

# Pulsar and FRB Radio Data Discovery and Access

- *an update from INAF and RIG activities* -



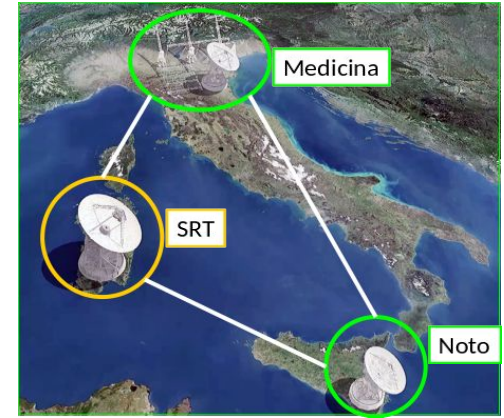
**V. Galluzzi**, A. Zanichelli (INAF Radio Data Archive WG Coordinator), C. Knapic (INAF-IA2 Coordinator),  
A. Bignamini, R. Butora, M. Burgay, M. Molinaro



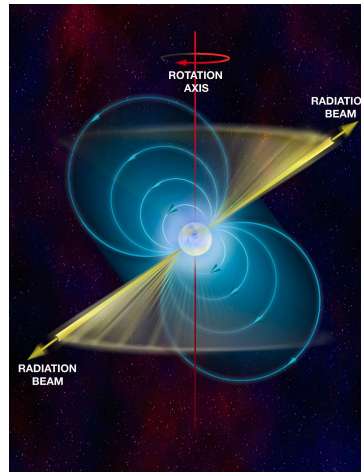
IVOA RIG people involved: Ada Nebot-Gomez, Brent Miszalski, Mireille Louys, Alan Loh, Mark Lacy,  
Jean-Matthias Griessmeyer, Yann Grange, Mark Cresitello-Dittmar, Baptiste Cecconi  
*Editor:* François Bonnarel

# Summary

1. Pulsar/transients observations and data formats at INAF
2. Data archiving workflow and internal data model
3. Mapping onto ObsCore DM
4. Proposal for an extension table
5. Summary and discussion points



Credit: Astron/Jive/Evn, Daniëlle Futselaar (artsource.nl)

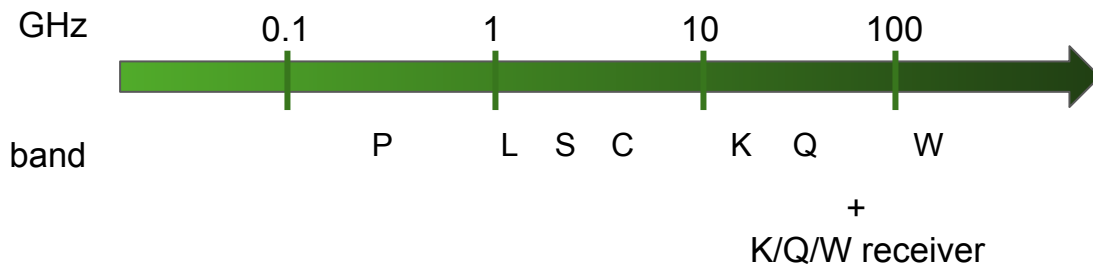


Credit: B. Saxton, NRAO/AUI/NSF

# Pulsar/transients data formats

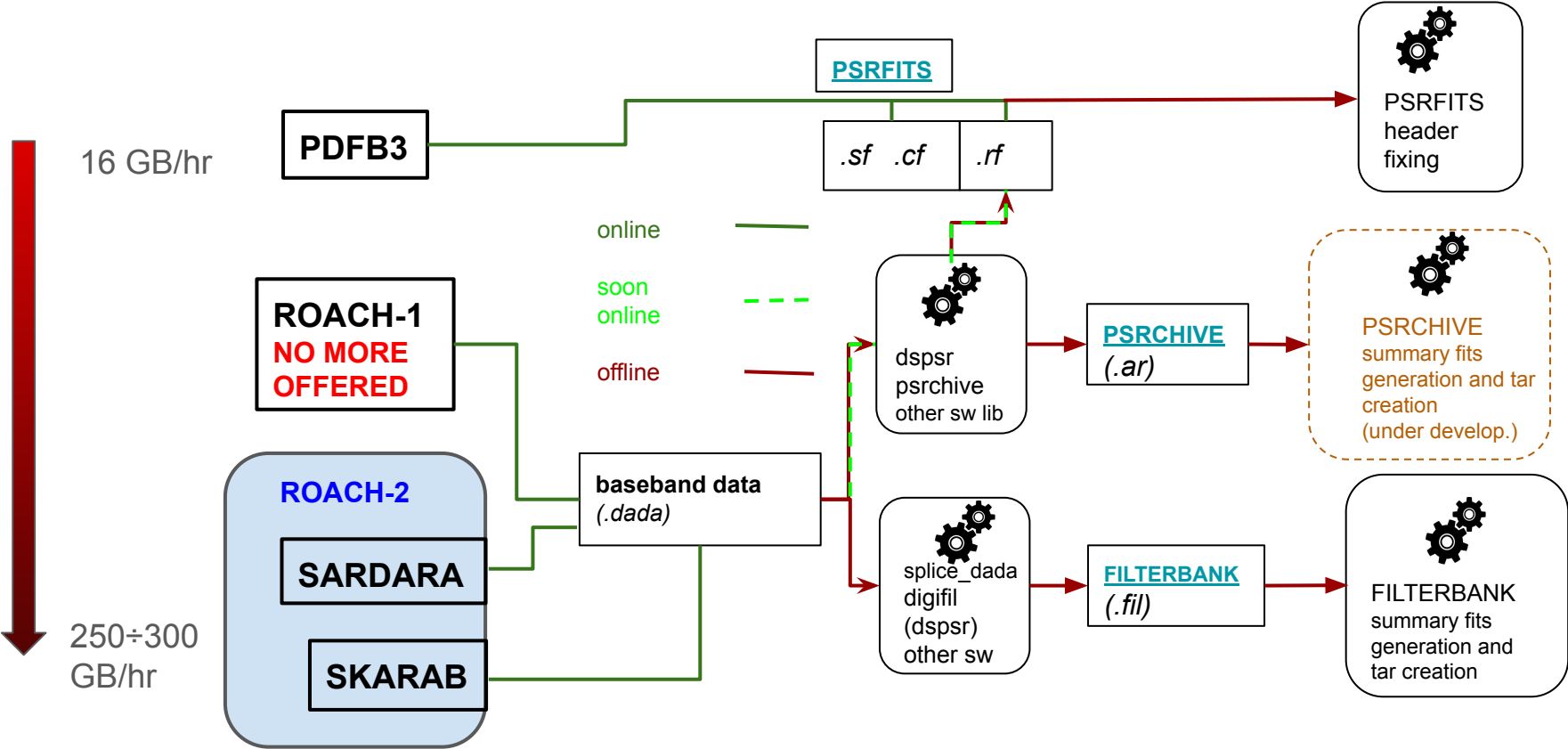
Name	Type	Integr. time	Max frequency bins
available bandwidths (MHz)			
<b>TP</b> 250, 680, 1200 (C and K bands only), 2000 (C and K bands only)	Analog total power	1-1000 ms	1
<b>XARCOS</b> Narrow band spectrometer; up to four (for single-feeds) simultaneous bandwidths: 0.5, 2.0, 7.8, 62.5	Spectro- polarimeter	10 s	2048
<b>SARDARA</b> 420, 1000 (L-band only, no f-track), 1500	Spectro- polarimeter	Down to 5 ms	1024 or 16384
<b>DFB3</b> 1024	Correlator for pulsars	Typical: 100 $\mu$ s	2048
<b>SKARAB</b> 93.75, 187.5, 1400 (only 2048 bins, full Stokes)	Spectro- polarimeter	1 ms to 1 s	65536 spectra only 32768 full Stokes
<b>DBBC</b> 512	VLBI	---	---

Thanks to recent backends time resolutions and frequency bandwidths (as well as the number of channels) are significantly improving. This comes with increasing data rates (20÷30 GB/hr for SARDARA, even a factor  $\sim 10$  is foreseen for SKARAB), posing challenges for several archival operations.



Thanks to high-frequency receivers available in the next future, INAF telescopes are going to offer bands from P to W (i.e. from 305 MHz up to 116 GHz).

# Pulsar/transients data formats

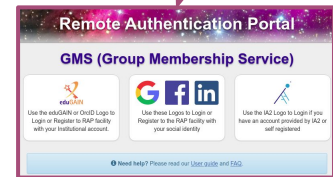
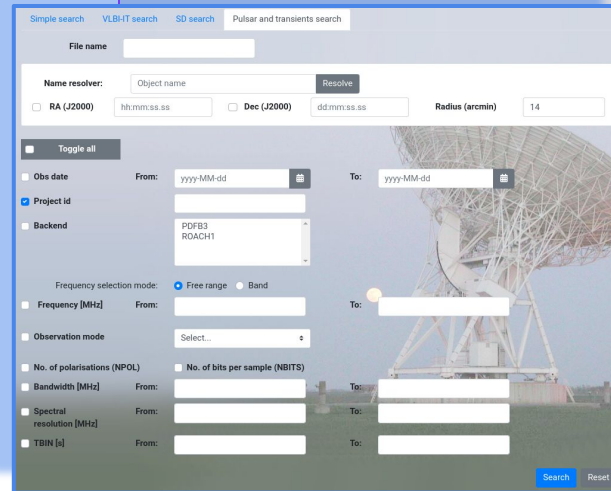
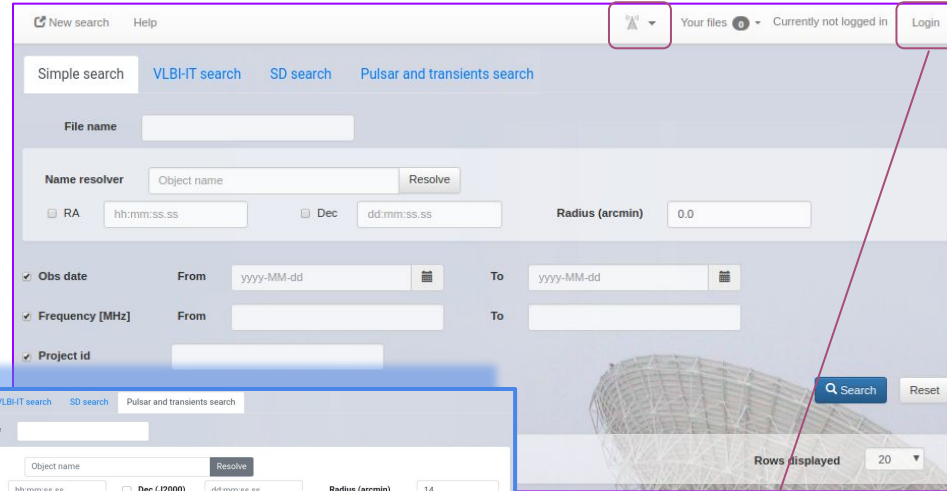
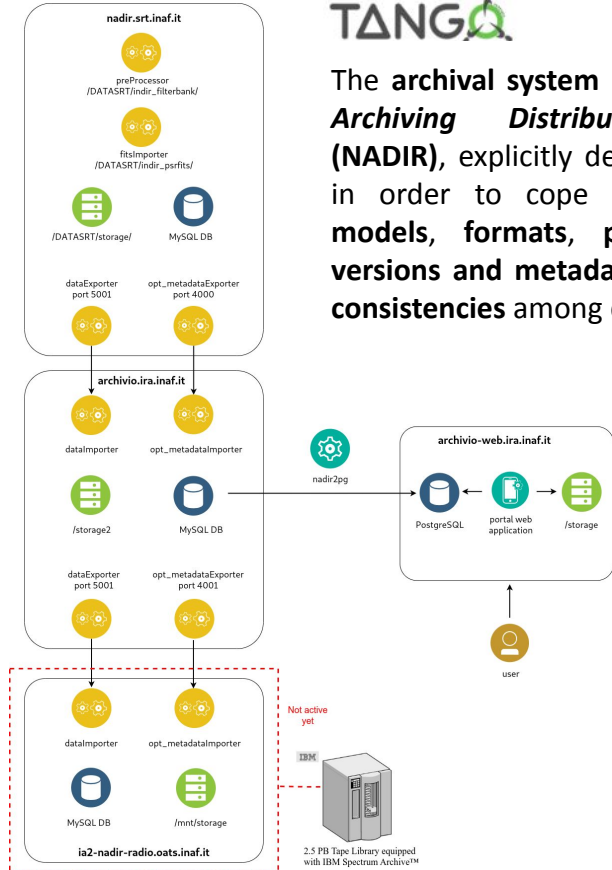


# The archival system

SAMP broadcast



The **archival system** is based on the **New Archiving Distributed InfrastructuRe (NADIR)**, explicitly designed to be **flexible** in order to cope with **evolving data models, formats, publication policies, versions and metadata contents, keeping consistencies** among different sites.



# Internal data model

- The metadata of an observation are all written in the header of the primary HDU of a FITS file (in case of FILTERBANK or PSRCHIVE, we produce an accompanying FITS file containing only a primary header PSRFITS-like).

id	column_name	column_type	fits_key_hdu	fits_key_pri	fits_key_sec	
1	TELESCOP	vchar	0	TELESCOP	TELESCOP	Telescope name
2	DATE_OBS	vchar	0	DATE-OBS	DATE-OBS	Date of observation (YYYY-MM-DDThh:mm:ss UTC)
3	SRC_NAME	vchar	0	SRC_NAME	SRC_NAME	Source or scan ID
4	OBSERVER	vchar	0	OBSERVER	OBSERVER	Observer name(s)
5	OBS_MODE	vchar	0	OBS_MODE	OBS_MODE	PSR, CAL, SEARCH)
6	BACKEND	vchar	0	BACKEND	BACKEND	Backend ID
7	RA_C	vchar	0	RA	RA	Right ascension (hh:mm:ss.ssss)
8	DEC_C	vchar	0	DEC	DEC	Declination (-dd:mm:ss.sss)
9	EQUINOX	double	0	EQUINOX	EQUINOX	Equinox of coords (e.g. 2000.0)
10	PROJID	vchar	0	PROJID	PROJID	Project name
11	OBSFREQ	double	0	OBSFREQ	OBSFREQ	[MHz] Centre frequency for observation
12	OBSBW	double	0	OBSBW	OBSBW	[MHz] Bandwidth for observation
13	SCANLEN	double	0	SCANLEN	SCANLEN	[s] Requested scan length (E) (N.B.: diff. from MBFITS)
14	NPOL	int	0	SUBINT.NPOL	SUBINT.NPOL	Nr of polarisations
15	TBIN	double	0	SUBINT.TBIN	SUBINT.TBIN	[s] Time per bin or sample
16	NBITS	int	0	SUBINT.NBITS	SUBINT.NBITS	Nr of bits/datum (SEARCH mode 'X' data, else 1)
17	CHAN_BW	double	0	SUBINT.CHAN_BW	SUBINT.CHAN_BW	[MHz] Channel/sub-band width
18	OBSDATAFORMAT	vchar	0	OBSDATAFORMAT	OBSDATAFORMAT	Data format of the observation

# Mapping onto ObsCore DM from PSRFITS/PSRCHIVE

- PSRCHIVE is the most comprehensive and most used for data exchange
- PSRFITS implementation depends on the telescope/data provider

## OK

calib\_level

facility\_name, instrument\_name

obs\_collection, obs\_id, obs\_publisher\_did

access\_url, access\_format, access\_estsize

target\_name, s\_ra, s\_dec, s\_resolution

t\_min, t\_max, t\_exptime

em\_min, em\_max

## Focus on

dataprodct\_type

o\_ucd

t\_resolution, t\_xel

pol\_states, pol\_xel



# Proposal for an extension table

- First of all, how to provide such an extension (additional table or additional columns for an ObsCore table)?

ObsCore extension column name	origin	PSRFITS mapping	comment	
t_delt	time extension	TBIN	or mean(OFF_SUB[n+1]- OFF_SUB[n])	} temporal sampling cadence
t_delt_min	time extension	TBIN	or min(OFF_SUB[n+1]- OFF_SUB[n])	
t_delt_max	time extension	TBIN	or max(OFF_SUB[n+1]- OFF_SUB[n])	
t_res_min	time extension	min(TSUBINT)	TSUBINT more appropriate than TBIN	} effective temporal resolution
t_res_max	time extension	max(TSUBINT)	TSUBINT more appropriate than TBIN	
t_mode		folded/search		
t_fold_period			NULL for search mode	
t_scale	time extension	TAI, TT, UTC, ...	any of IVOA timescale vocabulary <sup>9</sup> not available in PSRFITS metadata	



# Proposal for an extension table

t_refPosition	time extension	time measured at this position	not available in PSR-FITS metadata
t_origin	time extension	origin of relative time	not available in PSR-FITS metadata
t_refDirection	time extension	should be position of source in sky	not available in PSR-FITS metadata
t_format	time extension	ISO, JD, MJD, julian year, ...	not available in PSR-FITS metadata
s_fov_min	radio extension	$1.22 * \text{em\_min}/D$	where is D antenna diameter ?
s_fov_max	radio extension	$1.22 * \text{em\_max}/D$	where is D, antenna diameter
s_resolution_min	radio extension	$1.22 * \text{em\_min}/D$	doesn't make sense or equal to s_fov (no source separation inside this range)
s_resolution_max	radio extension	$1.22 * \text{em\_max}/D$	doesn't make sense or equal to s_fov (no source separation inside this range)
f_resolution	radio extension	$\text{mean}(\text{DATA\_FREQ}[n+1] - \text{DATA\_FREQ}[n])$	Proposed by INAF because em_res_power changes too much along the spectral band
f_max	radio extension	$c / \text{em\_min}$	Radio astronomers prefer frequencies
f_min	radio extension	$c / \text{em\_max}$	Radio astronomers prefer frequencies

Already in discussion for the ObsCore DM extension for radio data

# Mapping onto ObsCore DM from FILTERBANK

obscore column name	filterbank keyword	comment
t_min	tsart	
t_max	tstart + tsamp * nsamples	
t_exptime	tsamp * nsamples	assuming there is no interruption during observation
t_delt	tsamp	tsamp looks more a cadence than a resolution
t_resolution	tsamp ???	tsamp looks more a cadence than a resolution
t_mode	folded/search	
t_fold_period	period	
s_ra	src_raj	
s_dec	src_dec	
s_fov	?	depends from telescope and frequency
em_min	$c/(fch1+foff*nchans)/10e6$	
em_max	$c/fch1/10e6$	
f_max	$(fch1+foff*nchans)*10e6$	radioastronomers prefer frequencies
f_min	$fch1*10e6$	radio astronomers prefer frequencies
f_resolution	foff	sampling, resolution or both ?


obscore column name	filterbank keyword	comment
pol_xel	nifs	
pol_states	?	doesn't appear in keywords ?
o_ucd	phot.flux.density	is that always true for filterbank ?
dataprodct_type	data_type	filterbank is dynamic spectrum ? otherwise lightcurve , spectrum ?
facility_name	inferred from telescope_id value	
instrument_name	inferred from machine_id value	
target_name	source_name	
calib_level	1	

Table 3: ObsCore and Onscore extension mapping proposal for filterbank datasets

FILTERBANK is used for search-mode data, and it consists in a textual header + a binary table and is readable only by a specific software ([sigproc](#)): **facilities/telescopes and backends appears only with a numeric id the software is able to interpret.**

# Summary and Draft Document discussion points

- The INAF radio data archive now contains more than 1450 PSRFITS files acquired observed between 2018 and 2020, and ready for handling FILTERBANK data as well.



**Pulsar and FRB Radio Data Discovery and Access**  
Version 1.0

**IVOA Note 2022-09-22**

Working group  
DAL

This version  
<https://www.ivoa.net/documents/PulsarRadioDiscoveryAndAccess/20220922>

Latest version  
<https://www.ivoa.net/documents/PulsarRadioDiscoveryAndAccess/20220922>

Previous versions

Author(s)  
Alessandra Zanichelli, Ada Nebot-Gomez, Brent Miszalski, Gilles Louys, Alan Loh, Mark Lacy, Jean-Matthias Griessmeyer, Yann Grange, Vincenzo Galluzzi, Mark Cresitello-Dittmar, Baptiste Cecconi, François Bonnarel

Editor(s)  
François Bonnarel

**DRAFT DOCUMENT**

- Re-trigger the discussion in the Semantics WG (i.e. for `dataprodct_type`, `o_ucd`), considering also the usage of `dataprodct_subtype`
- complete/integrate the vocabulary for polarization states
- finalize of the proposal for the extension table (how to provide the information, how to tackle with the absence of info `t_scale` & Co. in pulsar data files (e.g. defaults values?), as well as maintaining a redundancy in energy (frequency and wavelength)
- `s_fov` and `s_resolution` quantities (*cf.* discussion about the extension for radio data)
- FILTERBANK-specific issues (mostly lack of information wrt PSRFITS/PSRCHIVE)