The INAF radio data archive: towards FAIR data handling

A Zanichelli - INAF IRA

From Science Gateways to Papers - May 24th, 2022

© Paolo Golfarini 2021

The INAF radio telescopes

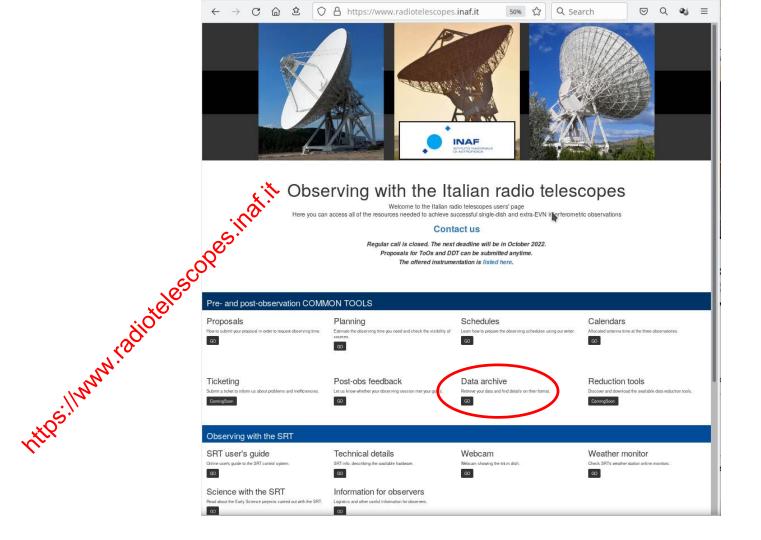


Observing modes: SD Pulsar VLBI

Share the same control software and data format

In operation: 300 MHz – 26(+) GHz Total intensity, spectropolarimetry





The INAF radio archive Working Group

INAF IRA

Alessandra Zanichelli Francesco Bedosti Vincenzo Galluzzi Mauro Nanni Andrea Orlati Simona Righini Matteo Stagni Franco Tinarelli

INAF OATs

Nicola Calabria Cristina Knapic Marco Molinaro Massimo Sponza

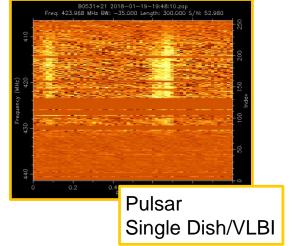
INAF OAC Marta Burgay Raimondo Concu Antonietta Fara Sergio Poppi Andrea Possenti

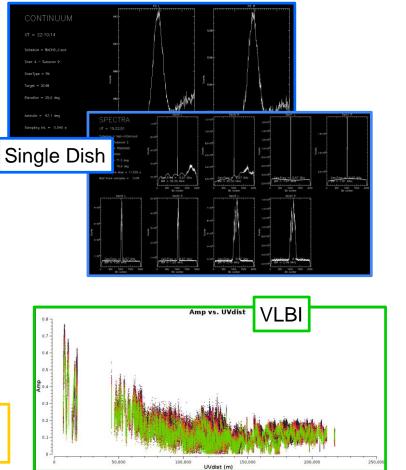
:	Help								[®] A ^d ▼ Your files	∎+ Logi
		Simple search VL	BI-IT search	SD search Pu	lsar search					
		File name								
RNING!		Name resolver:	Object n	ame		Resolve				
y raw a, at the		🗆 RA (J2000)	hh:mm:ss.ss	De	ec (J2000)	dd:mm:ss.ss	Radius	arcmin) 14		
nent		Toggle all								
		Equinox	Select	4	2 2					
	. X.	🗹 Obs date	From:	уууу-MM-dd		To:	yyyy-MM-dd	ŧ		
	ð.	Frequency [MHz]	From:			To:	CASE IN			
P.IIradioarchive.in		Project id					(CEL)	NAT 1		
Xilde		Telescope		Select	\$		N. Andrew	NILLE A	13/570	
ator		Frontend Bandwidth	From:	Select	•	To:	Backend	Select		A
dio		[MHz] Exposure time	From:			To:	10		AK LA	
INAC		[s] LST	From:	hh.mm.ss.sss		To:	hh.mm.ss.sss	13 14 14		1
; [,] \`		Scan type	TTOM.	Select	•	10.	nn.mm.ss.sss	- 1170		
		Spectral resolution	n (MHz)			Re	st frequency [MHz]	T		
		Frequency bi	ns	-						
		Wobbler		Select	+					
					20.00	And Annual Constants	and the second s		the start of the second start when	

Progressively being populated with SD, pulsar and VLBI-IT data (VLBI-IT =software correlated in Bologna)

WARNING! Only raw data, at the moment

Variety of data types





Data formats

		(1	2 1919	écunit-se (-/ <u>9</u> 55	:chivikiin:	libifits	gius	u15/1	31.9	1. 61	Й	$\overline{)}$		-		×
	Eile	Edit	Jools	≦ynta	x <u>B</u>	uffers	Wind	dow	He	lp							
			8	-		16 D	0	2	-	4	6		0	63	•	H	
	<data <ta < < < <</ta </data 	base> ble> PROJIO TELESO PINAMO OBJEC RA>30	>1313 OP>VL >Marc >0212 29189	.0' end fBITello Gi +735 (<br 26036 </th <th>ID> ELESC BJEC RA></th> <th>COP> etti<th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th>() () () () () () () () () () () () () (</th></th>	ID> ELESC BJEC RA>	COP> etti <th></th> <th>() () () () () () () () () () () () () (</th>											() () () () () () () () () () () () () (
VLBI FITS	-	D	I	d-Mc+ /FREC	> OBS_	9:43 5<br TENNAS: _MODE>		>									200
		TELES PINAMI OBJEC RA>45 DEC>30 START:	0P>VL >Marc D=0340 .72470 5.3701 >2014-	fBITello 61 +362 0<br 90019 <br 19325 0<br 03-28	ELESC Irole BJEC (RA> (DEC: 14:19	etti /</th <th></th> <th></th> <th></th> <td></td> <th></th> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>											

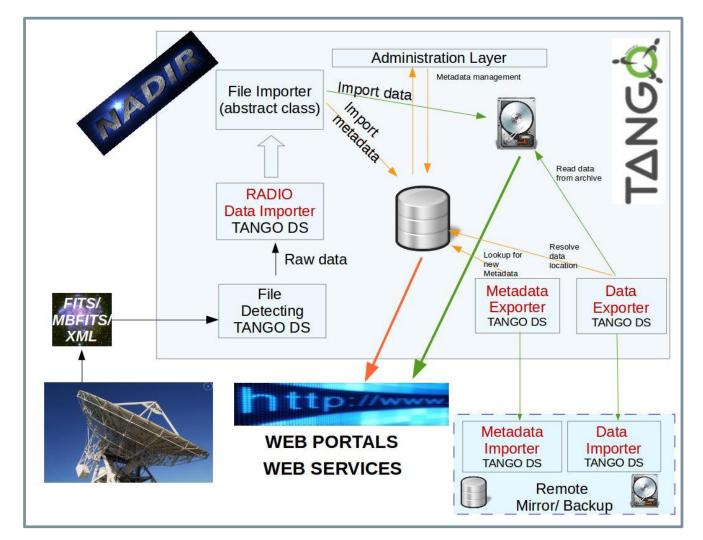
le Edit	Tools							He
Index	Extension	Туре	Dimension		View			
□ 0	Primary	Image	0	Header	lma	age		Table
□ 1	ARRAY_GEOMETRY	Binary	7 cols X 2 rows	Header	Hist	Plot	Ali	Select
□ 2	SOURCE	Binary	26 cols X 4 rows	Header	Hist	Plot	All	Select
□ 3	ANTENNA	Binary	13 cols X 18 rows	Header	Hist	Plot	All	Select
□ 4	FREQUENCY	Binary	6 cols X 1 rows	Header	Hist	Plot	Ali	Select
□ 5	INTERFEROMETER_MODEI	. Binary	20 cols X 210 rows	Header	Hist	Plot	All	Select
□ 6	CALC	Binary	11 cols X 5 rows	Header	Hist	Plot	All	Select
□7	MODEL_COMPS	Binary	21 cols X 210 rows	Header	Hist	Plot	All	Select
□ 8	UV_DATA	Binary	13 cols X 8410 rows	Header	Hist	Plot	Ali	Select
□ 9	SYSTEM_TEMPERATURE	Binary	10 cols X 0 rows	Header	Hist	Plot	Ali	Select
🛄 10	PHASE-CAL	Binary	17 cols X 378 rows	Header	Hist	Plot	All	Select

		î		dARY ADER		Si	ngl	e-d	is	h F	ITS	S		
			ĮD	ATA]		••(*	iv Header of	20140308-1	150612-M	antenance-En	npisy_001_0	02.fits[0]1		×
				NSION		File Edit	Tools						Hel	lp
		Extension 1	1 D/	TA		Search for:	[1						
		Extension 2		NSION ADER		SIMPLE = BITPIX = NAXIS = EXTEND =	FITS (Flexib	8 / r 0 / r T / 8	number o number o TTS dat	s conform to f bits per d f data axes aset may con	lata pixel tain exter	nsions		
			EXTI	ATA INSION ADER		COMMENT COMMENT V. HISTORY V. HISTORY V.	and Astrophy 1.0 Created 0.8 First o 0.82 The tsy	by S. Right utput stands column in	e 376, ni, M Ird for data ta	page 359; bi Bartolini & Italian radi ble raplaced	bcode: 200 A. Orlati otelescope with the	01ASA376 i So Tant table	6359H	
• + 💌	fy: Summary	Extension 3	1	ATA	001	HISTORY V.	0.9 The sect inputs tabl	ion table ha	for ea been	ch input of splitted int	each secti o two tabl	ion Les: sectio	ons and	
File Edit	Tools		<u> </u>			HISTORY V.	0.91 Added t 0.92 SubScan	he flux colu	an in s	ection table	word			
Index	Ext	ension	Туре	C	imension	HISTORY V.	1.0 Added ne ceivers: SER	w table to a	tore po	sition of su	briflector	: e primary	y focus	
0	Pr	imary	Image		0	HIERARCH P	'2014-03-08T roject Name	= 'GZT06C '	/ Nane	of the proj	ect)DThh:ma:s:	5 UT)	
	SECTI	ON TABLE	Binary	5 C	ols X 2 rov	OBSERVER= ANTENNA =	'SRT '	/ 1	lame of	the observer the station				
2	RF	INPUTS	Binary	9 c	ols X 2 rov	HIPPSDOU C	iteLongitude WiteLatitude WiteHeight =	= 0.68928357 650. / F	9621821	/ Latitude f the site (of the sit	site (radia Se (radiana	ana) a)	
_ 3	FEEI	D TABLE	Binary	4 c	ols X 1 rov	SECTIONS=	ample Size =	2 / 1	otal nu	mber of sect f bytes of a				
⊒ 4	DAT	A TABLE	Binary			HIERARCH B SOURCE =	Expty	= 'CCB / 5	' / Key	word that id dentifier	entifies t	maion at :	12000 (r	
_ 5	ANTENNA	TEMP TABLE	Binary	2 co	ls X 364 ro	VLSK =	eclination =	0. / 2	ource r	adial veloci	CV .		radians)	
6	SERV	O TABLE	Binary	8 co	ls X 364 ro	HIERARCH E	zimuth Offse Levation Off LightAscensio Meclination O	set = 0. / 1 n Offset = 0	atitude	offset in h gitude offse	orizontal t in equat	frame torial fram	ne	
						HIERARCH G HIERARCH G SCANID - HIERARCH S HIERARCH S HIERARCH L	alacticLon 0 alacticLat 0 subScanID = cheduleName ogFileName = subScanType =	ffset = 0. / ffset = 0. / 2 / 5 = '02T06C_20 '02T06C_201	Longit Latitu can Ide ubscan 140308- 40308-1	ude offset i de offset in ntifier Identifier 105028' 05930.log'	n galactic galactic	frame frame	escope m	



The radio archive architecture

New Archiving Distributed InfrastructuRe



Towards FAIRness

- Archived raw data must be discoverable and (re)usable
- Variety of observing projects and heterogeneity of the data: accurate characterisation scientific exploitation of the Archive
- General users must be able to address
 - ✓ if the data are suitable their own research
 - ✓ if all the necessary information for data exploitation is available
- Global data discovery and access requires to expose a uniform, standard data model

The IVOA Radio Interest Group



"Enhancing interoperable data access to radio data has become a science priority within the International Virtual Observatory Alliance (IVOA). This lead to the foundation of the IVOA Radio astronomy Interest Group...Together they are paving the way to a better integration of their services in the virtual observatory (VO) infrastructure and propose extension of IVOA standards to help achieving this goal." (Louys et al 2020)

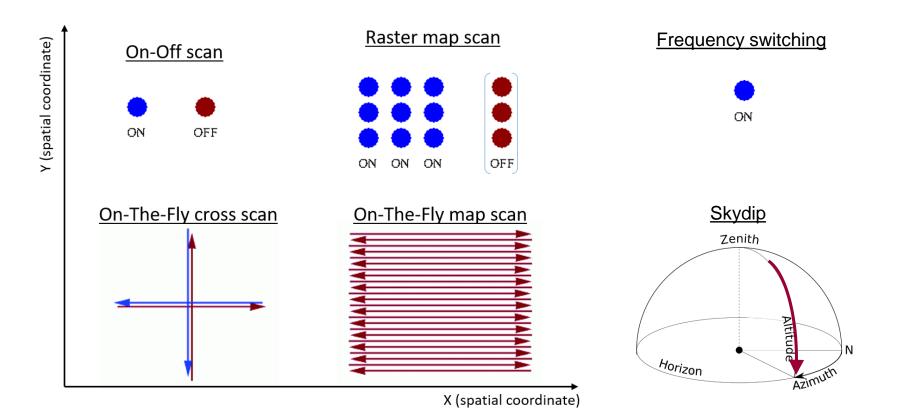
Main tasks:

- development of use cases for data discovery, access and visualization
- identification of metadata concepts needed by radio astronomy data that are not currently supported by the VO

Mapping INAF radio data into VO data models

- INAF radio data are compliant with the mandatory components of the VO ObsCore Data Model for discovery purposes
- However, to improve the discoverability some peculiarities have to be accounted for and better described (especially for SD and VLBI data)
- Example: spatial (geometrical) properties, frequency setup, time domain, UV coverage & c.

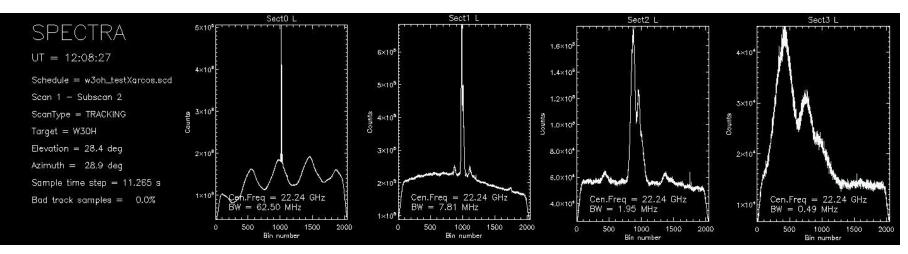
Some examples: single dish & geometry



Some examples: single dish & frequency setup(s)

It is possible to have more than one spectral window in the same scan, differing in bandwidth and spectral resolution. Examples:

- each spectral window is centered at a different frequency
- spectral windows at same frequency with increasing resolution ("zoom mode")



Multifrequency setup

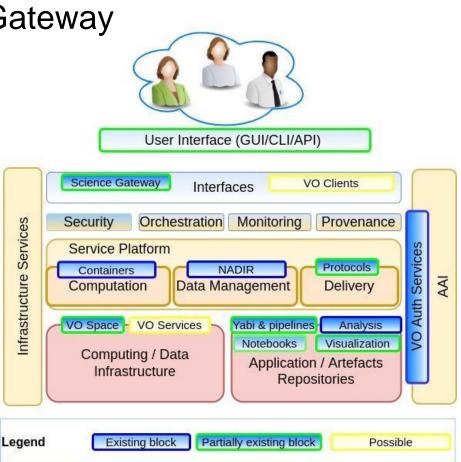
	Wobbler Se	elect		¢				Freq min	Freq max	Bandwidth	Spec res	Rest freq	_	
	Creator						105	[MHz]	[MHz]	[MHz]	[MHz]	[MHz]	Freq bins	
	Format version						(All and a second se	22235.18475	22242.99725	7.8125	0.0038146973	22235.07985	2048	LCP
	lonnac version							22235.18475	22242.99725	7.8125	0.0038146973	22235.07985	2048	RCP
							N.	22238.1144375	22240.0675625	1.953125	9.5367432E-4	22235.07985	2048	LCP
								22238.1144375	22240.0675625	1.953125	9.5367432E-4	22235.07985	2048	RCP
	ownload -						Rows displa	22238.8468594	22239.33514065	0.48828125	2.3841858E-4	22235.07985	2048	LCP
								22238.8468594	22239.33514065	0.48828125	2.3841858E-4	22235.07985	2048	RCP
	File name	Po	olicy P	Proj id	RA	Dec	Obs date F	22207.841	22270.341	62.5	0.030517578	22235.07985	2048	LCP
0	20190406-040544-13-19-G358.93-0	.03.tar FF	REE 1	13-19	17:43:10.020	0-29:51:45.80	2019-04-06T04:05:55	22207.841	22270.341	62.5	0.030517578	22235.07985	2048	RCP
	20190419-025725-13-19-G358.93-0	.03.tar FF	REE 1	13-19	17:43:10.020	0-29:51:45.80	2019-04-19T02:57:36 2	2207.631.22270.1						
	20190422-013655-13-19-Skydip tar	FF	REE 1	13-19	13:14:18.508	27:15:49.20	2019-04-22T01:36:56 2	2135.178.22385.1						
	20190422-014516-13-19-3C286.tar	FF	REE 1	L3-19	13:31:08.280	30:30:33.12	2019-04-22T01:45:17 2	2135 177 22385 1	1 200	No. of Concession, Name				
	20190422-015546-13-19-G358.93-0	.03.tar FR	REE 1	L3-19	17:43:10.020	0-29:51:45.80	2019-04-22T01:55:57 2	2207.575.22270.0		K				

Multifrequency setup

	✔ Wobbler	Select		;		Freq min [MHz]	Freq max [MHz]	Bandwidth [MHz]	Spec res [MHz]	Rest freq [MHz]	Freq bins	Pol
	Format version					22235.18475	22242.99725	7.8125	0.0038146973	22235.07985	2048	LCP
						22235.18475	22242.99725	7.8125	0.0038146973	22235.07985	2048	RCP
D	ownload - 🧹 🎸 Edit query					22238.1144375	22240.0675625	1.953125	9.5367432E-4	22235.07985	2048	LCP
	File name		Policy	Proj id	R/	22238.1144375	22240.0675625	1.953125	9.5367432E-4	22235.07985	2048	RCP
-	20190406-040544-13-19-G358.93	0.02 tor	FREE	13-19	17	22238.8468594	22239.33514065	0.48828125	2.3841858E-4	22235.07985	2048	LCP
	20190419-025725-13-19-G358.93		FREE	13-19	17	22238.8468594	22239.33514065	0.48828125	2.3841858E-4	22235.07985	2048	RCP
	20190422-013655-13-19-Skydip.ta	ar	FREE	13-19	13	22207.841	22270.341	62.5	0.030517578	22235.07985	2048	LCP
	20190422-014516-13-19-3C286.ta	ar	FREE	13-19	13							
	20190422-015546-13-19-G358.93	3-0.03.tar	FREE	13-19	17	22207.841	22270.341	62.5	0.030517578	22235.07985	2048	RCP

An Italian radio Science Gateway

- Processing pipelines, quality metrics etc.
- Visualisation & graphical information
- Definition of additional metadata to describe the process (reproducibility)
- Build on SKA DC experience/infrastructure



Summary

- Data from INAF radio telescopes discoverable and accessible in a public, webbased archive.
- Almost ready to be discoverable in the VO: publishing coming soon
- Enhance interoperability: the VO RadioIG
- Further steps towards FAIRness
- Towards an infrastructure for discoverability and usability of data: first prototype of Italian Science Gateway