HPSDPC @ Wits















The Wits Astro Data Center



Sergio Colafrancesco

Wits University - DST/NRF SKA Research Chair

Email: Sergio.Colafrancesco@wits.ac.za

Outline

- ◆ Astronomy & Astro-Particle Physics at Wits
 The challenge of Multi-v and Multi-disciplinary
- ♦ Science with Multi-v and Multi-disciplinary
- Data with Multi-v and Multi-experiments
- The Wits Astro Data Center

- The Wits approach to Data science
- The future of the WADC
 - SA SA
 - Africa
 - ... and beyond

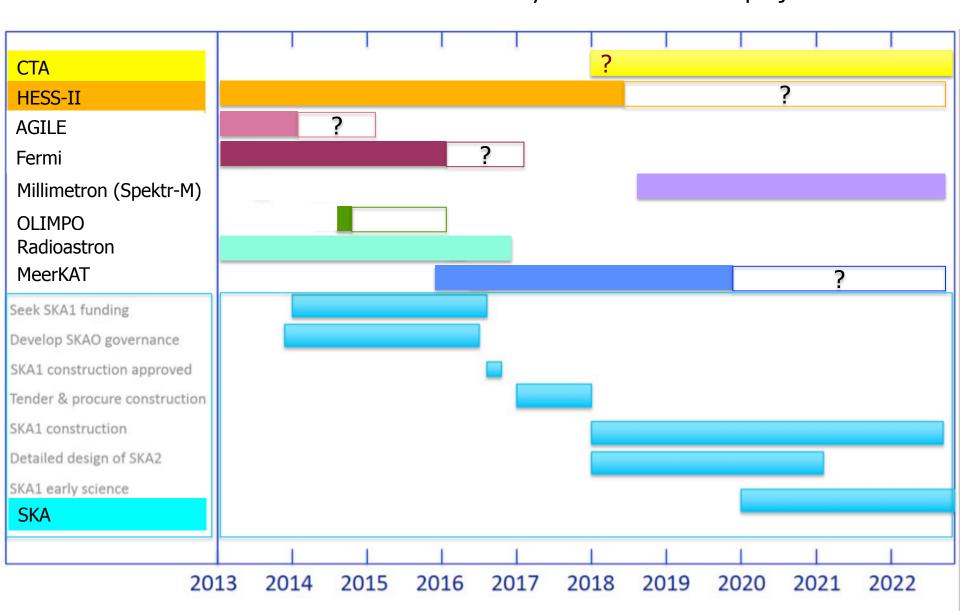
Outline

- ◆ Astronomy & Astro-Particle Physics at Wits
 The challenge of Multi-v and Multi-disciplinary
- ♦ Science with Multi-v and Multi-disciplinary
- Data with Multi-v and Multi-experiments
- The Wits Astro Data Center

- The Wits approach to Data science
- The future of the WADC
 - SA SA
 - Africa
 - ... and beyond

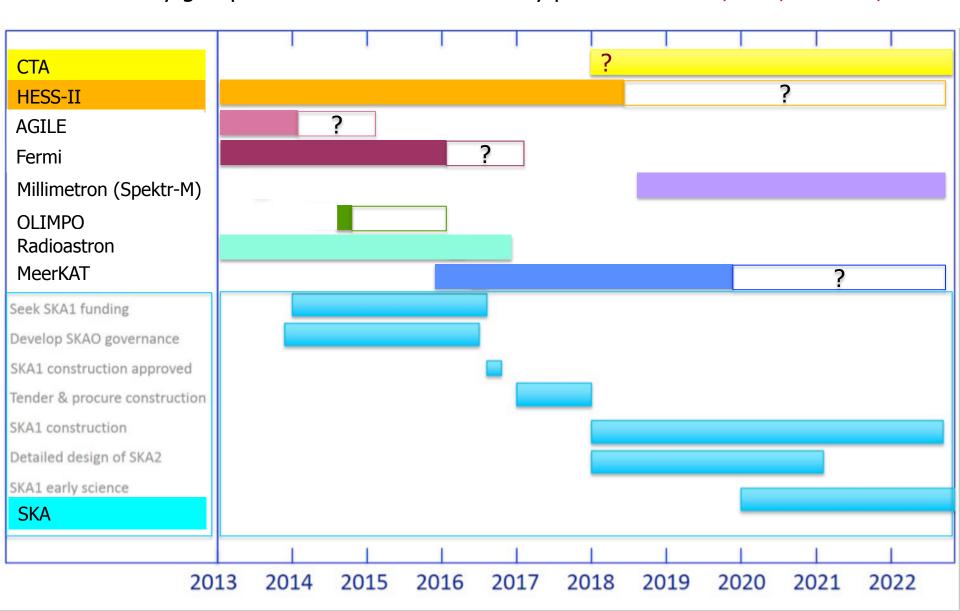
Astronomy & Astro-Particle Physics @ Wits

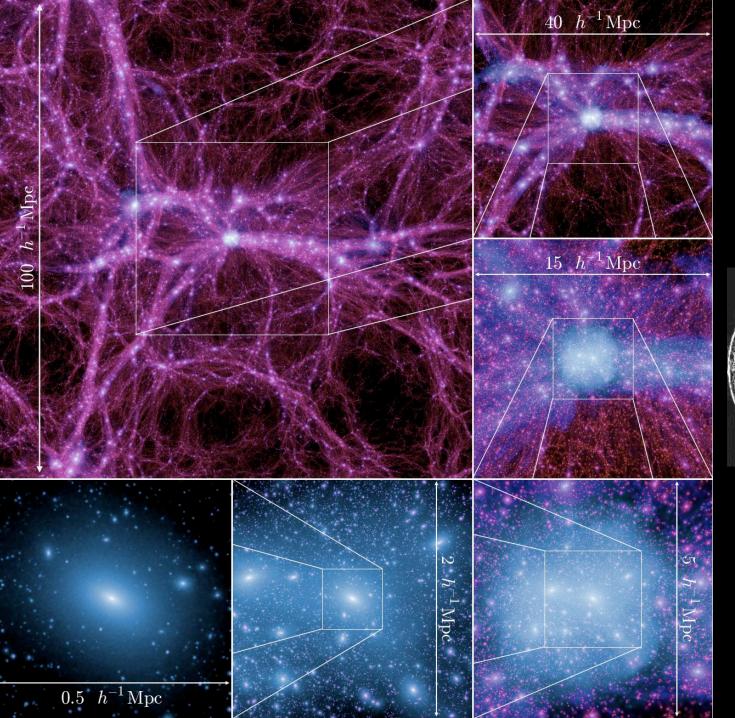
Wits official involvement in Astronomy and Astro-Particle projects



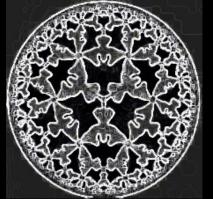
Astronomy & Astro-Particle Physics @ Wits

Wits is the only group in the World that is officially part of MeerKAT, SKA, HESS-II, CTA

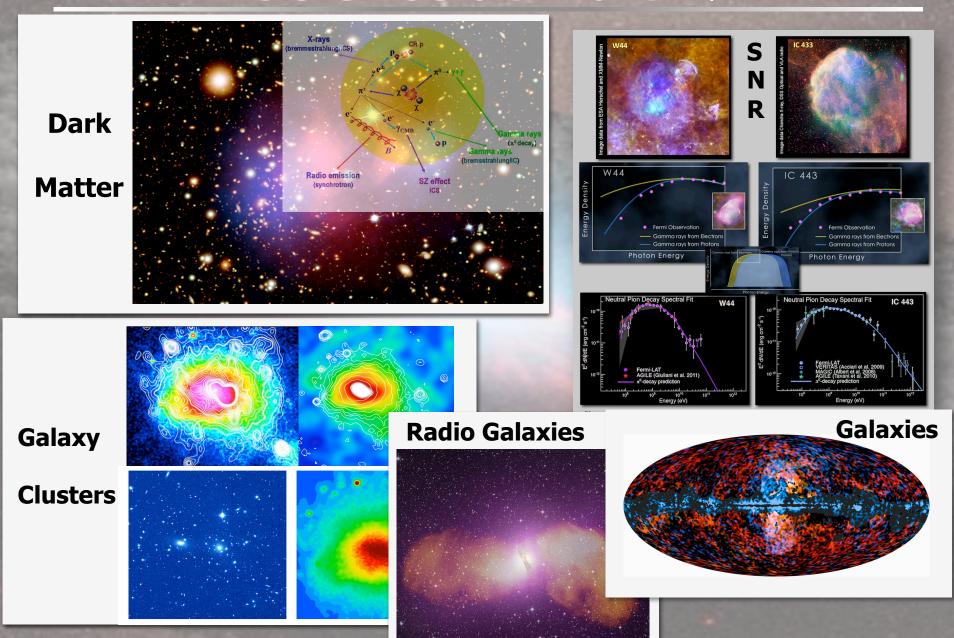




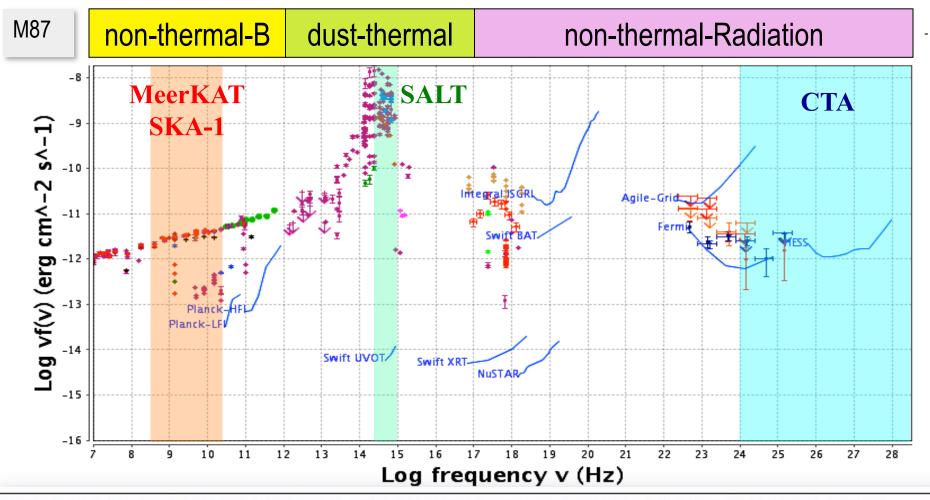
Science at Multi Scales



Science at Multi-v



Science at Multi-experiments



• DIXON • FIRST • KUEHR • NVSS • PKSCAT90 • VLSS • NED • IRASFSC 12 • IRASFSC 60 • IRASPSC 60 • ERCSC030 • ERCSC044 • ERCSC070
• ERCSC100 • ERCSC143 • ERCSC217 • ERCSC353 • ERCSC545 • GB6 • NORTH20CM (flux 20 cm) • NORTH20CM (flux 6 cm) • PCCS1F030 • PCCS1F044
• PCCS1F070 • PCCS1F100 • PCCS1F143 • PCCS1F217 • PCCS1F353 • PCCS1F345 • PCCS1F857 • WMAP5 (Freq. 23e9 Hz) • WMAP5 (Freq. 33e9 Hz)
• WMAP5 (Freq. 41e9 Hz) • WMAP5 (Freq. 61e9 Hz) • WMAP5 (Freq. 94e9 Hz) • ARIEL3A • IPC • UHURU4 • MAXIGSC • Fermi1FGL (200 MeV)
• Fermi1FGL (2GeV) • Fermi1FGL (600 MeV) • Fermi2FGL (200 MeV) • Fermi2FGL (2GeV) • Fermi2FGL (600 MeV) • Fermi2FGL

The problem(s)

Next generation Astronomical telescopes will all be Digital Instruments

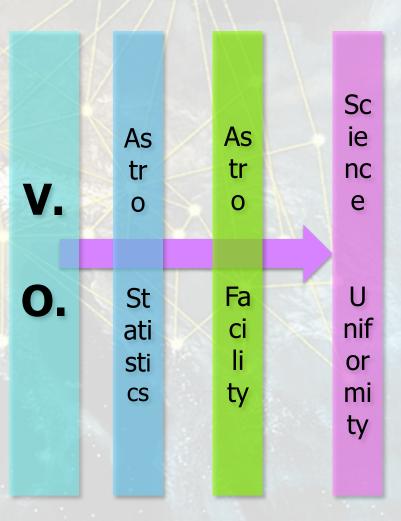
A change of approach is required:

- At the Telescope
 - Modes of scheduling, observing, calibrating, etc...
- At the Data
 - Modes of data flow, analysis, archiving, interacting, etc...

A generic approach

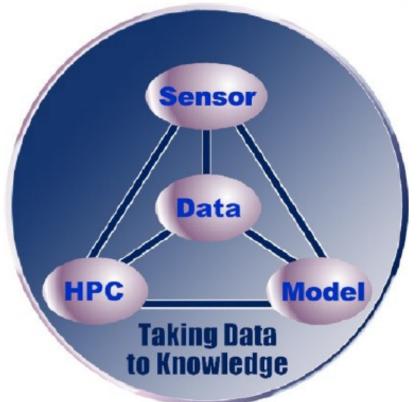
In ≈2030 a typical day for an astronomer on a university campus will have her start work by looking at:

- A list of science papers that have been intelligently selected by a software tool that surfed the Internet overnight.
- She clicks on an object in an online science paper and the Virtual Observatory (V.O.) database delivers views in γ-ray, X-ray, Optical, InfraRed and Radio observations.
- She queries the archive to do an automatic search to pull up information relevant to the questions the she asking about the object.



Astro Information Science

The V.O. and other astronomical community structures (i.e., ADASS, WGAS, and FITS WG) have been essential but incomplete steps toward an Astronomical Info/Data Science.



Need data-intensive **science research** tools to mine & discover **new knowledge** from the distributed data repositories & experiments.

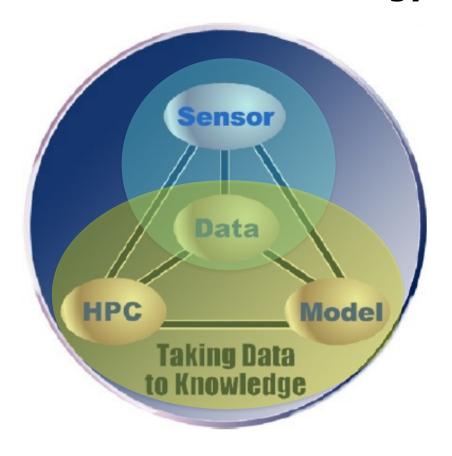
Outline

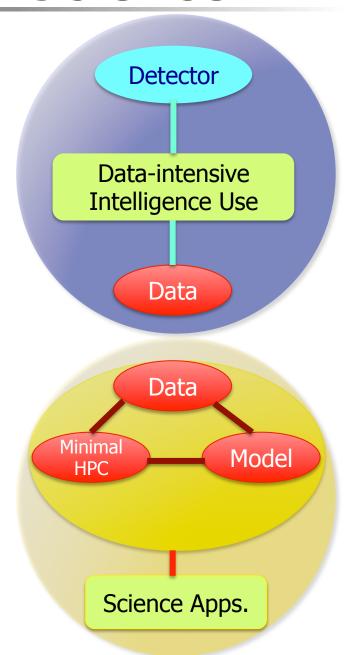
- ◆ Astronomy & Astro-Particle Physics at Wits
 The challenge of Multi-v and Multi-disciplinary
- Science with Multi-v and Multi-disciplinary
- Data with Multi-v and Multi-experiments
- The Wits Astro Data Center

- The Wits approach to Data science
- The future of the WADC
 - SA SA
 - Africa
 - ... and beyond

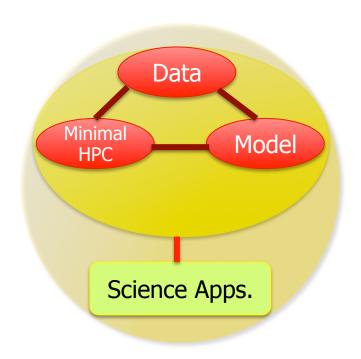
Wits Astro Data Science

A breakdown of specific tasks in a modulated strategy





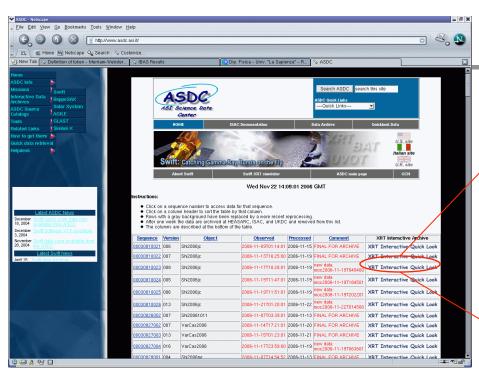
Research-oriented Data Center

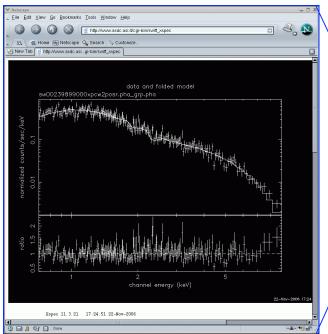


The Wits Astro Data Center

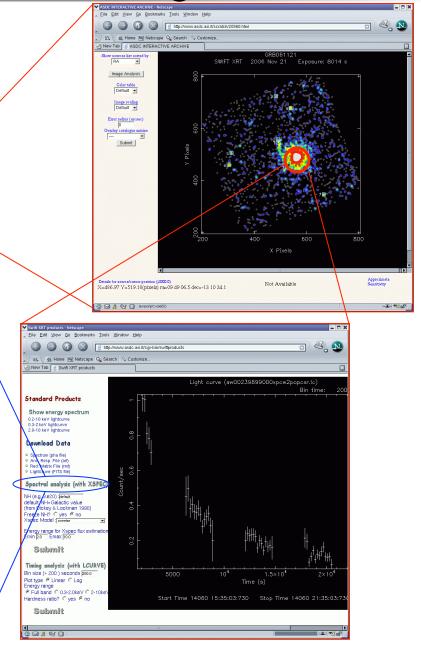
The WADC aims at conducting research in Astro Data Applications and make data and Apps available to the world-wide community through a Multi-Experiment Interactive (hands-on) Archive Example from a similar Data Center that I managed (2007-2011)



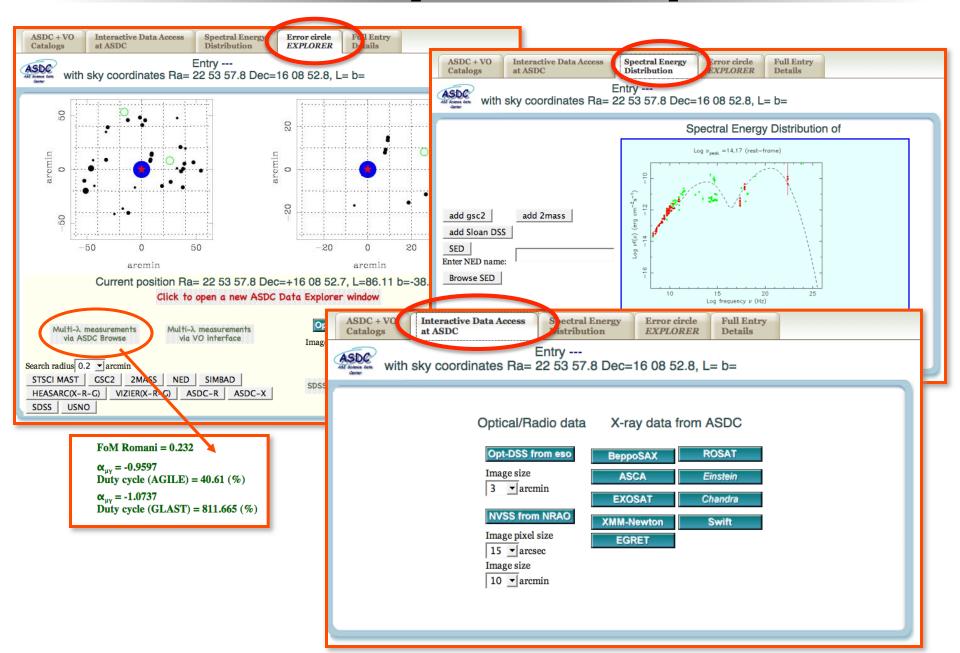




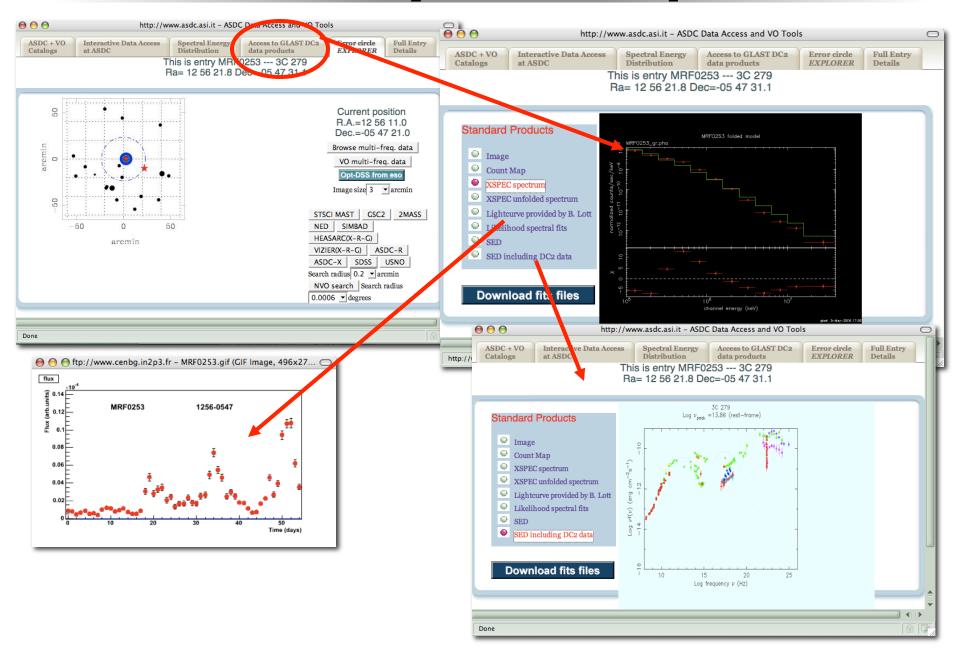
Surfing the WADC



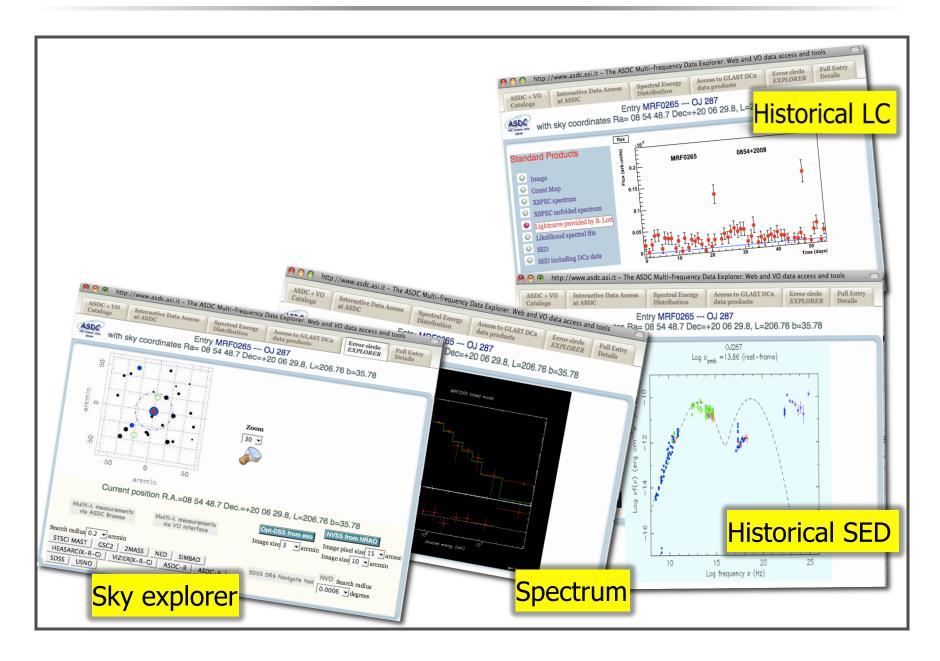
WADC Sky-Data Explorer



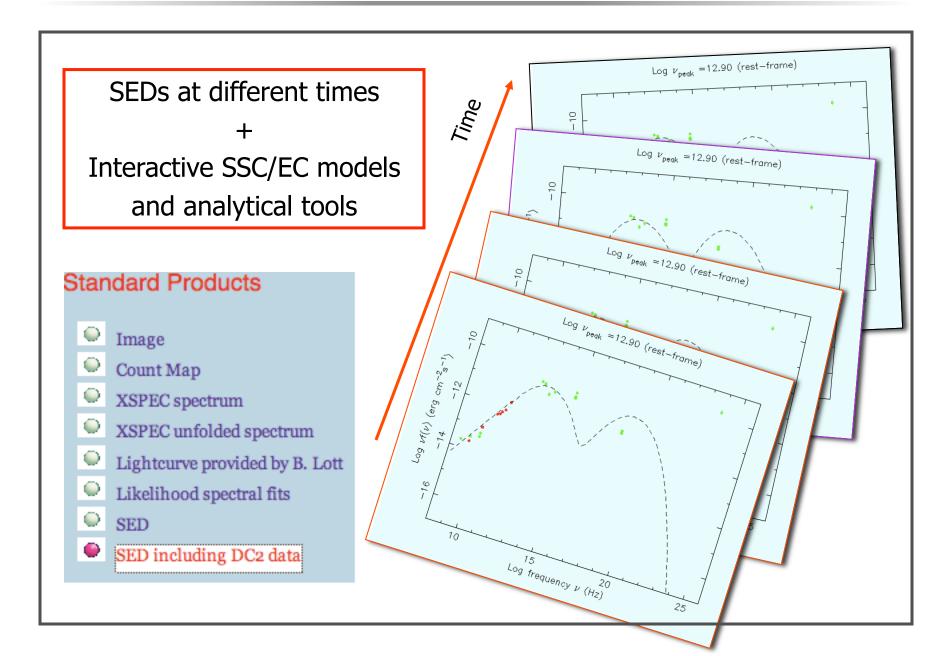
WADC Sky-Data Explorer



WADC Multi-v Tools



WADC Multi-v Tools



WADC Catalogs Tools

The 3-year WMAP Point Sources Catalog

@ ASDC

Help
Show/hide columns
Advanced filtering
Print current view of table
Print complete table
Reset all filters

P.Giommi, S. Colafrancesco, P. Padovani, D. Gasparrini, E. Cavazzuti and S. Cutini A&A, 508, 107, 2009 (based on G. Hinshaw, et al., 2007, ApJS, 170, 288)

Complete sample [f(41GHz) > 1 Jy] Full sample

Export Current view of Table in: Latex format FITS format Raw text format CSV text format

Previous Page Next Page Page Size (# of lines) 50

Refresh page Reset all filters Show all entries

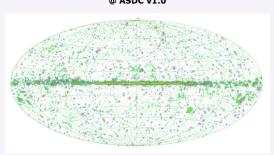
	Entry number		Source name	RA (J2000.0)	Dec (J2000.0)	LII	BII	Redshift	classification	Rflux(5GHz)	Rflux(41GHz)	
Life y maniber			IAU naming convention \$	hh mm ss.d ‡	dd mm ss.d ‡		511	Reasilit	Browse Classif ‡	mJy ‡	mJy ‡	
Selection mode: Include All			1	1	1 U	1	1	↑ • Stats	1	↑ V Stats	↑ Stats	
	1 🗹 Select	ASDC Data Explorer	WMAP3 J0006-0622	00 06 13.8	-06 23 35.0	93.5	-66.65	0.347	Blazar Q-type	2463	2300	
	2 🗹 Select	ASDC Data Explorer	WMAP3 J0012-3953	00 12 59.8	-39 54 24.9	332.46	-74.95	0	Blazar B-type	921	1000	
	3 🗹 Select	ASDC Data Explorer	WMAP3 J0019+2019	00 19 37.7	+20 21 45.0	112.89	-41.89	0	Blazar B-type	710	900	
	4 🗹 Select	ASDC Data Explorer	WMAP3 J0019+2602	00 19 39.1	+26 02 44.9	114.06	-36.28	0.284	Blazar Q-type	435	900	
	5 🗸 Select	ASDC Data Explorer	WMAP3 J0025-2604	00 25 49.0	-26 02 12.9	42.29	-84.17	0.322	Misd. NT-AGN RGal-type	3745	0	
	6 ✓ Select	ASDC Data Explorer	WMAP3 J0026-3511	00 26 16.3	-35 12 48.9	335	-80.3	0	Unidentified	121	1200	
	7 🗹 Select	ASDC Data Explorer	WMAP3 J0029+0555	00 29 45.7	+05 54 41.0	113.12	-56.55	1.317	Blazar Q-type	348	1200	
	8 🗹 Select	ASDC Data Explorer	WMAP3 J0042+5208	00 43 08.7	+52 03 33.9	121.63	-10.79	0.174	Misd. NT-AGN RGal-type	4124	0	
	9 🗹 Select	ASDC Data Explorer	WMAP3 J0047-2515	00 47 33.1	-25 17 17.0	97.39	-87.97	0.001	Starburst galaxy	2433	1100	
	10 🗹 Select	ASDC Data Explorer	WMAP3 J0049-5740	00 49 59.4	-57 38 26.9	303.31	-59.49	1.797	Blazar Q-type	1338	1000	
	11 🗸 Select	ASDC Data Explorer	WMAP3 J0051-0927	00 50 41.2	-09 29 04.9	122.32	-72.36	0.537	Blazar B-type	931	800	
	12 🗹 Select	ASDC Data Explorer	WMAP3 J0050-0647	00 51 08.2	-06 50 02.0	122.71	-69.71	1.975	Blazar Q-type	841	800	
	13 🗸 Select	ASDC Data Explorer	WMAP3 10050-4223	00 51 09.4	-42 26 33.0	303.12	-74.69	1.749	Blazar O-type	926	900	

ASI Science Data Center
VO mode: off (turn on)
Cone Search Source Name
Resolve name
RA, Dec O L,B Clean
(e.g. 00 02 34.6,-53 01 10.2 or 0.64417,-53.0195)
radius 5 \$ arcmin Search Reset filter

WADC Catalogs Tools

Planck Catalogue of Compact Sources @ ASDC v1.0







30GHz 44GHz 70GHz 100GHz 143GHz 217GHz 353GHz 545GHz 857GHz

This is an interactive version of the First Planck Compact Source Catalogue, released by ESA on March 21th, 2013.

This catalogue is also available from ESA Planck Legacy Archive and IPAC website.

A general description of the PCCS is available from The Planck Collaboration, 2013, while more details can be found in the Explanatory Supplement.

Flux density has been estimated using four methods. The appropriate photometry to be used depends on the nature of the source considered.

See The Planck Collaboration, 2013, for recommendations about photometry.

Export Current view of Table in: Latex format FITS format Raw text format CSV text format

Previous Page Next Page Page Size (# of lines) 50 \$ Refresh page Reset all filters Show all entries

Entry number		Source name	RA (J2000.0) hh mm ss.d ‡	Dec (J2000.0) dd mm ss.d ‡	Detection pipeline flux density +/- err (mJy) Flux density +/- err ‡	Extended Extended source? \$ 1	ERCSC source name
Selection mode: Include All		1	1	T.	1	1	1 U
1 🗸 Select	ASDC Data Explorer	PCCS1 030 G118.20+04.89	00 02 11.8	+67 18 28.0	2.1e+4+/-483.6	No	PLCKERC030 G118.11+04.94
2 🗸 Select	ASDC Data Explorer	PCCS1 030 G118.62+06.06	00 04 14.5	+68 32 11.0	5587.8+/-510.5	No	
3 ✓ Select	ASDC Data Explorer	PCCS1 030 G084.71-71.14	00 04 15.4	-11 51 20.0	736.1+/-111.6	No	PLCKERC030 G084.65-71.12
4 🗹 Select	ASDC Data Explorer	PCCS1 030 G323.93-67.58	00 04 40.0	-47 35 38.0	456.2+/-100.1	No	
5 🗹 Select	ASDC Data Explorer	PCCS1 030 G108.31-41.20	00 04 41.5	+20 20 26.9	454.3+/-103.5	No	

WADC Catalogs Tools

The TeGeV Catalogue @ ASDC (v1)

Authors: A. Carosi, F. Lucarelli, L.A. Antonelli, M. Capalbi, C. Leto, F. Verrecchia



Help

Show/hide columns

Advanced filtering Print current view of table

Print complete table

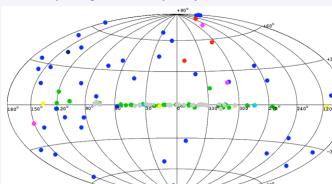
Reset all filters







The TeGeV Catalogue @ ASDC is a catalogue of VHE sources observed by ground-based Cherenkov telescopes. The TeGeVcat is collecting all the information publicly available about TeV sources observed by the past generation and current generation of imaging Cherenkov telescopes. The catalog contains also the public light curves and spectra (soon available in the v2 version).





















Export Current view of Table in: Latex format | FITS format | Raw text format | CSV text format

Previous Page Next Page Page Size (# of lines) 50 Carefresh page Reset all filters Show all entries

Entry	number		TEV NAME	# OBS.	OTHER NAMES	TYPE	RA (J2000) hh mm ss ‡	Dec (J2000) dd mm ss ‡	x E-12 cm2/s ‡	THR ENERGY [GeV]	Distance z ‡	Observatory	START TIME [MJD]	LIGHT CURVE
- Ir	ion mode: nclude All		1	1		1	↑ •	1	↑ V Stats	↑ V Stats	↑ U Stats	1	↑ V Stats	↑ Stats
1 6	Select	ASDC Data Explorer	TeV J0025+6410	1	TychoSNR/VERJ0025+641	SNR	00 25 27.0	+64 10 50	0.187	1000	-	VERITAS	54740	-
2 6	Select	ASDC Data Explorer	TeV J0240+6115	6 }	LSI+61_303	XRB	02 40 34	+61 15 25	2.2	400	-	MAGIC	53644	Show
3 6	Select	ASDC Data Explorer	TeV J0534+2200	6 ▶	Crab/G184.6-5.8/3C144/SN1054	PWN	05 34 31.1	+22 00 52	-	25	-	MAGIC	54374	-
4 6	Select	ASDC Data Explorer	TeV J0616+2230	2 }	IC443/SNR189.1+03.0/MAGIC0616+223	SNR	06 16 51	+22 30 11	4.63	300	-	VERITAS	54132	-
5 6	Select	ASDC Data Explorer	TeV J0632+1722	1	MGROJ0632+17/GemingaPWN	PWN	06 32 28	+17 22 00	-	35000	-	MILAGRO	-	-
6	Select	ASDC Data Explorer	TeV J0632+0548	4 >	HESSJ0632+057	XRB	06 32 58.3	+05 48 20	-	200	-	MAGIC	55599	-

WADC theory apps.

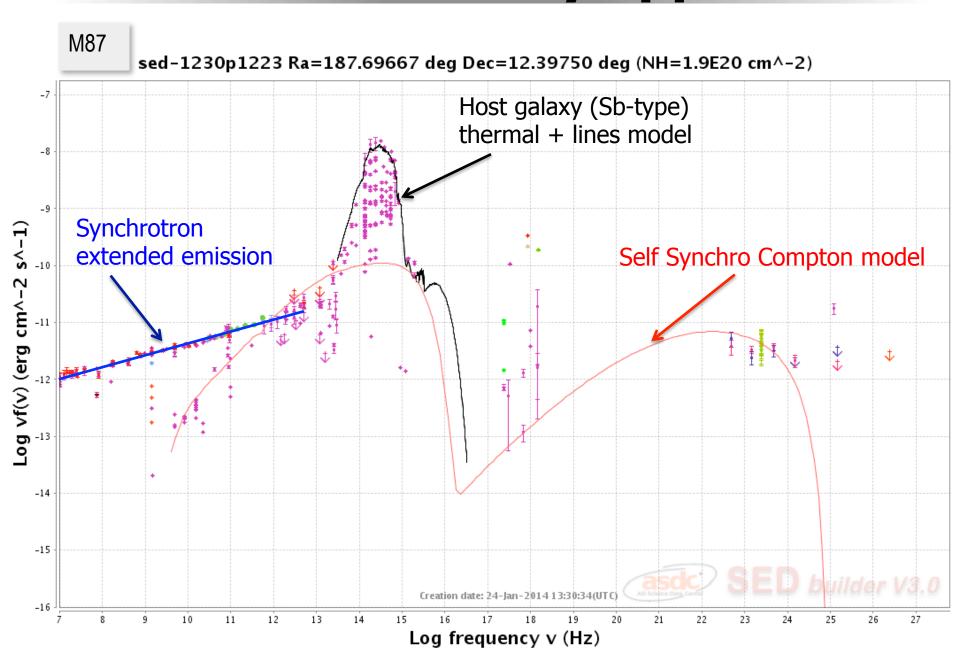


Search radius 0.2 arcmin

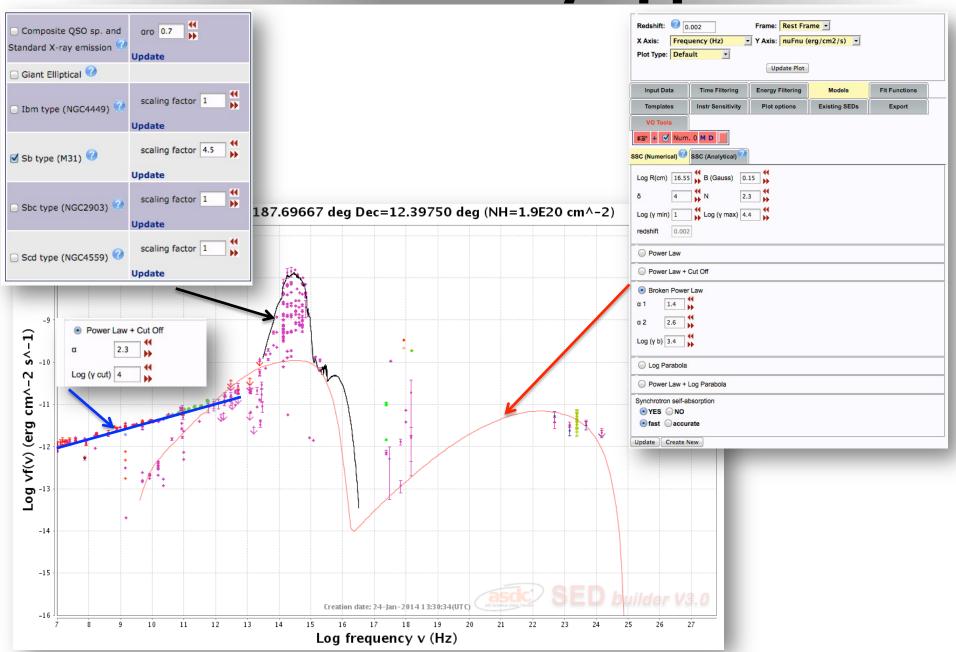
WADC: theory apps.



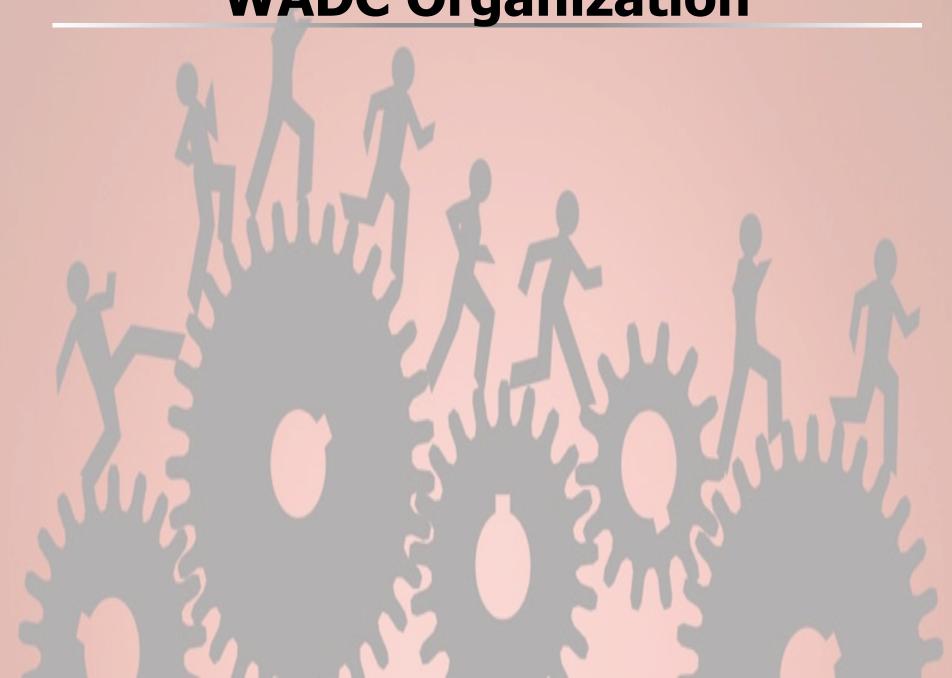
WADC theory apps.



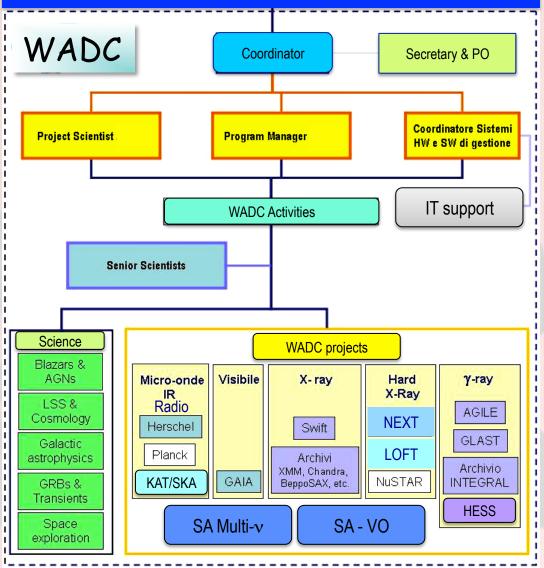
WADC theory apps.

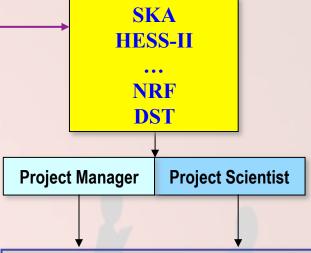


WADC Organization



Wits Astro Group





- Coordination of scientific activities
- Favor collaboration between WADC and the scientific community
- Collaboration with African countries
- Collaboration extra-Africa

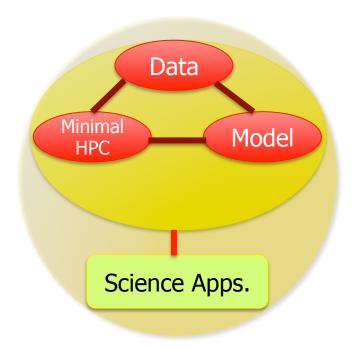
Outline

- ◆ Astronomy & Astro-Particle Physics at Wits
 The challenge of Multi-v and Multi-disciplinary
- Science with Multi-v and Multi-disciplinary
- Data with Multi-v and Multi-experiments
- The Wits Astro Data Center

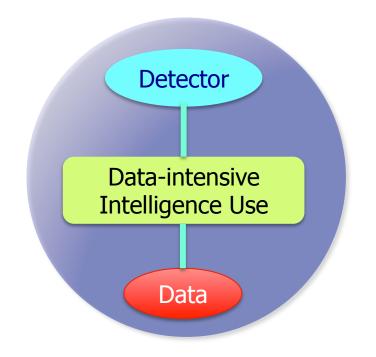
- The Wits approach to Data science
- The future of the WADC
 - SA SA
 - Africa
 - ... and beyond

Wits Astro Data approach

Fast, Interactive, Hands-on Research-oriented data management



Data-intensive, Intelligence-use Decision-making software Powerful electronics



WADC

SKA SA-CERN

Opportunities

WADC: (ex)portable, research-oriented, multi-experiment, multi-v, Data Center

Scientific Research

Science Apps

MeerKAT / SKA
HESS-II/CTA
Spektr-M, OLIMPO,
Survey telescope fast analysis

AstroInfo Research

Innovation

Flexibility

Scientific return

HESS-SA data center (@ Wits)

Scientific Research

Needs HPC+WADC

An example: LSST survey

Data rate: 30 terabytes of data obtained each night

- \rightarrow \approx 10⁷ sources/night
- → ≈2 10⁸ sources/month
- → $\approx 2 \cdot 10^9$ sources/year

Assume:

 \approx 10% of these sources are unknown (need rapid identification)

We will have:

 $\approx 2\ 10^7$ sources/month available for rapid multi-v identification/study (possible only with fast data analysis tools like at the WADC)

A Digital Copy of the Universe, Encrypted

As physics prepares for ambitious projects like the Large Synoptic Survey Telescope, the field is seeking new methods of data-driven discovery.



WADCS: specs, costs, upgrade

Present day specifications for the existing structure of the WADC

Specs

1 INTEL SERVER 4 Intel Worker Nodes		150 381.96 340 795.00
Synology NAS 40TB Storage Solution	R	95 056.62
UPS 24 Port Ethernet Switch 1Gb	R R	6 039.72 5 318.00
42U Rack	R	8 076.90
Total cost	R	605 668.20

Planned upgrade for the WADC (April-Dec 2014)

Upgrades

Total cost	R	536 056.40
42U Rack	R	8 076.90
Synology NAS 40TB Storage Solution	R	95 056.62
INTEL XEON SERVER	R	159 615.96
INTEL SERVER NODES	R	368 267.88

WADC full operational mode

Modes

- Multi-frequency, Multi-experiment Data Center
- HESS-II data pipeline (pre-)analysis