Image & Image Set Metadata Fields

Below are two sets of metadata which will be stored in the new archive in regard to Images and the Image Sets to which they belong. Included where relevant are the equivalent fields in the Virtual Observatory ObsCore Data Model, a potential source from which to obtain the data, and a comment. The images table was written specifically with the idea of spatial images in mind. Other data products (spectra, time series, etc) would have different amounts of granularity in the header information available.

Images Database Table:

Column Name	Units	VO Equivalent	Source	Comments
sourcename		target_name	OBJECT Keyword	
ra	deg	s_ra	Center of Ra Axis	
dec	deg	s_dec	Center of Dec Axis	
image_field_of_view	deg	s_fov	Quadratic Mean of the ra & dec extents (NAXISn*CDELTn)	
spatial_region		s_region	Derived from spatial Data	STC-S String Defined in the TAP (Dowler, et al 2010)
ra_element_count		s_xel1	Relevant NAXISn Keyword	
ra_pixel_size			abs(Relevant CDELTn)	
dec_element_count		s_xel2	Relevant NAXISn Keyword	
dec_pixel_size			abs(Relevant CDELTn)	
spatial_resolution	arcsec	s_resolution	sqrt(BMAJ*BMIN)	Geometric mean of the synthesized beam axes.
beam_axis_ratio			BMAJ/BMIN	Ratio of the synthesized beam axes.
starttime	MJD	t_min		These values should be obtained by some larger-scale process. VlassMgr, in the case of quicklook images.
endtime	MJD	t_max		
exposure_time	s	t_exptime		
min_frequency	Hz	em_min	Minimum of Spectral Axis	
max_frequency	Hz	em_max	Maximum of Spectral Axis	
rest_frequency	Hz		RESTFRQ Keyword	Used for transformations
band_code			weblog	Requires outside information source for accuracy.
polarization_id		pol_states	Polarization CRVALn value	CASA uses the CRVALn value to convey polarization information, with [1,2,3,4] mapped to [I,Q,U,V]. Default to 'None' in case of other values.
telescope		instrument_name	TELESCOP Keyword	This must accommodate images for multiple instruments (i.e. VLA + Single Dish)
file_id			automatically generated	link to information about the physical file
image_id		obs_id	automatically generated	Unique Identifier for the Image
image_units		o_ucd	BTYPE & BUNIT Keywords	Description of the physical quantity measured in the image
max_intensity	image_units		image table values	
min_intensity	image_units		image table values	
rms_noise	image_units		weblog	Both of these will need to come from the last stage of the imaging process, to maintain accuracy. For quicklook images, that is stage7. That may change (and will likely be different for the Single Epoch products).

thumbnail			weblog, find the matching thumbnail image generated.	
tags				Internal tagging system to facilitate searches.
ra_pixel_size			Appropriate CDELTn	
dec_pixel_size			Appropriate CDELTn	

FITS Data Description Keywords:

For the purpose of generality, FITS provides a detail-independent method of data access. It's easier to think of the data axis descriptors in groupings by their axis number. The NAXIS keyword provides the total number of dimensions within the data. For the n^{th} dimension of the data, we have a set of descriptor keywords which should be considered and used together:

- NAXISn - Total data size along this axis
- CRPIXn - Our reference location
- CRVALn - The physical value at our reference location
- CDELTn - The increment along the axis
- CTYPEn - The axis label
- CUNITn - The axis units

The CTYPEn and CUNITn values provide information about the axis to which this group of values applies. The rest of the keywords can then be used to calculate points of interest upon that axis. For instance, in axes longer than 1, we have:

Minimum: CRVALn + CDELTn*(1-CRPIXn)

Center: CRVALn + CDELTn*(NAXISn/2 - CRPIXn)

Maximum: CRVALn + CDELTn*(NAXISn - CRPIXn)

For the Frequency axis, which only has a single point (NAXISn=1), our calculations are simpler:

Minimum: CRVALn - CDELTn/2

Center : CRVALn

Maximum: CRVALn +CDELTn/2

Image Sets Database Table:

The Image Set information will need to come from outside sources, as most of the information is not guaranteed to be in the FITS files themselves. Vlass Manager holds all the needed information for their images, but future development will need to provide the relevant metadata as image sources broaden beyond VLASS.

Column Name	VO Equivalent	Source	Comment
image_set_id	obs_id	automatically generated	Unique Identifier for the Image Set
project_code		Required to facilitate Ingestion	
configuration		VlassMgr	This will need to hold the entire list used for the imaging.
collection_name	obs_collection	VlassMgr	
calibration_level	calib_level	VlassMgr	As defined by the VO in their 0-4 system
product_file_id		automatically generated	Link to the imaging products tar file
tags			internal tags which apply to all images of the set

VO ObsCore Remaining Fields:

VO Requirement	Value	Source
access_url		
access_estsize		files.filesize, or combined value for an image set
dataproduct_type	'image'	default
access_format	'fits'	default

obs_publisher_did		Obtained upon registering with the Virtual Observatory
facility_name	'NRAO'	default
t_resolution		images.exposure_time
t_xel	1	default
em_res_power	null	default
em_xel	1	default
pol_xel	1	default

Thumbnails

We won't store thumbnails in NGAS as such, for each we will:

1. compute the sha1 hash of the file
2. store the file in a filesystem at $ROOT/$1/$2/$3/$FILENAME, where $ROOT is a CAPO property that maps to the root of the filesystem, $1 is the first two characters of the sha1 sum, $2 the second two, and $3 the third two.
3. in the metadata database we will store the $1/$2/$3/$FILENAME path