Data sharing prototype of Pulsar data at CDN (Centre de Données de Nançay)

Cecconi, B; Le Sidaner, P.; Grießmeier, J.-M., Bondonneau, L. — May 2023
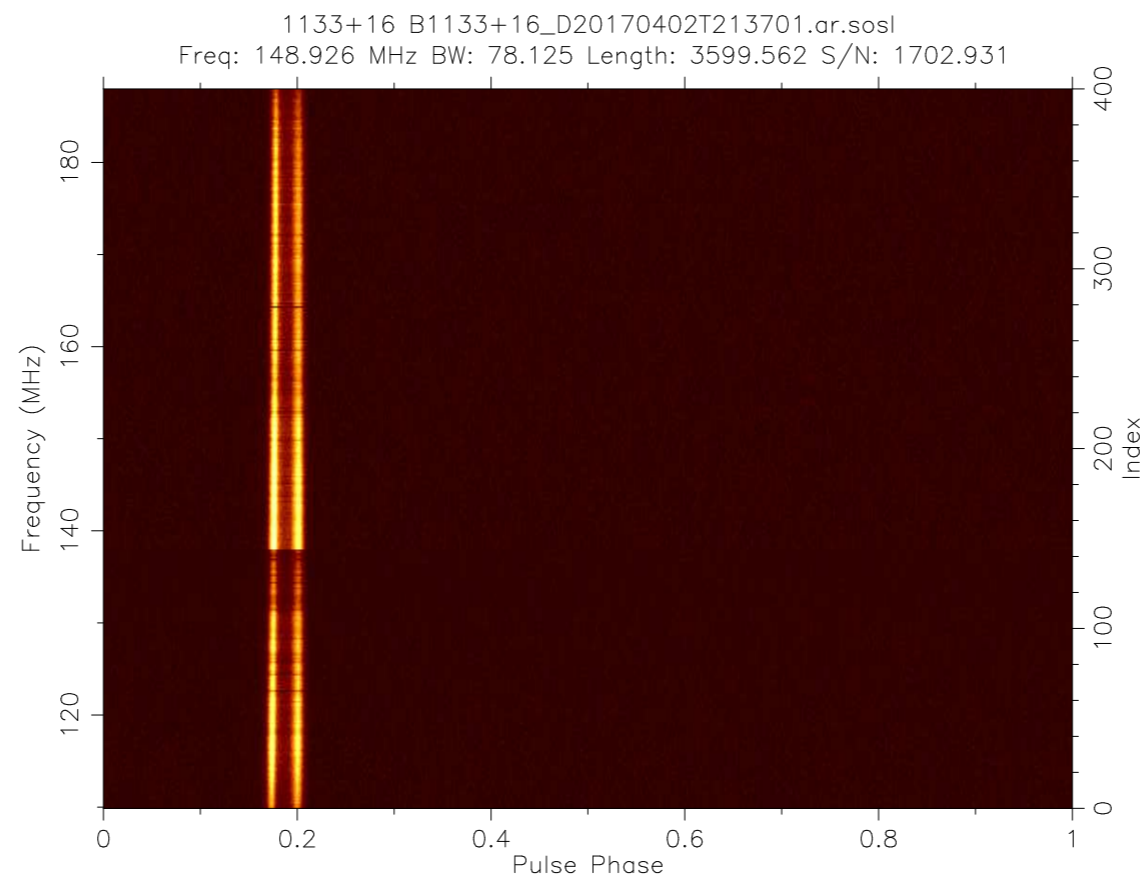
Radio Interest Group document
Pulsar and FRB Radio Data Discovery and Access

- Document drafted by (2022-09):
 *Alessandra Zanichelli, Ada Nebot-Gomez, Brent Miszalski, Mireille Louys, Alan Loh, Mark Lacy, Jean-Mathias Grießmeier, Yann Grange, Vincenzo Galluzzi, Mark Cresitello-Dittmar, Baptiste Cecconi, François Bonnarel*

- Review by Baptiste Cecconi, Jean-Mathias Grießmeier
  External contact with Joern Kuensemoeller (Uni. Bielefeld, DE)

Pulsar Dataset
Content

- dedispersed data polarisation or flux measurement

- a set of light curves (usually resulting from data folding at the pulsar rotation period), in a temporal-spectral grid.

  e.g.: for each epoch:
  a spectral-temporal phase plot

- many metadata (coverage, processing history, configuration…)

observatory	FR606
obs.id	B1133+16_D20170402T213701
PSRNAME	B1133+16
JNAME	J1136+1551
P0	1.18796334845016
DM	4.86110639572144
length	3599.56189695998
nsubint	379
center freq.	148.92578125
BW	78.125
S/N	2669.93
%RFI	13.34
quicklook created	April 13, 2017
by	process-and-calculate.sh (version 1.52.03, 10.12.2016)
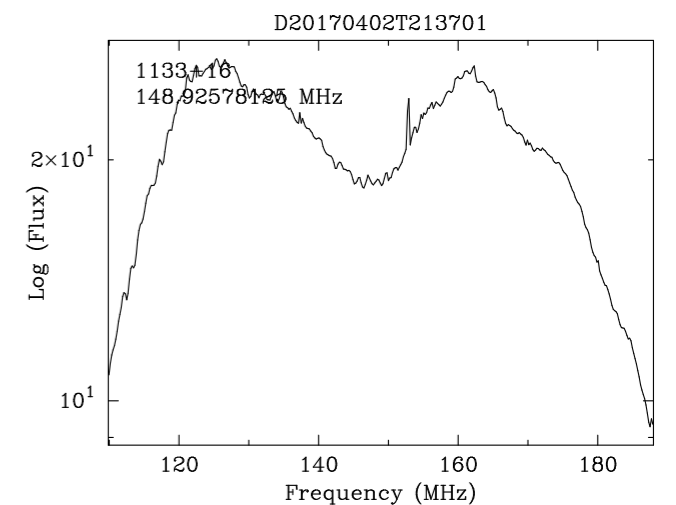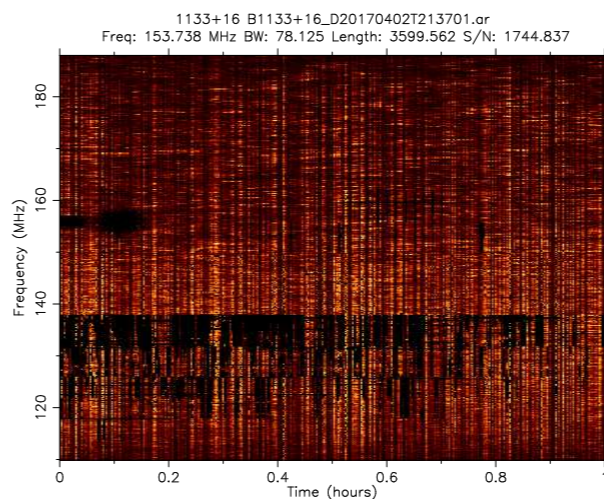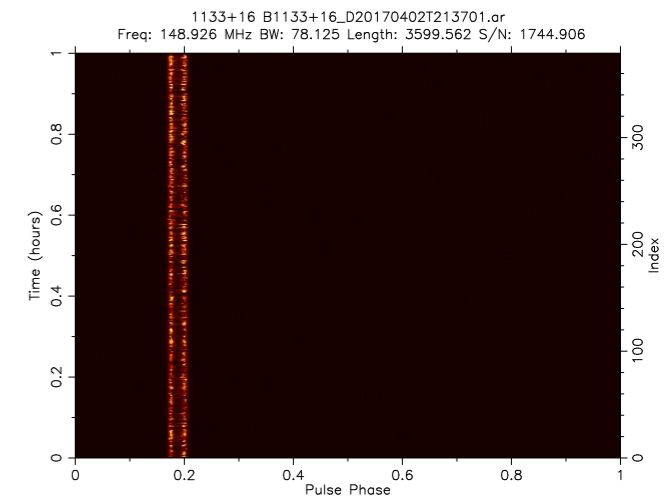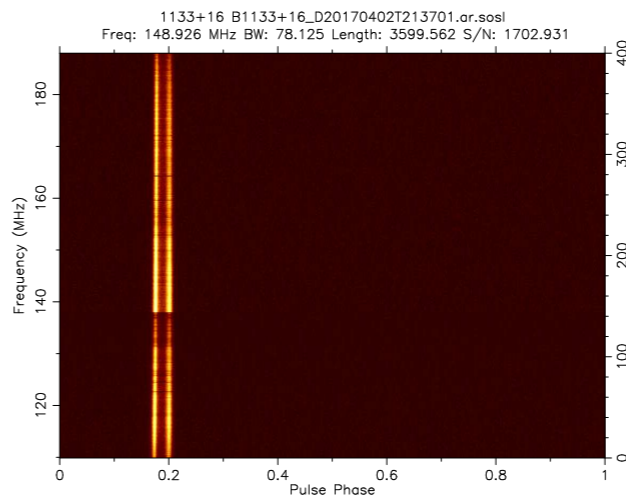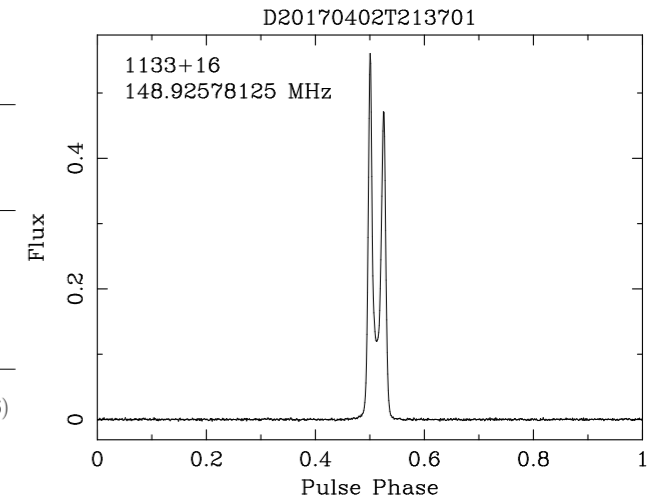
Pulsar Dataset
PSRCHIVE & PSRFITS

- Pulsar data are processed using many standard libraries, used by the pulsar community (PRESTO, TEMPO, TEMPO2…)

- Most comprehensive data format is *psrchive* (*.ar), which comes with a CLI: *psredit*, *psrstat*, *psrplot*…

- The PSRFITS format (*.psrfits) is a FITS format with a set of specifications. It is used for data distribution, but content is somewhat loosely defined.
  Most exchange of data is done with PSRCHIVE format.

- Data are usually not open and accessible.

- Sharing metadata is a first step.

ObsCore	PSRFITS FILE (.psrfits or .fits)	PSRCHIVE FILE (*.ar through psredit)
t_min	STT_IMJD + (STT_SMJD+STT_OFFS - TSUBINT[0]/2)/86400	int[0]:mjd - 1/2*int[0]:duration/86400
t_max	STT_IMJD + (STT_SMJD+STT_OFFS + OFFS_SUB[-1] + TSUBINT[-1]/2)/86400	int[nsubint-1]:mjd - 1/2*int[nsubint -1]:duration/86400
t_exptime	t_max - t_min	length
t_resolution	f[4].data["TSUBINT"][1]	int[1]:duration
t_xel	f[1].data["NSUB"][0]	nsubint
s_ra	f[4].data['RA_SUB'][0]	coord
s_dec	f[4].data['DEC_SUB'][0]	coord
s_xel1	1	
s_xel2	1	
em_min	c / f_max	c / f_max
em_max	c / f_min	c / f_min
em_xel	f[1].data["NCHAN"][0]	nchan
o_ucd	based on npol	
pol_states	based on f[4].header["POL_TYPE"]	based on state
pol_xel	f[4].header["NPOL"]	npol
facility_name	based on f[0].header["TELESCOP"]	based on site
instrument_name	based on f[0].header["FRONTEND"] and f[0].header["BACKEND"]	based on "rcvr:name" and "be:name"
t_scale	UTC	
t_refposition	TOPOCENTER	
t_origin	??	
f_resolution	f[1].data["CHAN_BW"][0]	bw/nchan
f_min	f[4].data["DAT_FREQ"][0][0] - f[1].data["CHAN_BW"][0]/2	int[0]:freq[0] - bw/nchan/2
f_max	f[4].data["DAT_FREQ"][0][-1] + f[1].data["CHAN_BW"][0]/2	int[0]:freq[nchan - 1] + bw/nchan/2

Implementation at CDN
Centre de Données de Nançay

- Using DaCHS server (http://vogate.obs-nancay.fr
  - schema = fr606 (name of french LOFAR station)
  - table: fr606.obscore

- Mapping from a set of Nenufar and LOFAR PSRFITS files.
  - PSRCHIVE files: not yet implemented at CDN

- NenuFAR data: Original observation (OBS_ID) contains several « beams », observing several pulsars in parallel => 1 row per pulsar beam.

- Mapping choices:
  - dataproduct_type => new term: *hyper-dynamic-spectrum*
  - facility_name => LOFAR/FR606, NenuFAR
  - instrument_name => backend name (LuMP, LUPPI)
  - pol_states => « XX/YY/ReXY/ImZY »

ADQL Query

Parameters

- ADQL query: select * from fr606.obscore

Result

Matched: 100

[Send via SAMP] [Quick Plot]

Query result probably incomplete due to the match limit kicking in. Queries not providing a TOP clause will be furnished with an automatic TOP 2000 by the machinery, so adding a TOP clause with a higher number may help.

Product key	Owner	Embargo ends	Type	File size [byte]	Access_estsize [kbyte]	Access_format	Access_url	Calib_level[calib]	Dataproduct_type	Target_class	Target_name	Obs_collection	Obs
B0031-07_D20200103T1719_58851_000181_0030_BEAM0_0001.fits	N/A	2025-01-01 00:00:00	application/x-fits-psrfits	7.8GiB	8356357440	application/x-fits-psrfits		2	N/A	pulsar	B0031-07	nenufar/LUPPI	

Obs_publisher_did	T_min	T_max	T_exptime [s]	T_resolution [s]	T_xel	S_ra [deg]	S_dec [deg]	S_xel1	S_xel2	Em_max [m]	Em_min [m]	Em_xel	O_ucd	Pol_states	Pol_xel	Fa
01	2020-01-03T17:20:01Z	2020-01-03T17:49:43Z	1782.41138	10.7374182	166	2.234970796946071	-1.928109503681259	1	1	6.0669461856126485	24.75705459612903	192	phot.flux.density	XX/YY/Cr(XY)/Im(XY)	4	

Facility_name	Instrument_name	S_fov [deg]	T_scale	T_refposition	File_name	Thumbnail_url	F_min [Hz]	F_max [Hz]	F_resolution [Hz]	Profile_nbin	Process_p0 [s]	Process_rm [rad/m**2]	Pr[p
nenufar													

F_min [Hz]	F_max [Hz]	F_resolution [Hz]	Profile_nbin	Process_p0 [s]	Process_rm [rad/m**2]	Process_dm [pc/cm**3]	Profile_rfi_fraction	Profile_snr	Src_elevation [deg]	Obs_owner	Obs_contact	Obs_creation_date	Obs_release_date
12.0117188	49.5117188	0.1953125	1	0.000655359996	N/A	10.9160004	N/A	N/A	N/A	ES03	LB	2020-01-03T17:49:56	N/A
49.5117188	87.0117188	0.1953125	1	0.000655359996	N/A	10.9160004	N/A	N/A	N/A	ES03	LB	2020-01-03T17:49:55	N/A
12.0117188	49.5117188	0.1953125	1	0.000655359996	N/A	10.9160004	N/A	N/A	N/A	ES03	LB	2020-01-04T17:45:55	N/A

Open Issues
and other prospects

- PSRFITS implementation is not homogeneous.
  - NenuFAR PSRFITS (.fits) => one extension (SUBINT)
  - LOFAR PSRFITS (.psrfits) => four extensions (HISTORY, PSRPARAM, T2PREDICT, SUBINT)
  - some parameters are not filled (or not located in the same extension).

- New *dataproduct_type* term? « **hyper-dynamic-spectrum** »
  (also applicable to space radar data, like Mars-Express/MARSIS)

- *pol_states*: not fully satisfactory (ReXY and ImXY are not listed in allowed values)

- Extra remark: **Seems to be missing a PSRCat name resolver**:
  - Should include temporary names
  - Possible to implement from successive versions of PSRCat