

Conception d'une GUI pour mapping de fichiers FITS avec le DM ObsCore (Annexe « Obscore »)

Le contexte

Les publications scientifiques en astronomie qui font office de catalogues VizieR (Annexes (3)) contiennent des données tabulaires mais aussi de “données associées”. Ces données associées sont constituées de série temporelles, de SED, d'images ou de spectres. Les données associées dont le contenu est exploitable (i.e. : contenant des méta-données) sont préservées dans le Système d'information VizieR (il s'agit ici des ressources au formats VOTable et surtout FITS). L'accès à ces ressources se fait via FTP et par des liens dans les pages WEB VizieR (Annexe (3)). Aujourd'hui, le système de gestion de bases de données Saada (Annexes (1)+(5)) est utilisé par VizieR pour l'indexation des données associées dont le format est FITS ou VOTable selon des méta-données basées sur le Data-model ObsCore de l'IVOA (Annexes (2)).

Des outils existent au CDS pour générer et ingérer les nouveaux documents : il s'agit d'API, ou d'applications en mode texte exécutées dans une console. L'usage d'une GUI destinée au CDS et en particulier aux documentalistes du CDS permettrait une meilleure prise en main de ce mapping.

Chiffres

- ~150 catalogues contiennent des images FITS
- ~150 catalogues contiennent des spectres/série temporelle/SED au format FITS
- Le nombre de Fichier FITS par catalogue ainsi que leur contenu est hétérogène (de 1 à plusieurs milliers de fichiers pour un catalogue)
- Les catalogues les plus importants en nombre de ressource FITS : Lamost (~2.5millions), CoRoT (~150,000)

Contenu et spécificité de l'existant

- un catalogue peut contenir à la fois des images et des spectres/séries-temporelles/SED.
- Le mapping est propre à chaque fichier FITS. Il est souvent similaire à l'intérieur d'un catalogue (mais pas toujours). Il est possible de regrouper une collection de fichiers FITS avec un même mapping au sein d'un catalogue.
- Un mapping peut avoir été fourni par le créateur des fichiers.

Choix techniques

- L'application sera développée en Java/Swing
- l'application sera hébergée sur le serveur Linux contenant le FileSystem où sont déposées les données associées
- Utilisation de l'API Saada pour générer/tester le mapping ObsCore-FITS
- Utilisation de l'API VizieR « vizier.catana » comme surcouche de l'API Saada, notamment pour la sérialisation du mapping
- Utilisation d'un SGBD propre à l'API Saada. On pourra utiliser la base de données PostgreSQL Saada « saadavizierlocal »

TODO

La GUI se focalisera sur les données au format FITS. L'outil permettra :

- de rechercher les ressources présentes dans un répertoire
- de proposer un mapping pour un fichier/une collection de fichiers :
une collection peut être déterminée par :
 - un répertoire
 - un sous ensemble : ex : img/a*.fits
 - par une liste de documents FITS données dans un fichier
- de charger/modifier un mapping existant
- de tester un mapping
- de sauvegarder un mapping
- de proposer des outils de visualisation des entêtes FITS pour les documents mappés
- de proposer un « quick-view » du mapping

Outils pour l'aide au mapping

Note : On pourra s'inspirer de la page web de soumission des données Vizier (Annexes (6))

- Visualisation des entêtes FITS pour les documents mappés avec possibilités d'ajouter les mots-clés FITS aux mappings
- proposer un test pour valider le mapping avec remontée des informations, de valeurs ou d'erreurs
- proposer la liste des bandes spectrales (requête SQL à la base de données « saadavizierlocal »)
- proposer les couples telescope/instrument/bande-spectrales (em_band/em_min/em_max) selon la base de connaissance « saadavizierlocal »
- proposer une résolution des noms d'objets (target_name) en positions (s_ra, s_dec) (service CDS sesame, voir API)
- proposer un « quick-view » du mapping
- proposer la mise à jour du Système d'information (voir API « vizier.catana »)

Annexes

- (1) Saada : <http://saada.unistra.fr/saada/>
- (2) IVO ObsCore : <http://www.ivoa.net/documents/ObsCore/>
- (3) Vizier : <http://vizier.u-strasbg.fr/viz-bin/Vizier>
- (4) Vizier données associées : <http://cdsarc.u-strasbg.fr/assocdata/>
- (5) Saada accès WEB: <http://cdsarc.u-strasbg.fr/saadavizier/>
- (6) Page de soumission des données Vizier : <http://cdsarc.u-strasbg.fr/vizier.submit/>
-

Annexe : ObsCore Data-model

dataprodukt_type	Logical data product type (image etc.)
calib_level	Calibration level {0, 1, 2, 3}
obs_collection	Name of the data collection
obs_id	Observation ID
obs_publisher_did	Dataset identifier given by the publisher
access_url	URL used to access (download) dataset
access_format	File content format (see in App. BB.5.2)
access_estsize	Estimated size of dataset in kilo bytes
target_name	Astronomical object observed, if any
s_ra	(deg)Central right ascension, ICRS
s_dec	(deg)Central declination, ICRS
s_fov	(deg)Diameter (bounds) of the covered region
s_region	Region covered as specified in STC or ADQL
s_resolution	(arcsec)Spatial resolution of data as FWHM
t_min	Start time in MJD
t_max	Stop time in MJD
t_exptime	(s)Total exposure time
t_resolution	(s)Temporal resolution FWHM
em_min	Start in spectral coordinates
em_max	Stop in spectral coordinates
em_res_power	Spectral resolving power
o_ucd	UCD of observable (e.g. phot.flux.density)
pol_states	List of polarization states or NULL if not applicable
facility_name	Name of the facility used for this observation
instrument_name	Name of the instrument used for this observation

Architecture

