



The Italian radio data archive: *data content and provenance information*

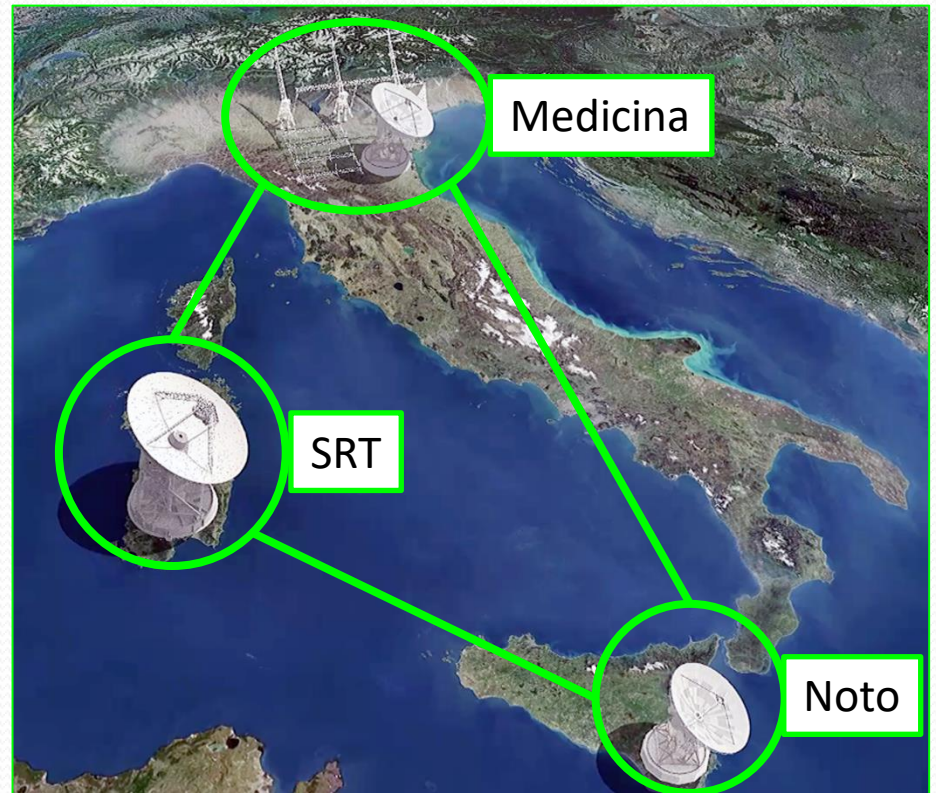
Alessandra Zanichelli

INAF – Istituto di Radioastronomia

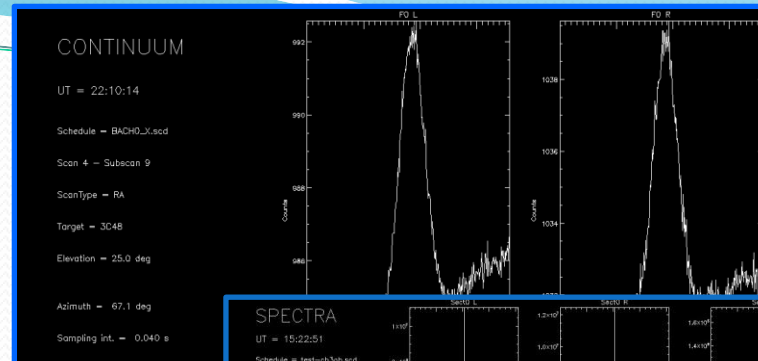
The INAF Radio Telescopes

Observing modes:

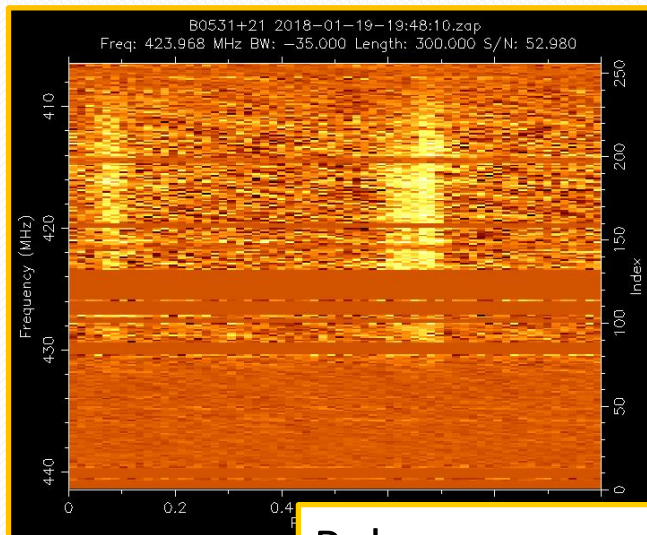
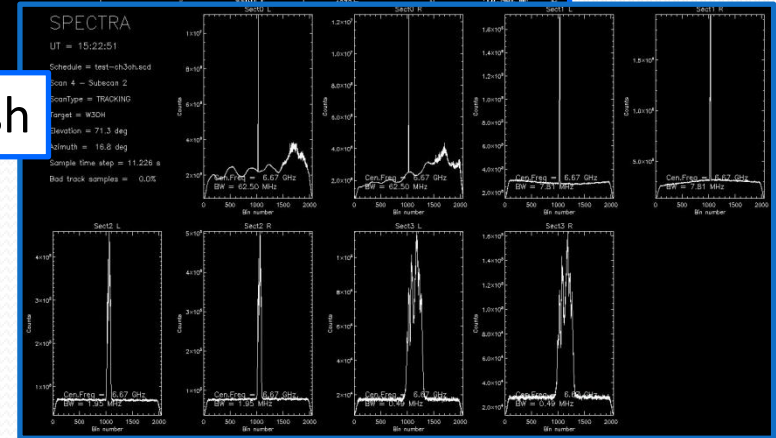
- Single-dish
- VLBI (EVN, IVS, EAVN)
- **VLBI-it**: *the «EVN lite» concept applied (+software correlator)*



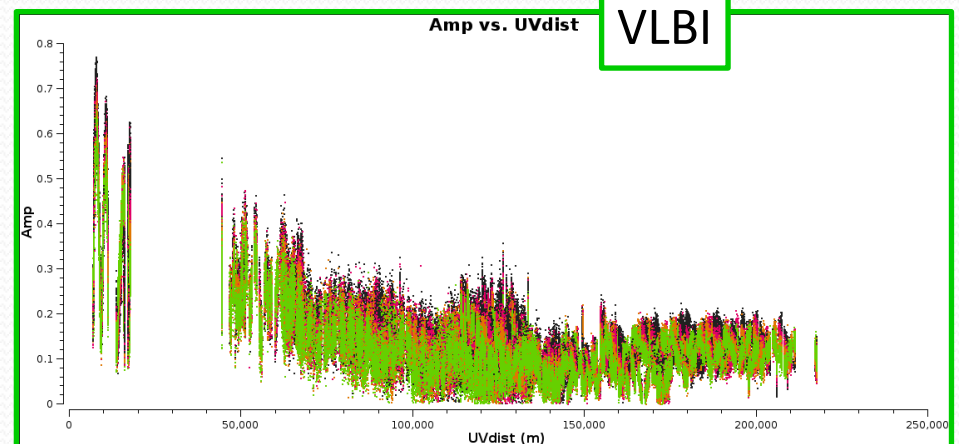
Different data types



Single Dish



Pulsar
Single Dish/VLBI



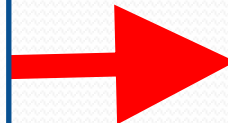
Data formats

MBFITS

```

alex@honeybee: /escs/mbfits/medicina/20121008/20121008-120852-140526-libardi_3C286$ ls -la
total 12
drwxr-xr-x 2 alex alex 4096 2013-02-21 13:01 .
drwxr-xr-x 2 alex alex 4096 2013-02-21 13:01 ..
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0002
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0003
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0004
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0005
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0006
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0007
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0008
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0009
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0010
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0011
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0012
-rw-r--r-- 2 alex alex 4096 2013-02-21 13:01 0013
-rw-r--r-- 1 alex alex 17280 2013-02-21 13:01 CCC-Backend-FEBEPAR.fits
-rw-r--r-- 1 alex alex 23040 2013-02-21 13:01 GROUPING.fits
-rw-r--r-- 1 alex alex 12096 2013-02-21 13:01 MONITOR.fits
  
```

The terminal window shows a directory listing for the 'mbfits' directory. It contains several FITS files, including '0002' through '0013', and three larger files: 'CCC-Backend-FEBEPAR.fits', 'GROUPING.fits', and 'MONITOR.fits'. A red box highlights the 'MBFITS' label.



Single-dish FITS

The screenshot shows a FITS file header and extension table. The header includes fields like 'PRIMARY HEADER', '[DATA]', 'EXTENSION HEADER', 'DATA', and 'EXTENSION 1'. The extension table lists various extensions and their dimensions.

Index	Extension	Type	Dimension
0	Primary	Image	
1	SECTION TABLE	Binary	5 cols X 2 rows
2	RF INPUTS	Binary	9 cols X 2 rows
3	FEED TABLE	Binary	4 cols X 1 rows
4	DATA TABLE	Binary	12 cols X 364 rows
5	ANTENNA TEMP TABLE	Binary	2 cols X 364 rows

The screenshot also shows a 'SUMMARY' window with fields like 'File', 'Edit', 'Tools', 'Index', 'Extension', 'Type', 'Dimension', and 'View'. It includes a table with columns for 'Header', 'Hist', 'Plot', 'All', and 'Select'.

The screenshot shows a FITS file header and extension table. The header includes fields like 'PRIMARY HEADER', '[DATA]', 'EXTENSION HEADER', 'DATA', and 'EXTENSION 1'. The extension table lists various extensions and their dimensions.

Index	Extension	Type	Dimension
0	Primary	Image	
1	ARRAY_GEOMETRY	Binary	7 cols X 2 rows
2	SOURCE	Binary	26 cols X 4 rows
3	ANTENNA	Binary	13 cols X 18 rows
4	FREQUENCY	Binary	6 cols X 1 rows
5	INTERFEROMETER_MODEL	Binary	20 cols X 210 rows
6	CALC	Binary	11 cols X 5 rows
7	MODEL_COMPS	Binary	21 cols X 210 rows
8	UV_DATA	Binary	13 cols X 8410 rows
9	SYSTEM_TEMPERATURE	Binary	10 cols X 0 rows
10	PHASE-CAL	Binary	17 cols X 378 rows

The screenshot also shows a 'SUMMARY' window with fields like 'File', 'Edit', 'Tools', 'Index', 'Extension', 'Type', 'Dimension', and 'View'. It includes a table with columns for 'Header', 'Hist', 'Plot', 'All', and 'Select'.

VLBI FITS (UVFITS)

Pulsar data (PSRFITS)

The INAF Radio Data Archive



In collaboration with the INAF Astronomical Archives infrastructure

- Continuum and spectropolarimetric raw data from single-dish, pulsar and VLBI observations
- Archived:
 - ✓ The scientific exposure with [data+instrument+site] metadata
 - ✓ The observing schedule and logfiles

(Future: storage of processed data; Science Gateway and User Space)

Web portal

New search Help Your files 0 Currently not logged in Login

Simple search **VLBI-IT search** SD search Pulsar search

File name

Name resolver Resolve

☐ RA ☐ Dec Radius (arcmin)

☒ Obs date From To

☒ Frequency [MHz] From To

☒ Project id

Telescope

☐ Antennas

☐ Data rate [Mbit/s]

☐ Spectral channels

☐ Spectral resolution [MHz (")]

Remote Authentication Portal

Login to Radio Archive

eduGAIN Use the eduGAIN Logo to Login or Register to the RAP facility if you belong to an eduGAIN idP.

Google f in Use these Logos to Login or Register to the RAP facility with your social identity

X.509 Use the X.509 Logo to Login with your personal certificate (GTF and TERENA-TACAR are allowed).

IA2 Use the IA2 Logo to Login if you have an account provided by IA2 or self registered

Need help? Please read our User guide and FAQ.

This software has been adapted by the IA2 team from the Remote Authentication Portal written by Franco Tinarelli at INAF-IRA.

Powered by IA2

Provenance

- Archived raw data must be discoverable and (re)usable
- Variety of observing projects and heterogeneity of the data
➡ accurate characterisation of the dataset is mandatory for the scientific exploitation of the Archive
- A «generic» Archive user (not the PI) must be able to address
 - ✓ if the data are suitable for her/his own research
 - ✓ if all the necessary information for data processing is available (e.g. calibration observations)

Provenance information

- Provenance DM not yet used but provenance information already stored for *raw data* and included in the Archive search parameters
- Metadata in the headers of archived files, for instance:
 - ✓ unique project ID, observer name, software version
 - ✓ telescope/instrument, etc.
- Ancillary information univocally associated to the dataset, for instance:
 - ✓ weather parameters
 - ✓ receiver performance, etc.

Provenance info

- Provenance DM not stored for *raw data* a

- Metadata

✓ unique

✓ telesc

- Ancillary in

✓ weather

✓ receive

```

PROJECT:          GZT05C
OBSERVER:         Scicom
SCANLIST:         GZT05Cextrasho
PROCEDURELIST:    GZT05Cextrasho
BACKENDLIST:      GZT05Cextrasho
MODE:             SEQ
SCANTAG:          1
INITPROC:         INIT

SC: 1 0006+397-A1
1_1 0.000000 1
1_2 12.000000 2
1_3 12.000000 3
1_4 12.000000 4
1_5 12.000000 5
1_6 12.000000 2
1_7 12.000000 3
1_8 12.000000 4
1_9 12.000000 5

SC: 2 0006+397-A2
2_1 0.000000 1
2_2 12.000000 2
2_3 12.000000 3
2_4 12.000000 4
2_5 12.000000 5
2_6 12.000000 2
2_7 12.000000 3
2_8 12.000000 4
2_9 12.000000 5

SC: 3 0006+397-A3
3_1 0.000000 1
3_2 12.000000 2
3_3 12.000000 3

```



1313f.xml + (~/.jescs/vlbit/vlbit)

File Edit Tools Syntax Buffers

File Edit Tools Syntax Buffers

```
<?xml version='1.0' encoding='UTF-8'>
<database>
  <table>
    <PROJID>1313f</PROJID>
    <TELESCOP>VLBIT</TELESCOP>
    <PNAME>Marcello Giroletti</PNAME>
    <OBJECT>0212+735</OBJECT>
    <RA>30. 2918926036</RA>
    <DEC>73. 8257282611</DEC>
    <START>2014-03-28 14:19:43</START>
    <ANTENNAS>Nt- Md- Mc</ANTENNAS>
    <FREQ>4966. 49</FREQ>
    <OBS_MODE>16x8MHz</OBS_MODE>
    <RATE>512. 0</RATE>
    <ONTIME>4430</ONTIME>
  </table>
  <table>
    <PROJID>1313f</PROJID>
    <TELESCOP>VLBIT</TELESCOP>
    <PNAME>Marcello Giroletti</PNAME>
    <OBJECT>0340+362</OBJECT>
    <RA>45. 7247090019</RA>
    <DEC>36. 370119325</DEC>
    <START>2014-03-28 14:19:43</START>
    <ANTENNAS>Nt- Md- Mc</ANTENNAS>
    <FREQ>4966. 49</FREQ>
    <OBS_MODE>16x8MHz</OBS_MODE>
    <RATE>512. 0</RATE>
    <ONTIME>4430</ONTIME>
  </table>
</database>
</xml>
```

-- INSERT --

File Edit Tools Help

Search for: Find Case sensitive? No

```

SIMPLE      =                               T / file does conform to FITS standard
BITPIX      =                               8 / number of bits per data pixel
NAXIS       =                               0 / number of data axes
EXTEND      =                               T / FITS dataset may contain extensions
COMMENT     FITS (Flexible Image Transport System) format is defined in 'Astronomy
COMMENT     and Astrophysics', volume 376, page 359; bibcode: 2001AS...376..359H
HIERARCH    Observation Date   = '2015-09-17T13:11:15' / file creation date (YYYY-
COMMENT     V 1.1 Created by   S. Righini, M. Bartolini & A. Orlati
HISTORY     V 0.8 First output standard for Italian radiotelescopes
HISTORY     V 0.82 The tsys column in data table replaced with the Tant table, it re
HISTORY     ports the tsys measurement for each input of each section
HISTORY     V 0.9 The section table has been splitted into two tables: sections and
HISTORY     rf inputs table
HISTORY     V 0.91 Added the flux column in section table
HISTORY     V 0.92 SubScanType added as primary header keyword
HISTORY     V 1.0 Added new table to store position of subreflector e primary focus
HISTORY     receivers: SERVO TABLE
HISTORY     V 1.01 New keywords in FEED TABLE header to describe derotator configura
HISTORY     tion
HISTORY     V 1.1 Summary.fits file included in order to describe the scan configura
HISTORY     tion
HIERARCH    BackendName = 'XARCOS' / Backend name
CREATOR     = 'ESCS v0.5' / Software (incl. version)
HIERARCH    Declination = 1.07990841507474 / Target declination (radians)
EQUINOX     = 2000 / Equinox of RA, Dec
EXPTIME     = 120. / Total integration time (seconds)
FITSVER     = 'V.1.1' / FITS version
LST         = '14:15:16.7' / Local sidereal time
HIERARCH    LogFileName = 'Xarcos-Test-w3oh_20150917_131115.log' / Name of the log
HIERARCH    NUSEBANDS = 3 / Number of sections
OBJECT      = 'w3oh' / Target source name
OBSID       = 'Moscadelli' / Observer or operator initials
PROJID      = '15-01' / ProjectID
CHAN3       = 2048
CHAN2       = 1024
CHAN1       = 256
BWD2        = 450.0
BWD3        = 300.0
BWD1        = 600.0
FREQ3       = 22150.0
FREQ2       = 22100.0
FREQRES3    = 146484.0
FREQRES2    = 439453.0
FREQRES1    = 2343750.0
FREQ1       = 22000.0
HIERARCH    RESTFREQ1 = 22235.07985 / Rest frequency (MHz)
HIERARCH    RESTFREQ2 = 22235.07985 / Rest frequency (MHz)
HIERARCH    RESTFREQ3 = 22235.07985 / Rest frequency (MHz)
HIERARCH    ReceiverCode = 'KKC' / Receiver name
HIERARCH    RightAscension = 0.641706660521798 / Target right ascension (radians)
HIERARCH    SCANSTART = 0.0 / Scan starting position (deg)
  
```

Provenance info

- Provenance DM not stored for *raw data* a

- Metadata

✓ unique

✓ telescope

- Ancillary info

✓ weather

✓ receiver

```
PROJECT: GZT05C
OBSERVER: SCICOMM
SCANLIST: GZT05C_extraso
PROCEDURELIST: GZT05C_extraso
BACKENDLIST: GZT05C_extraso
MODE: SEQ
SCANTAG: 1
INITPROC: INIT
```

```
SC: 1 0006+397-A1
1.1 0.000000 1
1.2 12.000000 2
1.3 12.000000 3
1.4 12.000000 4
1.5 12.000000 5
1.6 12.000000 2
1.7 12.000000 3
1.8 12.000000 4
1.9 12.000000 5
```

```
SC: 2 0006+397-A2
2.1 0.000000 1
2.2 12.000000 2
2.3 12.000000 3
2.4 12.000000 4
2.5 12.000000 5
2.6 12.000000 2
2.7 12.000000 3
2.8 12.000000 4
2.9 12.000000 5
```

```
SC: 3 0006+397-A3
3.1 0.000000 1
3.2 12.000000 2
3.3 12.000000 3
```

```
<?xml version='1.0' encoding='UTF-8'>
<database>
  <table>
    <PROJID>1313f</PROJID>
    <TELESCOP>VLBIT</TELESCOP>
    <PINAME>Marcello Giroletti</PINAME>
    <OBJECT>0212+735</OBJECT>
    <RA>30.2918926036</RA>
    <DEC>73.8257282611</DEC>
    <START>2014-03-28 14:19:43</START>
    <ANTENNAS>Nt-Md-Mc</ANTENNAS>
    <FREQ>4966.49</FREQ>
    <OBS_MODE>16x8MHz</OBS_MODE>
    <RATE>512.0</RATE>
    <ONTIME>4430</ONTIME>
  </table>
  <table>
    <PROJID>1313f</PROJID>
    <TELESCOP>VLBIT</TELESCOP>
    <PINAME>Marcello Giroletti</PINAME>
    <OBJECT>0340+362</OBJECT>
    <RA>45.7247090019</RA>
    <DEC>36.370119325</DEC>
    <START>2014-03-28 14:19:43</START>
    <ANTENNAS>Nt-Md-Mc</ANTENNAS>
    <FREQ>4966.49</FREQ>
    <OBS_MODE>16x8MHz</OBS_MODE>
    <RATE>512.0</RATE>
    <ONTIME>4430</ONTIME>
  </table>
  <table>
    <PROJID>1313f</PROJID>
  </table>
-- INSERT --
```

```
File Edit Tools Search for: Find Case sensitive? No Help

SIMPLE = T / file does conform to FITS standard
BITPIX = 8 / number of bits per data pixel
NAXIS = 0 / number of data axes
EXTEND = T / FITS dataset may contain extensions
COMMENT FITS (Flexible Image Transport System) format is defined in 'Astronomy
COMMENT and Astrophysics', volume 376, page 359; bibcode: 2001A&A...376..359H
HIERARCH Observation Date = '2015-09-17T13:11:15' / file creation date (YYYY-
COMMENT V. 1.1 Created by S. Righini, M. Bartolini & A. Orlati
HISTORY V. 0.8 First output standard for Italian radiotelescopes
HISTORY V. 0.82 The tsys column in data table replaced with the Tant table, it re
HISTORY ports the tsys measurement for each input of each section
HISTORY V. 0.9 The section table has been splitted into two tables: sections and
HISTORY rf inputs table
HISTORY V. 0.91 Added the flux column in section table
HISTORY V. 0.92 SubScanType added as primary header keyword
HISTORY V. 1.0 Added new table to store position of subreflector e primary focus
HISTORY receivers: SERV0 TABLE
HISTORY V. 1.01 New keywords in FEED TABLE header to describe derotator configura
HISTORY tion
HISTORY V. 1.1 Summary.fits file included in order to describe the scan configura
HISTORY tion
HIERARCH BackendName = 'XARCOS' / Backend name
CREATOR = 'ESCS v0.5' / Software (incl. version)
HIERARCH Declination = 1.07990841507474 / Target declination (radians)
EQUINOX = 2000 / Equinox of RA, Dec
EXPTIME = 120. / Total integration time (seconds)
FITSVER = 'V. 1.1' / FITS version
LOCALTIME = '14:15:16.7' / Local sidereal time
HIERARCH LogFileName = 'Xarcos-Test-w3oh_20150917_131115.log' / Name of the log
HIERARCH NITERANDS = 3 / Number of sections
OBJECT = 'w3oh' / Target source name
OBSID = 'Moscadelli' / Observer or operator initials
PROJECT = '1313f' / ProjectID
CHAN3 = 2048
CHAN2 = 1024
CHAN1 = 256
BWID2 = 450.0
BWID3 = 300.0
BWID1 = 600.0
FREQ3 = 22150.0
FREQ2 = 22100.0
FREQRES3 = 146484.0
FREQRES2 = 439453.0
FREQRES1 = 2343750.0
FREQ1 = 22000.0
HIERARCH RESTFREQ1 = 22235.07985 / Rest frequency (MHz)
HIERARCH RESTFREQ2 = 22235.07985 / Rest frequency (MHz)
HIERARCH RESTFREQ3 = 22235.07985 / Rest frequency (MHz)
HIERARCH ReceiverCode = 'KKC' / Receiver name
HIERARCH RightAscension = 0.641706660521798 / Target right ascension (radians)
HIERARCH SCANSTART = 0 0 / Scan starting position (deg)
```


Provenance info

- Provenance DM not stored for *raw data* a

- Metadata

✓ unique

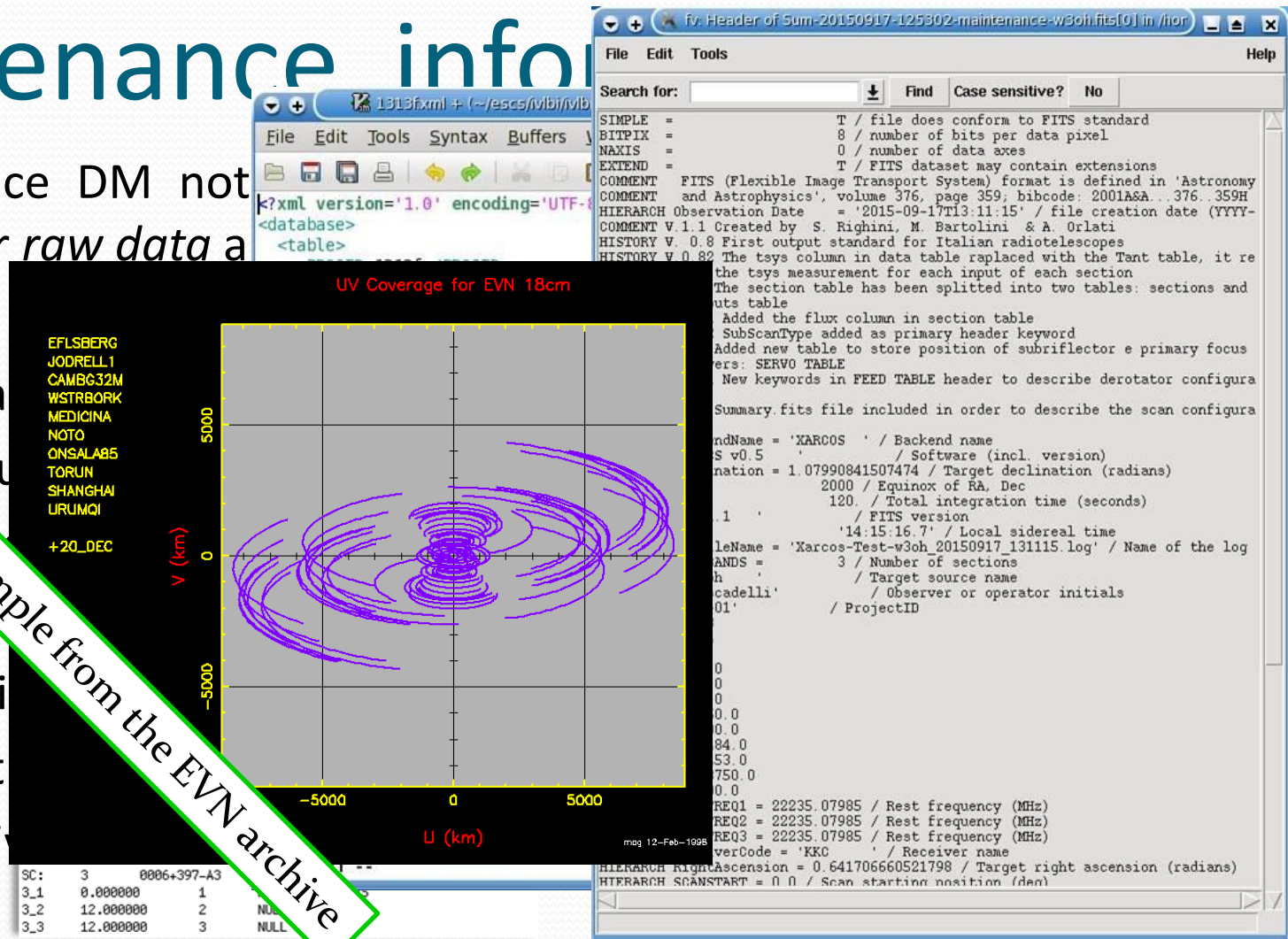
✓

- Ancillary i

✓ weat

✓ receiv

Example from the EVN archive



- (Graphical: instrument characterisation like UV coverage)

Provenance information (next)

- For *processed data*, provenance will include information necessary to fully describe the reduction history of the original dataset
- Processing pipelines, calibration and processing information as well as some level of quality metrics to be provided.
- Graphical information to visually describe specific characteristics
- Additional metadata to describe the reduction process with focus on the calibration steps (like RFI and atmospheric opacity removal)

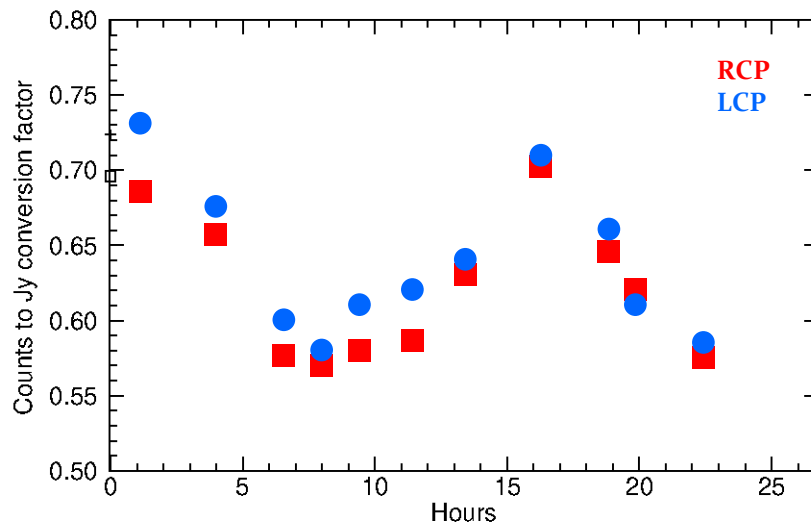
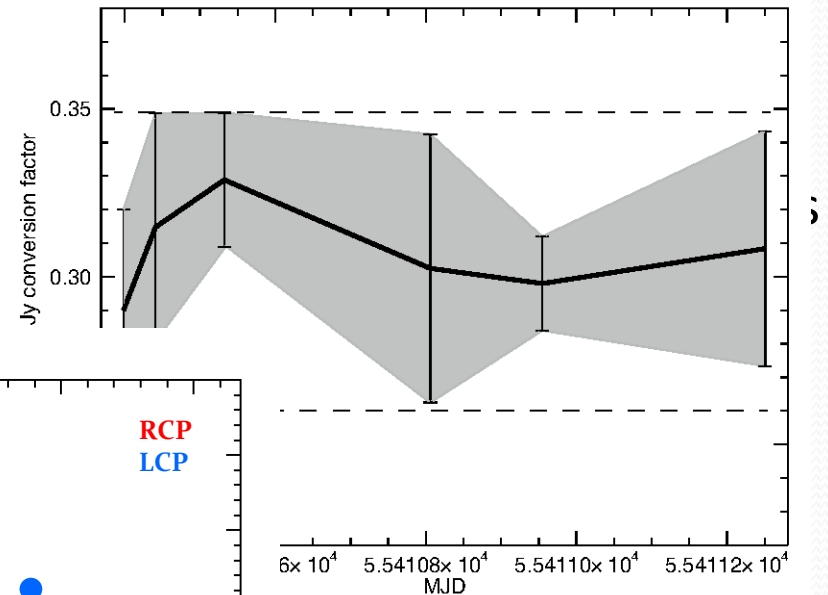
Provenance information (next)

- For *processed data*, provenance will include information necessary to fully describe the reduction history

- Processing pipelines, calibration and some level of quality metrics to be provided

- Graphical interface

- Additional information in the calibration



frequency removal)

Summary and future steps

- Different data types (and products) in the Archive
- Provenance information already present in raw data
- Raw datasets characterised by metadata and ancillary information
- *Archive processed data products and processing information. Effective quality metrics.*
- *IVOA Provenance DM for the interoperability of data.*
- *Level of granularity for provenance: critical for reproducibility*