



Actualité du groupe DAL : accès aux Cubes de données Évolutions ADQL et TAP

F.Bonnarel



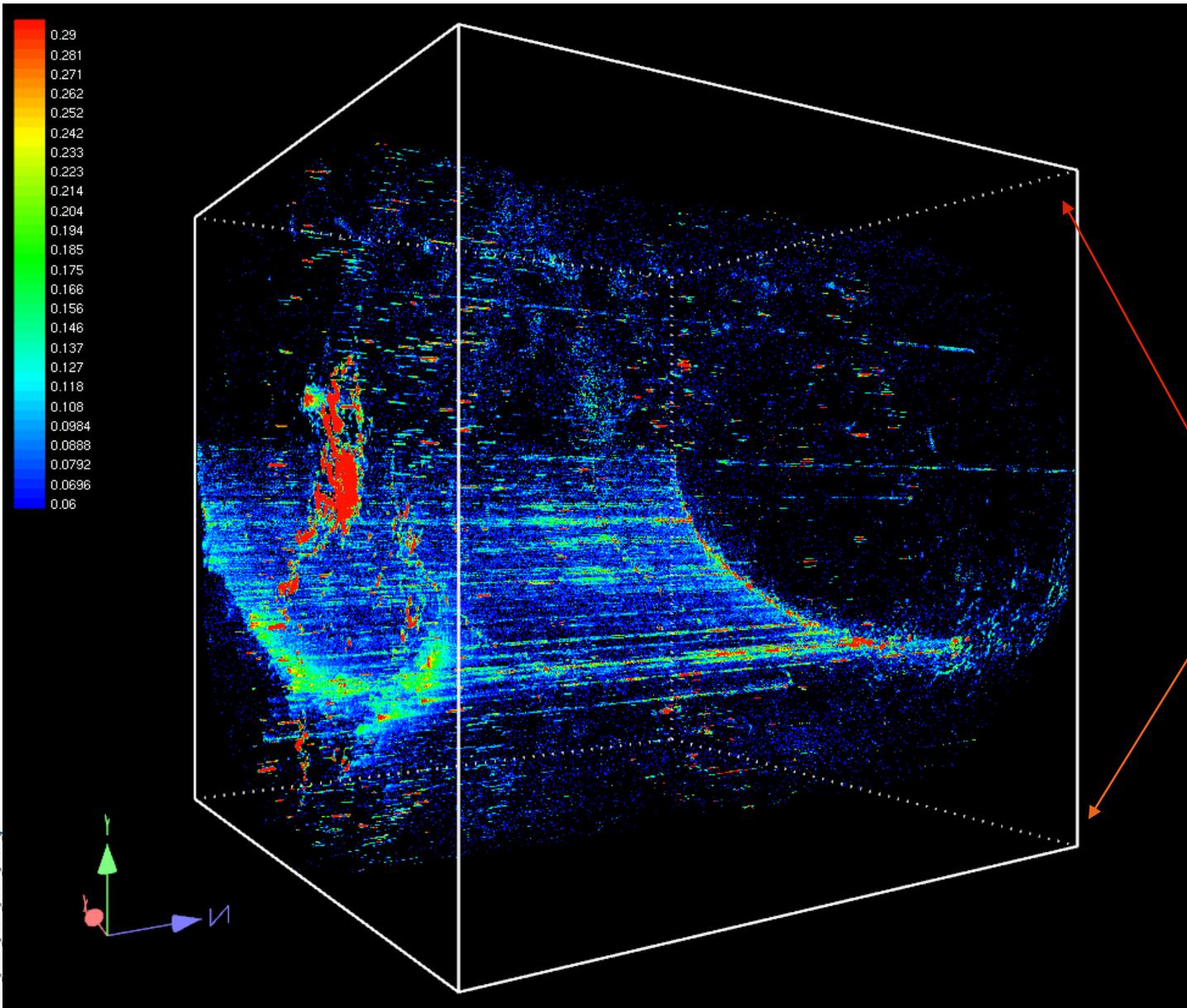
A propos des cubes de données en astronomie.

- Définition : un ensemble de mesures (flux relatif ou calibré, vitesse radiale, couleur, etc...) régulièrement échantillonnées dans l'espace physique des paramètres:
 - Spatiaux (2D)
 - Spectraux
 - Temporels
 - Polarimétriques
- Exemples : image 2D classique, cube radio avec dimension spectrale, IFU, séquence temporelle
- Event lists ? (vu comme des cubes « à trous »)
- Nombreuses expériences en radio (ALMA, LOFAR, SKA...) en optique (MUSE) et en X
- Très gros volumes : teras-octets/jour dans certains cas. Cubes individuels jusqu'à cent gigas-octets

Cubes de données et observatoire virtuel

- Quelques exemples de demandes à satisfaire en terme de cubes de données:
 - A un répertoire : Quels sont les services répertoriés renvoyant des cubes spectraux dans le domaine radio ?
 - A un service de découverte: Quelle est la liste des cubes disponibles pour la raie 21 cm de l'hydrogène dans l'amas de Virgo? Avec description de leur domaine spatial et spectral exact et de leur résolution spatiale et spectrale.
 - A un service d'extraction de données : demande d'un sous-cube centré sur la galaxie M87 et avec une largeur spectrale de 1Ghz

ETAPE 1 : découverte



Cube de données radio

- Axes rouge et vert :
- Dimensions spatiales
 - axe bleu
 - Dimension longueur d'onde

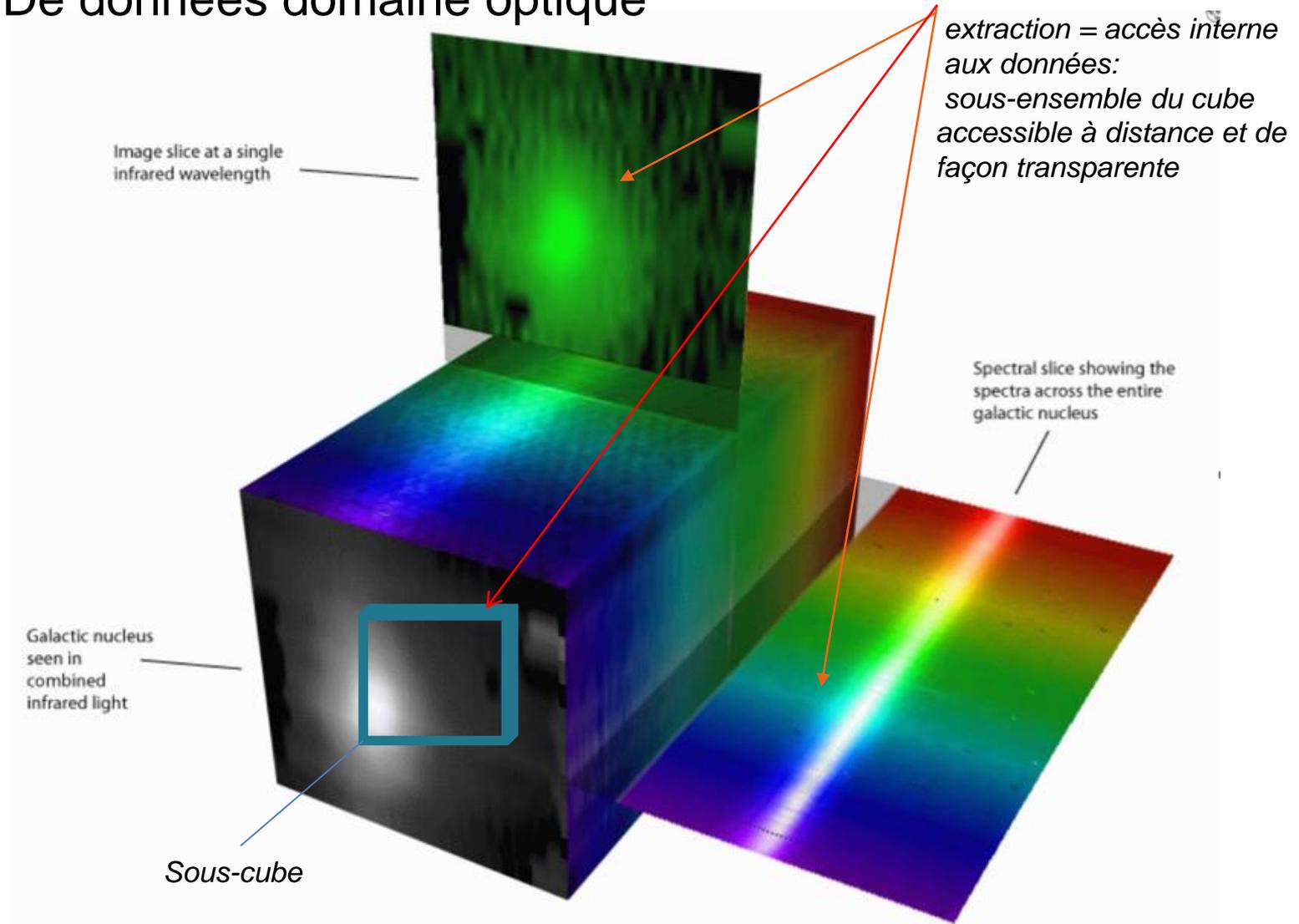
ETAPE 2:

description :

- nature des axes
- Extension
- Valeur typique
- Date
- Producteur de données
- Fournisseur d'accès

« Integral Field Unit » cube De données domaine optique

ETAPE 3



Détail du scénario IVOA simple: (première version des protocoles)

- I) Trouver les services de cubes dans un répertoire IVOA: les services ObsTap ou SIAV2
→ *Observation Table Access Protocol et Simple image Access protocol version 2.*
- II) Requête de découverte à un service ObsTap
« *select * from Obscore where dataproduct_type = cube* »
en Syntaxe ADQL
- II bis) Requête à un service SIAV2 (paramètres de requête http)
« *http://.....?request=query&pos=CIRCLE 12 34 0.5&band=....* »

Découverte de services dans un répertoire

Cette page est en anglais Voulez-vous la traduire ? Traduire Non Toujours traduire les pages en anglais Options

VAO Search All Virtual Observatory Collections: Radius: Arcmin
[User Guide](#) | [Discovery Tool v1.5 \(6846\)](#)... Examples: [M101](#), [14.03.12.6](#) +[54.20.56.7](#), [more...](#)

Start Page NGC 6946 (RA: 20:34:52.322, Dec: +60:09:14.08), radius: 0.01667°

Displaying 46 of 358 Total Rows

Filters

Clear Filters Edit Facets... Help...

Filter All Record Fields

Type

- Catalog (0 of 307)
- Image (46 of 46)
- Spectra (0 of 5)

Waveband

- EUV (2 of 5)
- Gamma-ray (3 of 12)
- Infrared (20 of 104)
- Millimeter (2 of 5)
- Optical (14 of 154)
- Radio (2 of 71)
- UV (11 of 39)
- X-ray (9 of 66)

Publisher

- Canadian Astronomy Data Centre (2 of 2)
- CDS (0 of 205)
- Chandra X-ray Observatory (2 of 4)
- ESO (1 of 1)
- European Space Agency (1 of 1)
- German Astrophysical Virtual Observatory (1 of 2)
- MAST (5 of 6)
- NASA/GSFC HEASARC (15 of 82)
- NASA/HEASARC (1 of 1)
- NASA/IPAC Infrared Science Archive (10 of 12)
- National Optical Astrono (0 of 2)
- Observatory of Strasbourg, SSC Team (1 of 1)

Actions	Short Name	Type	Title	Waveband	Records For
	Spitzer Level 1		Spitzer Level 1 / Basic Calibrated Data	Infrared	3127
	CADC		CADC Image Search	Millimeter, Infra...	869
	CADC/SIAv1		CADC Image Search (SIA)		869
	Spitzer Level 2		Spitzer Level 2 / post Basic Calibrated Data	Infrared	546
	WISE All-Sky L1B		WISE All-Sky 4-band Single-Exposure Images	Infrared	177
	ST-ECF/HLA/SIA		ST-ECF Hubble Legacy Archive Images		89
	hdap_siap [1]		HDAP -- Heidelberg Digitized Astronomical Plates	Optical	63
	SkyView		SkyView Virtual Observatory	Radio, Infrared, ...	58
	NED(images)		The NASA/IPAC Extragalactic Database Image Data Atlas	Radio, Millimeter...	47
	HLA [1]		Hubble Legacy Archive	Optical, Infrare...	40
	ST-ECF/HST/SIA		ST-ECF Hubble Space Telescope Images		35
	HST Previews		Hubble Space Telescope Preview Images	Optical	29
	MAST-Scrapbook		The MAST Image Scrapbook	Infrared, Optica...	28
	IRTS		The Infrared Telescope in Space Data Atlas	Infrared	26
	ROSAT SIA		SIA Service for ROSAT Archive	X-ray	22
	2MASS QL		2MASS All-Sky Quicklook Image Service	Infrared	18
	2MASS ASKY AT		2MASS All-Sky Atlas Image Service	Infrared	18
	DSS ESO		Digitized Sky Survey		16

AstroView

20:35:44.469 +60:25:47.943
20:34:52.322 +60:09:14.076 RA DEC
hhmmss/deg

Détail du scénario IVOA simple: (première version des protocoles, mai 2014)

- III) la réponse:

 - en VOTABLE, compatible avec modèle Obscore = description des cubes de façon standardisée

- IV) un service DataLink associé permet de lier à un cube :

 - Des liens fixes,
 - des services de métadonnées,
 - Des services d'extraction de données (AccessData)
 - *Des services non standardisés (ex : fit de modèle)*

- IV bis) directement un service AccessData (extraction de sous-données)

Découverte de donnée via Un service ObsTap

The screenshot shows the TAPHandle web interface. At the top, the URL is `cadc` and the Node Selector is `cadc>ivoa>ObsCore>wlwclxqkwx29q5jq`. On the left, a tree view shows the hierarchy: `cadc` (selected), `TAP_SCHEMA`, `ivoa`, `ivoa.ObsCore`, `cfht`, `caom2`, `caom`, and `Goodies`. The main area displays a table with 13 columns: `access_estsize`, `target_name`, `s_ra`, `s_dec`, `s_fov`, `s_region`, `s_resolution`, `t_min`, `t_max`, `t_exptime`, `t_resolution`, and `em_min`. The table contains 13 rows of data. Below the table, there is a query editor with tabs for `Select What`, `Where`, `Plain Text Query`, and `Job Control`. The `Where` tab is active, showing the query: `SELECT TOP 100 * FROM ivoa.ObsCore WHERE ivoa.ObsCore.target_name = 'M31'`. A `Result Limit` input field is set to `100`. At the bottom left, the URL `saada.u-strasbg.fr/taphandle/#tapselect` is visible.

	access_estsize	target_name	s_ra	s_dec	s_fov	s_region	s_resolution	t_min	t_max	t_exptime	t_resolution	em_min
nk	null	M31	10.684280	41.268274	0.75172824	☀	NaN	54349.349	54349.380	13.858047	NaN	0.00086506567
nk	null	M31	10.684280	41.268274	0.75172824	☀	NaN	54349.349	54349.380	13.878724	NaN	0.00086497792
nk	null	M31	10.684942	41.268232	1.0513781	☀	NaN	54349.348	54349.380	2720.0000	NaN	0.00086497865
nk	null	M31	10.684652	41.268042	0.75315870	☀	NaN	54349.384	54349.415	13.912904	NaN	0.00086505104
nk	null	M31	10.684943	41.268233	1.0513813	☀	NaN	54349.384	54349.415	2737.0000	NaN	0.00086497865
nk	null	M31	10.684615	41.268007	0.75315870	☀	NaN	54349.384	54349.415	13.931988	NaN	0.00086497792
nk	null	M31	10.685881	41.269388	0.75173089	☀	NaN	54382.325	54382.357	11.208046	NaN	0.00086504622
nk	null	M31	10.685057	41.268589	0.75315870	☀	NaN	54382.325	54382.357	13.994818	NaN	0.00086497792
nk	null	M31	10.685161	41.269028	1.0527622	☀	NaN	54382.325	54382.357	2755.0000	NaN	0.00086497870
nk	null	M31	18.854066	35.020254	1.4872042	☀	0.82000000	52966.227	55567.244	4170.0000	NaN	4.0780000e-7

```

- <INFO name="QUERY" >
  <![CDATA[SELECT TOP 100 * FROM ObsCore AS ObsCore]]>
</INFO>
- <TABLE name="Results">
  <FIELD ID="oidsaada_100" name="oidsaada" arraysize="*" ref="oidsaada_100" datatype="char"/>
  <FIELD ID="sky_pixel_csa_100" name="sky_pixel_csa" ref="sky_pixel_csa_100" datatype="long"/>
  <FIELD ID="dataproduct_type_100" name="dataproduct_type" arraysize="*" ref="dataproduct_type_100" datatype="char"/>
  <FIELD ID="calib_level_100" name="calib_level" ref="calib_level_100" datatype="int"/>
  <FIELD ID="targe_name_100" name="targe_name" arraysize="*" ref="targe_name_100" datatype="char"/>
  <FIELD ID="target_class_100" name="target_class" arraysize="*" ref="target_class_100" datatype="char"/>
  <FIELD ID="obs_id_100" name="obs_id" arraysize="*" ref="obs_id_100" datatype="char"/>
  <FIELD ID="obs_collection_100" name="obs_collection" arraysize="*" ref="obs_collection_100" datatype="char"/>
  <FIELD ID="obs_creator_name_100" name="obs_creator_name" arraysize="*" ref="obs_creator_name_100" datatype="char"/>
  <FIELD ID="obs_publisher_did_100" name="obs_publisher_did" arraysize="*" ref="obs_publisher_did_100" datatype="char"/>
  <FIELD ID="bib_reference_100" name="bib_reference" arraysize="*" ref="bib_reference_100" datatype="char"/>
  <FIELD ID="data_rights_100" name="data_rights" arraysize="*" ref="data_rights_100" datatype="char"/>
  <FIELD ID="access_url_100" name="access_url" arraysize="*" ref="access_url_100" datatype="char"/>
  <FIELD ID="access_format_100" name="access_format" arraysize="*" ref="access_format_100" datatype="char"/>
  <FIELD ID="access_estsize_100" name="access_estsize" arraysize="*" ref="access_estsize_100" datatype="char"/>
  <FIELD ID="datalink_100" name="datalink" arraysize="*" ref="datalink_100" datatype="char"/>
  <FIELD ID="s_ra_100" name="s_ra" ref="s_ra_100" datatype="double"/>
  <FIELD ID="s_dec_100" name="s_dec" ref="s_dec_100" datatype="double"/>
  <FIELD ID="s_fov_100" name="s_fov" ref="s_fov_100" datatype="double"/>
  <FIELD ID="s_region_100" name="s_region" arraysize="*" ref="s_region_100" datatype="char"/>
  <FIELD ID="s_resolution_100" name="s_resolution" ref="s_resolution_100" datatype="double"/>
  <FIELD ID="s_ucd_100" name="s_ucd" arraysize="*" ref="s_ucd_100" datatype="char"/>
  <FIELD ID="t_min_100" name="t_min" ref="t_min_100" datatype="double"/>
  <FIELD ID="t_max_100" name="t_max" ref="t_max_100" datatype="double"/>
  <FIELD ID="t_exptime_100" name="t_exptime" ref="t_exptime_100" datatype="double"/>
  <FIELD ID="t_resolution_100" name="t_resolution" ref="t_resolution_100" datatype="double"/>
  <FIELD ID="t_cal_status_100" name="t_cal_status" arraysize="*" ref="t_cal_status_100" datatype="char"/>
  <FIELD ID="t_stat_err_100" name="t_stat_err" ref="t_stat_err_100" datatype="double"/>
  <FIELD ID="em_unit_100" name="em_unit" ref="em_unit_100" datatype="double"/>
  <FIELD ID="em_min_100" name="em_min" ref="em_min_100" datatype="double"/>
  <FIELD ID="em_max_100" name="em_max" ref="em_max_100" datatype="double"/>
  <FIELD ID="em_res_power_100" name="em_res_power" ref="em_res_power_100" datatype="double"/>
  <FIELD ID="em_resolpower_min_100" name="em_resolpower_min" ref="em_resolpower_min_100" datatype="double"/>
  <FIELD ID="em_resolpower_max_100" name="em_resolpower_max" ref="em_resolpower_max_100" datatype="double"/>
  <FIELD ID="em_resolution_100" name="em_resolution" ref="em_resolution_100" datatype="double"/>
  <FIELD ID="em_stat_err_100" name="em_stat_err" ref="em_stat_err_100" datatype="double"/>
  <FIELD ID="o_ucd_100" name="o_ucd" arraysize="*" ref="o_ucd_100" datatype="char"/>
  <FIELD ID="o_calib_status_100" name="o_calib_status" ref="o_calib_status_100" datatype="int"/>
  <FIELD ID="o_stat_err_100" name="o_stat_err" ref="o_stat_err_100" datatype="double"/>

```

Réponse en VOTABLE : définition des champs

```

- </TD>
- <TD>
  <![CDATA[3XMM]]>
</TD>
- <TD>
  <![CDATA[SSC XMM-Newton]]>
</TD>
- <TD>
  <![CDATA[ivo://xcatdb/3xmm/866379999790235705]]>
</TD>
<TD/>
- <TD>
  <![CDATA[public]]>
</TD>
- <TD>
  <![CDATA[http://saada.unistra.fr/3xmm/download?oid=866379999790235705]]>
</TD>
- <TD>
  <![CDATA[application/fits]]>
</TD>
- <TD>
  <![CDATA[100]]>
</TD>
- <TD>
  <![CDATA[http://saada.unistra.fr/3xmm/smartdatalink?oid=866379999790235705]]>
</TD>
<TD>214.2525</TD>
<TD>44.935</TD>
<TD>0.5</TD>
- <TD>
  <![CDATA[POLYGON ICRS 213.8925 44.575 213.8925 45.295 214.6125 45.295 214.6125 44.575]]>
</TD>
<TD/>
- <TD>
  <![CDATA[pos.eq]]>
</TD>
<TD>52616.7887384259</TD>
<TD>52617.0707986111</TD>
<TD>1015.41666666744</TD>
<TD/>
- <TD>
  <![CDATA[COARSE]]>
</TD>
<TD/>
<TD/>
<TD>6.2E-9</TD>
<TD>1.03E-10</TD>
<TD/>
<TD/>
<TD/>
<TD/>
- <TD>
  <![CDATA[phot.count]]>
  _

```

Réponse en VOTABLE : partie données

DataLink: service de liens

(extrait du document de référence: contenu de la table)

The list of links that is returned by the {links} resource can be represented as a table with the following columns:

name	description	required	UCD
ID	Input identifier	yes	meta.id;meta.main
access_url	link to data or service	one only	meta.ref.url
error_message	error if an accessURL cannot be created		meta.code.error
service_def	reference to the description of a service at access_url	no	meta.ref
description	human-readable text describing this link	no	meta.note
semantics	limited vocabulary describing this link	no	meta.code
content_type	mime-type of file the link returns	no	meta.code.mime
content_length	size of download the link returns	no	phys.size;meta.file

Cube moyenné (lien), extraction de données, service de métadonnées,
(*services « libres »*)

DataLink: service descriptor

```
<RESOURCE type="meta" utype="adhoc:service">
  <PARAM name="resourceIdentifier" datatype="char"
arraysize="*" value="ivo://example.com/mySIA" />
  <PARAM name="standardID" datatype="char"
arraysize="*" value="ivo://ivoa.net/std/SIA#1.0" />
  <PARAM name="accessURL" datatype="char"
arraysize="*" value="http://example.com/sia/query" />
  <GROUP name="inputParams">
    <PARAM name="POS" datatype="char"
arraysize="*" value="" />
    <PARAM name="SIZE" datatype="char"
arraysize="*" value="0.5" />
    <PARAM name="VERB" datatype="int" value="0" />
    <PARAM name="FORMAT" datatype="char"
arraysize="*" value="ALL">
      <VALUES>
        <OPTION value="ALL" />
        <OPTION value="image/fits" />
        <OPTION value="METADATA" />
      </VALUES>
    </PARAM>
  </GROUP>
</RESOURCE>
```

AccesData :

extraction de données

- « Cutout » (découpe) dirigé par des paramètres identiques à ceux de la requête de découverte
 - POS=CIRCLE 12 34 0.5
 - POS=RANGE 12/14 34/36
 - BAND=500/550
 - TIME=2012-01-01/2012-12-31
 - POL=Q,POL=.....

• États des protocoles DAL

(Data access layer)

- **ObsTap 1.0** : recommandation 2011
- **DataLink 1.0** : la revue TCG est terminée. Modifs publiées prochainement -> recomm immminente.
- **SIKV2.0** → revue TCG demande de modifs + implementation client
- **AccessData 1.0** -> Working draft, compromis à rédiger : description des paramètres d'entrée par 3 attributs.
- **Modèles de données** :
 - **Obscore** -> 1.1
 - **ImageDm 1.0** -> 1.1 (et spectre?)

Et ensuite ?

(seconde version des protocoles)

- SIAV2.1 : Requête de découverte à un service SIAV2 plus sophistiqué:
« *http://.....?request=metadata&pos=...&band=.....* »
 - Cubes archivés
 - Cubes virtuels : génération dynamique via mosaïque, Extraction, rééchantillonnage, etc...
- Accès direct à la méthode Metadata comme à « query »
fournit des métadonnées additionnelles compatibles avec ImageDM dans la réponse en VOTABLE.
- Accessdata (extraction de données) plus riche:
 - « *cutout* » (découpe),
 - *rééchantillonnage (changement de taille du pixel)*,
 - *recalage (changement de pixelisation)*,
 - *Transformations, calculs de moments (indice moyen , dispersion, etc....)*

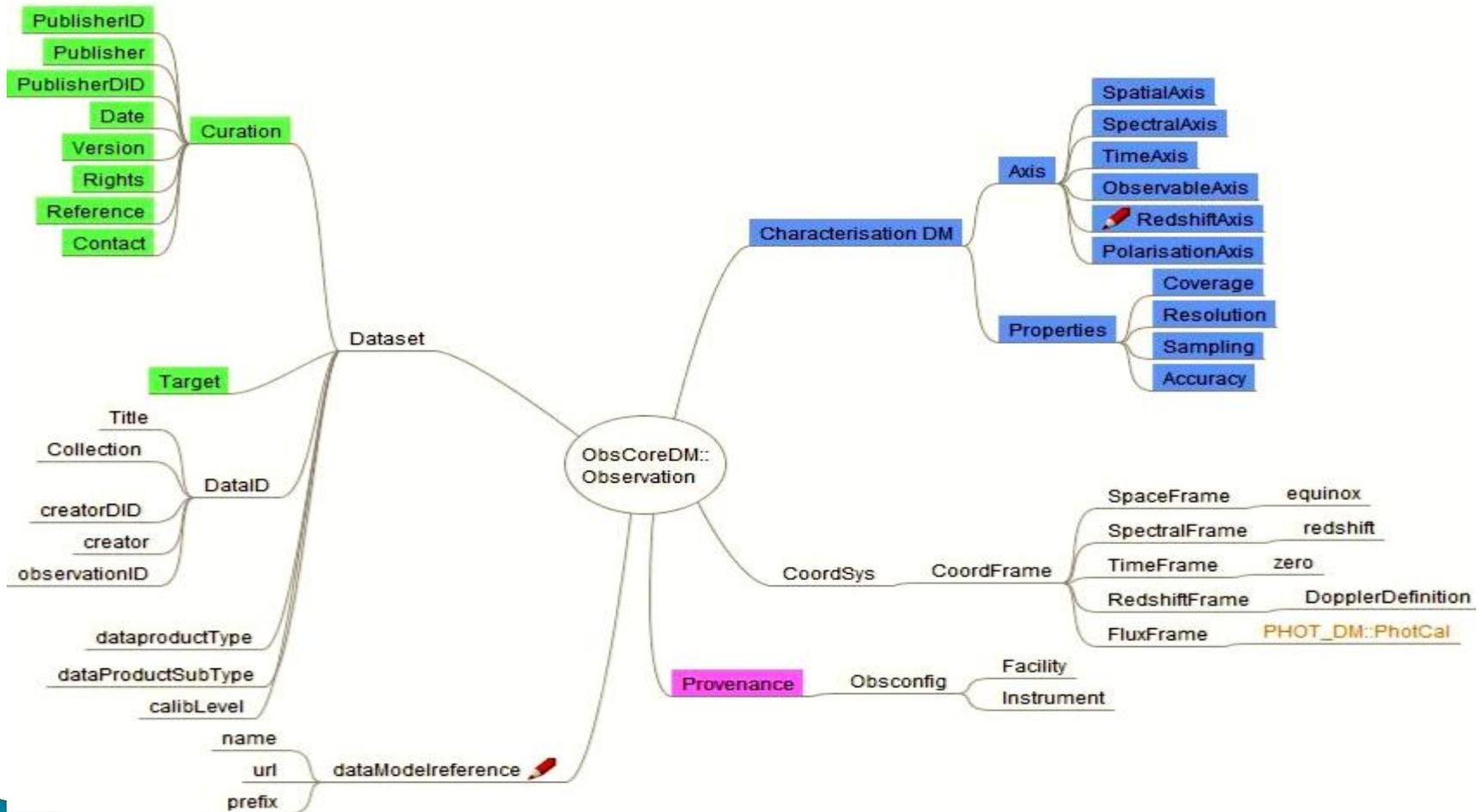
Problèmes en débat

- Syntaxe pour forcer les « cutout » (découpe) :
 - En pixel en plus de coordonnées réelles ?
 - Syntaxe « à la cfitsio »
- ObsCore a besoin d'une mise à jour pour description des array
 - Aller jusqu'au WCS ?
 - Quelle relation avec ImageDM ?
 - ObsCore « sous modèle » ou « vue » du modèle ImageDM
- Relation à HiPS (vue comme mode de découverte et mode d'AccessData particulier)
- Syntaxe Json pour la query étendue et relation à SimDal

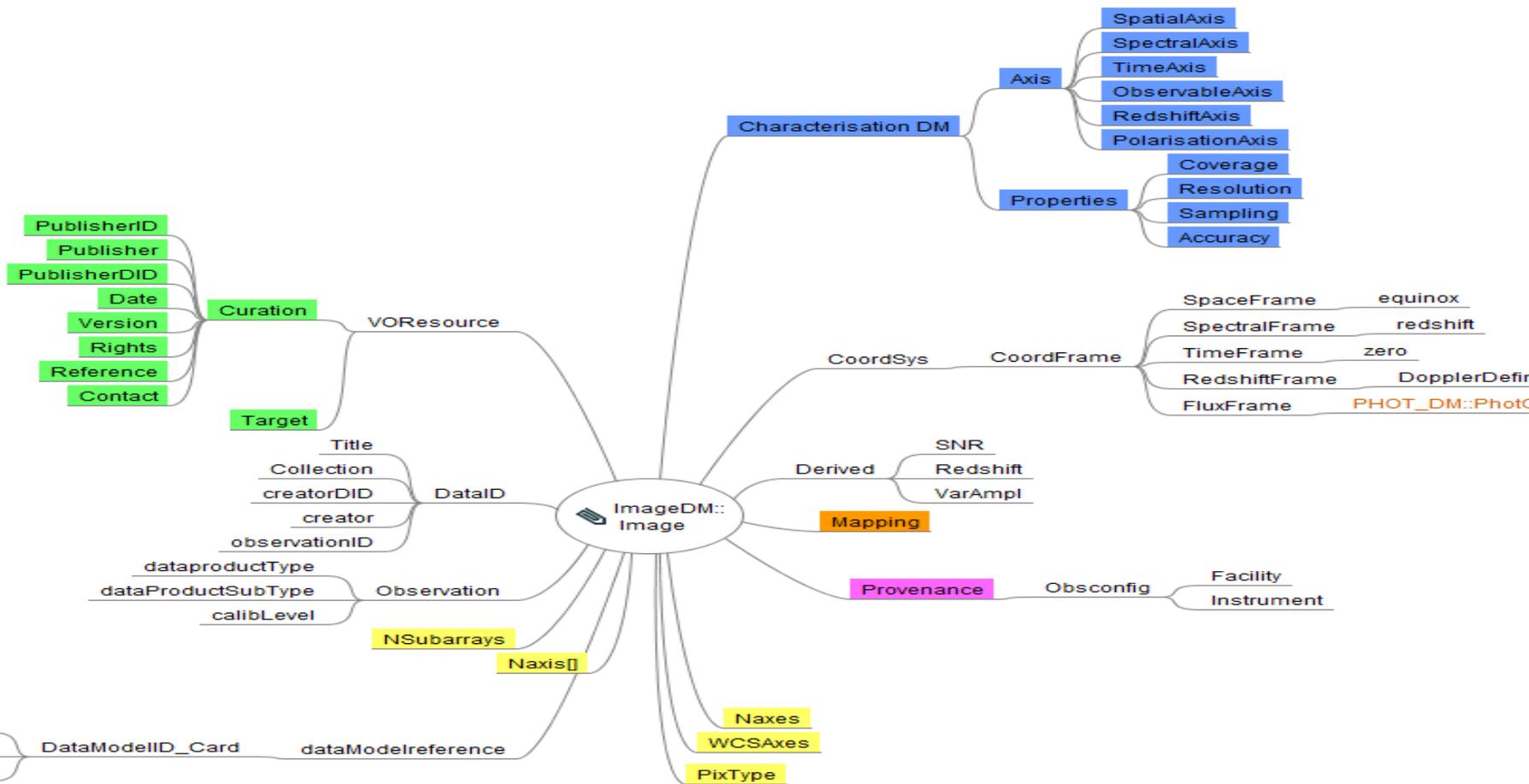
Evolution Obscore

- Ajout de certains champs nécessaires à la description de données n-d
- Nbre d'axes, nbre de pixels par axes,
- Jusqu'où aller :
 - Question des subarray
 - Wcs simple ?
 - Question non encore résolue
 - Découverte ou préparation à AccessData

Carte heuristique modèle de données ObsCore

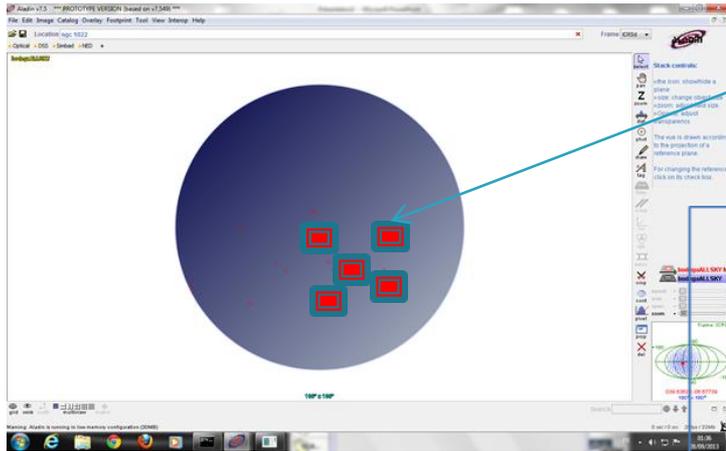


Carte heuristique modèle de données ImageDM

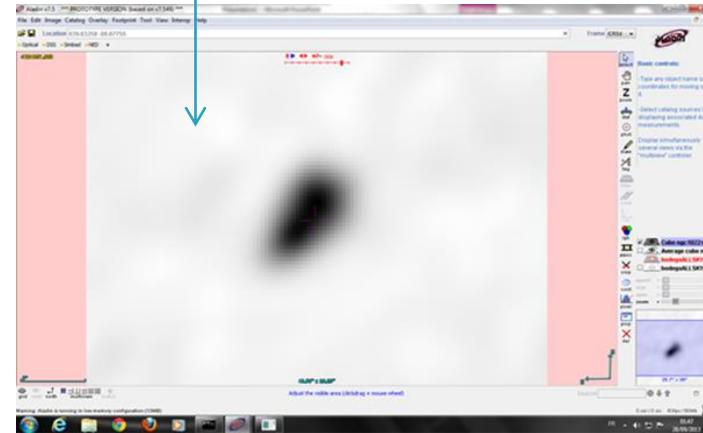
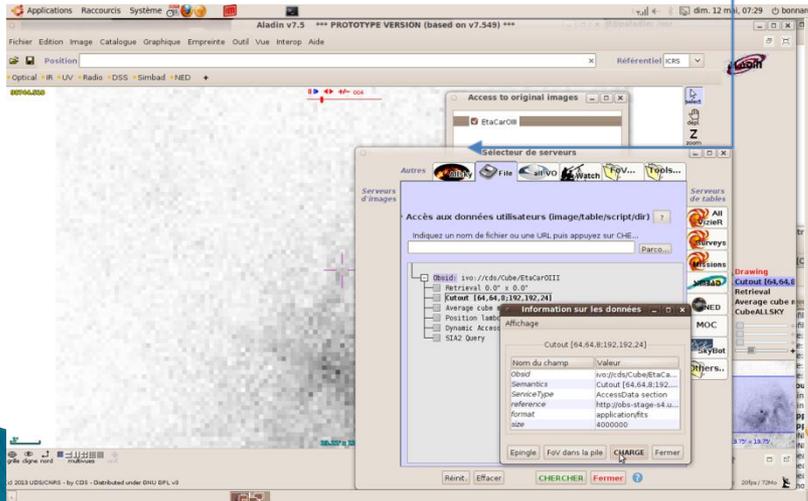


« DataLink » pour « AccessData » pour les cubes derrière HiPS dans Aladin

(prototype)



- 1) Centrage sur NGC 1022
- 2) Liens disponibles dans une fenêtre. Zoom sur le cube de données moyenné de NGC 1022 .
- 3) sous- cube en mode cinéma



ADQL

(suivi par Marco Molinaro)

- Suite de la discussion sur les items ADQL dans les TAPNotes .
 - Draft Errata en décembre 2014
 - Points d'évolution pour ADQL-2.1 et pour ADQL-3.0 identifiés
 - Rejet de certains items
- Accord atteint pour un ADQL-2.1 WD au printemps 2015
- Points Critiques
 - vérifier et tester nouvelles fonctionnalités dans toutes les plus grandes DBMS-es
 - Idée: ce qui est optionnel dans 2.1 serait obligatoire dans la version 3.0
- Inter standard
 - Définition pour les types dans ADQL, TAP pointera sur lui
 - Pagination: OFFSET dans ADQL

TAP

(suivi par Marco Molinaro)

- Points identifiés dans les TAP notes
 - errata (Draft en décembre 2014)
 - changements pour TAP-1.1 (et pour TAP-2.0:)
 - items rejetés
- Accord pas totalement atteint
 - Des détails d' implementation encore à discuter
 - TAP-1.1 WD cependant planifié pour Juin 2015
 - idées: optionelles 1.1 deviendront obligatoires en version 2.0
- Points Critiques
 - atteindre accord pour la gestion d'échelle des VOSI#tables
 - Discuter la pagination (voir ADQL OFFSET,)
- Inter standard
 - VOSI-1.1 de GWS peut être nécessaire pour la gestion d'échelle des VOSI#tables