# Les données Radio (et millimétriques) dans l'Observatoire Virtuel :
## état de l'art et questions ouvertes

F.Bonnarel (CDS)

remerciements à l'IVOA DAL WG, DM WG, TDIG, aux collaborations ASTERICS DADI, aux équipes XMM/SVOM et CDS de Strasbourg

CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

Astronomy ESFRI & Research Infrastructure Cluster

IVOA

# Résumé

- 0 ) Contexte politique de cette revue
- I ) rappel de ce qui fonctionne
  - Exemples de projets
  - Problèmes pour l' interopérabilité
  - Rappels des modèles de données utiles
  - Solutions pour la découverte: ObsCore, SIAV2
  - Solutions pour l'accès : DataLink et SODA
  - Discovery + Access + View : MOC et HiPS
  - Interfaces : Aladin, TOPCAT
- II) implémentation dans l'archive scientifique ALMA
  - Données disponibles dans l' archive
  - Interface ObsTAP
- III ) données ASKAP : implémentation CASDA
- IV ) Questions ouvertes et plus complexes
  - Caractérisation fine
  - Accès / traitement des données de type

# Cubes radio dans le VO contexte « politique »

- Aiguillon de Priorité IVOA « Multi-D »
- Précurseurs de SKA dans ASTERICS : ALMA,LOFAR, ASKAP. Résultats limités
- Participation de SKA dans le projet ESCAPE = changement d'échelle.
- Réunion du 28 février pendant le dernier Tech Forum ASTERICS/ DADI → lancement ESCAPE

# Radio data Cubes

- ALMA, CADC
- LOFAR, ASKAP, JIVE
- Vers SKA

→ découverte des données par sélection de critères portant sur les attributs de description ou métadonnées

→Accès aux données : extraction de données utiles dans les « datasets » découverts

# Accès aux données

Les Cubes radio sont « énormes »

Le chargement du cube complet est trop long

→ Nécessité des cutouts ?

→ Finalement veut-on du traitement à la volée (données virtuelles, rééchantillonage)?

→ nécessité de métadonnées détaillées au-delà d'ObsCore (Modèle de données Cube , qui peut venir plus tard)….

# Modélisation OV

## 1 ) Obscore : métadonnées pour la découverte

| | | | |
|---|---|---|---|
| obs_collection | unitless | String | Name of the data collection |
| obs_id | unitless | String | Observation ID |
| obs_publisher_did | unitless | String | Dataset identifier given by the publisher |
| access_url | unitless | String | URL used to access (download) dataset |
| access_format | unitless | String | File content format (see in App. BB.5.2 ) |
| access_estsize | kbyte | integer | Estimated size of dataset in kilo bytes |
| target_name | unitless | String | Astronomical object observed, if any |
| s_ra | deg | double | Central right ascension, ICRS |
| s_dec | deg | double | Central declination, ICRS |
| s_fov | deg | double | Diameter (bounds) of the covered region |
| s_region | unitless | String | Sky region covered by the data product (expressed in ICRS frame) |
| s_xel1 | unitless | integer | Number of elements along the first spatial axis |
| s_xel2 | unitless | integer | Number of elements along the second spatial axis |
| s_resolution | arcsec | double | Spatial resolution of data as FWHM |
| t_min | d | double | Start time in MJD |
| t_max | d | double | Stop time in MJD |
| t_exptime | s | double | Total exposure time |
| t_resolution | s | double | Temporal resolution FWHM |
| t_xel | unitless | integer | Number of elements along the time axis |
| em_min | m | double | Start in spectral coordinates |
| em_max | m | double | Stop in spectral coordinates |
| em_res_power | unitless | double | Spectral resolving power |
| em_xel | unitless | integer | Number of elements along the spectral axis |

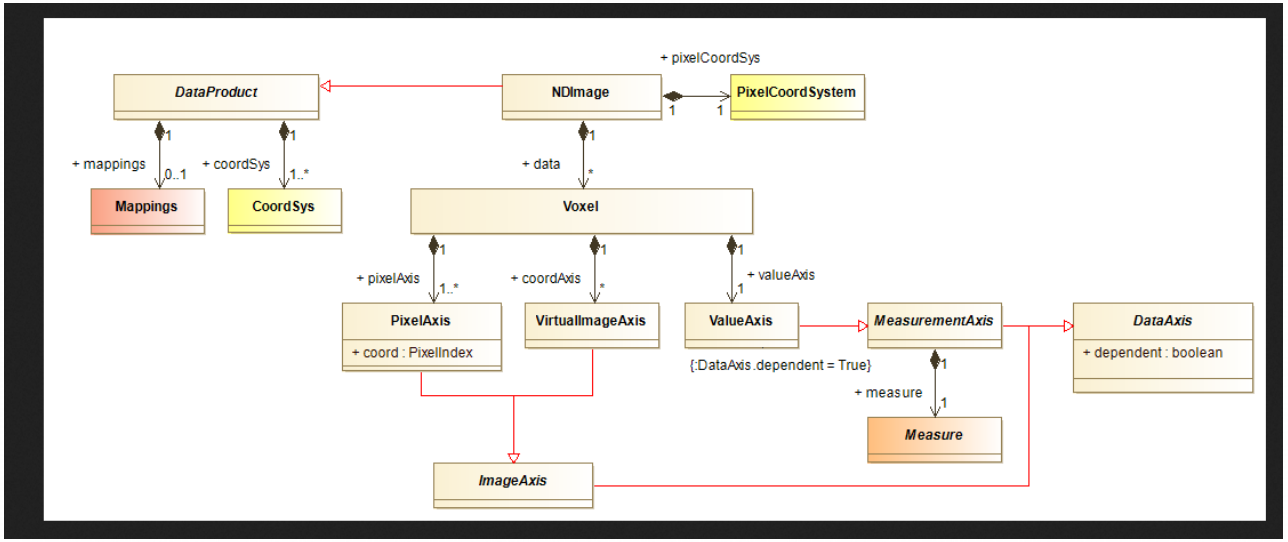
# Modélisation OV

## 2 ) Autres modèles utiles existants

- STC 1.0 : concepts utiles, actuellement en révision majeure (Coords, Measurements et Transformations)
- PhotDM : principalement pour l' astronomie optique
- Characterisation 1.0 : réutilisé par d'autres modèles de données (comme ObsCore, Cube) en révision
- Spectral Line Data Model : Des observations à la Physique
- Modèle de données Simulation et protocole DAL. Comparaison avec les Simulations

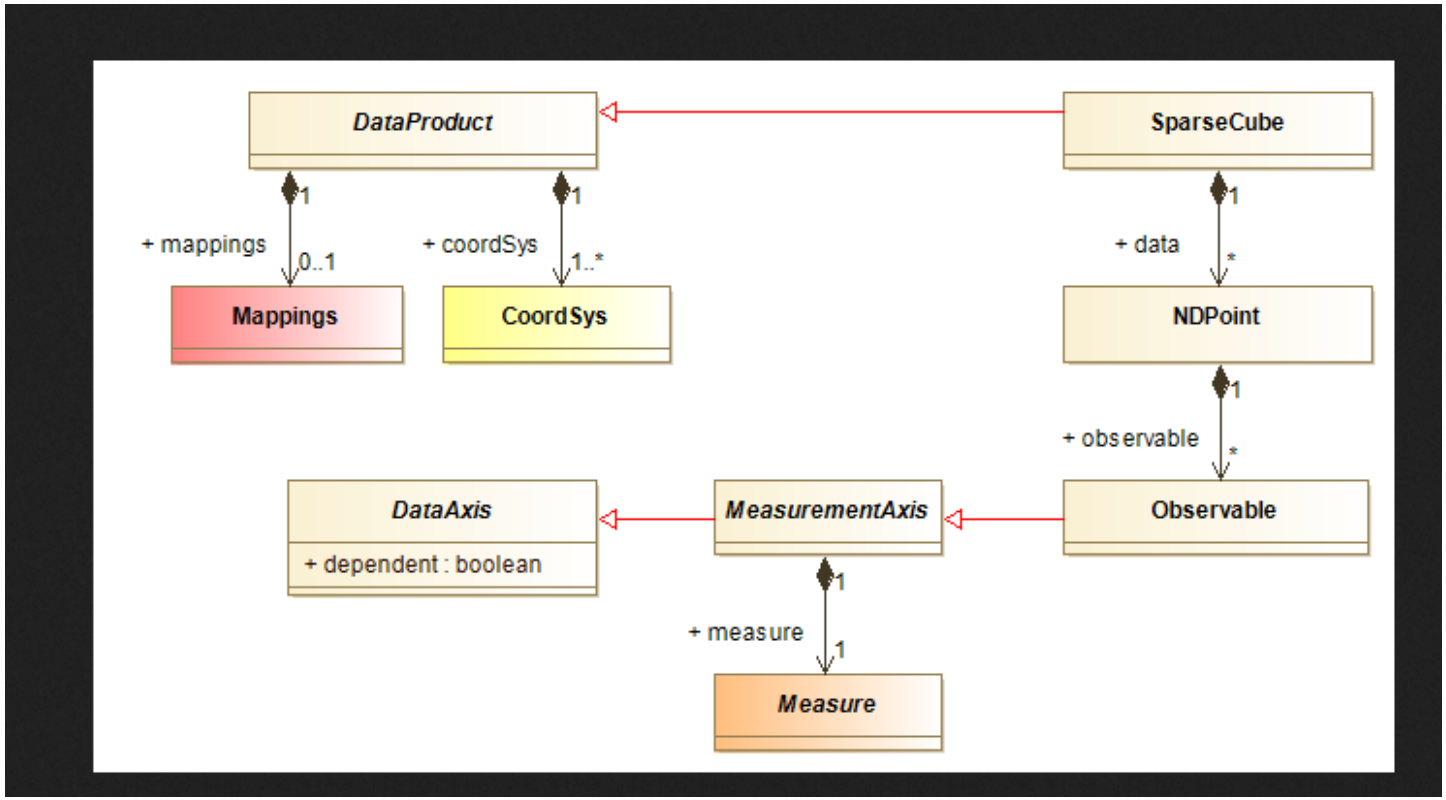
# Modélisation OV

## 3 ) Cube datamodel (travail en cours) : Image ND



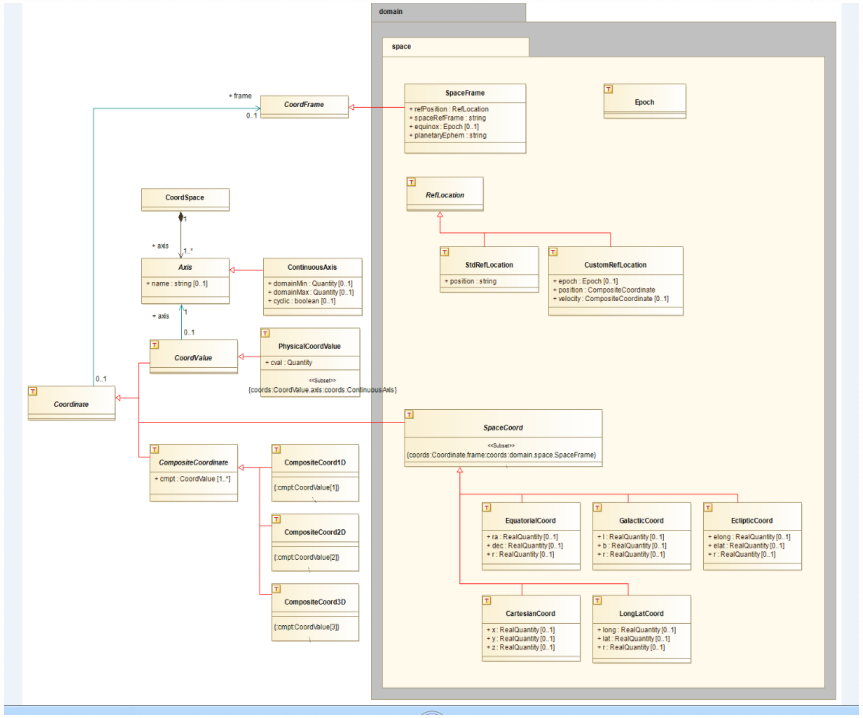
# Modélisation OV

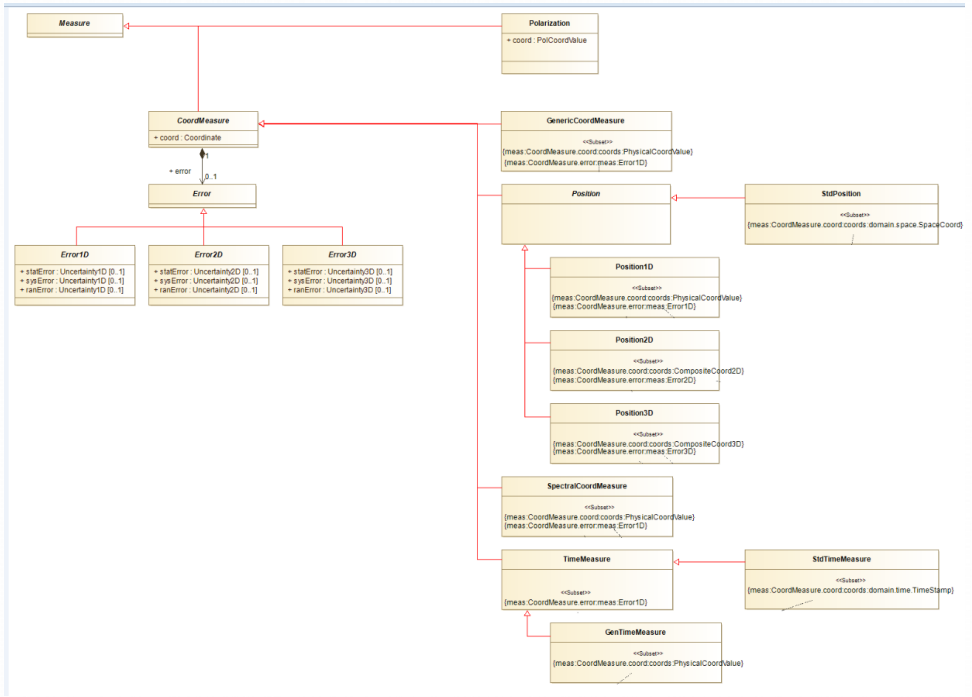
## 3 ) Cube datamodel (travail en cours) : sparse cube



# Modélisation OV

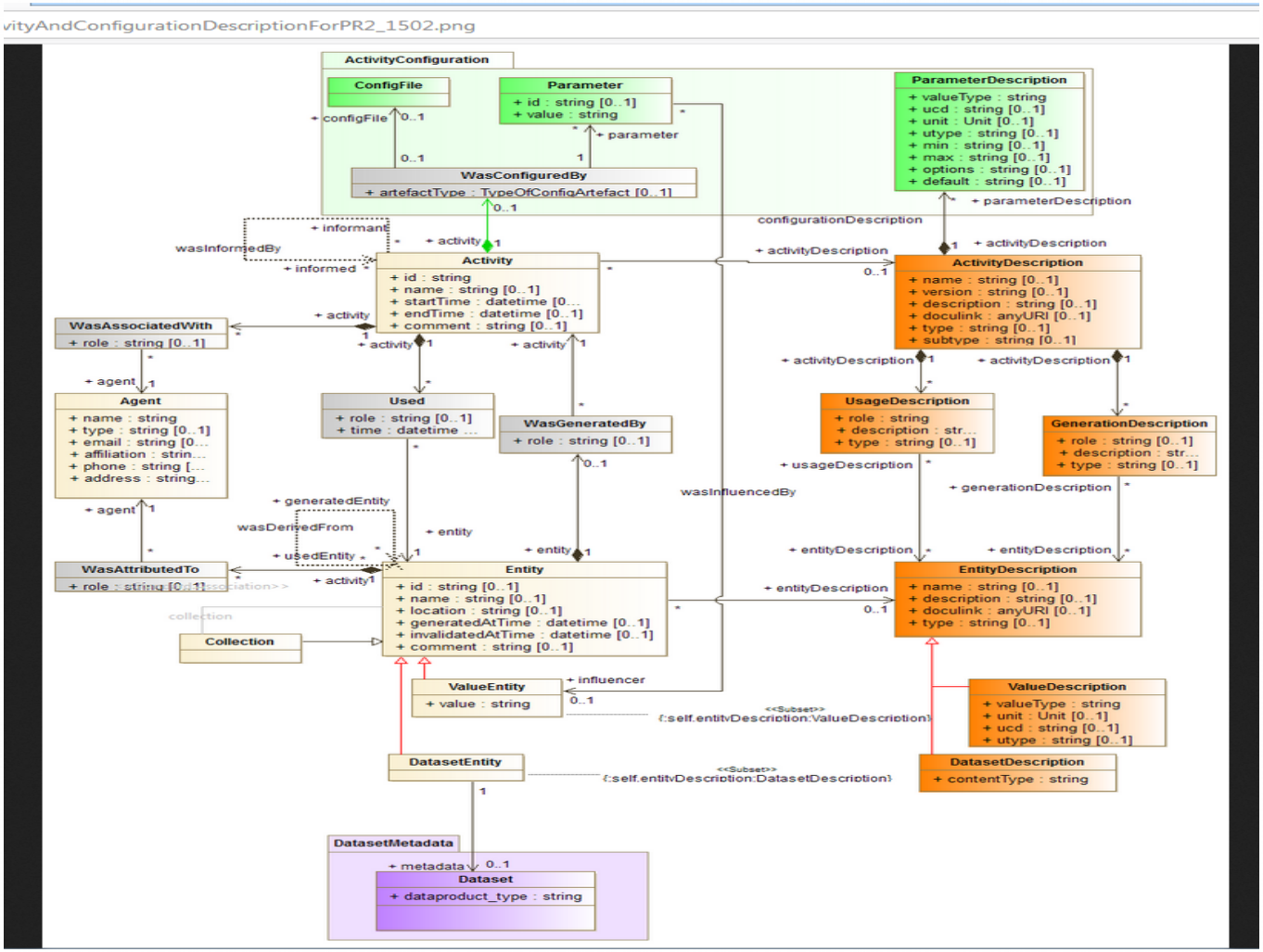
## 4 )STC coords et measurements (travail en cours)



# Modélisation OV

## 5) Provenance (travail en cours)

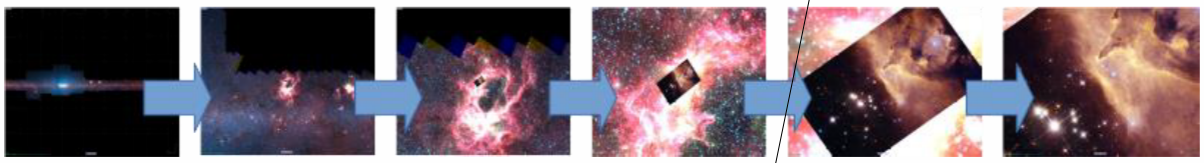


# Couche d'accès aux données OV Solutions

- ObsTAP
- SIAV2
- Chargement du dataset complet ou
  - DataLink et SODA
- Alternative : Découverte , Représentation et Visualisation de cubes en  HiPS

# HiPS ? What is it ?

- **Hi**erachical **P**rogressive **S**urvey
  *"The more you zoom in on a particular area,the more details show up"*
- Multi-resolution **HEALPix** data structure
- for **Images**, **Catalogues**, 3-dimensional data **cubes**, ...
- **Conserves scientific** data properties alongside visualisation considerations
- No databases or dedicated servers are required, just http



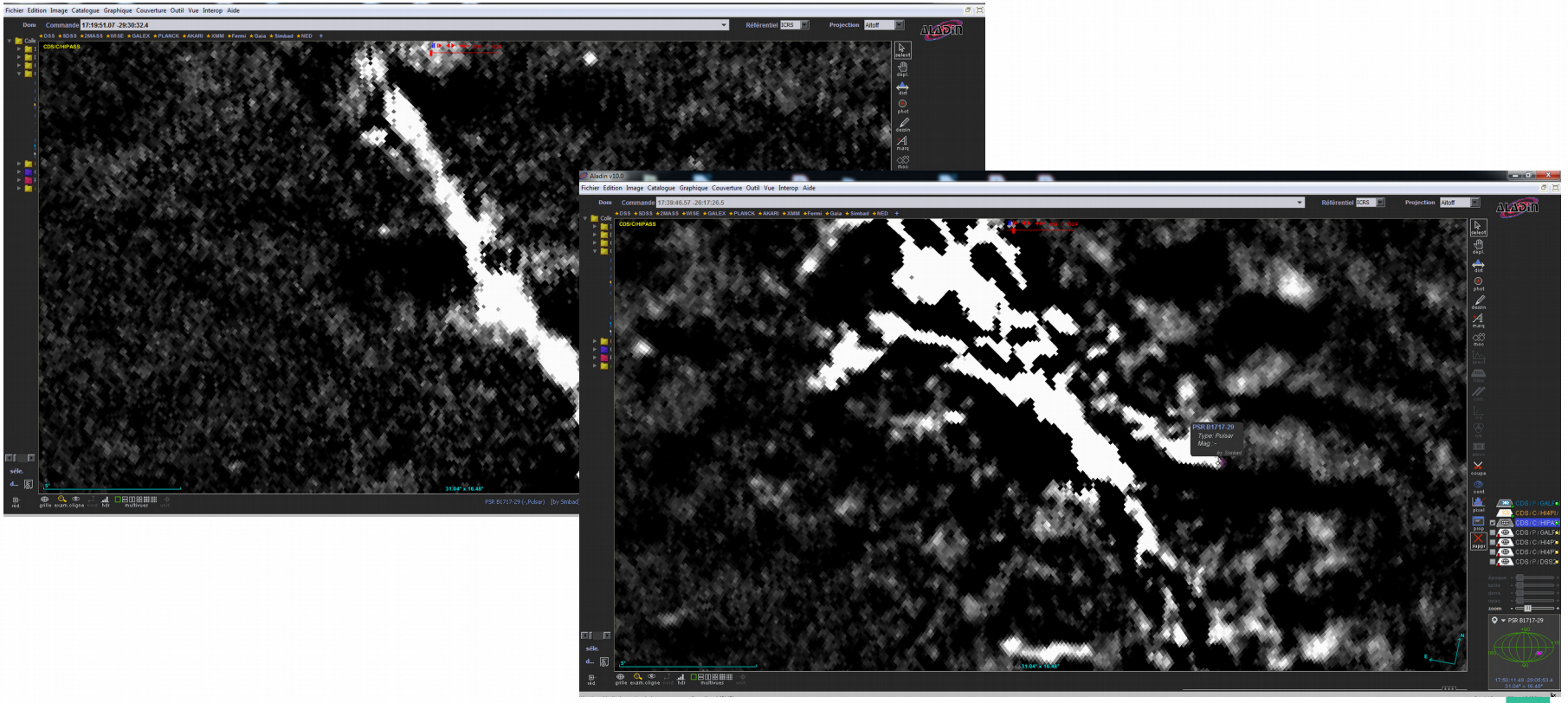


Et MOC pour découverte : indexation Par cellule HEALPIX

View, Access, Discovery of DataCube

# Cube HiPS

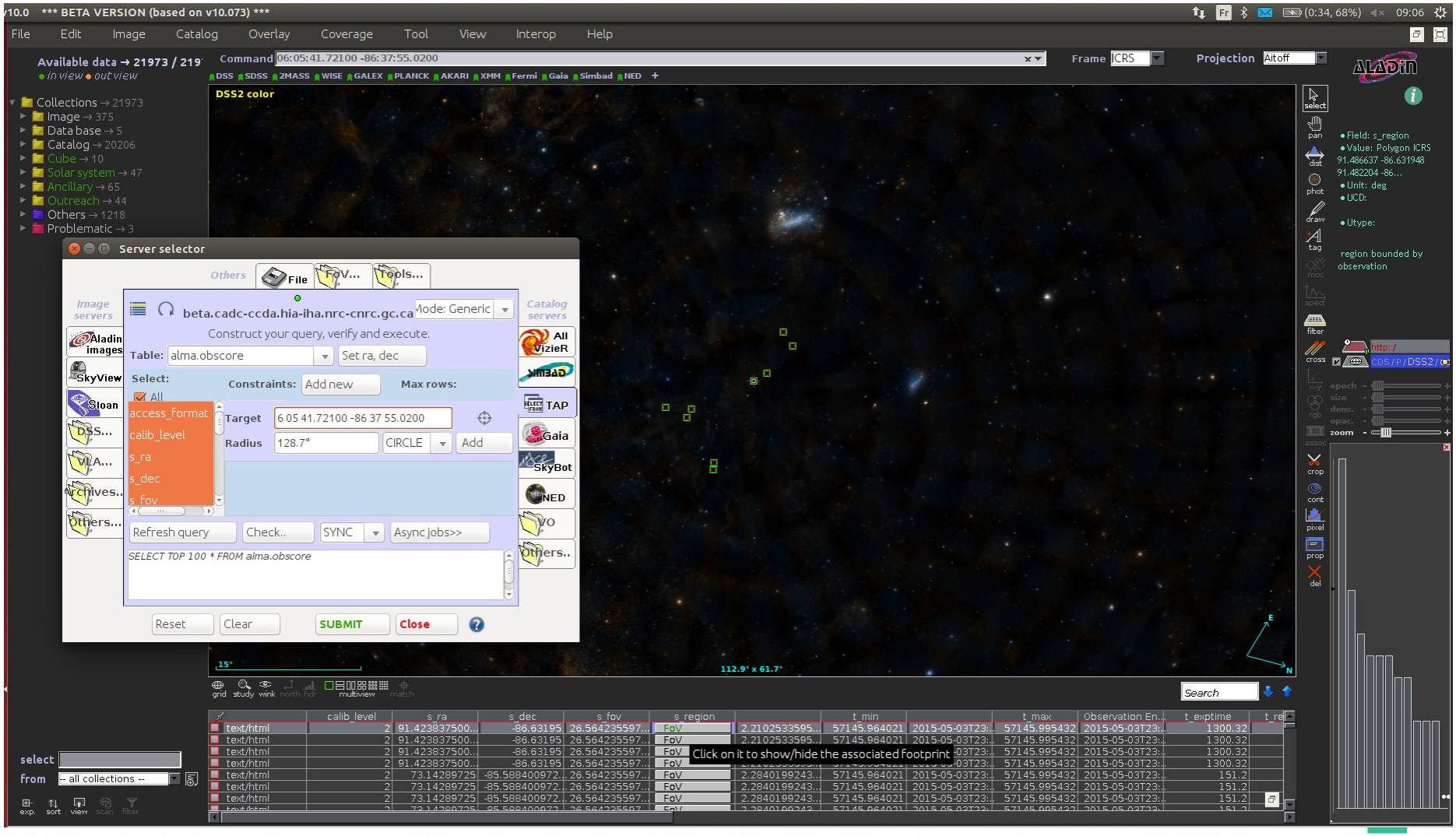
. HI4PI (Effelsberg), GALFAHI (Arecibo), et HIPASS (Australie)



# Archive scientifique ALMA

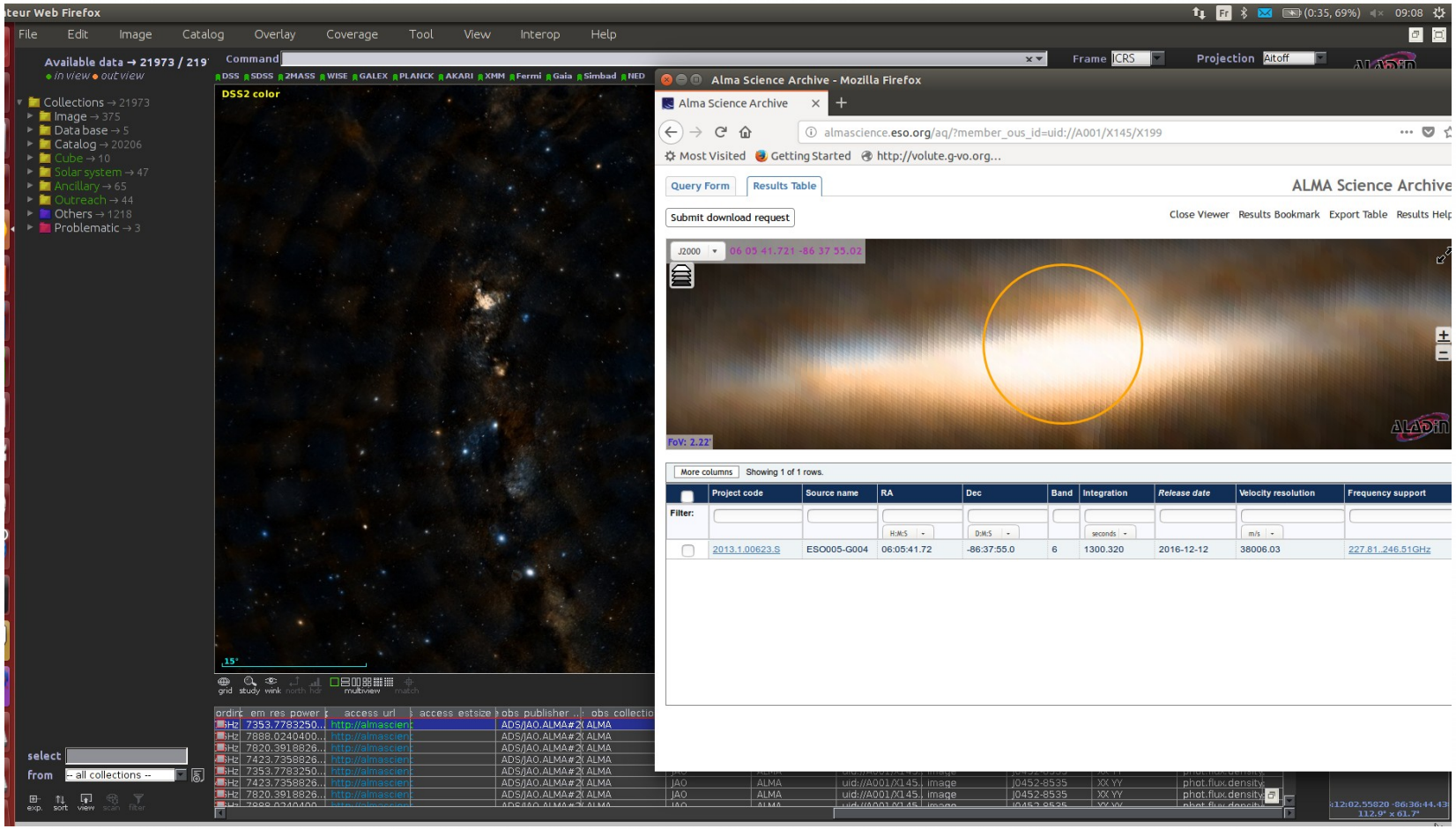
- ALMA fournit des données interférométriques dans le domaine millimétrique
- L'archive a un interface natif
- Cubes en imagerie pas disponibles partout
- L'archive ALMA délivre lentement des blocs d'interfaces OV :
  - Un service ObsTAP est dispo (voir diapos suivantes)
  - MOC est disponible
  - Intégré dans ESA Sky (avec AladinLite)
  - Projets d'intégration : DataLink, SIAV2, SAMP dans l'interface
  - Un projet de Service de retraitement installé en UWS

# Archive scientifique ALMA service (obs)TAP dans Aladin



# Archive scientifique ALMA

l'url dans la réponse ObsCore mène à une  page affichée dans le navigateur (pas encore de transfert direct de données)

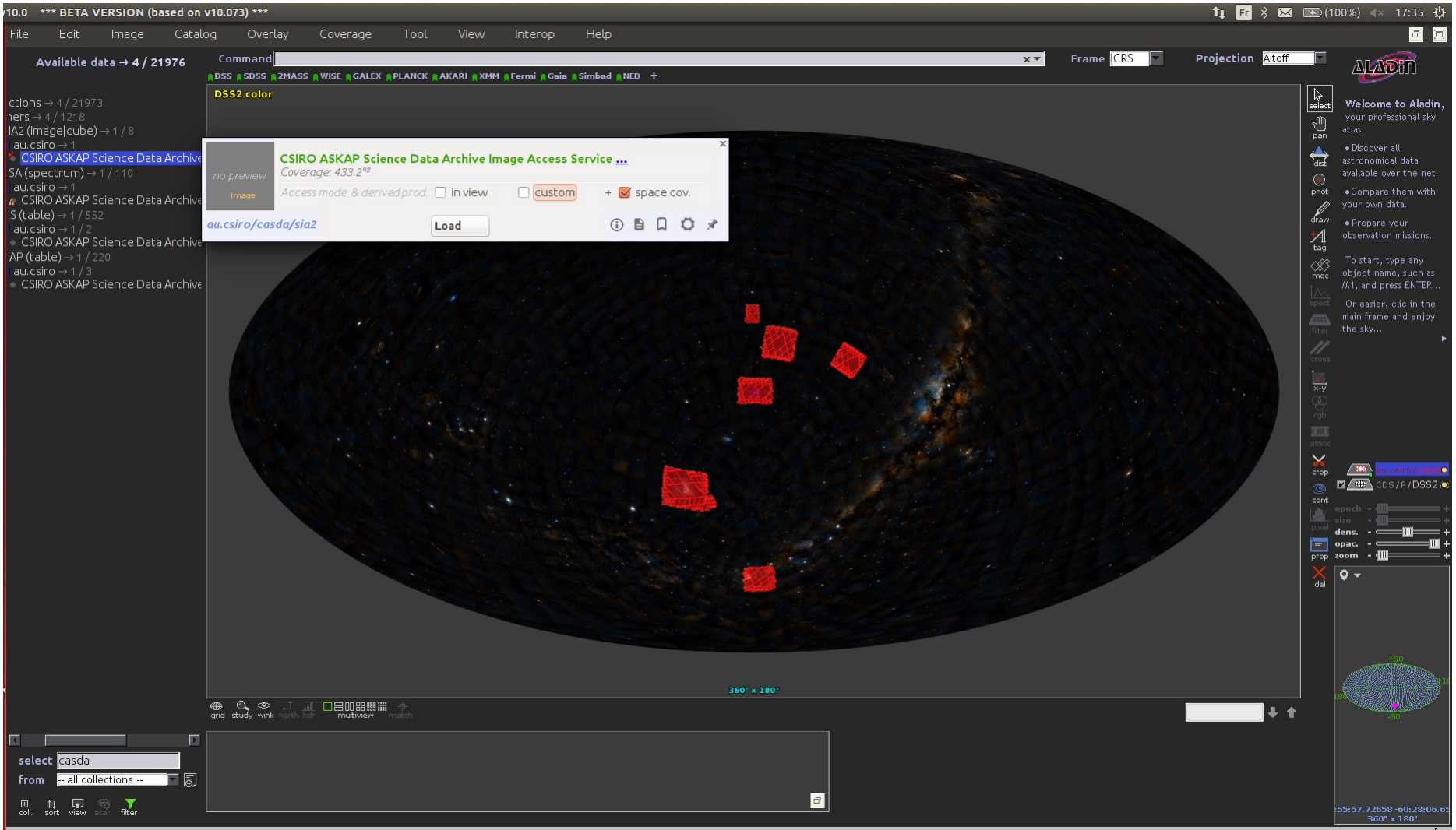


# Données ASKAP : services CASDA

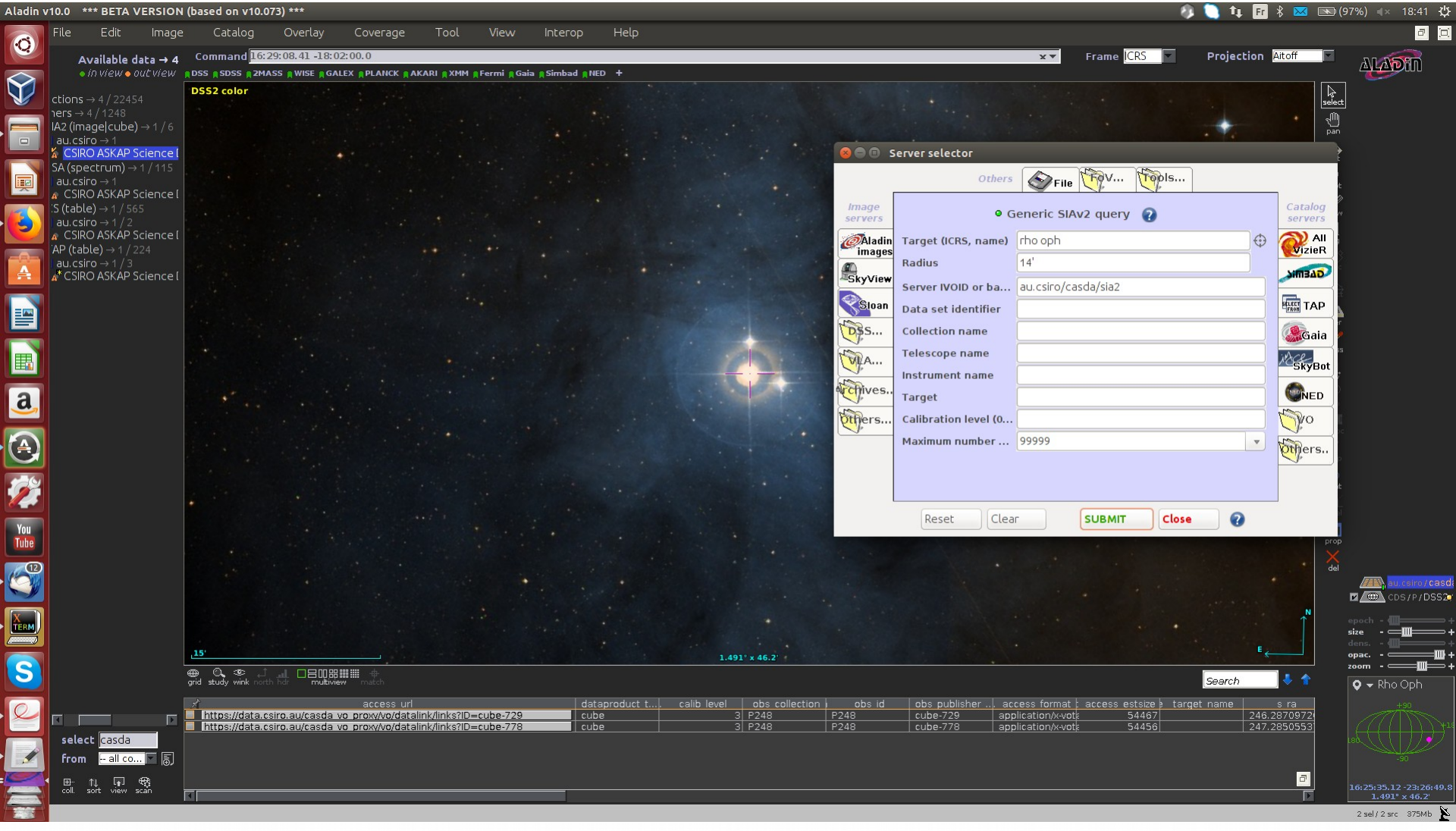
- CASDA a implémenté
  - MOC
  - ObsTAP, et services SIA disponibles
  - DataLink avec Cutouts
  - Interface SODA
  - Scripts

# CASDA (= données ASKAP)

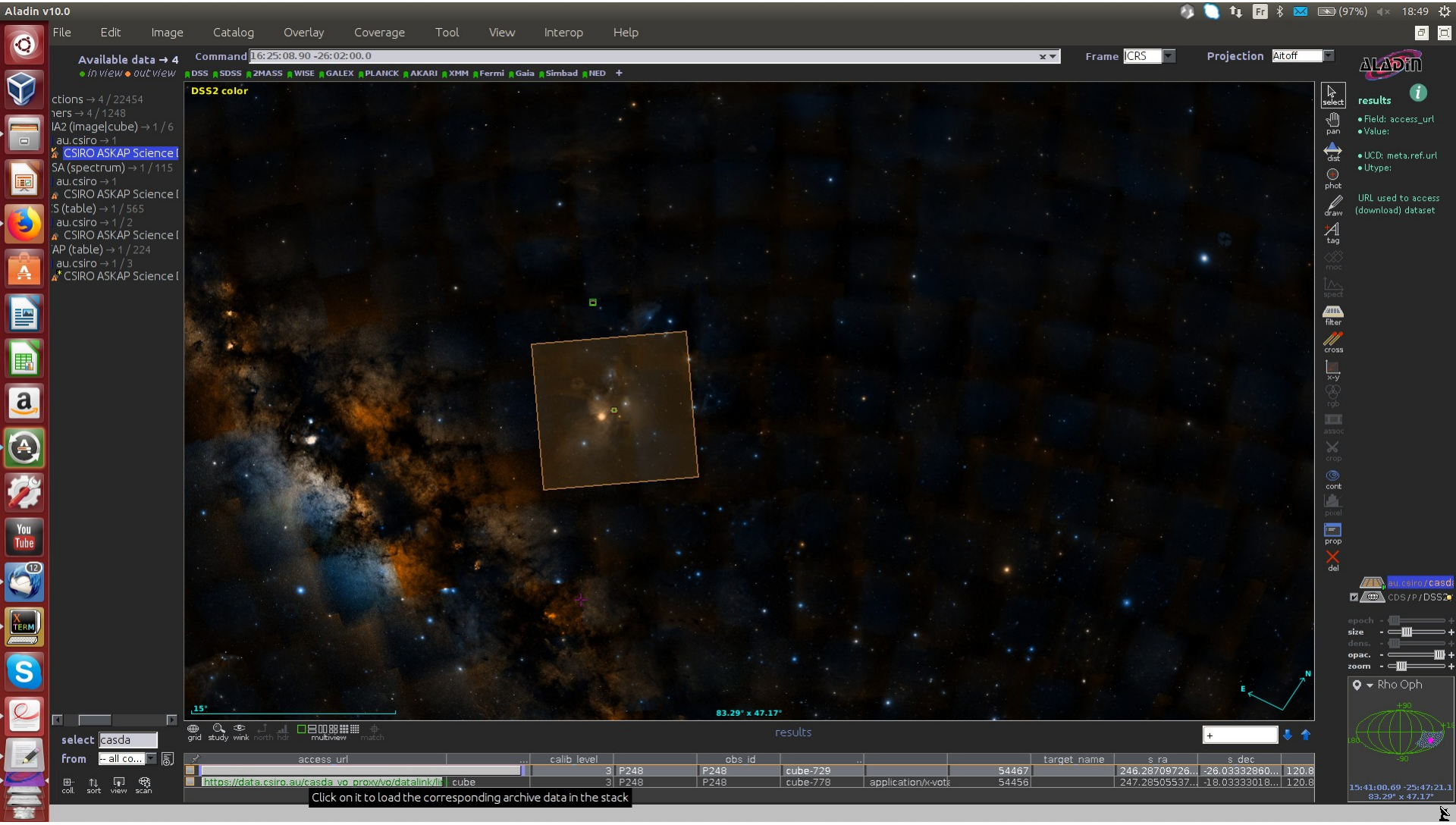
## ( distribution des cubes de données sur tout le ciel)



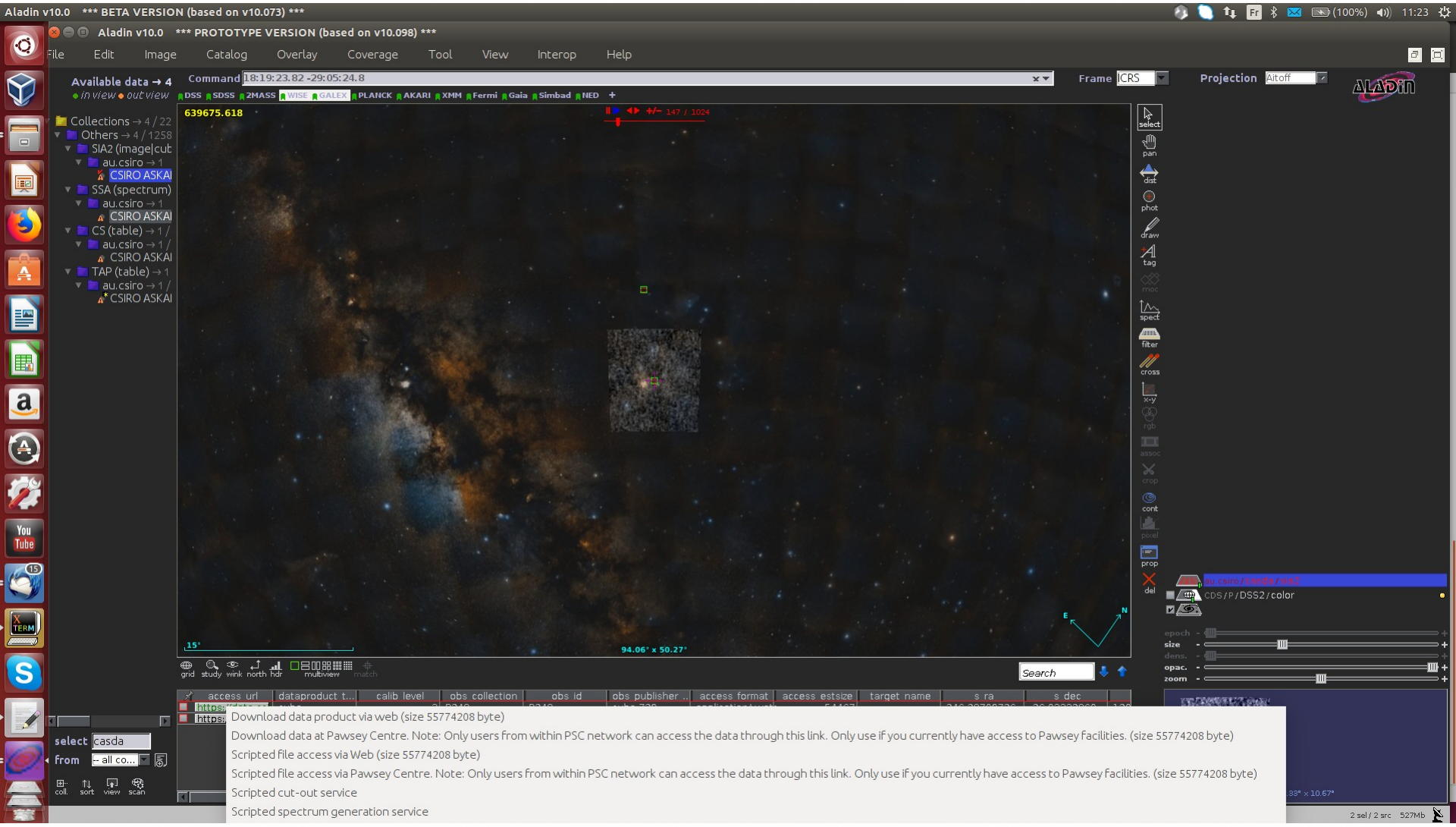
# Données ASKAP : services CASDA
# SIAP 2.0



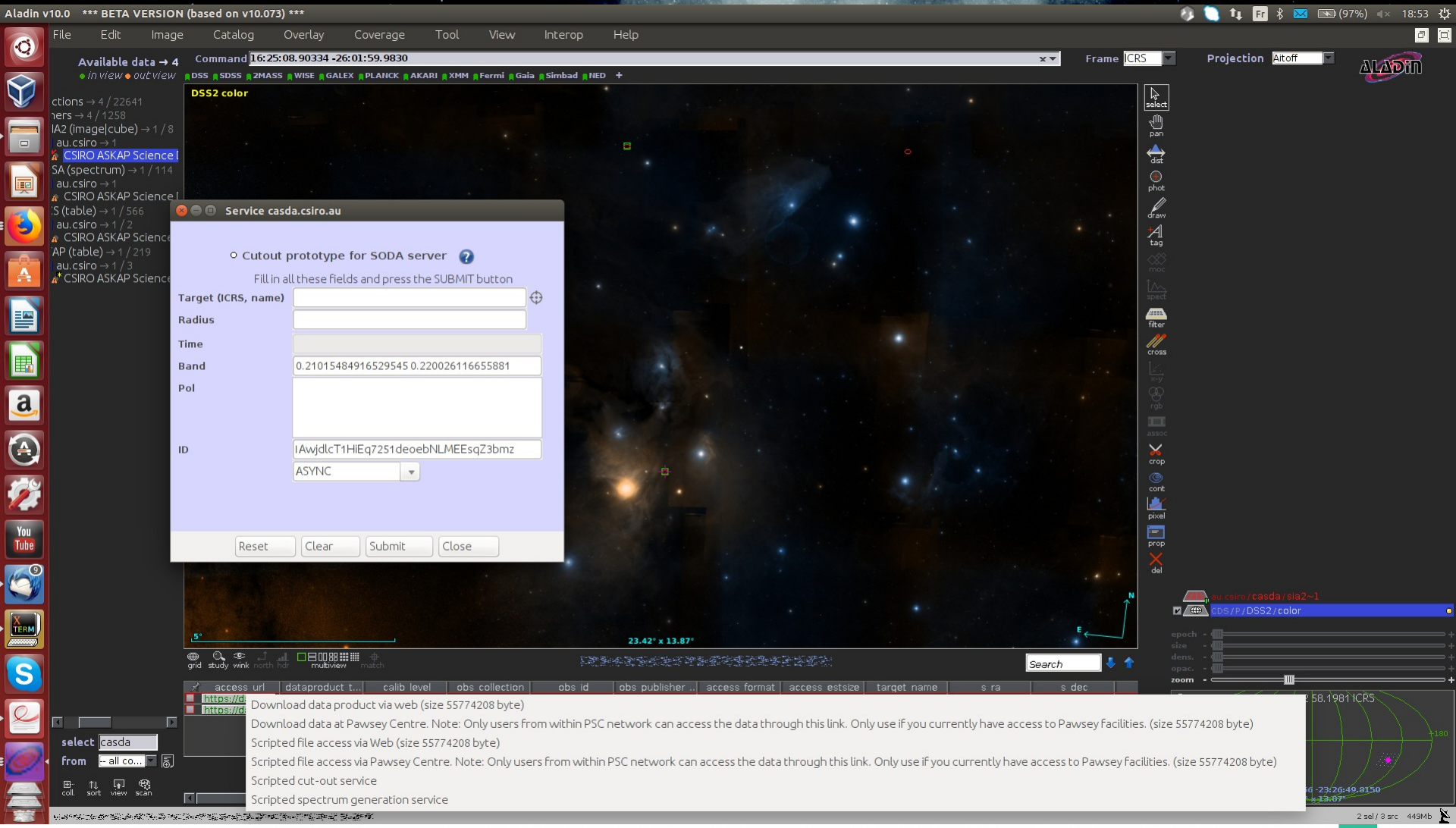
# Données ASKAP : services CASDA
# Champ de vue du dataset



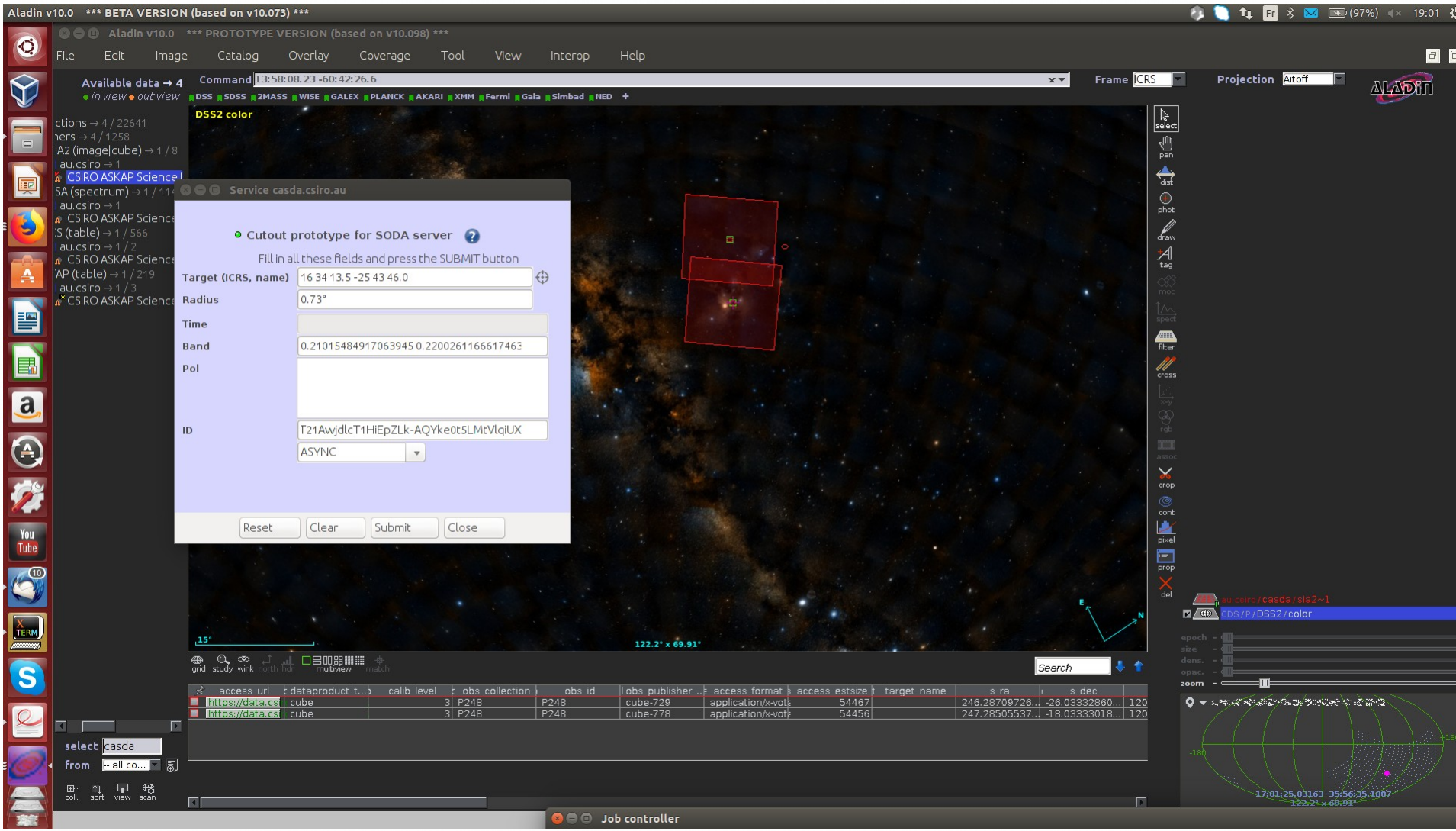
# Données ASKAP : services CASDA DataLink



# Données ASKAP : services CASDA interface SODA

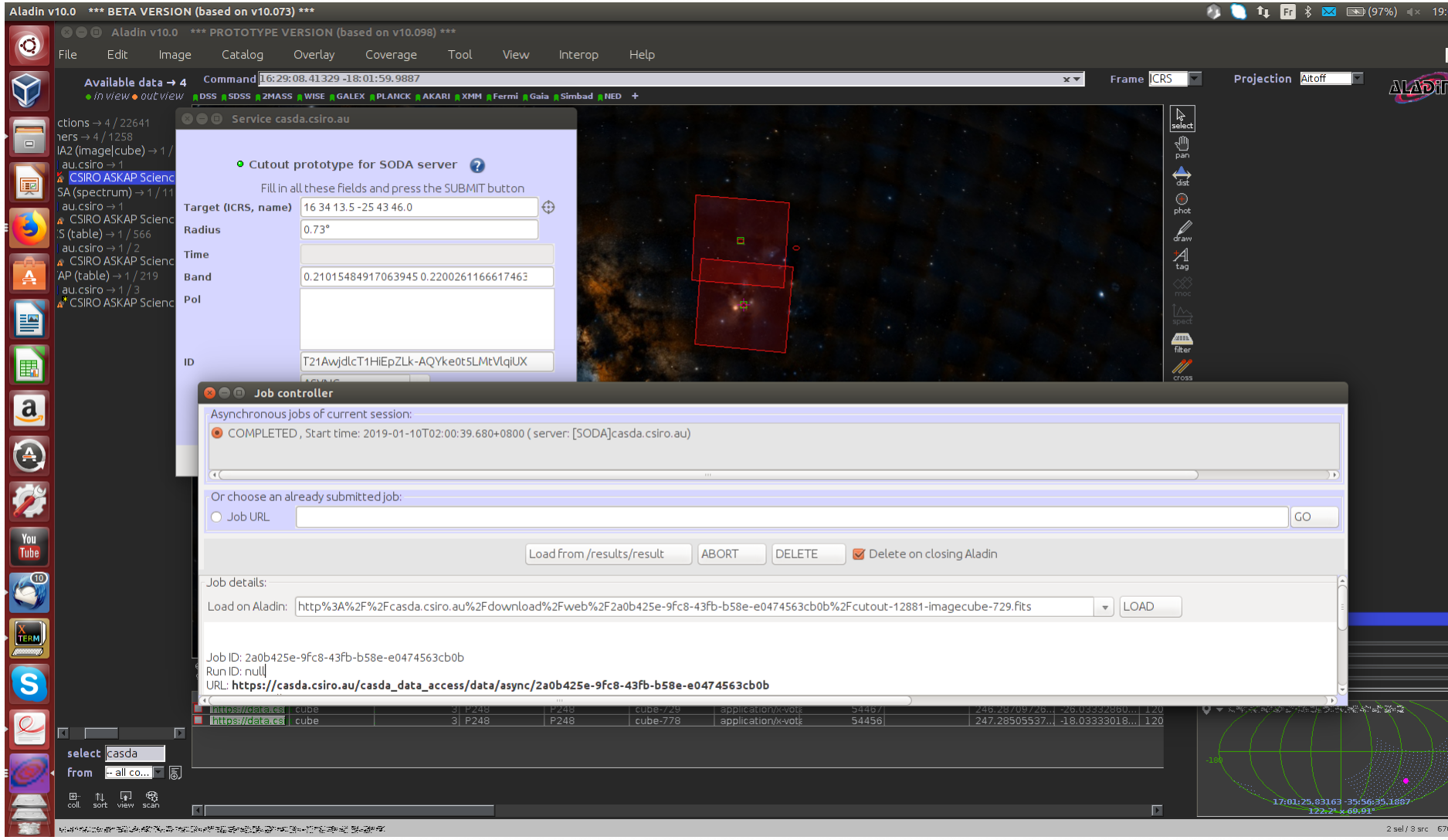


# Données ASKAP : services CASDA
# SODA : Prêt à lancer la requête

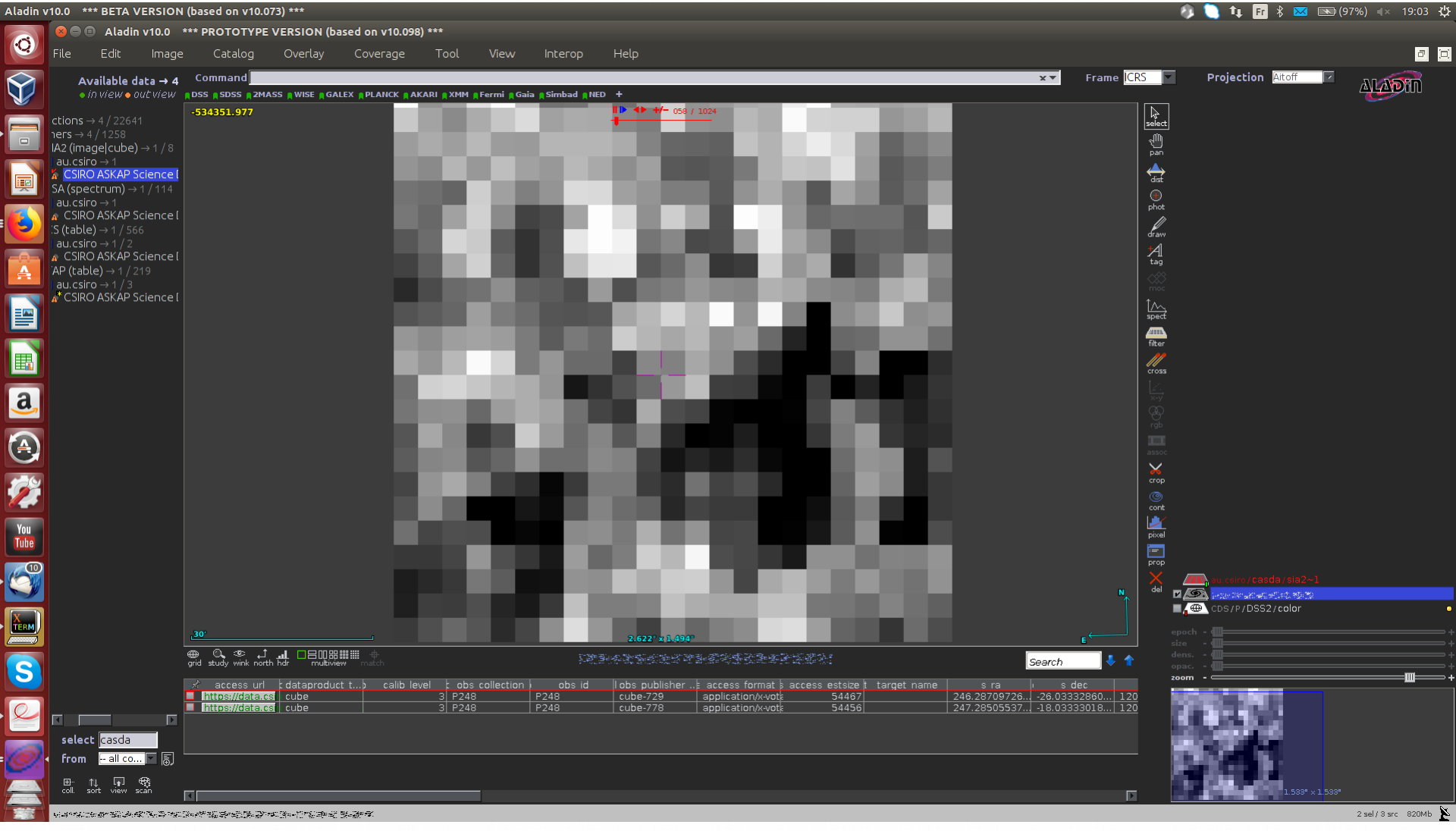


# Données ASKAP : services CASDA

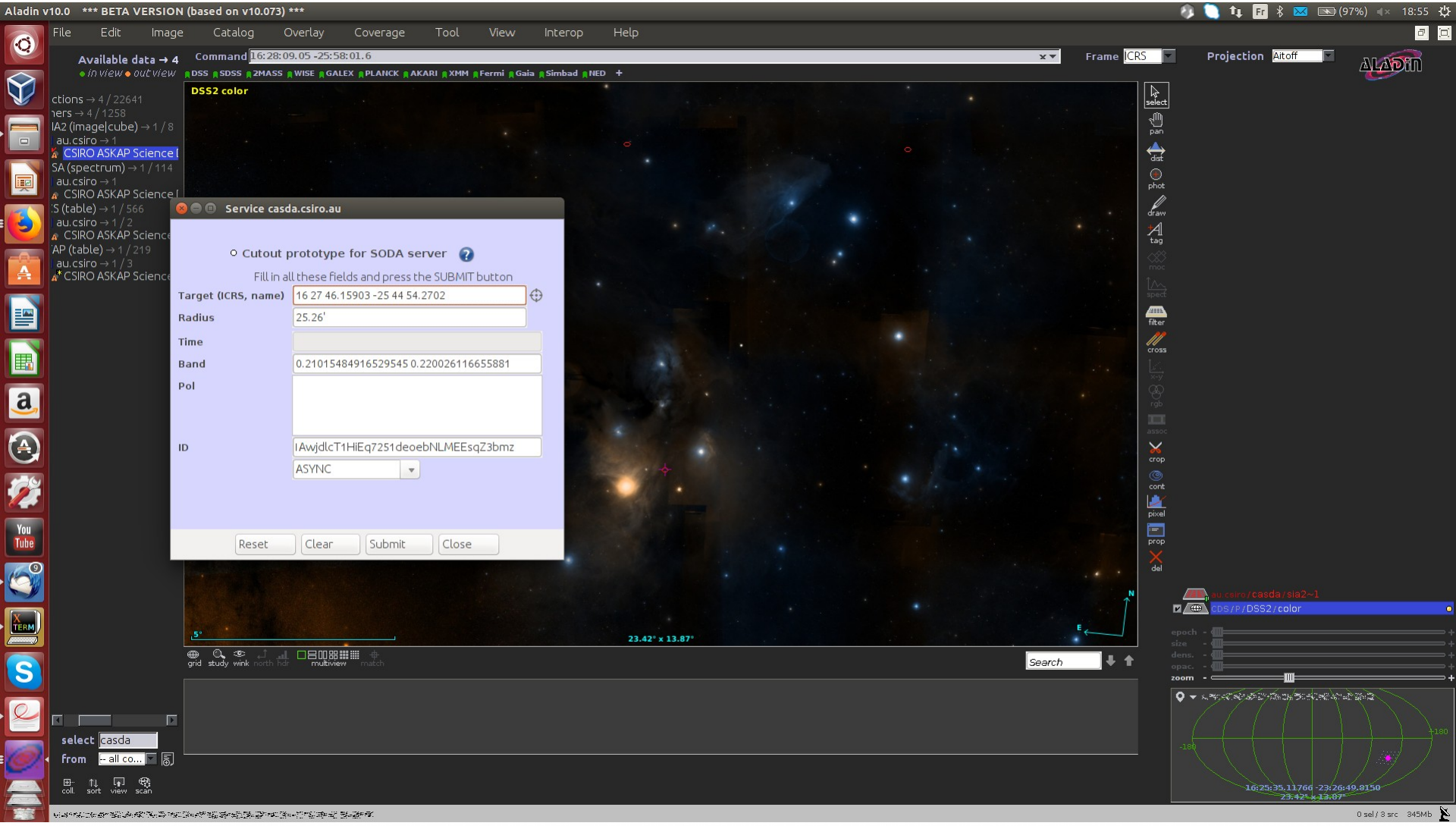
## La tâche SODA asynchrone terminée



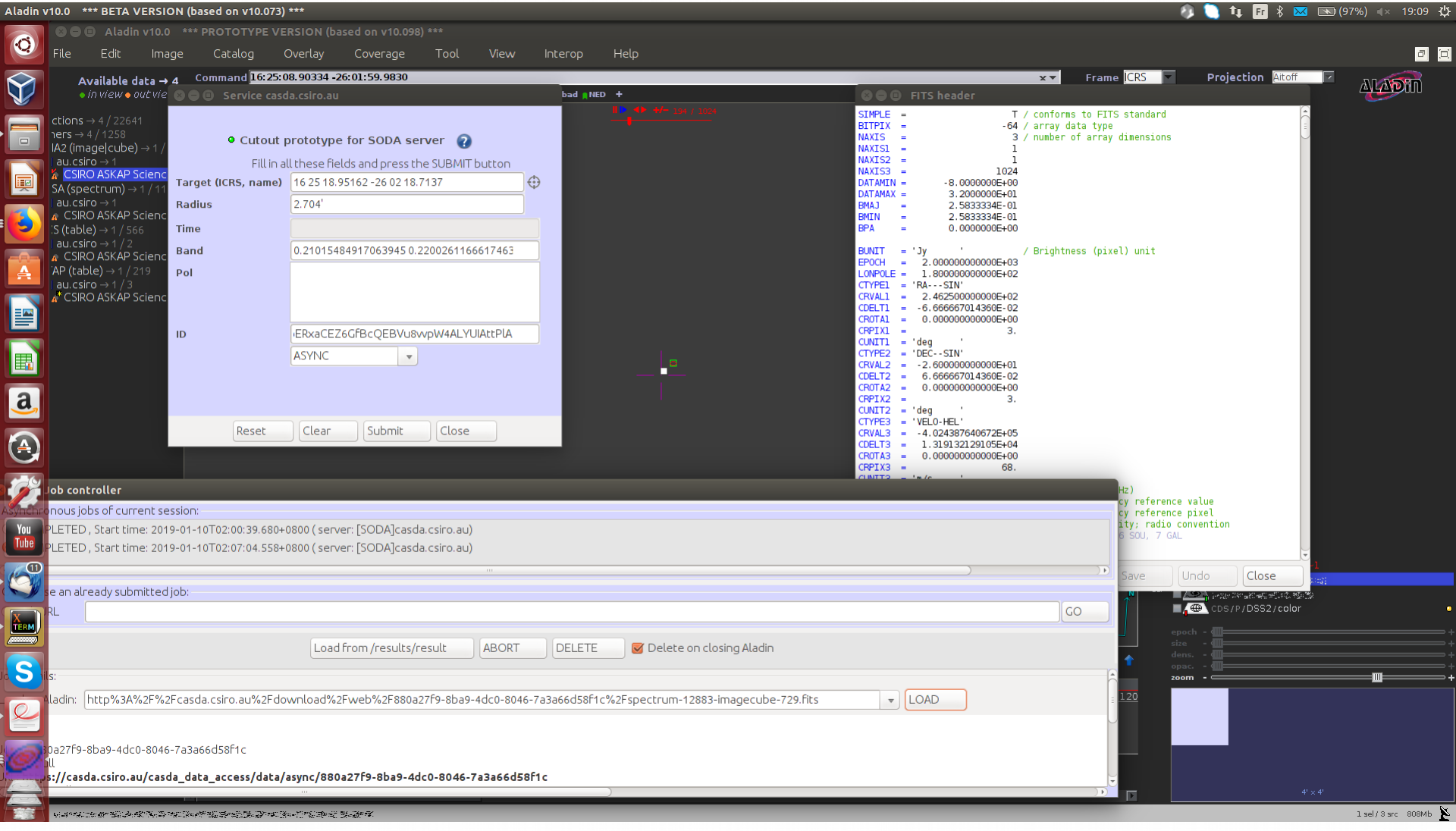
# Données ASKAP : services CASDA sous - cube extrait et chargé



# Données ASKAP : services CASDA spectre extrait



# Données ASKAP : services CASDA
# spectre extrait



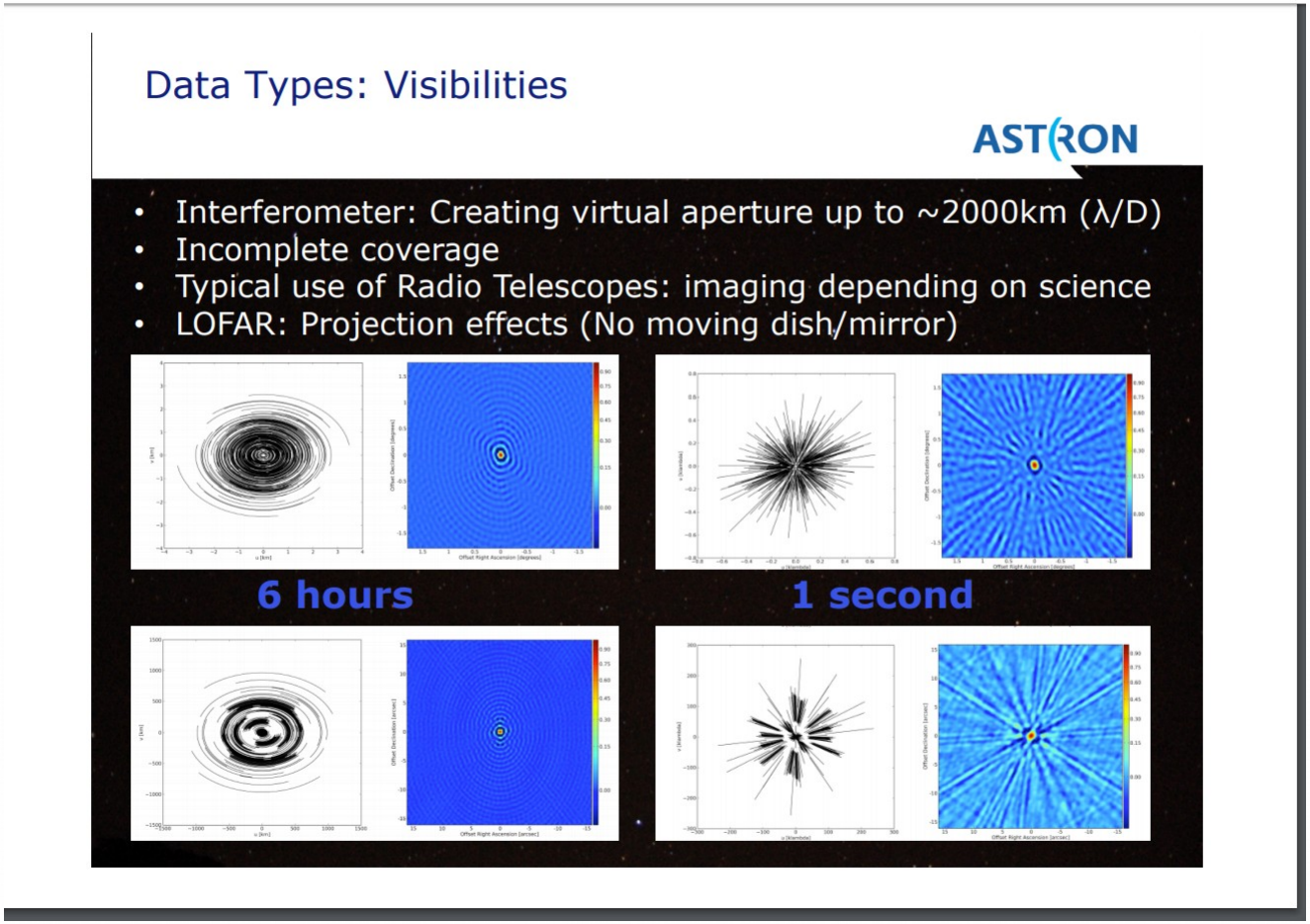
# Limitations

- Données complexes (plusieurs sous-parties) et volumineuses
- Caractérisation ObsCore insuffisante en terme de sensibilité, de couverture du plan UV, etc...
- Beaucoup de données non réduites dans les archives (« visibility » data).
- Réductions possibles non uniques

# Limitations. Comment les lever ?

- Données complexes (plusieurs sous-parties) et volumineuses
- Compléter Obscore avec une stratégie à la CAOM (parties et wcs accessibles)?
- Proposition de faire des HiPS des cartes continuum ou de cartes de moment, et d'accéder aux vraies données via l'interface « original data ».

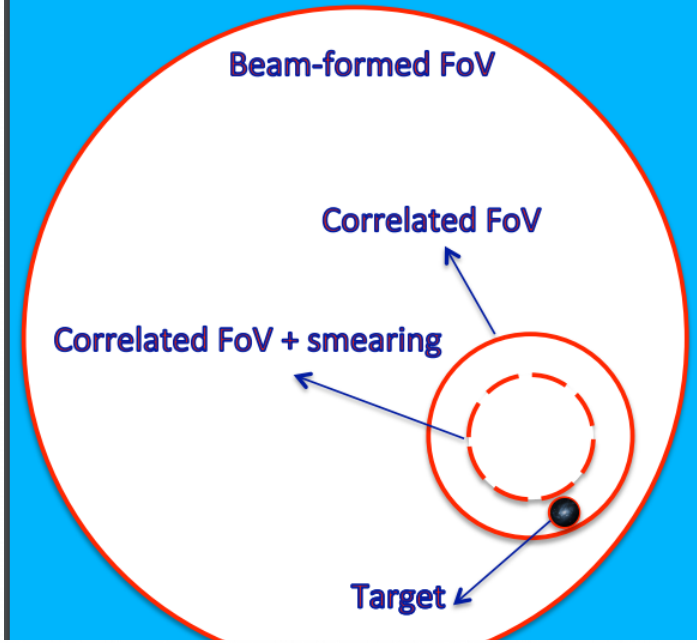
# Données Visibility ?
## (découverte directe ou via provenance)



# Portail de données ASTRON

## The ADP & new tooling

Characterization tools needed when dealing with level 0/1 data products from wide-field interferometers ➔ accounting for beam-shapes of different modes for searches and selection but also for processing / analysis




Observed / processed data characterization in terms of:

- uv-coverage
- beam pattern (FoV & sensitivity)

Accurate reconstruction:

- computationally expensive, need data ➔ not practical

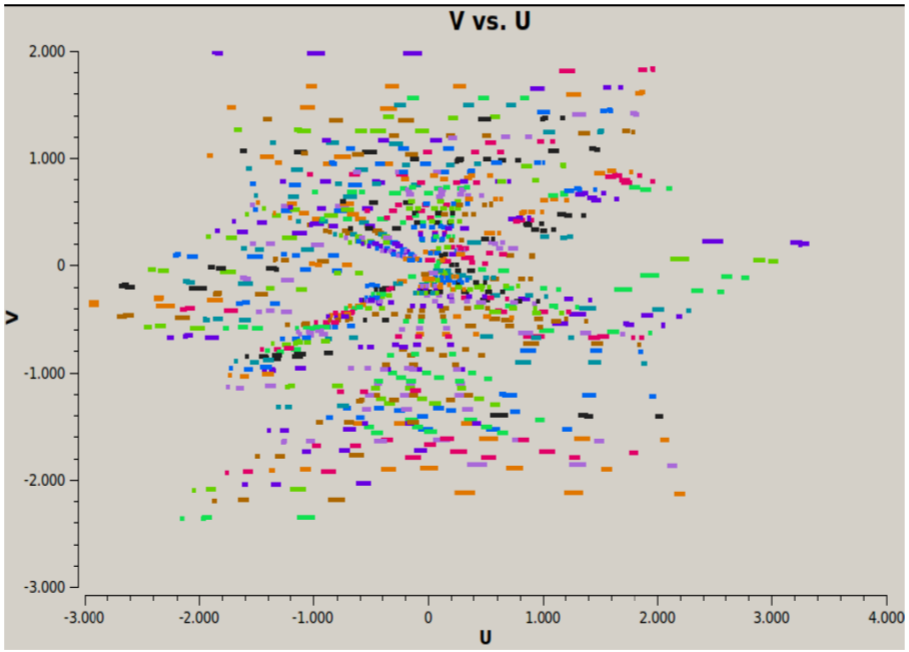
Approximated reconstruction:

- only metadata needed ➔ enough for data selection

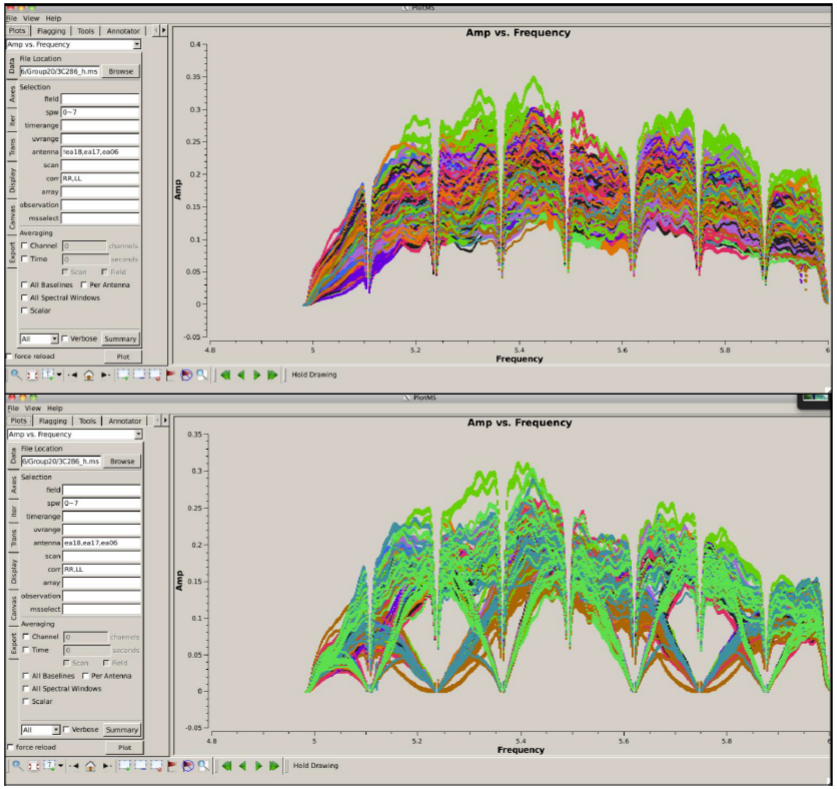
# Limitations : comment les lever ?

- Découverte « gros grain » via ObsCore + métadonnées additionelles (caractérisation fine ) via DataLink
  - Plot UV ( radialement, ou dans plan uv)
  - Courbes fréquence / amplitude, ou phase / amplitude
  - Forme et taille du « beam »
  - Nombres d'antennes, détails instrumentaux
- Caractérisation dynamique (quelles données brutes assembler pour un but recherché) --→ Web service ou Application dédiée

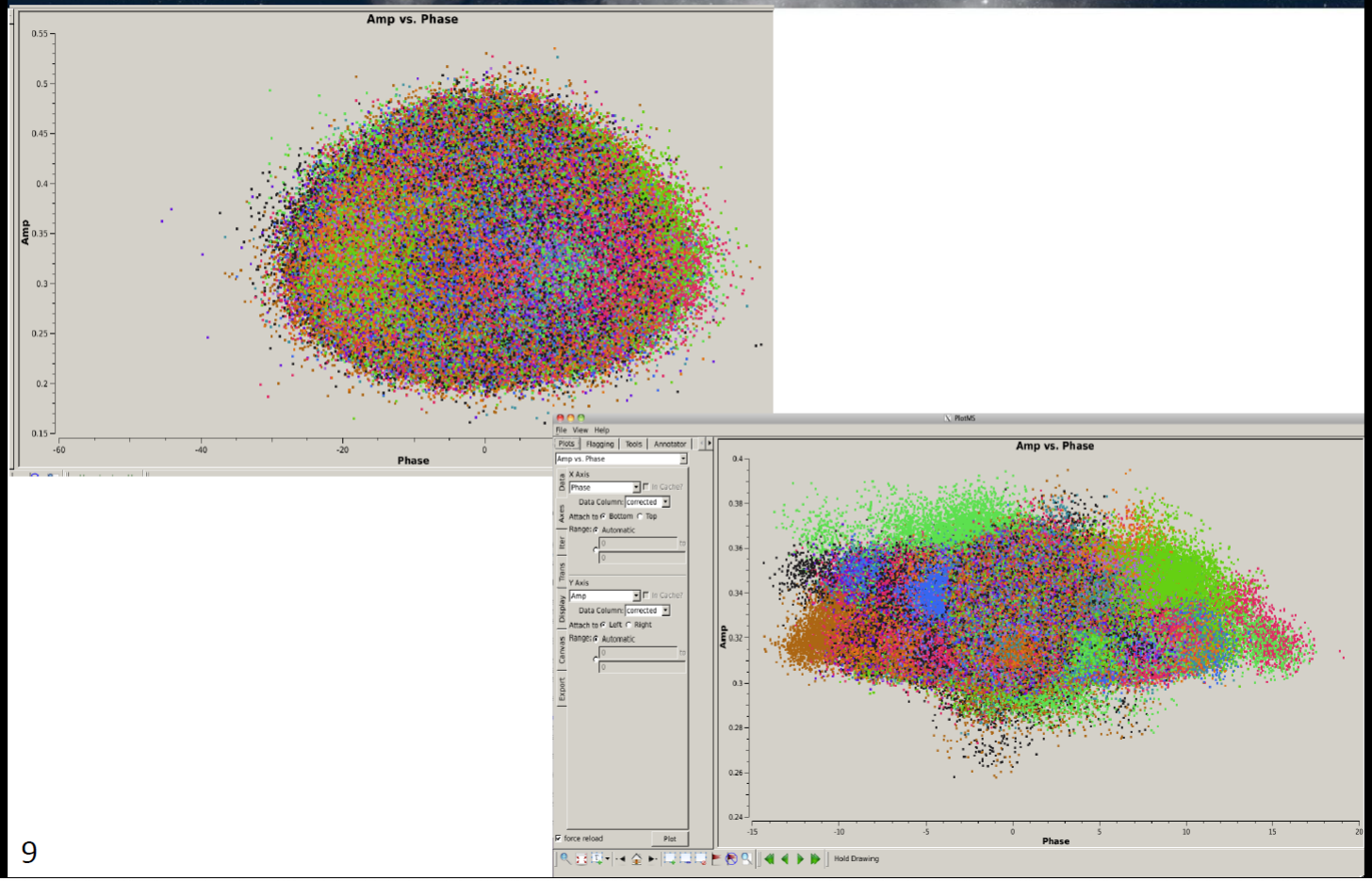
# Uv coverage


V vs. U
2.000
1.000
0
-1.000
-2.000
-3.000
V
-3.000
-2.000
-1.000
0
1.000
2.000
3.000
4.000
U

# Amp vs. Freq


Amp vs. Frequency
Amp
Frequency
Amp vs. Frequency
Amp
Frequency

# Amp vs. phase



# Limitations

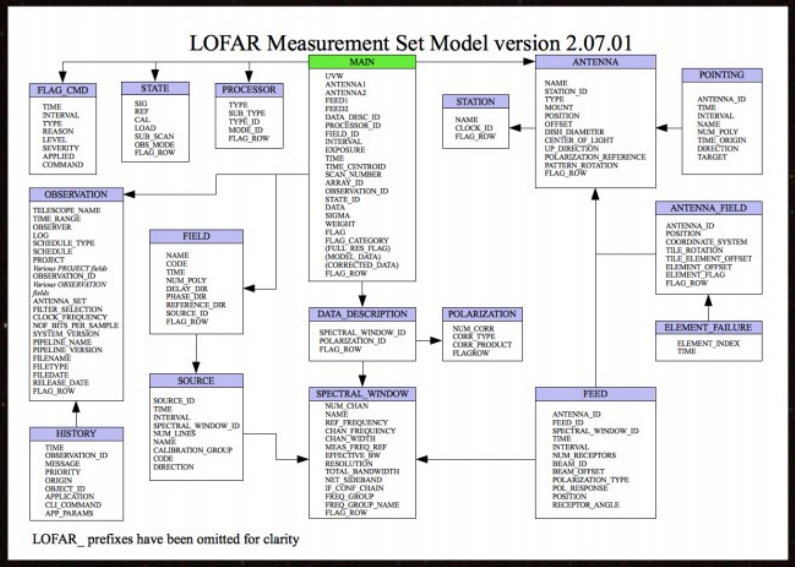
- Beaucoup de données non réduites dans les archives (« visibility » data).
- Réductions possibles non uniques

# données Visibility? (modèle des « measurement sets » cas LOFAR )

## Data Types: Visibilities

ASTRON

- Stored as CASA Measurement Sets
  - 10 MB - 100 GB in size, per 192 kHz frequency band
  - Usually 400-488 per observation
  - Relational structure
  - Each cell can have multiple dimensions (pol x freq)



# Modèle de données IVOA

# Visibility :

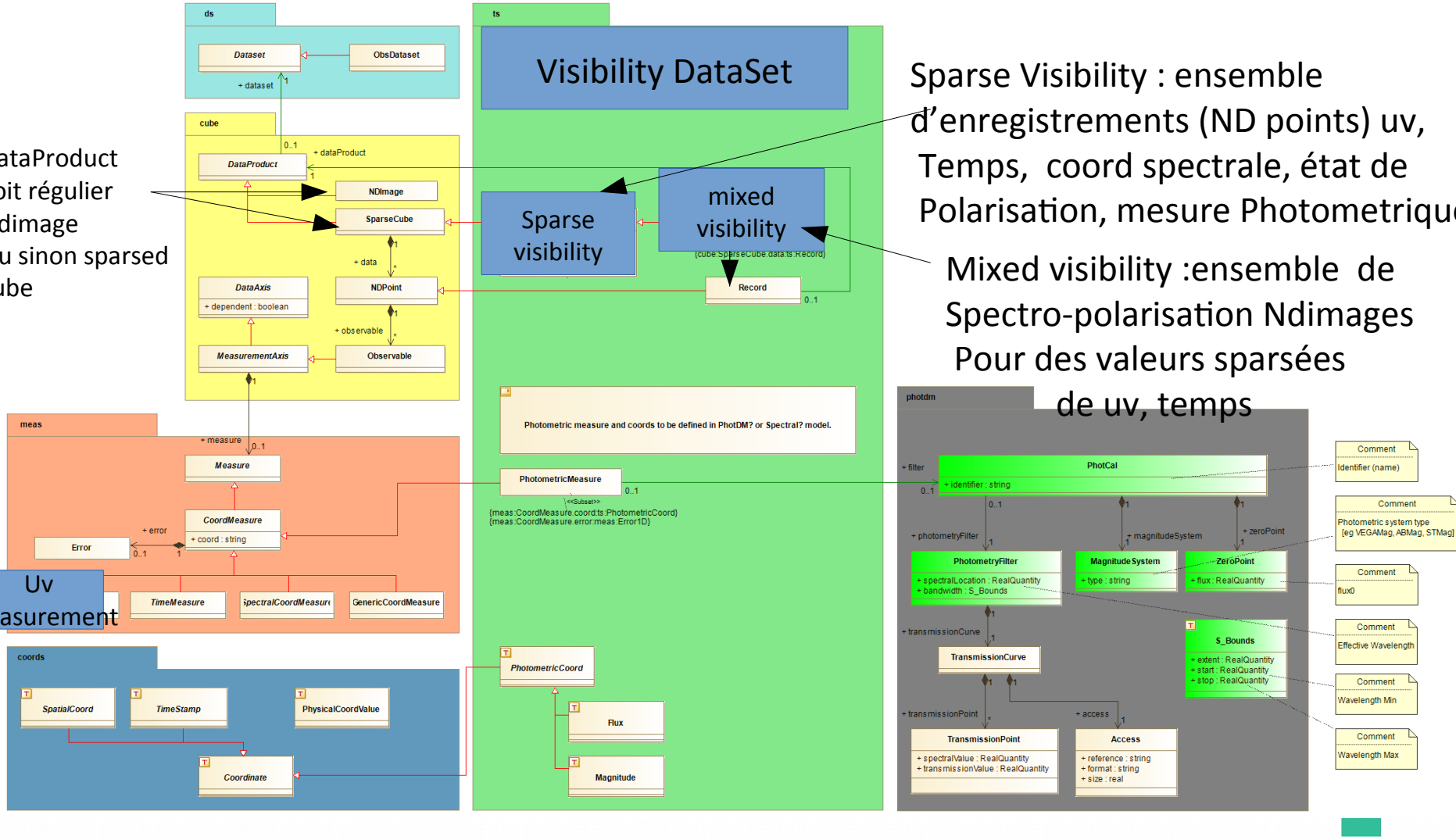
## derivé du modèle cube datamodel

DataProduct
Soit régulier
Ndimage
Ou sinon sparsed
cube

Sparse Visibility : ensemble d'enregistrements (ND points) uv, Temps, coord spectrale, état de Polarisation, mesure Photometrique

Mixed visibility :ensemble de Spectro-polarisation Ndimages Pour des valeurs sparsées de uv, temps

Uv measurement



# Limitations. Comment les lever ? construire des données « scientifiques »

- Retourner aux données d'origine via les métadonnées provenance
- Stocker et rendre accessibles les cubes réduits par les scientifiques utilisateurs des archives
- Proposer un service qui calcule lui même un cube en fonction des caractérisations souhaitées (ALMA).
- Une fois données découvertes, lancer les codes (CASA par exemple) à distance
- Amener le code vers les données (Jupiter Note book)

# Dans ESCAPE : comment commencer

- Faire connaître l'existant
- Collecter et décrire de façon détaillée les cas d'utilisation scientifique
- Expérimenter et prototyper de nouvelles standardisations et nouveaux services
- Retour d'expérience.