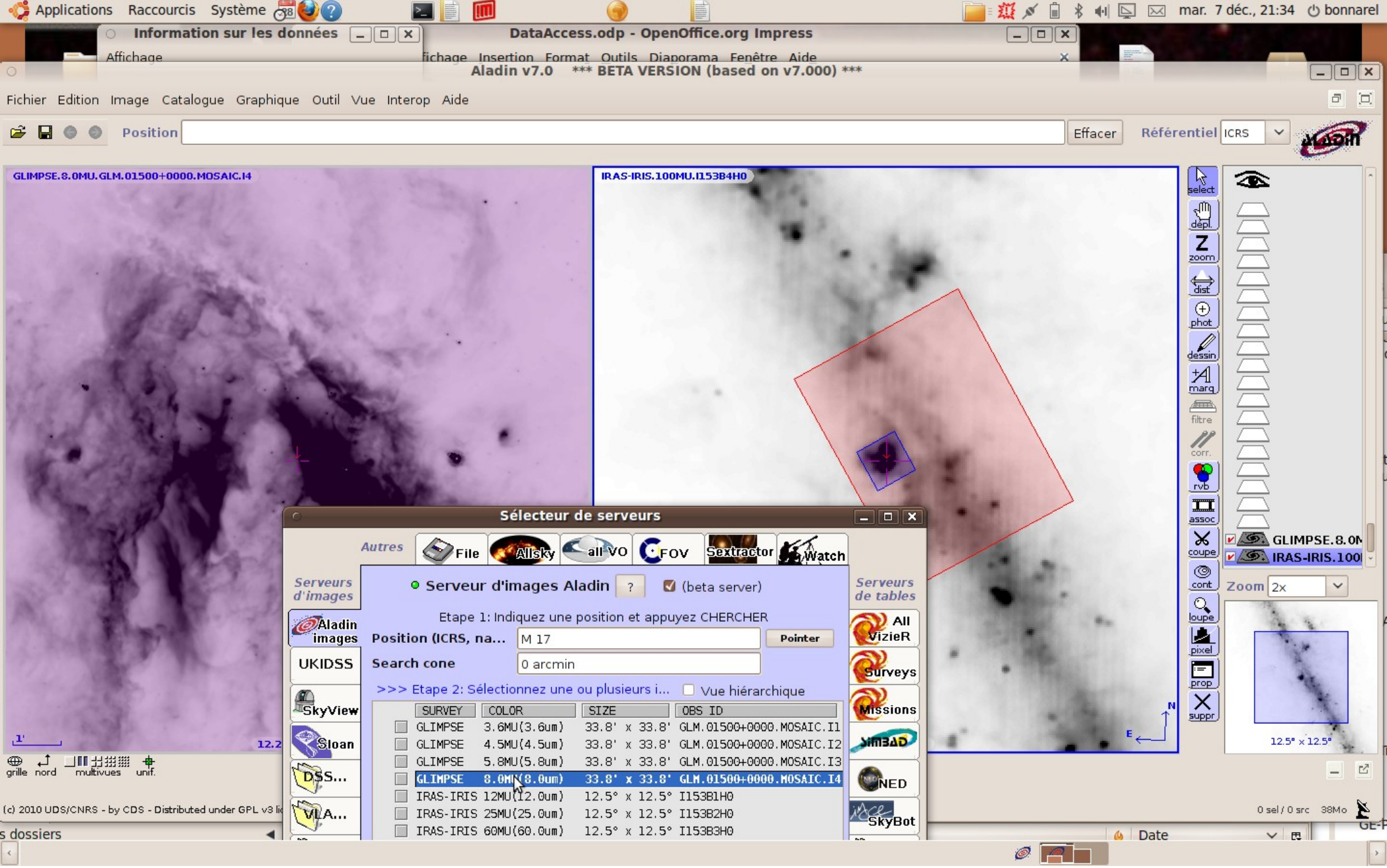
Data Link and Data Access

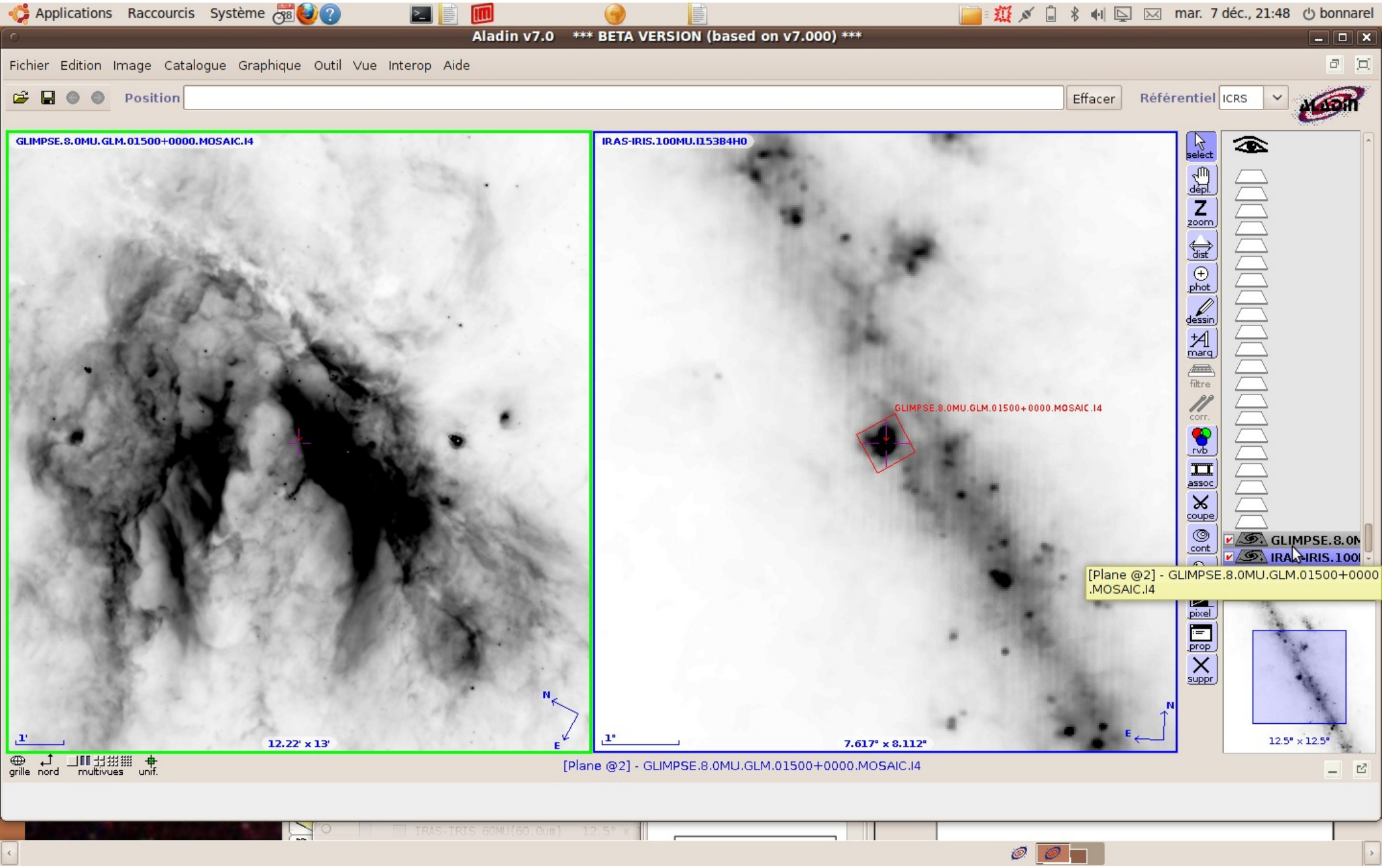
F.Bonnarel (CDS)

DAL and Data Access: beyond discovery

- DAL is not just about data discovery but also about accessing data.
- Accessing data is NOT JUST retrieving a full dataset... Need for
  - Filtering
  - Cutouts
  - Creation of virtual data (resampling/combining/etc...)
  - Spectral issues addresed by Peter
- ON LINE Access to images and cubes is both necessary
  - For discovery fine tuning.
  - For processing in some cases :multi wavelength correlation (color composition), on line SED extraction, etc...
  - Science is also on the desktop (see science requirements for image display)



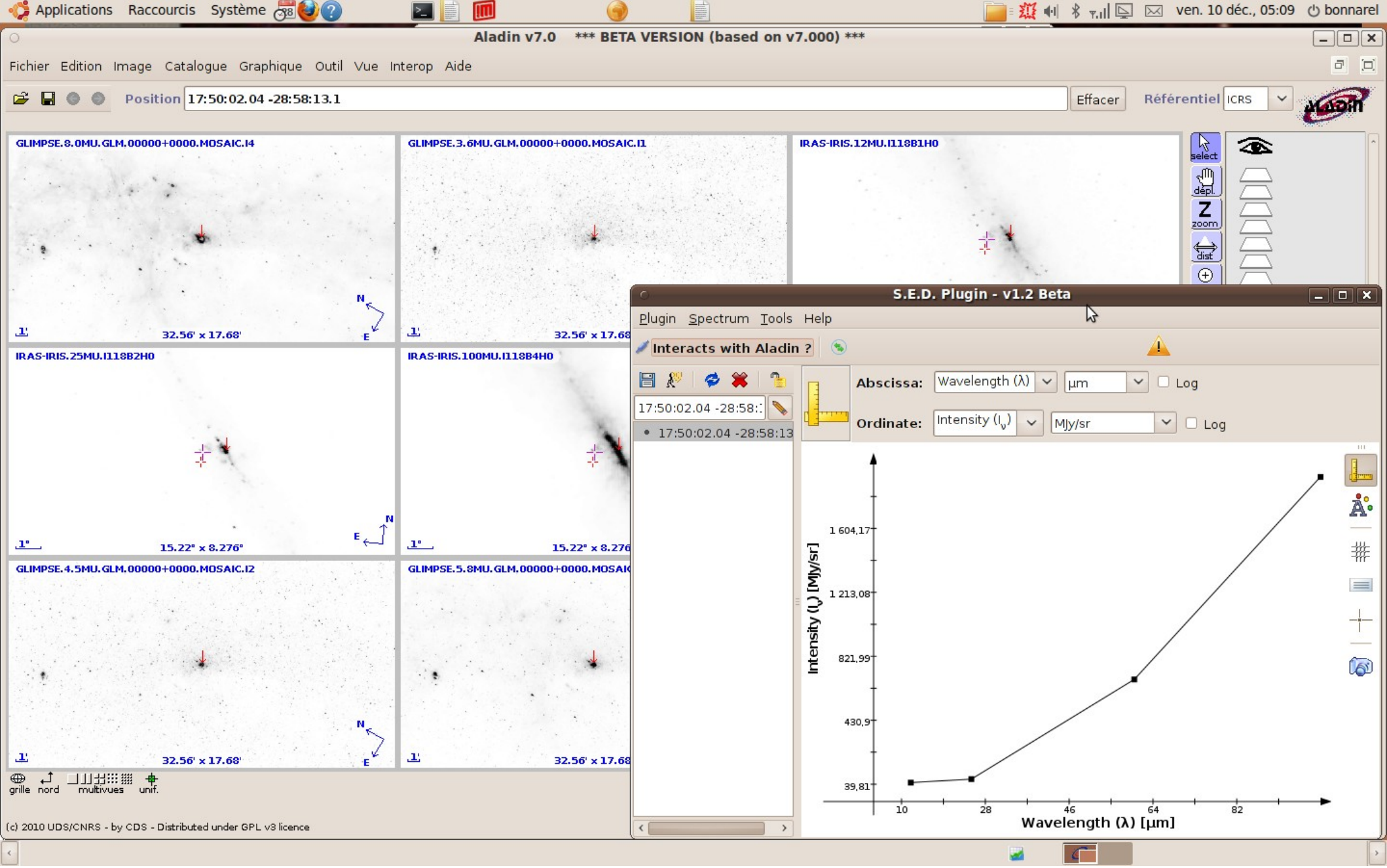
FOV (red) of a GLIMPSE/Spitzer plate(250 to 500 megabytes) on top of IRAS 100 micron image. In blue default cutout contour



FOV (red) of the GLIMPSE/Spitzer cutout on top of IRAS 100 micron image.

DAL and Data Access

- DAL is not just about data discovery but also about accessing data.
- Accessing data is NOT JUST retrieving a full dataset... Need for
  - Filtering
  - Cutouts
  - Creation of virtual data (resampling/combinig/etc...)
  - Spectral issues adresed by Peter
- ON LINE Access to images and cubes is both necessary
  - For discovery fine tuning.
  - For processing in some cases: multi wavelength correlation (color composition), on line SED extraction (SED plugin), etc...
  - Science is also on the desktop (see science requirements for image display)

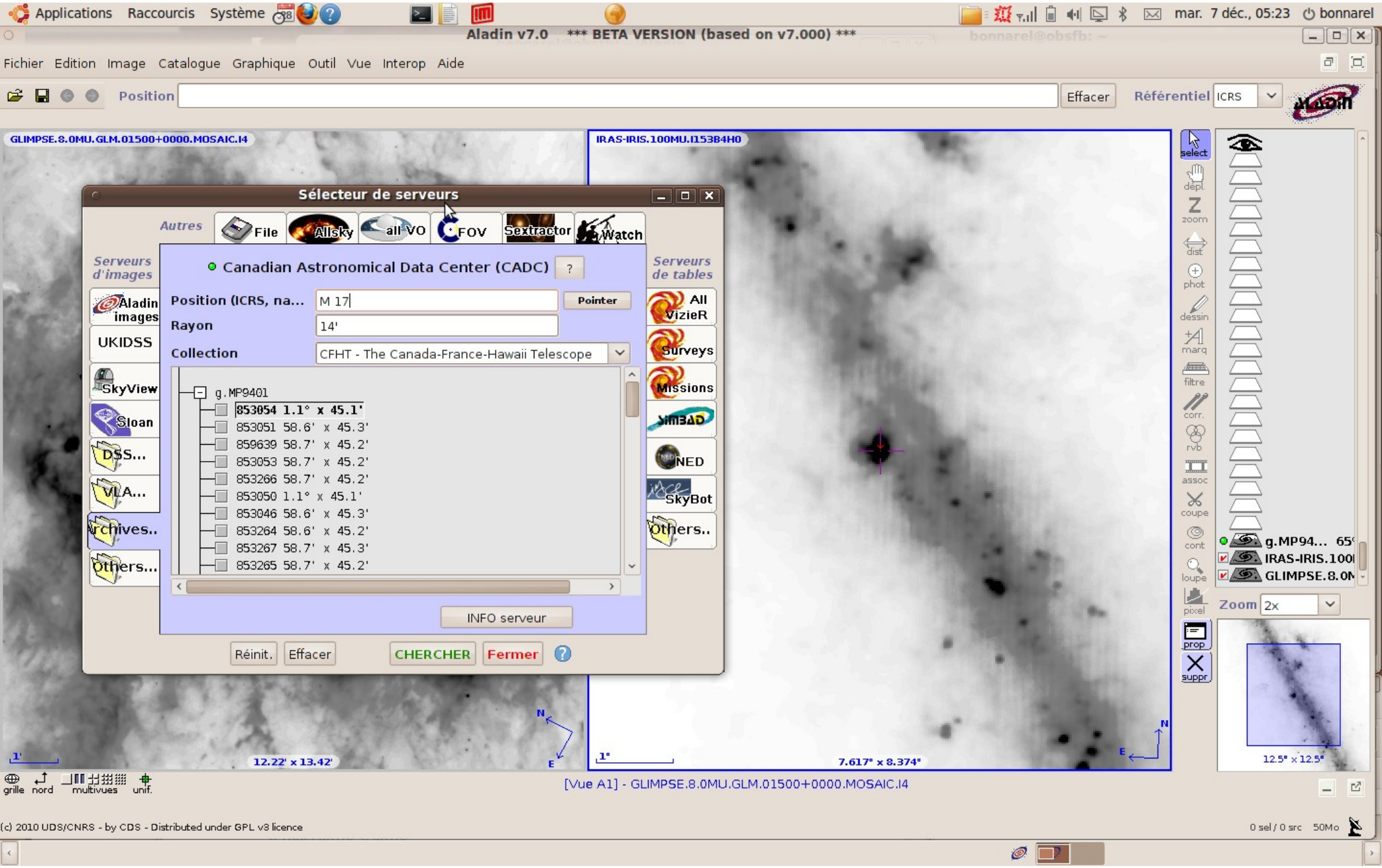


SED Plugin: computes region fluxes from calibrated image

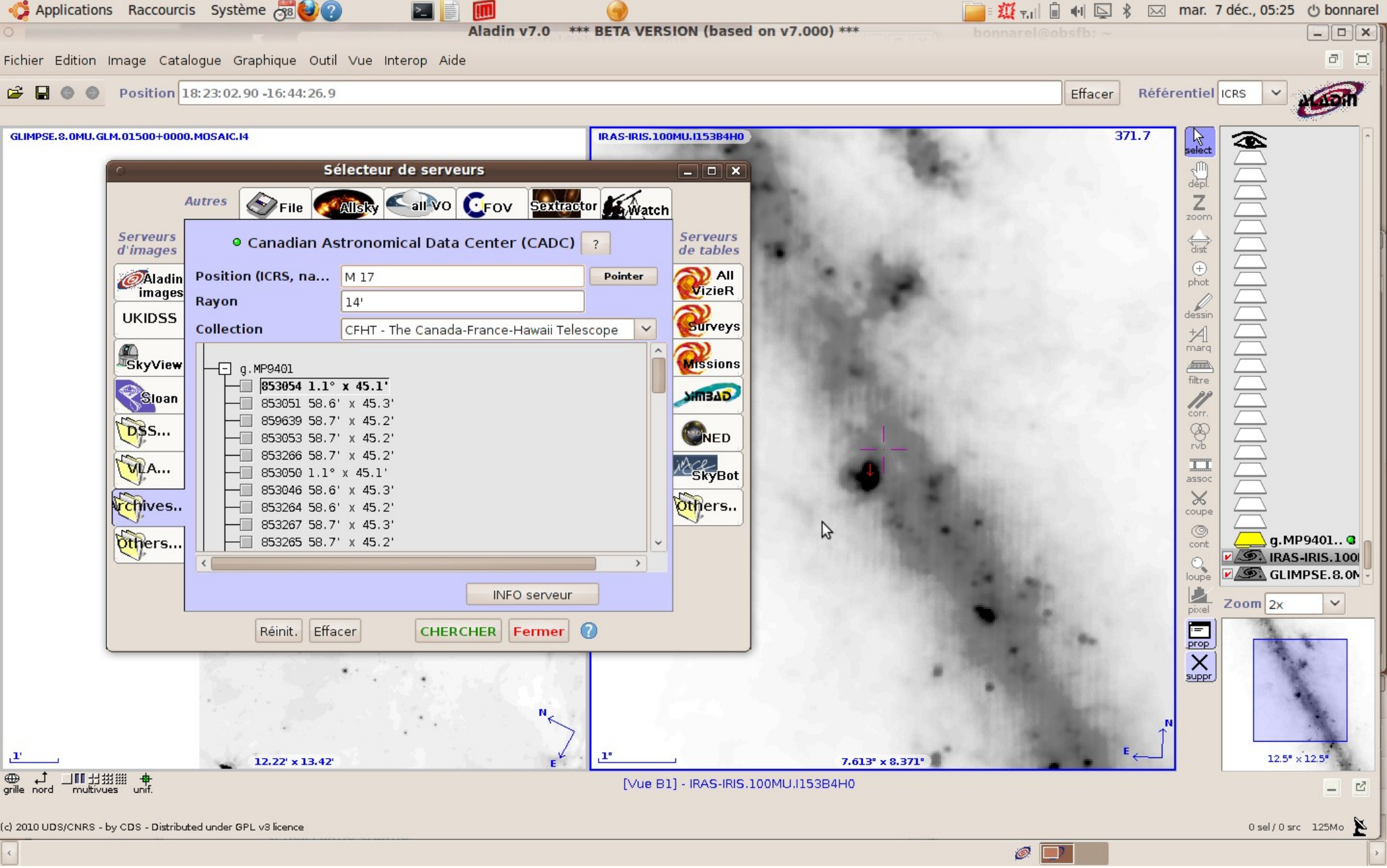
DAL and Data Access: Obstap and others

- Obstap helps to standardize data discovery but

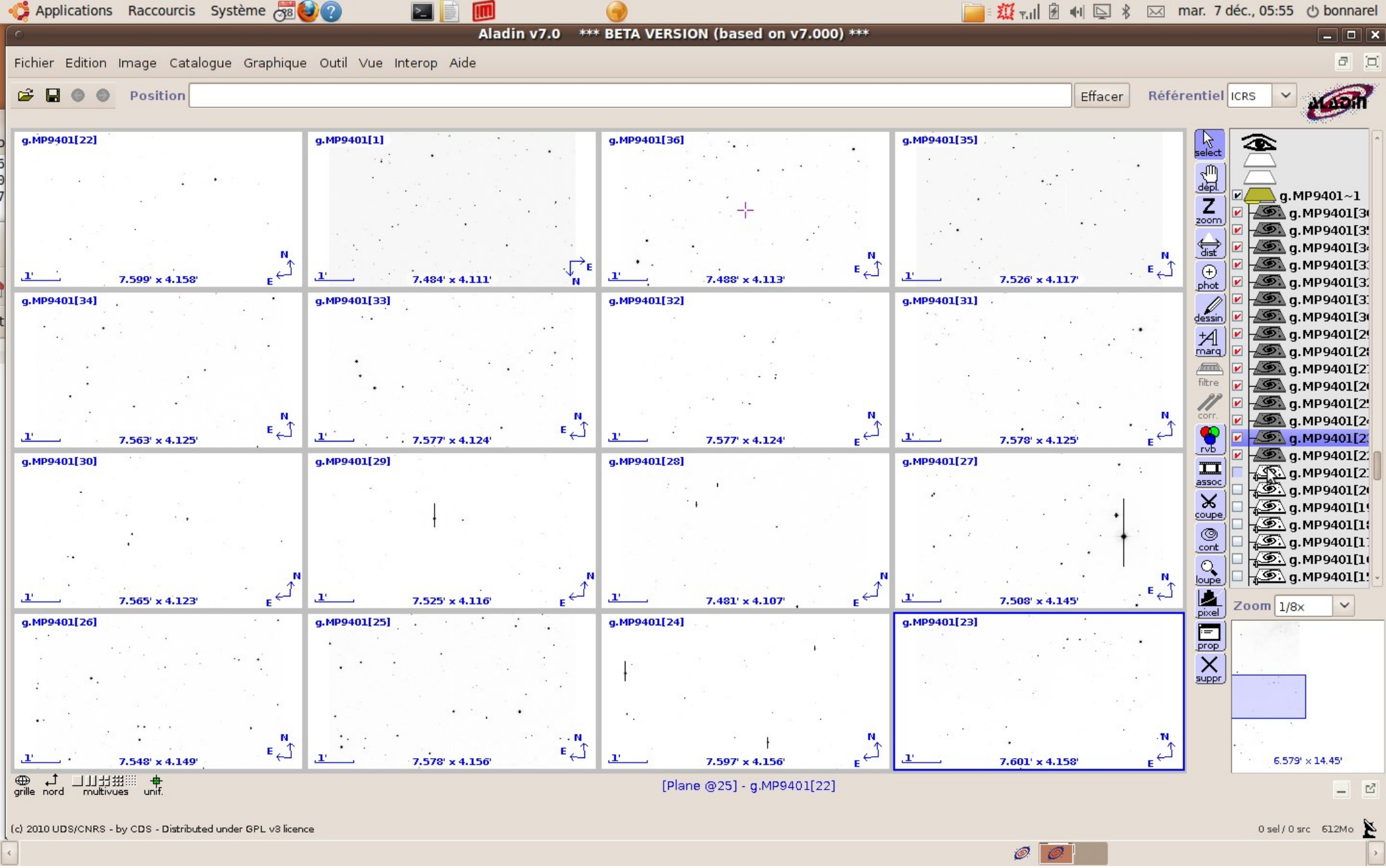
  don't tell us how to acces images, cubes, spectra, time series ....
- Some SIA1 « solutions » work for data acces . 2D Cutouts and mosaic/resampling ..
- SSA provides some filtering facility (but .... see Peter's talk)
- Lesson learnt : keep Old SIA1/SSA services beside Obstap untill they are replaced by new generation ones



TAP / PQL service for Full MEF discovery: choosing the dataset (700 mega bytes)



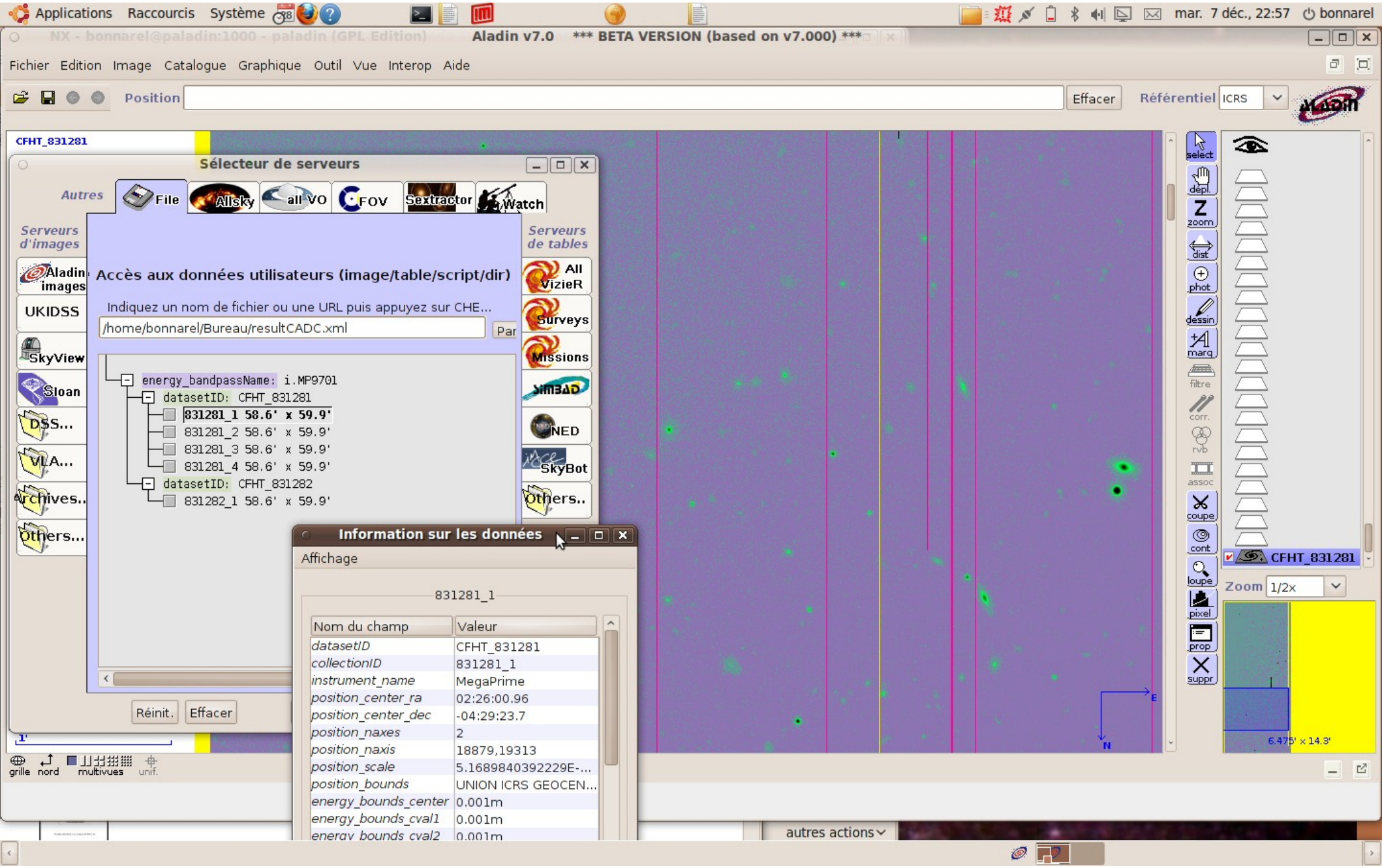
TAP / PQL service for Full MEF discovery:loading start of the MEFt



TAP / PQL service for Full MEF discovery: after 15 minutes/ not that convenient.

DAL and Data Access: Obstap and others

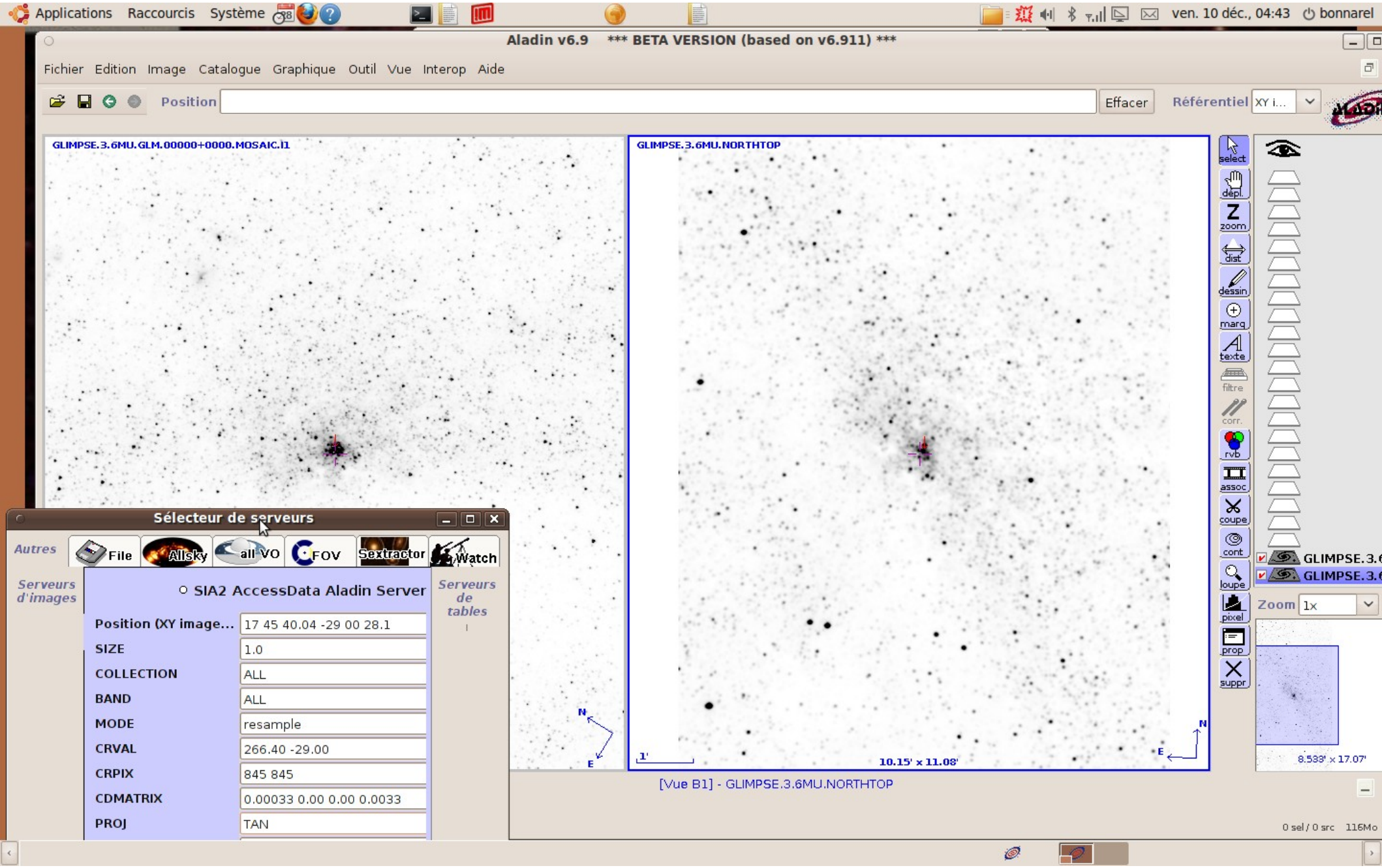
- Obstap helps to standardize data discovery but don't tell us how to acces images, cubes, spectra, time series ....
- Some SIA1 « solutions » work for data acces . 2D Cutouts and mosaic/resampling ..
- SSA provides some filtering facility (but .... see Peter's talk)
- Lesson learnt : keep Old SIA1/SSA services beside Obstap untill they are replaced by new generation ones



Classical SIA service: Individual CCDs (around 20 Mbytes).
Belonging to the same Dataset allows Associations and gathering under the same node in the tree

AccessData

- What is it?
  - Knowing that a dataset of some specific type exist: this method allows retrieving or generating « virtual » Data for a specific scientific purpose (may include easy display)
- For Spectra, and images in SIA1, some parameters drive generation of virtual data but there is no explicit AccessData method.
- SIAV2 introduces the concept of AccessData method, via new parameters.
  - « Filetring » cubes on bands , polarization states
  - Spatial and spectral « cutouts »
  - General resampling methods
- Other specific data (Time Series, etc …) need other parameters



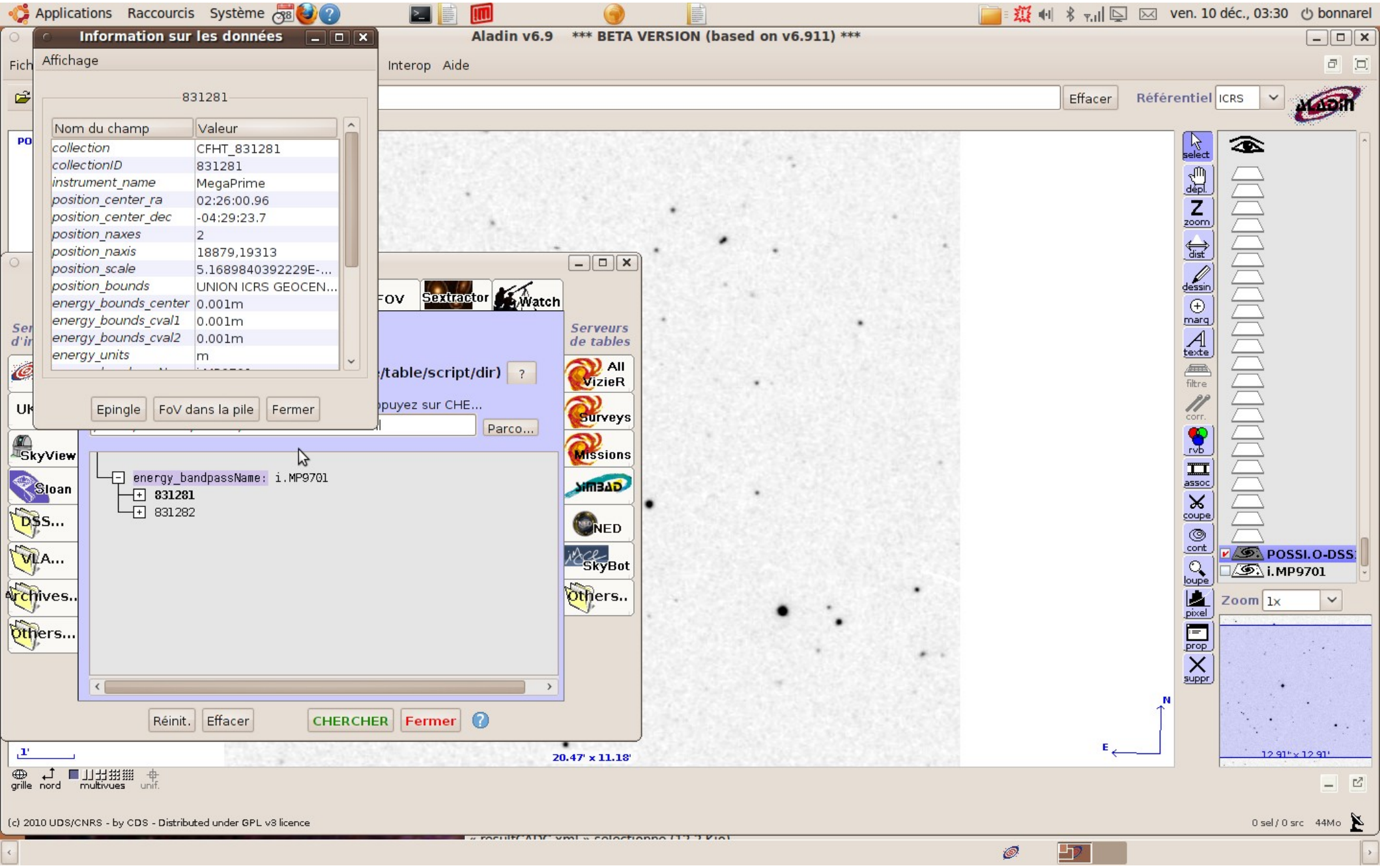
SIA2 prototype AccesData method for resampling: put the North on top

Data Link

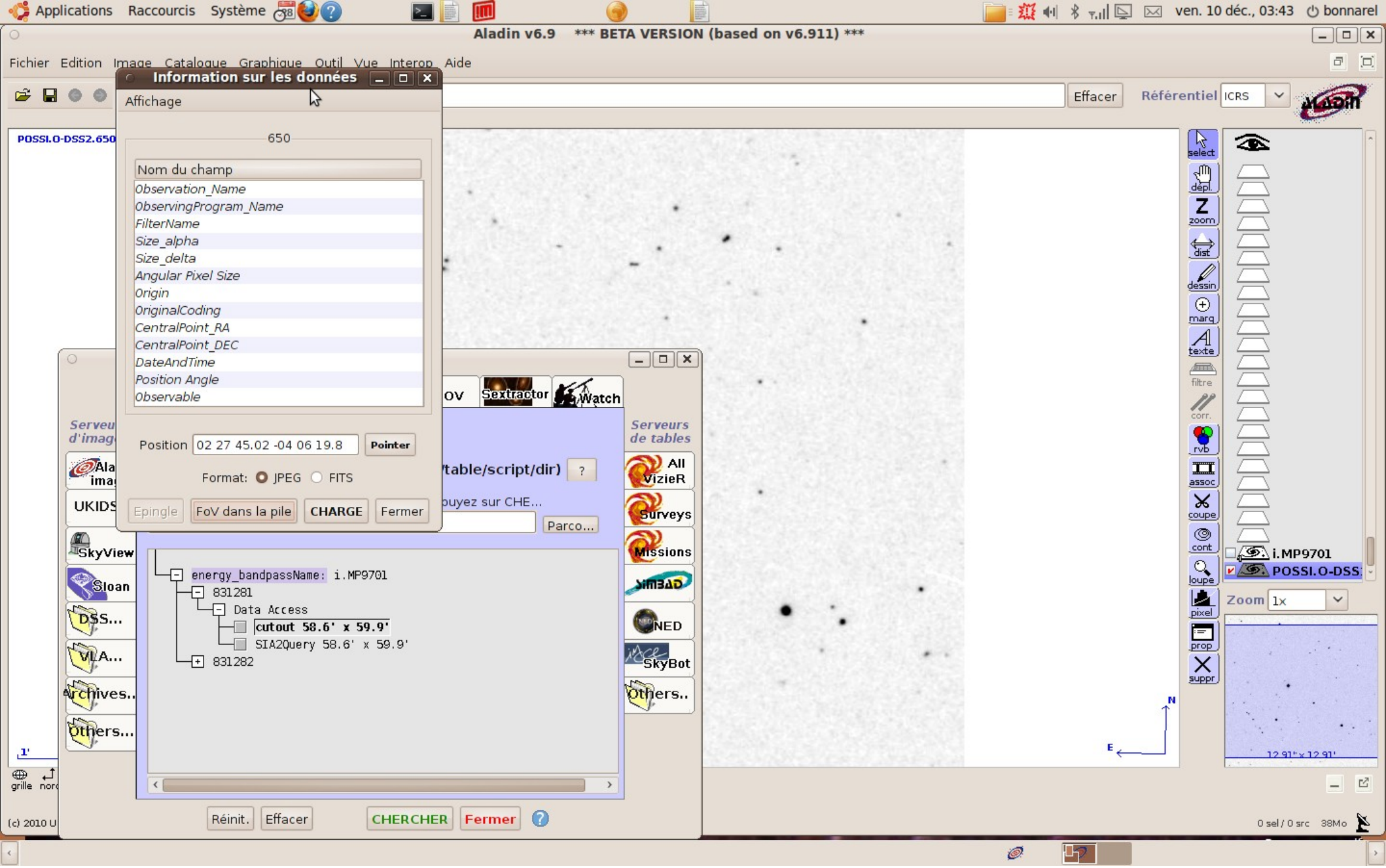
- How does that fit together?
  - Obstap cannot have AccessData method because it's generic. Datasets with complex content are not covered.
  - Obstap Could have additional optional column with URL to a specific data service AccessData method
    - One single possibility per dataset (or duplicate → bad)
    - Modification of the Obstap service definition
- Propose a new service instead: DataLink

Data Link

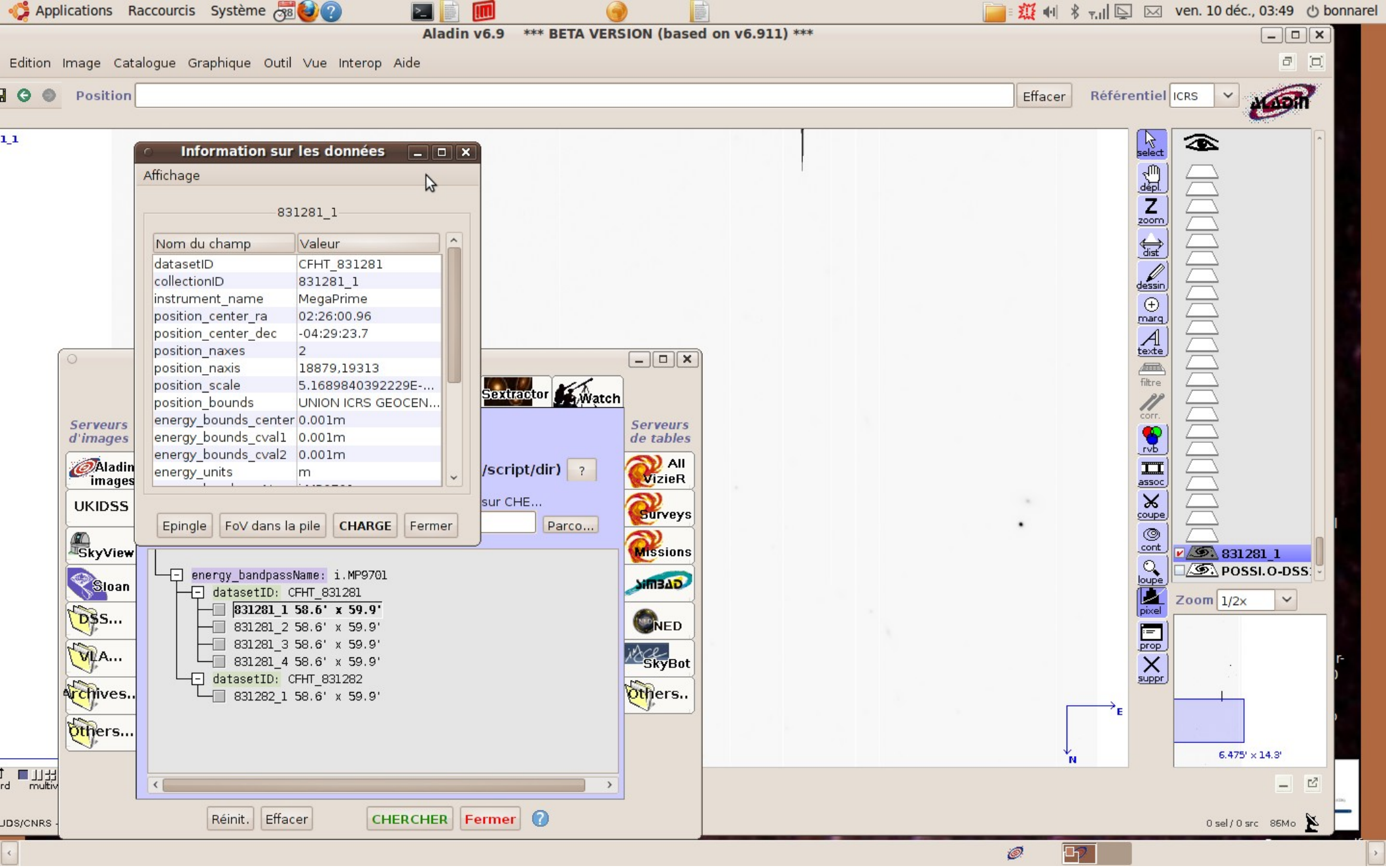
- A new service with one mandatory parameter creatorDID. (can be a list)
- May have other parameters to be defined
- The response is a votable with following columns:
  - CreatorDid
  - Relation name between two sides of the link (full metadata, spectra in the field, cutout, individual CCD in a mosaic, resampled spetrum, calibration files, ettc...)
  - URL type: Retrieval, S*A response, S*A AccesData (using creatorDID),
  - Access columns (like in Char 2) : format, size, url, extraction parameters (extension, column, cutout, etc …)
- This service may be implemented in TAP
- This service may be called after S*A sessions
- PQL queries to Obstap could drive classical and



Discovery of datasets with PQL/Obstap



Opening the data link node : Cutout interface



From second node: sia query response, then CCD image retrieval