1 - Background

Many years ago astronomers coined the expression "Virtual Observatory" to express the idea of pooling the data resources existing in different observatories and data centres scattered across the world, so that complementary information from different sources can be rapidly located, recovered and used to further scientific research. The astronomers were very strongly motivated, for numerous reasons :

- A1. Astronomical observations are of objects which are so far away that the point of observation on the Earth is of little importance.
- A2. The objects observed are generally not varying rapidly in time.
- A3. Their observations are object-oriented ; more precisely, the primary key for indexing an observation is it's location in the sky, *i.e.*, its celestial coordinates or the name of a known celestial object lying in that direction.
- A4. Astronomy benefits from widespread use of near standard formats.

Reasons A1 and A2 make it useful to compare observations made in different places at different times. Reason A3 enables the observations to be classified and hence be easily retrievable, while Reason A4 enables them to be readily compared.

Astronomy of the solar system, benefits from none of these practical advantages.

- S1. For solar system observations knowlege of the position of the instrument making the observation is generally essential. Plasma instruments are mounted on satellites which actually fly through the regions whose properties are being measured, while spacecraft are also sent to observe Sun from directions different from that of the Earth, or to observe and even explore the planets in situ.
- S2. The phenomena observed are time-varying. Phenomena such as collisionless shocks, discontinuities, waves and turbulence are immaterial and can be, and usually are, rapidly varying or transient. Nevertheless, these are the phenomena which determine the characteristics such as density, velocity and temperature of the plasma which fills most of the solar system.
- S3. The data are organised as a time-series of measurements performed along some trajectory in space-time (for in-situ observations), or as a set of one or more images of a unique event (solar eruption, aurora, or ring of Jupiter at a particular time).
- S4. Although the FITS format used by astronomers is also used for solar system imager data, the in-situ data come in almost as many different formats as there are laboratories which provide the instruments.

Thus it can be seen that the development of Virtual Observatory for the solar system starts from a set of conditions radically different from those of the astronomers. It is not surprising that development proceeded more slowly, and along different lines.

2 - Solar System VO today

At first solar system data was archived by different agencies who realised that data not archived would disappear. The advent of the Internet soon showed that archived data which is readily accessible is better exploited. Data became available from many sites, all with different access and query procedures. There was a call from the community that something been done to alleviate these problems.

In 2003 an international consortium SPASE, Space Physics Archive Search and Exchange, was formally set up to coordinate development of a Virtual Observatory for space physics. In

2004 the NASA Living With a Star program organised an international workshop at which the formal organisation of Solar System VO activity was discussed. The organisation must take account of many factors and many compromises must be found :

- The best data are found in the laboratories of the teams who build the instruments, but the primary interest of these teams is not in running a data server !
- Whilst new data centres may be set up respecting technical norms designed to facilitate interoperability, today the vast majority of data lies in the data centres designed before any such norms had been invented.
- It must be sufficiently easy for data centres to participate in VO for the return on investment to be clearly worthwhile.
- To be useful, the data must be readily useable by the scientific community: only data which is physical units should be made available.

Today, two years after that workshop, the structure of the nascent solar system observatory is:

End users are the scientists in the field who submit their user requirements to the virtual observatories. This is probably the group which is the least organised, yet which stands to gain most from a well-organised global data management and access service.

Data providers supply the data made available. In space physics it is generally the Principal Investigator (PI) who is finally responsible for both the nature and the quality of the data products from his instrument. PI collaboration is essential, and everybody (including funding agencies) must recognise the effort required to generate the products requested by the end users, in adequate number and quality, and in a timely way.

Data repositories preserve data and metadata and make it accessible over the Internet. This archiving service includes ensuring the provision of adequate catalogues, the respect of certain rules for the description both data and metadata, maintenance of the network access, {\it etc}, plus attending to the long-term evolution of formats and physical support medium technology. These activities can be undertaken by the PI himself, but it is more reasonable that he concentrate on his research while entrusting archiving to a specialist centre. Data repositories vary from instrument specific to broadly based, and are generally supported by national or international agencies.

Service providers supply software which may be used to perform a service, such as special operations on the data files, comparison of data, of manipulation of auxiliary data, such as finding orbital conjunctions.

Service centres make available services. Examples are the Satellite Situation Centre, http://sscweb.gsfc.nasa.gov/}, and the Collaborative Sun-Earth Connector, http://cosec.lmsal.com/}.

Virtual observatories provide uniform access to data and services for some particular group of users, whose perimeter is defined by considerations of scientific interest, resources, and geo-politics. Specific communities have terms adapted to their own needs depending, for example, on what they observe and the way they observe it. But often the source of funding of a virtual observatory is better identified than its precise scientific scope.

SPASE promotes interoperability within and between Virtual Observatories (VOs) serving Solar System physics research, by elaboration of standards and implementation guidelines, and reference implementations. SPASE may occasionally provide infrastructure when this cannot be identified as being the responsibility of a particular virtual observatory or data repository.

There are presently several Virtual Observatories in the USA :

GAIA – Global Auroral Imaging Access (http://gaia-vxo.org/) A multi-national VO browsing/obtaining summary data from All Sky Imagers, Photometers, and Riometers
VSO – Virtual Solar Observatory (http://umbra.nascom.nasa.gov/vso/) Allows searching for solar data files from around the world.
VSPO – Virtual Space Physics Observatory (http://vspo.gsfc.nasa.gov/websearch/about.jsp) A first step towards building the ``Heliophysics Great Observatory", it includes the Coordinated Data Analysis Web.
VSTO – Virtual Solar-Terrestrial Observatory (http://vsto.hao.ucar.edu/) Supported by NOAA and NSF, initial efforts are focussed on solar and CEDAR data (Coupling, Energetics, and Dynamics of Atmospheric Regions). Uses an ``ontological" basis for queries.
EGSO - European Grid of Solar Observations (http://www.mssl.ucl.ac.uk/grid/egso/) The only European solar system virtual observatory
VHO – Virtual Heliospheric Observatory} (http://vho.gsfc.nasa.gov/)
VMO/G – Virtual Magnetospheric Observatory (http://vmo.nasa.gov/)
VMO/U – Virtual Magnetospheric Observatory (http://vmo.igpp.ucla.edu/) Includes groundstation, derived products, UCLA and other data centres.
ViRBO – Virtual Radiation Belt Observatory (http://virbo.scs.gmu.edu/) Information and services for satellite engineers, operators, and radiation belt researchers
VITMO – Virtual Ionosphere, Thermosphere, and Mesosphere Observatory (http://lwsde.gsfc.nasa.gov/VITMO_VO_Workshop.ppt) Focusses on multi-instrument science

All these virtual observatories are American. Are any needed in Europe ? Whenever there is a lot of data to manage, specific requirements to satisfy (national language support), or multiple sources of funding with which to interact, this may be beneficial. Indeed, within the continental USA, NASA deems it useful to have two Virtual Magnetospheric Observatories, one on the east coast, the other on the west coast. An independent central registry is most certainly desirable in Europe, especially when the central SPASE registry is on a network (nasa.gov) which is reliable but could be considered strategically sensitive : decisions concerning network security are not taken by scientists.

These questions need to be asked. But by whom ?

3 – The European Context

Europe has not been slow in developing data management, access, and processing tools for solar system research, indeed there is a rich diversity of initiatives. Using the above categorisation of each resource, which has been developed in response to operational experience. In nearly alphabetical order we may mention (list to be completed and probably reformatted) :

1. CSDS – http://sci2.estec.esa.nl/cluster/csds/csds.html
   Cluster Science Data System is a data repository operational since 2001.
   Run by a consortium piloted by ESA, and composed of the majority of national funding agencies contributing hardware to the Cluster mission, each financing its own contribution
2. CAA – `http://caa.estec.esa.int/`
   Cluster Active Archive is a data repository, operational since 2006

An archive which makes available the most recent high resolution data available from the instruments of the Cluster mission.

3. CDPP – http://cdpp.cesr.fr/
   Centre de Données de la Physique des Plasmas in a multi-mission data repository operational since 1999. It is financed by CNES and CNRS to preserve and encourage the exploitation of data from instruments financed by the French national research budget, or possibly other data requested by the French research community.
4. EGSO – http://www.mssl.ucl.ac.uk/grid/egso/ A virtual observatory
   European Grid of Solar Observations is true virtual observatory which has been operational since ???? It was developed as an ???? within the European 5th Framework programme.
5. ESAC – http://www.esa.int/SPECIALS/ESAC/
   European Space Astronomy Centre is a data repository operational for non-solar system data since ?????. It will soon be archiving data from planetary missions, and also serving as a science operations planning centre for solar system missions.
6. EuroPlaNet/IDIS – http://demo7.nexse.com/idis/home.htm Under development
   European Planetary Network Integrated Data and Information System is under development within a Coordination Action of the Sixth Framework program. The result of this activity should be an implementation plan for a future European Planetary Virtual Observatory, including a demonstrator which will use only a few selected existing data sets.
7. SPENVIS – http://www.spenvis.oma.be/
   SPace ENVironment Information System is a data repository and service centre operational since ?????
   It is funded by ESA to provide access to models of the space environment and its effects, including the natural radiation belts, solar energetic particles, cosmic rays, plasmas, gases, and "micro-particles"
8. SWEnet – http://esa-spaceweather.net/swenet/
   Space Weather EuroNetwork Operational since ?????
   The SWENET Infrastructure is a central resource centre for space weather activities, providing interested users access to space weather data and services.
9. WDC-C – `http://www.wdc.rl.ac.uk/`
   The World Data Centre for Solar-Terrestrial Physics – Chilton is part of the worldwide system of geophysical data centres, coordinated by the Panel on World Data Centres established by the International Council of Scientific Unions (ICSU).

10. and many others : (BASS2000, MEDOC, ONERA, etc., ...)

Most of these implementations can be classed as data repositories. The notable exception is EGSO, which serves the solar physics community and makes data available from several tens of relatively small data repositories. Looking at this list, one is impressed by the diversity of data repositories : single-mission / multi-mission, national / international, funding for the long-term or not. This is not a disadvantage, although it does highlight the necessity of coordination to make the all work together to serve their various user communities efficiently..

Turning to the potential user communities, we can identify :

a) Solar Physics. This community is served by EGSO, which needs, however, a more robust form of support.

b) Space Plasmas. This is the study of the ionised ambient environment of the Earth. It has a long history, the first application of spacecraft being to measure the local environment. Since 99% of the Universe is constituted of magnetised plasma, which it is well-nigh impossible to produce in the laboratory without complicated boundary effects, space plasma physics is useful for studying fundamental physical processes, such as wave-particle interaction, collisionless shocks and discontinuities, stochastic particle acceleration, turbulence, in a medium in which is nearly homogeneous and with instruments small enough not to overwhelmingly perturb what is being observed.

c) Space Weather. This is the joint use of solar and space plasma physics to obtain practical information of social and economic importance. Solar storms induce the well-known auroras, but also many other less agreeable manifestations such as geomagnetic storms which can cause widespread failure of electricity supplied, interruption of radio communications, interference and even permanent damage to operational spacecraft and aircraft navigation systems, and perturbation of global satellite positioning systems. Although VO activity is not necessary for the operational side of Space Weather, it is essential for the research aspects needed to improve the reliability of the services offered.

d) Planetology. I have not received but new of the EuroPlaNet IDIS (Integrated Data and Information System) recently, but I sicerely hope that it is still working towards the objective of laying the foundations for a European Planetary Virtual Observatory, which will include plasma.

e) Other scientific communities may be large enough (e.g., heliosphere) to make it worthwhile to create a dedicated VO for their disciple. If not, they can use, for example, the VOs in the USA.

In fact, all European scientists can use the VOs which exist in the USA. But NASA has made it clear that these VOs exist primarily to support US science, and the requirements specifications are therefore those of our US colleagues, as are the location and timing of workshops and other meetings of the various user groups.

4 – The Proposal

The current proposal is to set up a design study for a European Sun-Earth Virtual Observatory. This study would be expected to

1. Enhance the level of awareness of on-going activity in Europe. There is presently no European forum at which different communities of Solar System data management experts can share their experience, explain how their data repositories produce and manage the metadata.
2. Address the questions at the end of Sectioin 2.
3. Identify the scientific domains for which virtual observatories would be beneficial in Europe, and suggest the scientific perimeter of each of these virtual observatories.
4. Suggest the level of interoperability required between these virtual observatories (the level of functionality between virtual observatories is somewhat naturally less than the level of functionality with any particular virtual observatory.
5. Assist the various virtual observatories with the development of their data models and metadata dictionary, and its implementation. Create a forum for exchanging hands-on experience.
6. Promote the production of Outlook pages, and address the question of National Language Support.