



Data discovery in planetary science

Where it help in data preservation

Pierre Le Sidaner, Cyril Chauvin,
Regis Haigron, Chloé Azria, S. Erard, B.
Cecconi

Preservation came 20 years ago with OAIS

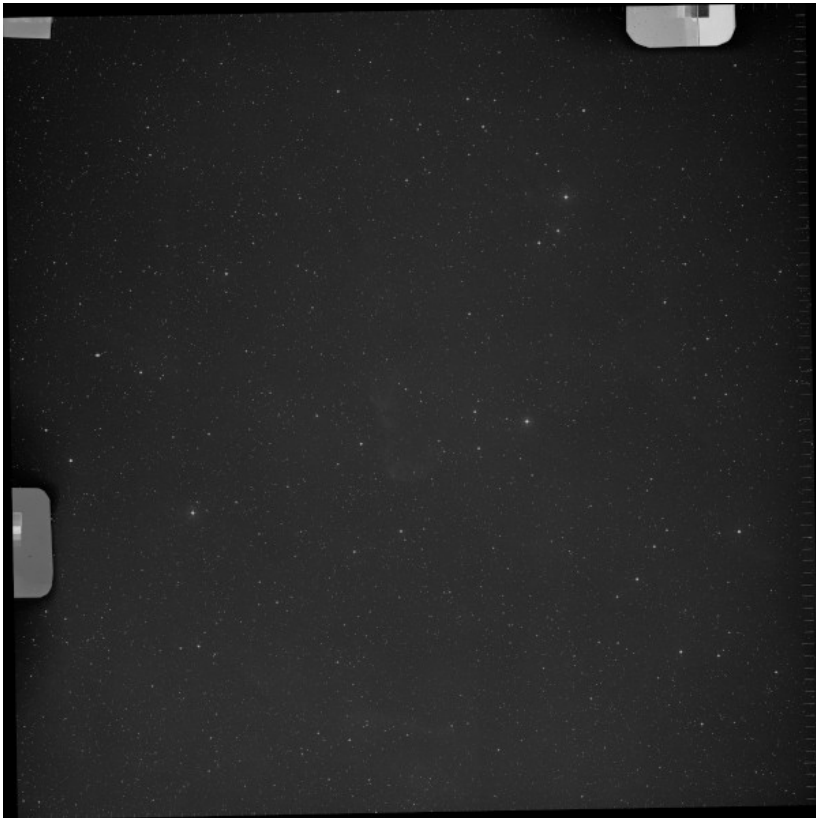


Valorisation come with FAIR

(interoperable but also findable and accessible)



- Preservation with metadata



What to do with png file without metadata

```

SIMPLE = T / This is a FITS file
BITPIX = 16 /
NAXIS = 2 /
NAXIS1 = 34481 / NUMBER OF ELEMENTS ALONG THIS AXIS
NAXIS2 = 34494 / NUMBER OF ELEMENTS ALONG THIS AXIS
EXTEND = T / This file may contain FITS extensions
EQUINOX = 2000.00000000 / Mean equinox
MJD-OBS = 4.301900000000E+04 / Modified Julian date at start
RADESYS = 'ICRS' / Astrometric system
CTYPE1 = 'RA--TAN' / WCS projection type for this axis
CUNIT1 = 'deg' / Axis unit
CRVAL1 = 3.468331946005E+01 / World coordinate on this axis
CRPIX1 = 1.707550E+04 / Reference pixel on this axis
CD1_1 = -1.865469821496E-04 / Linear projection matrix
CD1_2 = 0.000000000000E+00 / Linear projection matrix
CTYPE2 = 'DEC--TAN' / WCS projection type for this axis
CUNIT2 = 'deg' / Axis unit
CRVAL2 = -8.446640080320E+01 / World coordinate on this axis
CRPIX2 = 1.890150E+04 / Reference pixel on this axis
CD2_1 = 0.000000000000E+00 / Linear projection matrix
CD2_2 = 1.865469821496E-04 / Linear projection matrix
EXPTIME = 0.000000000000E+00 / Maximum equivalent exposure time (s)
GAIN = 0.000000000000E+00 / Maximum equivalent gain (e-/ADU)
SATURATE= 4.999782311225E+04 / Saturation Level (ADU)
COMMENT
SOFTWARE= 'SWarp' / The software that processed those data
SOFTVERS= '2.38.0' / Version of the software
SOFTDATE= '2014-05-27' / Release date of the software
SOFTAUTH= '2010-2012 IAP/CNRS/UPMC' / Maintainer of the software
SOFTINST= 'IAP http://www.iap.fr' / Institute
COMMENT
AUTHOR = 'lesidaner' / Who ran the software
ORIGIN = 'CAI-MAMA-VO-PARIS Observatoire de Paris' / Where it was done
DATE = '2014-06-30T15:35:09' / When it was started (GMT)
COMBINET= 'AVERAGE' / COMBINE_TYPE config parameter for SWarp
COMMENT
COMMENT Propagated FITS keywords
VOLTAGE = 8.210000000000E+00 / Lamp voltage of the scanning machine
FSTPAV = 'SRC003JPB' / Name of the original pave
INSTRUME= 'MAMA' / Microdensitometer
ATLAS = 'SRC Blue-Atlas' / Atlas name
FLD = '003' / SRC-J atlas field number
PLTLABEL= 'J2565' / Observatory plate label
EMULSION= 'IIIaJ' / Photographic emulsion
FILTER = 'GG395' / Filter
ALPHA_50= '02:24.0' / Approximate 1950 R.A. for the plate centre
DELTA_50= '-85:00' / Approximate 1950 Dec for the plate centre
DATE-OBS= '1976-08-29' / UT date of Observation
JDAY = '2443020.27321' / Julian date of observation at mid exposure
LST_SOE = '02:28:00' / Local sidereal time at start of exposure
PLTDATE = '1976.6626' / Decimal date of the plate exposure
EXPOSURE= '60.0' / Exposure time (minutes)
PLTGRADE= 'AI2' / Plate grade
TELNAME = 'Siding Springs' / Telescope location

```

- Preservation deal with FAIR

Preservation \Leftrightarrow futur usage of data

Reusable mean no preservation without metadata

Standardized metadata comes with IVOA

(interoperable but also findable and accessible)



How to make Planetary science data available without reinventing the wheel

It come from several communities :

- **Solar physics**
surface, activity, corona, wind
- **Planets**
Interior, surface (OGC), atmosphere
- **Plasma**
interplanetary medium, solar wind, planetary atmosphere
- **Small bodies & comets**
- **Exoplanets**
- **Minerals and samples**

- Technical issues

- **ASCII tables**
- **Images not only fits**
- **Spectrum (not only incident light, minerals ...).**
- **Dynamic spectrum**
- **OGC data and services**
- **Orbitographic data (spice kernel)**
- **Maps**
- **Events (VO)**
- **.....**



Vespa \subseteq IVOA

Long time maintenance => use existing eco-system

- **adoption of IVOA with its eco-system (registry, TAP, Applications)**
- **deviation of obscure into EPN-CORE**
- **list of mandatory parameters**
- **optional parameter list to satisfy sub-communities**
- **no more RA DEC access => time, objects, mission**



Work to be done

- **Define a flat data model**
- **Define a web client to access all data**
- **Double complexity**
 - **Each community have their usage**
 - **Vespa started with only few services**

Vespa client

Refine your search ADQL Query Data Services Help

Main Parameters

Target Name

[Target Class](#)

[Dataproduct Type](#)

Instrument Host Name

Instrument Name
=

Processing level

Time

Location

Spectral

Illumination

Data Reference

Optional

abs_cs - Data for numerical modeling of planetary atmospheres	13 results	●	○	○
AMDA - Planetary and heliophysics plasma data at CDDP/AMDA	2539504 results	●	○	○
APIS - Auroral Planetary Imaging and Spectroscopy	74535 results	●	○	○
BaseCom - The Nançay Cometary Database	6886 results	●	○	○
bass2000 - Bass2000 solar survey archive	370931 results	●	○	○
BDIP - IAU database of historical planetary Images	16906 results	●	○	○
cassini_jupiter - Cassini RPWS/HFR Calibrated Jupiter Flyby Dataset	7 results	●	○	○
CLIMSO - CLIMSO coronagraphs at Pic du Midi de Bigorre	2943730 results	●	○	○
cpstasm - CLUSTER STAFF-SA Spectral Matrix Data	11688 results	●	○	○
DynAstVO - Asteroid orbital database and ephemerides	31320 results	●	○	○
eit_syn - Synchronous synoptic maps of the solar corona from EIT/SoHO	18482 results	●	○	○
ExoPlanet - Extrasolar Planets Encyclopaedia	5335 results	●	○	○
Exotopo - Simulated Topography of Exoplanets	1800 results	●	○	○
expres - ExPRES Simulation Database	100581 results	●	○	○
Gaia-DEM - Thermal structure maps of the solar corona from SDO	746364 results	●	○	○
GEM_Mars - Profiles from Mars Global Climate Model	1399680 results	●	○	○
HFC1AR - Heliophysics Feature Catalog active regions	1194277 results	●	○	○
HFC1T3 - Heliophysics Feature Catalog type 3 radio bursts	90845 results	●	○	○
hrsc3nd - HRSC nadir Images of Mars	4093 results	●	○	○
hst_planeto - Planetary data from the Hubble Space Telescope	52768 results	●	○	○
IKS - IR spectroscopy of comet Halley	206 results	●	○	○



EPN-TAP Services Custom Service

Main Parameters

Target Name

[Target Class](#)

[Dataproduct Type](#)

Instrument Host Name

Instrument Name
=

Processing level

Time

Location

Spectral

Illumination

abs_cs - Data for numerical modeling of planetary atmospheres	13 results	●	○	○
AMDA - Planetary and heliophysics plasma data at CDDP/AMDA	1198088 results	●	○	○
APIS - Auroral Planetary Imaging and Spectroscopy	36091 results	●	○	○
BASECOM - The Nançay Cometary Database	15611 results	●	○	○
bass2000 - Bass2000 solar survey archive	305603 results	●	○	○
BDIP - Base de Données d'Images Planétaires	16906 results	●	○	○
BIRA-IASB TAP - Profiles from SPICAV-SOIR/VEX	13465 results	●	○	○
cassini_jupiter - Cassini RPWS/HFR Calibrated Jupiter Flyby Dataset	7 results	●	○	○
CLIMSO - CLIMSO coronagraphs at pic du midi de Bigorre	498038 results	●	○	○
cpstasm - CLUSTER STAFF-SA Spectral Matrix Data	11688 results	●	○	○
CRISM - CRISM georeferenced cubes	20722 results	●	○	○
DynAstVO - Asteroid orbital database and ephemerides	20035 results	●	○	○
ExoPlanet - Extrasolar Planets Encyclopaedia	4037 results	●	○	○
expres - ExPRES Simulation Database	77526 results	●	○	○
HFC1AR - Heliophysics Feature Catalog active regions	948627 results	●	○	○
HFC1T3 - Heliophysics Feature Catalog type 3 radio bursts	90845 results	●	○	○
hisaki - Hisaki Planetary Database	4112 results	●	○	○
hrsc3nd - HRSC nadir images of Mars	4093 results	●	○	○
hst_planeto - Planetary data from the Hubble Space Telescope	44260 results	●	○	○
litateHF - litate HF data	3614 results	●	○	○
IKS - IR spectroscopy of comet Halley	206 results	●	○	○

Plotting tools

- TOPCAT
- Aladin
- SPLAT
- CASSIS
- 3DView

Example queries

Saturn in March 2012

Help

Help

- **Each service appear indepently in the portal**
- **NASA (PPI) come with 170 services**
- **Vespa succes increase with 20 services / year**

<http://vespa.obspm.fr>

development version

<http://voparis-europlanet-dev.obspm.fr/>

- **How web portail can handle that?**
 - **Thematic sort ?**
 - **Other grouping of services ?**
 - **What is the response time to the query ?**
 - **When does the dynamic display break ?**



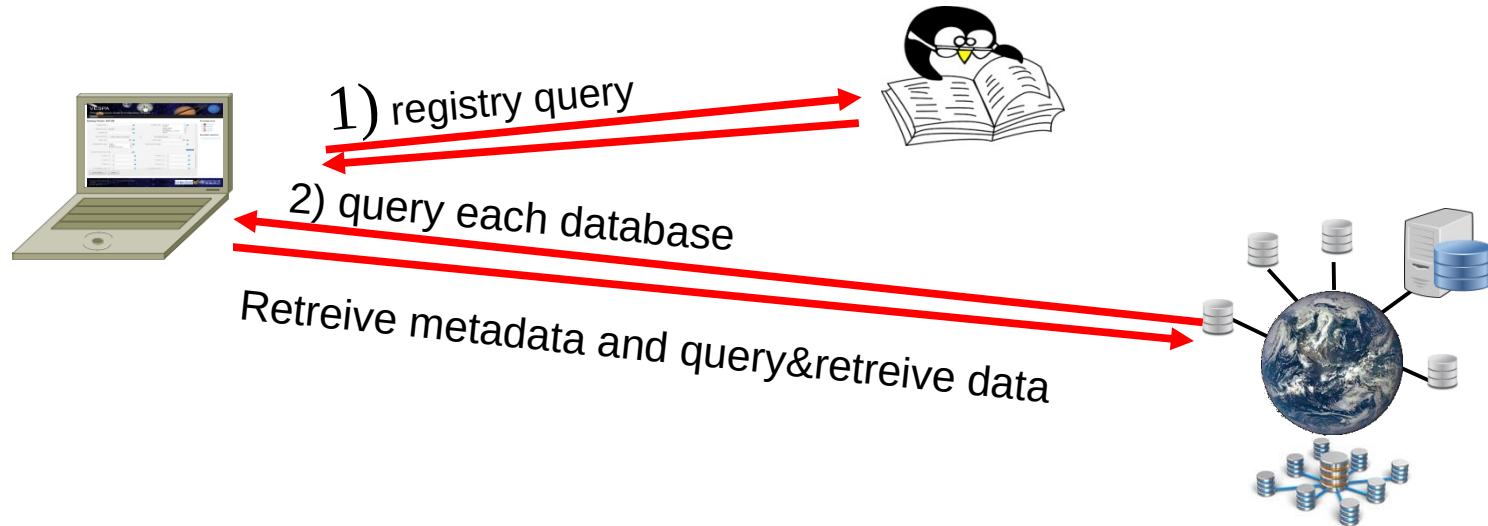


User feedback Complains

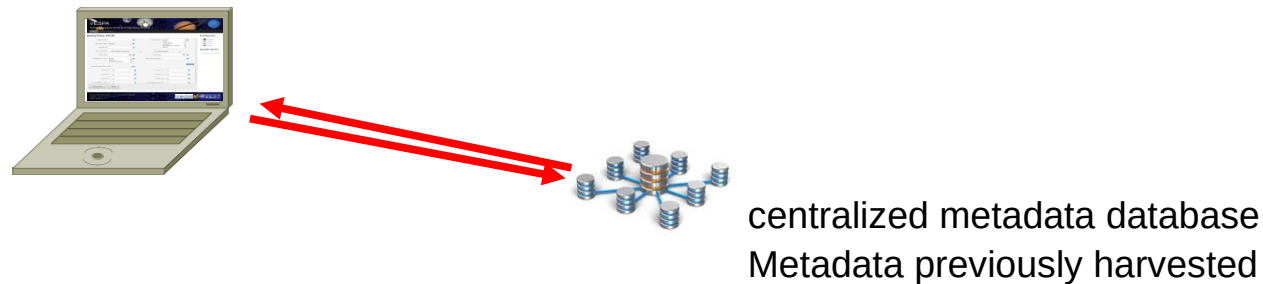
- **You can't find what you're looking for ?**
- **It's not intuitive !**
- **We don't work like that !**
- **Everyone has a point of view on ergonomomy !**

Local approach

Distributed architecture IVOA



Local metadata database



Local approach

- retrieval of the 34Millions metadata set
- ingestion in a search engine.
- Creation of spatial range zone (time, frequency)
to be treated separately
 - Polygons management problem**
 - Range long/lat**
 - Moc ?**

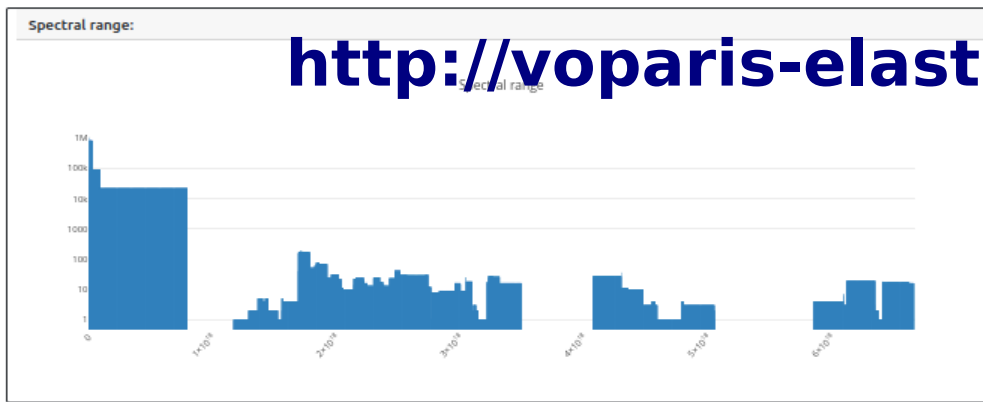
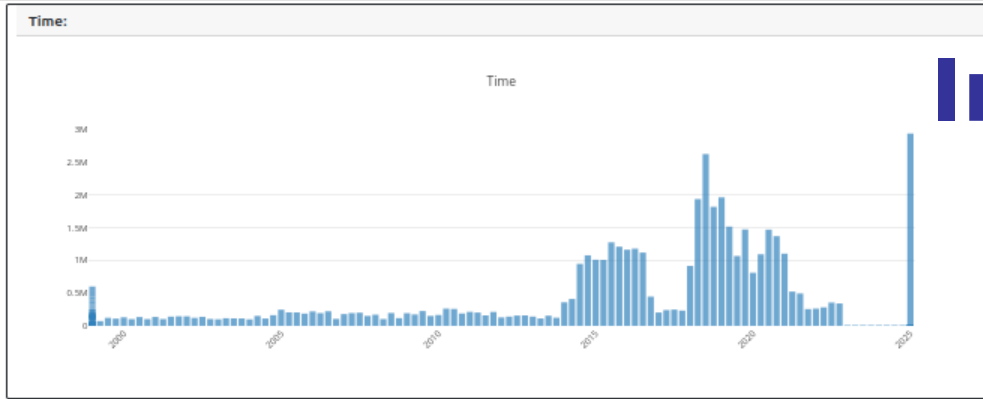


NoSQL facilities/difficulties

- **Easy way to scan meta datas**
- **Facets search make it intuitive**
 - **Search in large heterogeneous datasets**
- **Many difficulties on displaying heterogeneous outputs (mixing carrot and califlower) !**
- **range filed type are usefull but weighs down the queries.**
- **We have to optimized before clustering.**

Interface

<http://voparis-elasticsearch.obspm.fr/dev3>



Missions	Instruments	Services	Datatype
ExoMars 2016 18420402	CaSSIS 9811364	psa 28189162	ci 32874174
Simulation 8343694	ACS 8251877	mcd 5892272	pr 8369794
Rosetta 7204645	Mars Climate Database V3.3 6992272	climso 3928730	im 5164331
OMP, OA, Pic-du-Midi 3923969	ROSINA 4054847	AMDADB 2536522	ts 2477302
SOHO 1694681	GEM-Mars_V052 1387680	gem_mars 1387680	ds 1227727
SDO 1370452	AIA 1379453	mpc 1257234	sc 359402
Mars Express 1182071	OSIRIS 1363906	hfc1ar_majtrim 1173418	mo 62495
		hfc1ar 1165739	ma 58191

Target region	Target class	Target name
atmosphere 8354405	planet 31385992	Mars 27256532
2941463	star 7484383	67P 70375217
craters 384561	comet 7079380	Sun 4286280
chromosphere 181158	asteroid 1523856	sun 5933980
corona 132215	1322130	Earth 2095467
Magnetosphere 77090	satellite 865085	Venus 1637242
Aurora 40392	interplanetary_medium 749085	1127840
photosphere 37923	Planet 139681	Moon 633273



Interface

- **real progres to do**
- **Complementary to the calassical VO portal.**
- **Main question remain how to display results**
- **Take advantage of the Nasa portal ?**