



# Database Knowledge and the autoMAGIC Database

## autoMAGIC Workshop Padova

Jan Lukas Schubert

28<sup>th</sup> - 30<sup>th</sup> August 2024





- Database: Basic concept
- Database relationships
- The autoMAGIC database structure:
  - Organization of the Data
  - Organization of the MCs
  - Organization of the Jobs
  - Examples for relationships
- Basic and advanced database commands







#### **Databases – Basic Concept**

- Tables storing information
  - > Columns: parameters
  - Rows: Entries
- Each row is assigned a unique id
- Tables are linked to each other with relationships





#### **Databases - Relationships**

 One-to-One relationship: One object of a DB table can only be linked to one object of another table (uncommon in practice!)







### **Databases - Relationships**

• One-to-Many relationship: One object can be linked to many objects of another table







### **Databases - Relationships**

 Many-to-Many relationship: Many objects of one table can be linked to many objects of another table







### **Databases - Unique Constraints**

- DBs tables can use unique constraints to prevent doubling of the entries
- Example: superstar jobs
  - Unique combination of mars\_version and run\_number





#### **The autoMAGIC database Structure**

alembic_version		coach_types		corsika_versions		exit_codes		hv_settings		H_tables		13_tables		mars_versions		jobs_flute_mataju_exclude		jobs_sorcerer		jobs_selectmc_mataju_e	exclude_pixels
version_num V/	RCHAR(32)	id INTEGER		M INTEGER		id INTEGE	R	id INTEGER		H INTEGER		id INTE	SER	м	INTEGER	id	INTEGER	id	INTEGER	ы	INTEGER
		name VARCHAR		name VARCHAR		mars_executable VARCH4	R	name VARCHAR		name VARCHAR		name VARC	GAR	name	VARCHAR	start_date	DATETIME	run_id	INTEGER	monte_carlo_set_id	INTEGER
						code INTEGE	я							release_date	DATETIME	stop_date	DATETIME	output,M1	VARCHAR	output_file	VARCHAR
						description VARCHA	R							tarball	VARCHAR	coach_job_id	INTEGER	output_M2	WARCHAR	mars_version_id	INTEGER
																badmoness or	ROAT	mara_version_id	INTEGER	state_id	INTEGER
monte_carlo_production	ins .	monte_carlo_triggers		sources		states		telescopes		zenith_ranges		analysis_periods		jobs_add_data		theta2 out	FLOAT	exit code sorrerer M1 M	INTEGER	exit_code_id	INTEGER
ы п	NTEGER	id INTEGER		d INTEGER		id INTEGER		id INTEGE	R	id INTEGER		id	INTEGER	id .	INTEGER	badranness cut from efficiency	ROOLFAN	exit code sorcerer M2 id	INTEGER	unescamp	LATE THE
name V	ARCHAR	name VARCHAR		name VARCHAR		name VARCHAR		name VARCHA	R	name VAROIAR		name	VARCHAR	source_id	INTEGER	theta2_cut_from_efficiency	BOOLEAN	timestamp	DATETIME	parrang_o	
path V	ARCHAR			ra FLOAT				nom_hv_min FLOA	π	ad_min VARO4AR		description	VARCHAR	date	DATETIME	quantile_hadronness_cut	FLOAT				
identifier V	ARCHAR			dec FLOAT				nom_hv_max FLOA	π	zd_max VAROIAR		mc_production_id	INTEGER	max_subruns_number	INTEGER	quantile_theta2_cut	FLOAT	di3_obs_melibea_mataju_	exclude_associations		
								red_hv_min FLOA	ar .			start_date	DATETIME	mars_version_id	INTEGER -	binning_id	INTEGER	job_dl3_converter_id	INTEGER	Jobs_menbea	Umacaju
			_					red_hv_max FLOA	π			stop_date	DATETIME	state_id	INTEGER	run_fold	BOOLEAN	job_obs_melibea_id	INTEGER	Ha Ha	INTEGER
moon_conditions		observations		jobs_selectmc		monte_carlo_runs	_			coach_train_files	_			timestamp	DATETIME	mars_version_id	INTEGER			cuperity job i	M INTEGER
id IN	TEGER	id INT	reger	Id INTE	GER	id INTEG	ER			job_coach_id INTEGE	R					state_id	INTEGER			mara version in	M INTEGER
name vou	TECHNIC	source_id INI	TEGER	monte_cano_sec.id INTE	GER CER	calibrated_M1 W000H	AR			jab selectine id initiali	K					exit_code_id	INTEGER	dl3 obs mellbea matalu s	tandard associations	state.id	INTEGER
de max	TEGER	duration in a		and the second s	UAP.	moste cado set id invitio										timestamp	DATETIME	job dtl converter id	INTEGER	exit_code_id	INTEGER
d bd	FLOAT	mola duration in a	TOAT	outrue file 1 VARC	140	number VADO	MP.											job obs melibea id	INTEGER	timestamp	DATETIME
d.N2	FLOAT	mola_significance_le	FLOAT	output file 2 VARC	HAR	first_calibrated VARCH	AR									jobs_flute_mataju_standard					
noise_lv_mean	FLOAT	mola_significance_he	FLOAT	mars_version_id INTE	GER	last_calibrated VARCH	AR	coach_corsika_association	ins	view_cones		jobs_mc_mellbea		jobs_star_supersta		id li	INTEGER				
noise_lv_rms	FLOAT			state,id INTE	GER			job_coach_id	INTEGER	H INTEGER		id	INTEGER	id.	INTEGER	start_date	DATETIME	dl3_mc_mellbea_standar	d_associations	lobs mc supersta	ar mately standard
hv_settings_id IN	TEGER			exit_code_id INTE	GER			corsika_version_id	INTEGER	name VARCHAR		coach_job_id	INTEGER	runjid	INTEGER	stop_date	DATETIME	job_dl3_converter_id	INTEGER	id	INTEGER
version VA	RCHAR			timestamp DATET	TIME					rumin FLOAT		selectric_job_id	INTEGER	rv(Jb	FLOAT	coach_job_id	INTEGER	job_mc_melibea_standard	M INTEGER	is_test_set	BOOLEAN
										r_max FLOAT		mars_version_id	INTEGER	cl_W2	FLOAT	hadronness_cut	FLOAT			monte_carlo_run_id	INTEGER
												state_id	INTEGER	noise_lv_mean	FLOAT	theta2_cut	FLOAT			output_file	VARCHAR
			_	the second s						No. of Concession, Name		exit_code_id	INTEGER	noise_lv_rms	FLOAT	hadronness_cut_from_embency	IDOLLAN			mars_version_id	INTEGER
coach_off_files		jobs_melibea		selectmc_inputs		flute_mc_melibea_association		flute_obs_mellbea_assoc	lations	subruns		timestamp	DATETIME	mars_version_id	INTEGER	meta2_cut_from_embercy	BOOLDAN	dl3_mc_melibea_exclude	associations	state_id	INTEGER
job_coach_id	INTEGER	id INTE	GER	job_selectmc_id IN	ITEGER	job_flute_id INT	EGER	job_flute_id	INTEGER	id .	INTEGER			output_file	VARCHAR	mantle theta2 cut	EGAT	job_di3_converter_id	INTEGER	exit_code_id	INTEGER
projsuperstarja	INTEGER	concepto de la	CER.	project superstantio	ITEGER	Justin Charles Con	EVEN	jou, ous, menuea, o	INTEGER	calibrated file M2	MARCHAR			state_id	INTEGER	binning id	INTEGER	pro, mc, meioes, excluse, i	INTEGER.	timestamp	DATETIME
		soperate joojo inter	OER.							out M	INTEGER			exit_code_ittar_M1_id	INTEGER	run fold	BOOLEAN				
		state id INTER	GER							subsus number	VARCHAR			exit_code_star_M2_id	INTEGER	mars version id	INTEGER				
		exit_code_id INTE	GER							24L M1	FLOAT			Enclose Supersurja	DATETIME	state_id	INTEGER	flute_obs_melibea_mataj	u_exclude_associations		
irf_types		timestamp DATET	IME	dl3_obs_mellbea_associations		coach_source_associations		runs		2d_M2	FLOAT	jobs_coach				exit_code_id	INTEGER	job she melihes id	INTEG		
M INTEGER				job_dl3_converter_id IN	ITEGER	job_coach_id INTEGER		м	INTEGER	az_M1	FLOAT	id	INTEGER			timestamp	DATETIME	project interests			
name VARDIAR				job_obs_melibea_id IN	TEGER	source_id INTEGES	·	observation_id	INTEGER	az_M2	RLOAT	production_id	INTEGER								
								run_number	VARCHAR	L3t_rate_M1	FLOAT	zenith_id	INTEGER					coach_off_files_mataju		flute_mc_melibea_excl	lude_associations
								wobble_offset	FLOAT	L31,rate_M2	FLOAT	trigger_id	INTEGER			jobs_dl3_converter_mataju_excl	ude	job_ceach_id	INTEGER	job_flute_mataju_id	INTEGR
								wobble_angle	FLOAT	cloudiness	FLOAT	moon_condition_id	INTEGER			м	INTEGER	job_superstar_id	INTEGER	job,mc,melibea,mataju	exclude_id INTEGE
								L1_table_id	INTEGER	transmission3km_mean	FLOAT	mars_version_id	INTEGER			coach_job_id	INTEGER				
Jobs_mc_superscar		di3_mc_melloea_association	na	selecome_imputs_mataju_star	veare	pointing		L3_table_id	INTEGER	transmission6km_mean	FLOAT	state_/d	INTEGER	monte_cario_sets		irf_type_id	INTEGER	John my melling man	and a shade		
No.	INTEGER	job_dtl_converter_id	INTEGER	jab enectine mataju_standard	A INTEGER	INTEGER		zd	FLOAT	transmission5km_mean	PLOAT	exit_code_RFgh_id	INTEGER	ad earth	INTEGER	hadronness_out	FLOAT	jest inclinent a intra	INTEGER		
mono condition id	INTEGER	loning interesting	ITTE I	juojin juojin juoji juoj		wohble offset INTIGER		az	FLOAT	transmission12km_mean	FLOAT	exit_code_cols_id	INTEGER	production id	INTEGER	eneral_cor	ROAT	coards inde int	INTEGER		
outrad file	VADCHAR					wohble ande INTEGER		La cate Ma	ROAT	transmission/skim_stid	FLOAT	encrue in aspire	DATETALS	weth M	INTEGER		ROIT	selectric job exclude pixel	s id INTEGER		
mars version id	INTEGER							chudiness	ROAT	domean_M1	FLOAT	view cone id	INTEGER	corsika version id	INTEGER	hadrospar of for efficiency	BOOLEAN	pointing id	INTEGER		
state_id	INTEGER							transmission3km mean	FLOAT	ocmean_M2	ROM	source set	INTEGER	num_files	INTEGER	theta2 cut from efficiency	BOOLEAN	mars_version_id	INTEGER		
exit_code_id	INTEGER	selectmc_inputs_mataju_ex	clude_pixels	coach_source_associations_m	sataju	coach_train_files_mataju		transmission6km,mean	FLOAT	meanity M2	ROAT	off_mc_events_ratio	FLOAT	trigger_id	INTEGER	mars, version, id	INTEGER	state_id	INTEGER		
timestamp	DATETIME	job_selectmc_mataju_exclude	pixels_id INTEGER	job_coach_id	INTEGER	job_coach_id INTEGE	C	transmission9km_mean	FLOAT	obs start M1	DATETIME	underpopulated bins	INTEGER	num_events_zenith_bi	n INTEGER	state_id	INTEGER	exit_code_id	INTEGER		
		job_mc_superstar_mataju_exi	ude_pixels_id INTEGER	source_id	INTEGER	job_selectmc_id INTEGE	i	transmission12km_mean	FLOAT	obs_start_M2	DATETIME			view_cone_id	INTEGER	exit_code_id	INTEGER	timestamp	DATETIME		
								dcmean_M1	FLOAT	obs_stop_M1	DATETIME					timestamp	DATETIME				
								domean_M2	FLOAT	obs_stop_M2	DATETIME					pointing_id	INTEGER				
								duration_in_s	FLOAT	duration_in_s_M1	FLOAT							juusjincjineliueajina	aju_stantuaru		
labe much motols		much comits accordations	matala	John Bute		John dill comunities		number_events	INTEGER	duration_in_s_M2	FLOAT			Maxim Buts		jobs_dl3_converter_mataju_star	Idard	ia internet internet	INTEGER		
joba jeonen jina capa	Differe	concritor and annocations	Unitaria	jobs_note	NUTLEY A	jour_uns_converter	INTERES	c_p_id	INTEGER	number_starguider_reports_M1	INTEGER	C points	DODCOD.	contrag note		м	INTEGER	colorigación estadarent la	d INTEGER		
and the second s	INTEGER	por coacte of	INTEGER	attant data	DATETRA	search table of	INTEGER			number_starguider_reports_M2	INTEGER	a cur cumber	VIDCHAR			coach_job_id	INTEGER	mara version id	INTEGER		
production (no	INTEGER	Conna_rennen_o	Integer	stop data	DATETIME	inf turns id	INTEGER			number_identified_stars_M1	FLOAT	c nin number	VADCHAR	in the second		irf_type_id	INTEGER	state id	INTEGER		
view cone ld	INTEGER			coach iob id	INTEGER	mars version id	INTEGER			number_identified_stars_M2	FLOAT	c file M1	VARCHAR			hadronness_cut	FLOAT	exit_code_id	INTEGER		
trigger_id	INTEGER			mars_version_id	INTEGER	state_id	INTEGER			number_events	INTEGER	cifie.M2	VARCHAR			theta2_cut	FLOAT	timestamp	DATETIME		
source_set	INTEGER			state_id	INTEGER	exit_code_id	INTEGER			raw,file_M1	VARCHAR	p_file_M1	VARCHAR			quantile_hadronness_cut	FLOAT				
underpopulated_bins	INTEGER	jobs_star_superstar_mataju		exit_code_id	INTEGER	timestamp	DATETIME	jobs_mc_superstar_mata	ju_exclude_pixels	raw_file_M2	WARCHAR	p_file_M2	VARCHAR	jobs_selectmc_mat	iju_standard	quantile_theta2_cut	FLOAT				
off_mc_events_ratio	FLOAT	Id	INTEGER	simestamp	DATETIME	hadronness_cut	FLOAT	Id	INTEGER			observation_id	INTEGER	M	INTEGER	hadronness_cut_from_efficiency	BOOLEAN				
mars_version_id	INTEGER	mars_version_id	INTEGER	binning_id	INTEGER	hadronness_out_from_efficiency	BOOLEAN	monte_carlo_run_id	INTEGER	flute_mc_melibea_standard assoc	ciations			monte_carlo_set_id	INTEGER	theta2_cut_from_efficiency	BOOLEAN				
state_id	INTEGER	output_file	VARCHAR	run_fold	BOOLEAN	quantile_hadronness_cut	FLOAT	output_fle	VARCHAR	job_flute_mataju_id	INTEGER			output_file	VARCHAR	mars_version_id	INTEGER				
exit_code_RFgh_id	INTEGER	state_id	INTEGER	hadronness_cut	FLOAT	quantile_theta2_cut	FLOAT	mars_version_id	INTEGER	job_mc_melibea_mataju_standard_id	INTEGER			mars_version_id	INTEGER	state_id	INTEGER				
exit_code_LUTs_id	INTEGER	exit_code_star_M1_id	INTEGER	hadronness_out_from_efficiency	BOOLEAN	theta2_out	FLOAT	state_id	INTEGER			flute_obs_mellbea_mat	aju_standard_associations	state_id	INTEGER	exit_code_id	INTEGER				
exit_code_Rfdisp_id	INTEGER	evit_code_star_M2_id	INTEGER	quantile_hadronness_cut	FLOAT	theta2_cut_from_efficiency	BOOLEAN	exit_code_id	INTEGER			job_flute_id	INTEG	ER exit_code_id	INTEGER	umestamp	SWITETING				
timestamp	DATETIME	exit_code_superstar_id	INTEGER	quantile_theta2_cut	FLOAT			timestamp	DATETIME			job_obs_melibea_id	INTEG	ER timestamp	DATETIME						
		omestamp	DATETIME	theta2_cut	FLOAT			pointing_id	NATEGER					is_test_set	BUOLEAN						
		LOSS BOOCHER UN		where a state from a state of a state	The factor of the set																





#### **The autoMAGIC database Structure**

- Many tables are just for preventing writing too much redundant information into the DB
- Example: job states
  - Writing the string "created", "queued", etc. into the corresponding DB column is possible, but redundant
  - Better: just link the objects in the states table by their id
  - Instead of writing redundant informations, we just store one integer in the column and by the DB structure (relationships!), the information can be retrieved





#### **The autoMAGIC Database Structure**

Table Observation														
id	source	date	significance											
1	Crab Nebula	2020-05-30	15											
2	NGC 1275	2019-01-31	0.5											
3	Crab Nebula	2011-07-04	8											
4	NGC 1275	2016-08-11	1.2											
5	Perseus-MA	2017-06-23	2.9											

				Tab	ole Run	$\checkmark$			
	id	run number	zenith	azimuth	transmission	DC	observation id		
	1	500001	25	118	0.85	1244	1		
	2	500002	26	117	0.85	1240	1		
7	3	500003	27	116	0.83	1155	1		
	4	541241	48	50	0.65	508	4		
	5	555621	7	152	0.92	256	3		

Table SubRun									
id	subrun number	calibrated file M1	calibrated file M2	run id					
1	004	500001.004_M1.root	500001.004_M2.root	1					
2	051	500001.057_M1.root	500001.057_M2.root	1					
3	006	500001.006_M1.root	500001.006_M2.root	1					
4	122	524528.122_M1.root	524528.122_M2.root	7					
5	089	597952.098_M1.root	597952.098_M2.root	9					

Jan Lukas Schubert

autoMAGIC Workshop Padova 2024





#### **The autoMAGIC Database Structure**

Table Analysis Period												
id	name	start	stop	production id								
1	ST0314	2020-05-14	2020-09-14	1								
2	ST0313	2020-02-26	2020-03-13	2								
3	ST0307	2017-11-10	2018-06-29	3								
4	ST0307	2016-04-29	2017-08-02	3								
5	ST0306	2014-11-24	2016-04-28	5								

Γ	Table Monte Carlo Production												
	id	name	path	identifier									
5	1	ST0314	/pnfs/path/to/ST0314	M1_AD6.0_MF0.53_M2_AD4.5_MF0.65									
	2	ST0313	/pnfs/path/to/ST0313	M1_AD6.0_MF0.53_M2_AD4.5_MF0.65									
	3	ST0307	/pnfs/path/to/ST0307	M1_AD3.5_MF0.68_M2_AD5.0_MF0.71									
7	4	ST0310	/pnfs/path/to/ST0310	M1_AD3.5_MF0.612_M2_AD5.0_MF0.675									
	5	ST0306	/pnfs/path/to/ST0306	M1_AD5.5_MF0.63_M2_AD5.5_MF0.69									

				Table M	lonte Carlo S	et	$\checkmark$
	id	trigger	zenith	corsika	view cone	number files	production id
<b>&gt;</b>	1	standard	low	mmcs699	diffuse 1.5	5000	1
	2	standard	medium	mmcs699	diffuse 1.5	5000	1
	3	standard	high	mmcs699	diffuse 1.5	5000	1
	4	standard	low	mmcs6500	ringwobble	3999	5
	5	sumT	medium	mmcs699	ringwobble	4000	3
	6	EGAL sumT	high	mmcs699	diffuse 2.5	4999	6

${ m Table}$ Monte Carlo Run											
id	number	calibrated file M1	calibrated file M2	number events	MC set id						
1	854654	GA_854654_M1.root	GA_854654_M2.root	25417	1						
2	875429	GA_875429_M1.root	GA_875429_M2.root	14554	1						
3	845701	GA_845701_M1.root	GA_845701_M2.root	42542	1						
4	831477	GA_831477_M1.root	GA_831477_M2.root	15497	7						
5	842215	GA_842215_M1.root	GA_842215_M2.root	25344	9						

autoMAGIC Workshop Padova 2024

#### Jan Lukas Schubert



Jan Lukas Schubert



#### **The autoMAGIC Database Structure**

id MC prod.			d.	moon condition				zenith view		cone		exit	code	id	sta	te i	d	
	1	ST	0307	)307 no moon					low	ring	wobb	ole		1			5	_
7	2	ST	0312		no	moon			low	ring	wobb	ole		1			5	
	3	ST	0312		mod	erate	moo	n	medium diffuse1			1.5	1				5	
id         run           1         576432           5756432         5756432								ble Job Superstar exit code id st 1			state	state id 5						
						2	552	427	:	1		5						
						3	589	420		2		6						
					J	,				able .	Job	Meli	bea					
id coach job id						.d s	superstar job id ex:				it co	de id	sta	ate	id			
	1 3					1				1			5					
			2		2				3			1			5			
			3		L 2	-		4										
			5			2		5				4			6			
			5															
		_										<u>ፓ</u>			ſ			
											L	Ta	ble St	ate				
								id	name	e								
Table Exit Code									1	crea	ated							
	id exit code description									2	quei	ued						
5	1		0		6	every	thing	g fi	ne		5	3	idle	e				
	2		6		1 5	Starg	uideı	r is	sue		4 runni			ning				
	3		254		i	input	file	es no	ot found		5 success							
												6	erro	or				
												7	sto	ped				

Table Job Coach

autoMAGIC Workshop Padova 2024





#### **Database Commands**

- Start your DB session with sqlite3 /path/to/your/db.sqlite
- Sometimes not entering the sqlite shell gives a performance boost:
  - > sqlite3 /path/to/your/db.sqlite 'SELECT \* FROM table\_name'
- You can configure your DB to be more convenient with:
  - .header on
  - .mode column
    - Can also be written into a .sqliterc file as general setup
- You can display all available table names with .tables





#### **Basic Database Queries**

- Query the whole moon conditions table:
  - > SELECT \* FROM moon\_conditions
- Query only certain columns from moon conditions table:
  - > SELECT id FROM moon\_conditions
  - > SELECT id, dc\_min FROM moon\_conditions
- Query under certain condition:
  - > SELECT \* FROM moon\_conditions WHERE dc\_min<2200</pre>





#### **Advanced Database Commands**

- Joining Tables
  - SELECT \* FROM runs

JOIN observations ON observation\_id=observations.id JOIN sources ON source\_id=sources.id WHERE name='NGC1275';

- Updating the Database
  - > UPDATE jobs\_melibea SET state\_id=1 WHERE state\_id=4
- Deleting entries
  - > ALWAYS use with the WHERE command!
  - DELETE FROM jobs\_add\_data\_superstar\_pic WHERE id=1





#### **Explore the Database**

- Have a look into the organization of observations and runs:
  - > Table names: observations, runs
- Have a look into the organization of the MCs:
  - Table names: analysis\_periods, mc\_productions, mc\_sets, mc\_runs
- Have a look into the job tables (most of them will still be empty!):
  - jobs\_add\_data\_superstar\_pic, jobs\_star\_superstar, jobs\_coach, jobs\_melibea, jobs\_dl3\_converter