HiPS catalogue implementation at CDS

IVOA Interoperability Meeting

Shanghai, the 16th of May, 2017

F.-X. Pineau, T. Boch, P. Fernique and the CDS team

International

Virtual

Observatory

Alliance

HiPS – Hierarchical Progressive Survey

Version 1.0

IVOA Proposed Recommendation
6rd April 2017

This version:

1.0: Proposed Recommendation 2017-04-06

Previous version(s):

1.0: Proposed Recommendation 2017-04-03
1.0: Proposed Recommendation 2017-02-07
1.0: Proposed Recommendation 2016-11-22
1.0: Working Draft 2016-06-23

Interest/Working Group:

Applications: http://www.ivoa.net/twiki/bin/view/IVOA/IvoaApplications

What is a HiPS?

- http://www.ivoa.net/documents/HiPS/
- "*HiPS [is] a hierarchical* ***scheme for the description, storage and access of sky survey data****. The system is* ***based on hierarchical tiling of sky*** *regions* ***at finer and finer spatial resolution*** *which facilitates a progressive view of a survey, and supports multi-resolution zooming and panning. HiPS uses the* ***HEALPix*** *tessellation of the sky as the basis for the scheme and is implemented as a simple file structure [...].*"
- Purpose: **navigate through all-sky data** (image surveys, catalogues, ...) **à la Google Maps**
  - Explore the sky by zooming and panning, **no explicit query**
  - **The more you zoom, the more finer details you get**

Make and explore your own HiPS

- CDS offers **2 tools** allowing users to build their own HiPS
  See http://aladin.u-strasbg.fr/hips/
  - **Hipsgen** for images:
    http://aladin.u-strasbg.fr/hips/HipsIn10Steps.gml
  - **Hipsgen-cat** for catalogues:
    http://aladin.u-strasbg.fr/hips/HipsCat.gml
- and 2 HiPS visualizers: **Aladin** and **Aladin Lite**
- (other visualizers: **Mizar** and **prototype extension of the MAST portal**)

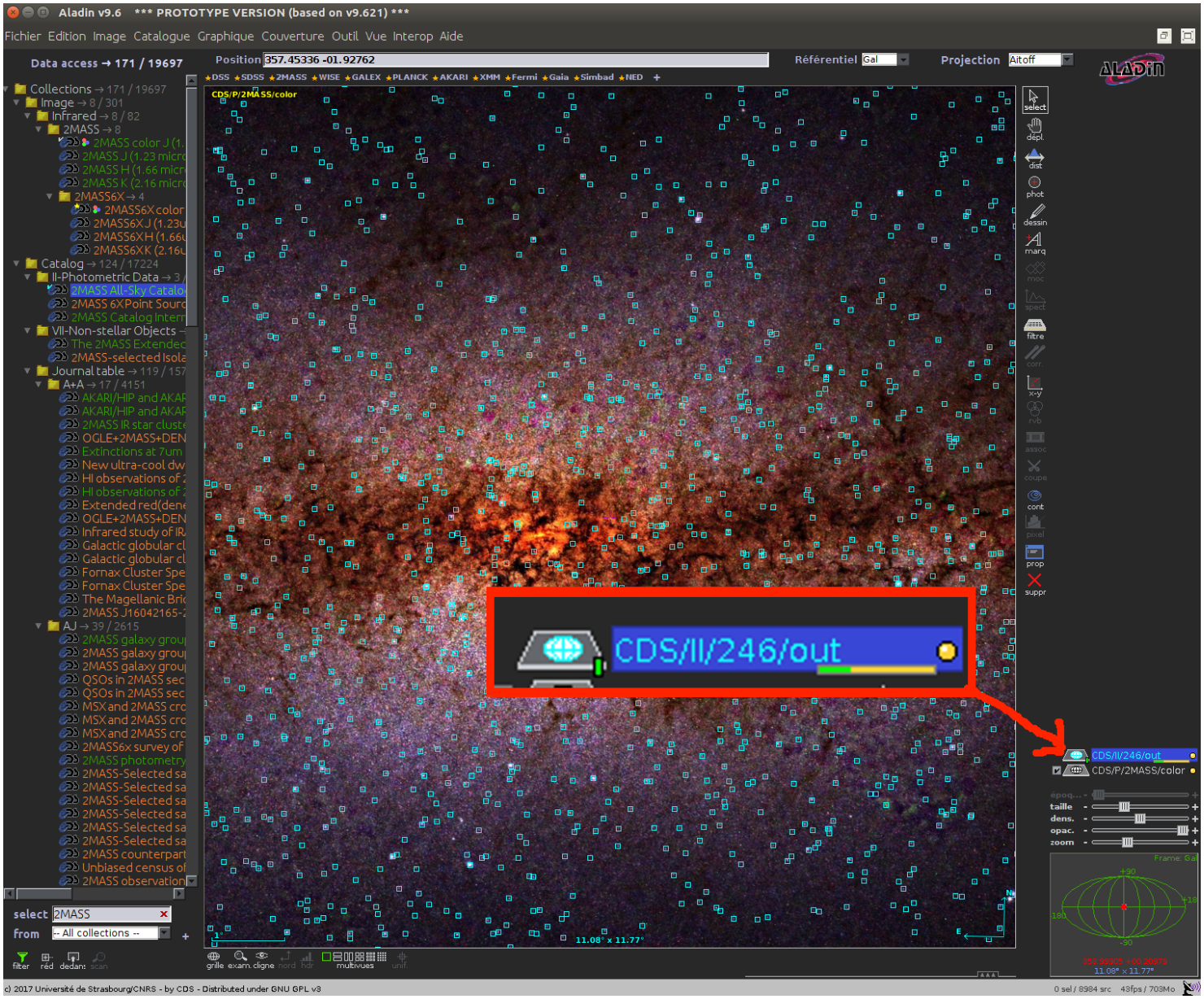
Example of directory structure in output of the Hipsgen-cat tool (public)

```
> ls -l MyHiPS

metadata.xml      # Std: metadata in VOTable format
Moc.fits          # Std: MOC of the table at order max
Moc.json          # Not Std: MOC in JSON format
Norder1           # Std: dir containing order 1 tiles
Norder2           # Std: dir containing order 2 tiles
Norder3           # Std: dir containing order 3 tiles
Norder4           # Std: dir containing order 4 tiles
Norder5           # Std: dir containing order 5 tiles
properties        # Std: HiPS meta information
index.html        # Std: webpage embedding AladinLite
arguments         # Not Std: HiPSgen-cat input args
densmap_o0.fits   # Not Std: order 0 HEALPix density map
densmap_o1.fits   # Not Std: order 1 HEALPix density map
densmap_o2.fits   # Not Std: order 2 HEALPix density map
densmap_o3.fits   # Not Std: order 3 HEALPix density map
densmap_o4.fits   # Not Std: order 4 HEALPix density map
densmap_o5.fits   # Not Std: order 5 HEALPix density map
```

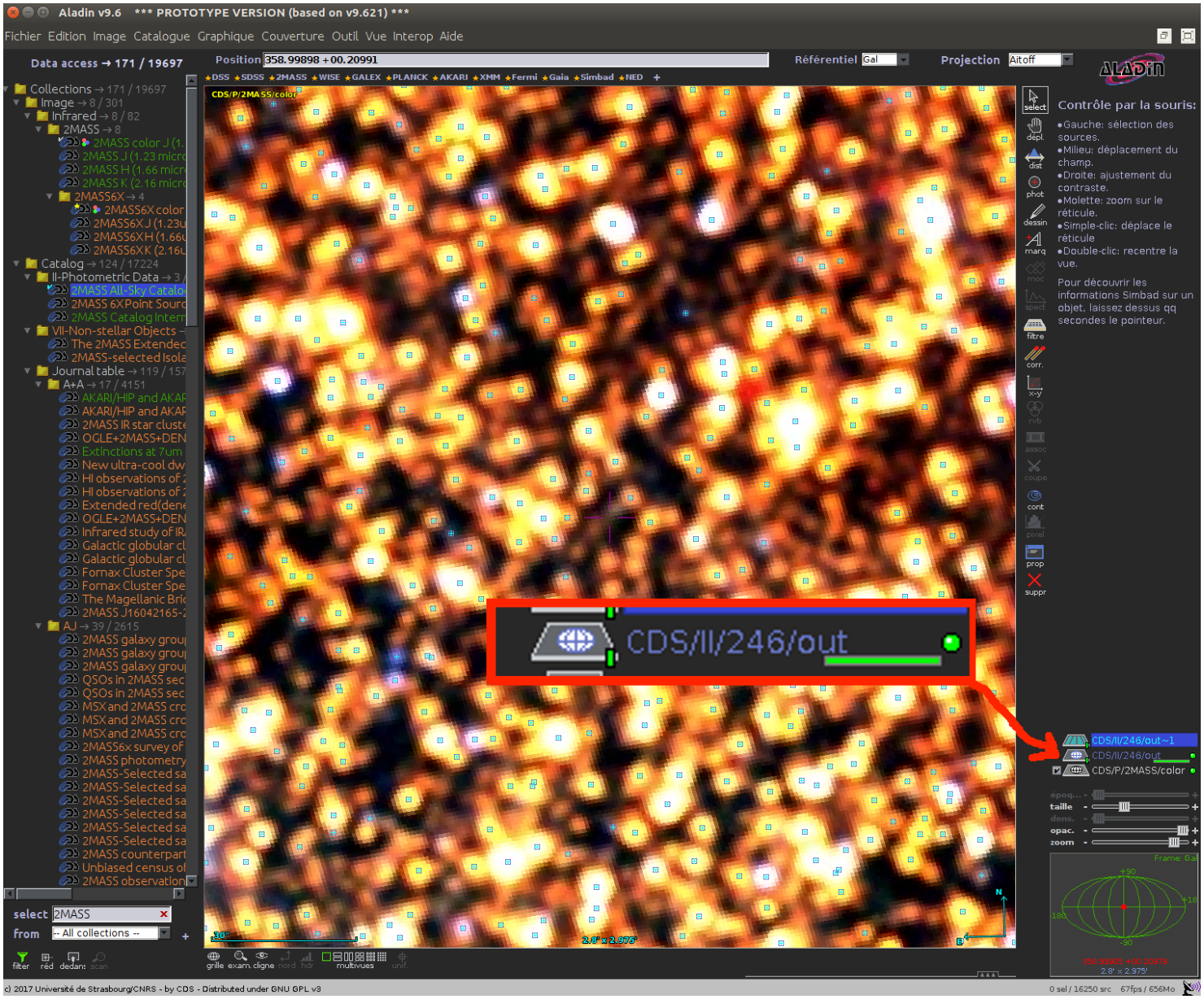
Example of a HiPS catalogue tile

```
> more MyHiPS/Norder1/Dir0/Npix4.tsv

# Completeness = 1969 / 2031
_RAJ2000        _DEJ2000        HIP     RAhms    DEdms    Vm
138.723590      4.442900        45383   091453.72        +0
133.781760      1.546508        43790   085507.60        +0
135.582844      8.468490        44376   090219.52        +0
135.209707      5.241545        44263   090050.47        +0
142.114170      9.056778        46454   092827.38        +0
142.645828      10.599953       46634   093035.11        +1
137.443434      11.564377       44984   090946.45        +1
145.287637      9.892308        47508   094109.12        +0
130.685548      9.556699        42748   084244.40        +0
130.717620      9.553099        42762   084252.10        +0
129.961632      11.522672       42499   083950.86        +1
128.963632      6.622776        42173   083551.34        +0
134.269528      11.646984       43948   085704.71        +1
135.322818      15.265768       44295   090117.55        +1
```

Example of HiPS in Aladin V10



Example of HiPS in Aladin V10



HiPS catalogue available at CDS

- 24 HiPS so far: http://axel.u-strasbg.fr/HiPSCatService/hiplist

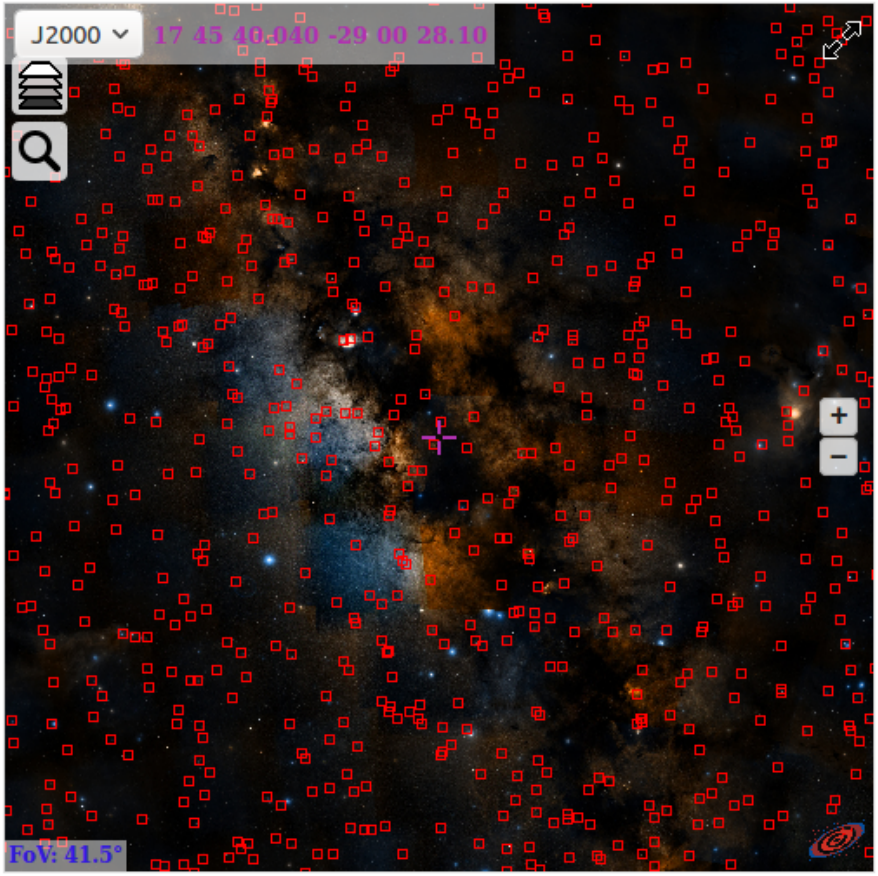
```
creator_did         = ivo://CDS/II/341/vphasp
dataproduct_type    = catalog
obs_collection      = VPHAS+
obs_description     = The VPHAS+ catalogue
hips_service_url    = http://axel.u-strasbg.fr/HiPSCatServi
hips_release_date   = 2016-10-16T21:04Z
hips_status         = public master unclonable
hips_tile_format    = tsv
hips_order          = 11
hips_frame          = equatorial
client_category     = Catalog
client_application  = AladinDesktop
client_sort_key     = 02-341-06

...
```

- Goal: put the 16000 VizieR tables in the HiPS format

http:/axel.u-strasbg.fr/HiPSCatService/I/339/hsoy

The I/339/hsoy progressive table

This Web resource contains HiPS(*) components for the **I/339/hsoy** progressive table.

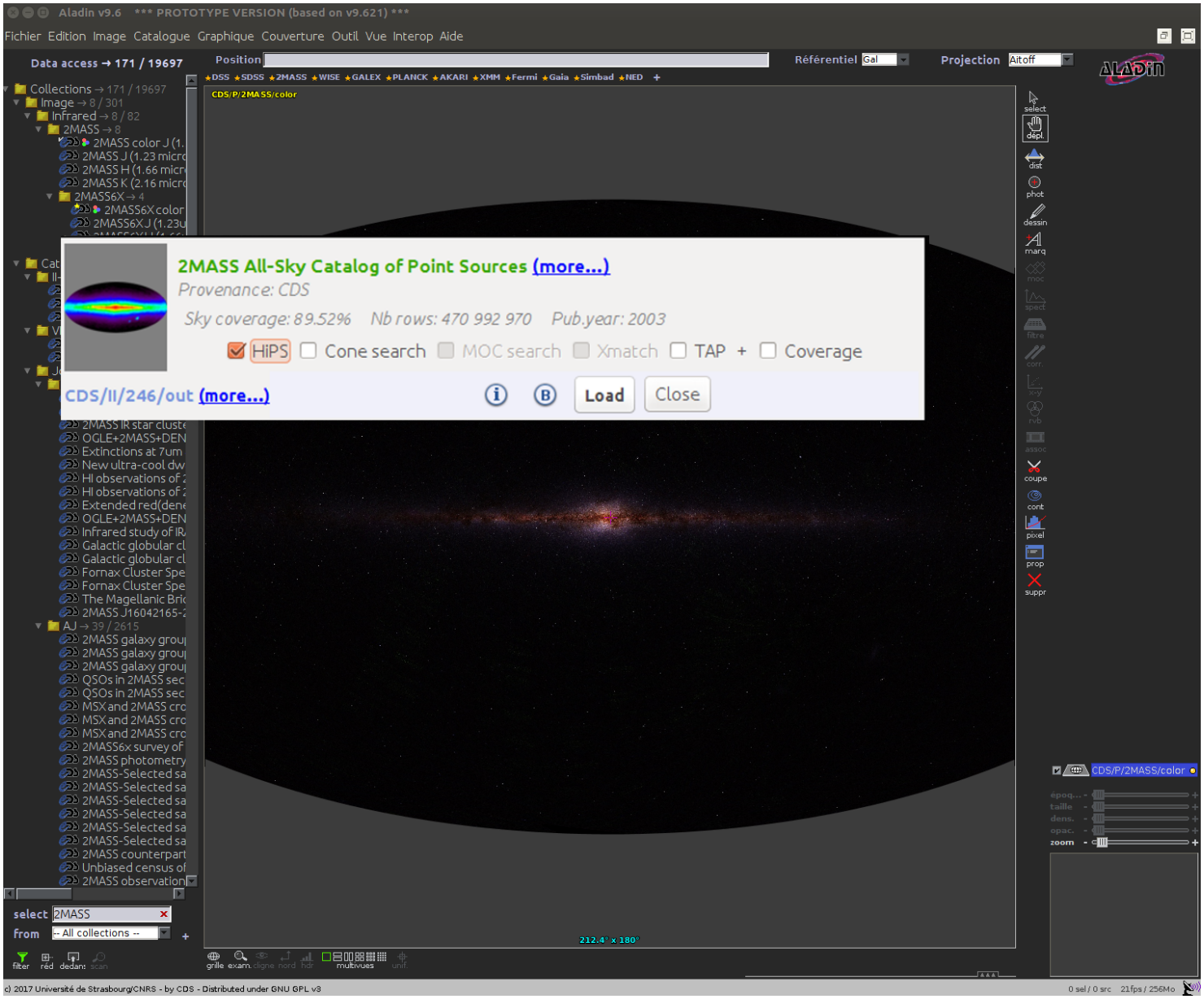


- **Label:** I/339/hsoy
- **Type:** HiPS table
- **Raw property file:** properties
- **Metadata:** metadata.xml
- **Associated coverage map:** Moc.fits
- **Base URL:** http://axel.u-strasbg.fr /HiPSCatService/I/339/hsoy/
- **Allsky level 1:** allsky1.tsv
- **Allsky level 2:** allsky2.tsv
- **Tiles URL format:** http://axel.u-strasbg.fr /HiPSCatService/I/339/hsoy/Norder*[1-orderMax]*/Dir*[0-9]**/Npix*[1-ipixMax]*.tsv

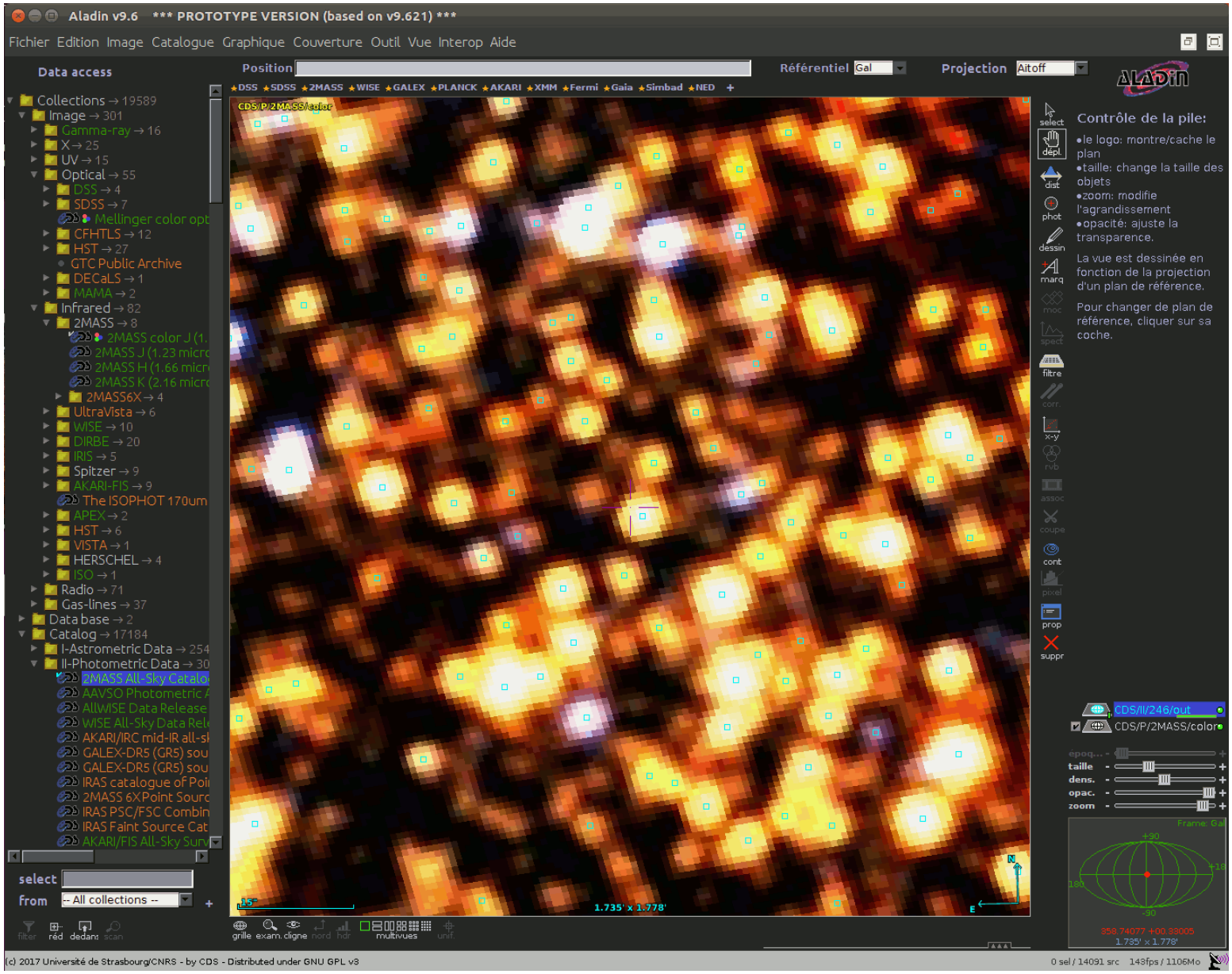
This progressive catalogue can be displayed by the Aladin Desktop client (just open the base URL) or any other HiPS aware clients.

(*) The HiPS technology allows a dedicated client to access an astronomical table at any location and at any scale. HiPS is based on HEALPix sky tessellation and it is designed for astronomical scientifical usages. HiPS technical documentation is available here

Example of CDS HiPS in Aladin V10



CDS HiPS in Aldin V10 (animated gif)



HiPS: not necessarilly physical files

- On page 8, the standard says:
  "*[T]he actual implementation of HiPS as directories and files is not an obligation, only the view as directories and files is required (see HiPS distribution section). Internally, a HiPS **may** be stored in a database, or any other appropriate packaging (tar or zip files...) rather than in a basic file system directory structure.*"
- **Hipsgen-cat** = 2 tools in one
  - 1 public: FITS, VOTable, CSV, ... $\rightarrow$ Standard HiPS
  - 1 private: CDS Xmatch Binary file $\rightarrow$ HiPS binary files
    - **HTTP API** mimicking the file hierarchy
    - Access through **Tomcat Servlet + rewriting rules**
  - large parts of the code in common

Structure of HiPS catalogues at CDS

```
> ls -l MyHiPS

header.bin   # Not Std: metadata in a proprietary format
Moc.fits     # Std: MOC of the table at order max
Moc.json     # Not Std: MOC in JSON format
l1.prg       # Not Std: binary file of order 1 tiles
l2.prg       # Not Std: binary file of order 2 tiles
l3.prg       # Not Std: binary file of order 3 tiles
l4.prg       # Not Std: binary file of order 4 tiles
l5.prg       # Not Std: binary file of order 5 tiles
l6.prg       # Not Std: binary file of order 6 tiles
arguments    # Not Std: HiPSgen-cat input args
```

- **.prg** file: index + concatened tiles
  - index: tile number $\rightarrow$ (starting row, No. rows) + Tot. No. rows in the HEALPix cell
  - tiles: concatenated in a format similar to FITS BINTABLE

Advantages of CDS HiPS catalogues

Pros	Cons
Light for the file-system	More complex access (need a specific tool)
Fast copy on another machine	Harder to debug
Output columns can be chosen on-the-fly (not in the std)	Less rich metadata (to be solved)
Very fast generation: few hours for largest tables	Same date for all tiles

- In practice, the Apache Tomcat Servlet is very robust
- Large parts of the code in common with CDS Xmatch / VizieR large catalogues

CDS Algorithm: beyond basic ones

- Threshold on the source brightness
  - not fixed, depends on each tile
- Number of sources per tile:
  - not fixed, $\in [50, 500]$ by default (from depth > 3)
  - depends on:
    - the number of sources in the HEALPix cell
    - the catalogue coverage in the HEALPix cell
  - specific treatment at level 1, 2, 3 to better mimic the density of sources
- $\Rightarrow$ The depth of the HiPS is variable
  - area of higher density $\rightsquigarrow$ deeper hierarchy ($\sim$ QuadTree)

Last words

- Reminder:
  - main purpose: easy exploration of large dataset (no queries)
  - the users must understand how it works (which sources first?, when are all sources loaded?)
- Recent improvements from user feedbacks (JAXA, USA):
  - now **compatible with V1.0 of the standard**.
  - web page added (**index.html**) (but need to be access through an HTTP server for security reasons)
  - default positional column metadata for AladinLite
  - few bugs corrected
- Future plan:
  - public version of **Hipsgen-cat**: improve performances

Thank you!