catalogues. From version 4.0 (7/2013) it also supports CDF files, which is the de facto standard in plasma physics.

• VOPlot is a light and flexible Java display for 1D vectors or histograms (VO India).

• VOSpec is a Java interactive viewer for 1D spectral data (ESA).
It is specialized in specific spectral data analysis. It implements the Simple Line Access Protocol (SLAP) to query databases of atomic and molecular data.

• Specview is a Java interactive viewer for 1D and 2D spectral data (STSI).
It provides some spectral data analysis tools. It integrates the LineView protocol to query databases of atomic and molecular data (related to the VAMDC program). It is used in the APIS service to display spectra of giant planets from HST: http://lesia.obspm.fr/apis/

• Euro3D client deals with datasets in the Euro3D FITS format (VO-Paris). It is currently restricted to the older PLASTIC protocol.

• VisiVO is a general-purpose visualization tool for large 3D astronomical data sets, either observational or theoretical (model outputs). It includes support for volume data (INAF/CINECA).

B2 Data exchange: protocols and exchange formats

• SAMP is the modern IVOA protocol to exchange data between VO applications, to be implemented in new applications. SAMP supersedes the older PLASTIC protocol.

• VOTable is an exchange format for metadata. In Planetary Science, it is already used by the PDAP protocol for the PSA and PDS, and by EPN-TAP.

• Cone Search, Simple Image Access, Simple Spectra Access, Table Access Protocol, and Simple Line Access Protocol are various protocols to search data of specific types in catalogues. Some are used in other contexts (e.g. Simple Line Access Protocol and LineList are also used by the VAMDC program).
VO-Event is a protocol derived from RSSD/ESA to inform about transient events.

⌘ It is proposed to promote the inclusion of data as VOTables together with scientific publications, so as to facilitate the ingestion of such data in the VO.

B3 Registries management

Registries are the “yellow pages” of IVOA services and data. They are organized by data type and access protocol (e.g., images, spectra, catalogues…) rather than by thematic entry. The IVOA registry implemented at VO-Paris has been slightly modified to accommodate Planetary Science resources (as well as VAMDC services).

Several applications can manage the IVOA registries, which can be extended to other services:

• AstroGrid server is a complete VO platform to manage data and services. It contains software for registries management, publishing, and browsing.

Among those, Ravioli is a web-based client that can browse and search VO registries (developed by AstroGrid, http://astrogrid.roe.ac.uk/ravioli/).

• EuroVO has built another system to handle the same registries, using different software (developed and maintained at ESAC).

• VOdesktop is a client application to visualize the contents of the IVOA registry and the available services. It includes the VOExplorer tool, which handle IVOA resources of various kinds. This tool could possibly be used to handle the present resource list of Europlanet, provided those are correctly defined in an extension of the IVOA system.

• VODb is an SQL query builder providing an easier way to access data services. It includes a version of TopCat for data visualization.

B4 Frameworks

Using these environments allows the data providers to set up VO services with minimal efforts. Such environments handle VO queries and answers, and manage local databases.

• DaCHS [RD3] is used in VO-Paris EPN-TAP services to provide a generic interface for the TAP protocol, including VOSI functions. It contains a client to query local databases.
• VO-Dance has been used to set up an EPN-TAP service at INAF / IAPS.

Appendix C - Other tools of interest

C1 Geographic Information Systems

Concerning software:

• The most used free access GIS for Planetary Science include Pigwad (USGS, web-based), GRASS (open source), Jmars (ASU), Marsweb (UCL), onMars and onMoon (JPL). Compatibility of VO standards with these tools should be studied.

• GoogleEarth and its variations also use open source protocols and currently accept many data sets from the science community. The standard used to describe footprints in this context (KML) is certified by the OGC and is already used for some planetary applications. KML files also provide a URL to data files, so they can in principle be interfaced easily with VO tools, provided the file format is VO compatible.

• The NASA World Wind Java tool is a multiplatform tool (in contrast with initial versions of WorldWind) which can be used as an open source library to build GIS applications. It uses the WMS and KML standards from the OGC.

• Openlayers is an opensource javascript library to build dynamical maps on the web.

• Quantum GIS is an open source application to plot geographic data. A demonstrator of use in a Planetary Science context was developed at VO-Paris using AMIE/Smart-1 images (image mosaics plotted on top of the Clementine basemap).

Concerning standards:

• GML (http://www.opengeospatial.org/standards/gml) or KML RETURN_TYPE could be useful for geographical planetary information systems. GML and KML format for vectorial data are supported by all GIS softwares (both free and commercial). A standard planetary nomenclature e.g. is distributed by the IAU Working Group for Planetary System Nomenclature in KML format. Many space data footprints are available on the PDS Geoscience node: http://ode.rsl.wustl.edu/mars/indextools.aspx (click Mars footprint explorer).

• GeoTIFF format (http://trac.osgeo.org/geotiff/) for raster images is a project of the Open Source Geospatial Foundation. It is not yet supported by all GIS software but will be in the future. Web Processing Service standard (http://www.opengeospatial.org/standards/wps) is for instance requested to evolve in that direction. The GeoTIFF description of projections is in any case an interesting starting point for standard definition of projections in planetology: it is flexible and, in principle, easy customizable (see also http://www.opengeospatial.org/standards/ct for OGC projection service standards).

C2 Data handling and visualization

• SItools2 is an application developed by CNES to help handling space experiment data during operations. This version (public release 2011) includes basic VO standards (cone search) and can send queries on data sets. It also implements a WMS service in support of Earth observation experiments. http://sitools2.sourceforge.net

C3 Dynamical visualization

• MM3DView is a multi-mission version of the 3DView tool developed by CNES for the Rosetta mission. It can make movies of space mission flybys using adequate Spice kernels.

• The open source Celestia software also includes the usage of Spice kernels (from version 1.6), and can make movies of space missions activities.

• The Matisse tool (ASI) provides a high-level visualization interface to PDS databases (in development).

Although these tools are probably difficult to interface with a VO environment, they provide ability to generate 3D material for both research and public outreach, and are based on field standards.

C4 Computational environments

Although there is a tendency to include computational functions in the display tools (e.g. in Aladin), heavy or accurate computations are best handled in specialized environments.

• IVOA developments include the possibility to launch workflows from VO tools. The UWS system in particular can handle computational requests with parameters.

• The Orfeo Tool Box is a very large open source library of image processing algorithms provided by CNES (developed for its Earth observation program). Information is provided here:
http://smsc.cnes.fr/PLEIADES/lien3_vm.htm

• The USGS ISIS software is able to project data in various coordinate systems, to build image mosaics, and to perform many types of processing on imaging data.
http://isis.astrogeology.usgs.gov/

• IRAF is a classic astronomy data reduction and analysis environment from NOAO, which normally uses DS9 as an image viewer. Since v2.16 (March 2012), it includes some VO functions (support for VOTable, registry, SAMP...), and can be used e.g. with Aladin.

• IDL is the standard working environment of many researchers in the field. Among many other things, it provides access to cartographic tools (including projections) and to Spice computations. Some of those are functional in the GDL opensource environment.
Many external libraries exist, some of which provide specialized tools for Planetary Science. The most complete list is available here:
http://idlastro.gsfc.nasa.gov/other_url.html

Some external libraries provide a VO interface to IDL:

- VOlib 0.2 enables VOTable access and implements some IVOA protocols (with some difficulties, apparently)
  http://www.ctio.noao.edu/%7Echrism/VOlib
- Mark Buie's library provides ability to query SQL servers
  ftp://ftp.lowell.edu/pub/buie/idl/buie_idlv62.tar.gz
- The SSW library includes many VO interface functions.
- IDL_Aladin_interface permits to exchange data between IDL and Aladin (on Unix systems only, requires IDL ≥6.0, not compatible with GDL, not working on MacOS 10.6).

- In addition, the readpds library from PDS Small Bodies nodes can read most PDS dataset in IDL. The virtispds/v_readpds library
(available here: http://voparis-europlanet.obspm.fr/othertool.shtml) is similar, but is also functional under GDL. The later has been used in a demonstrator of PDS on-line visualization at VO-Paris: http://voplus.obspm.fr/samp/SAMPWebProfile+FITS/demo.php
The library is used to convert VIRTIS PDS cubes to FITS files and extract individual spectra from Aladin.