From Science Operations to instrument and Survey Management: Exploring Machine-learning and AI techniques

Credit: C. Durán/ESO

Francisco Montenegro

Universidad Complutense de Madrid

IPARCOS Workshop, 16-17 June 2022

About

0000-0002-7430-3771
franciscomontenegro
fmontene@ucm.es

Former Physics student @ UCM

UNIVERSIDAD COMPLUTENSE MADRID

Séneca mobility programme UCM-ULL

INAF ISTITUTO NAZIONALE DI ASTROFISICA

Ph.D. @ ULL
International Research fellow @ INAF Istituto di Radioastronomia (Bologna)

Licenciado en Física @ ULL

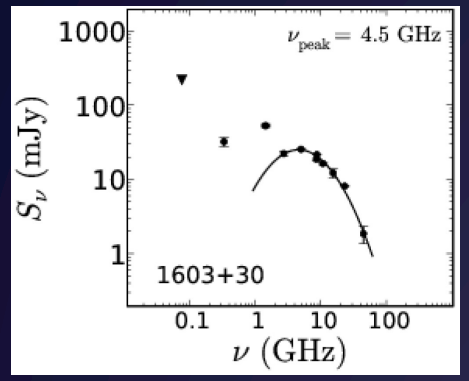
ULL Universidad de La Laguna

INSTITUTO DE ASTROFISICA DE CANARIAS IAC

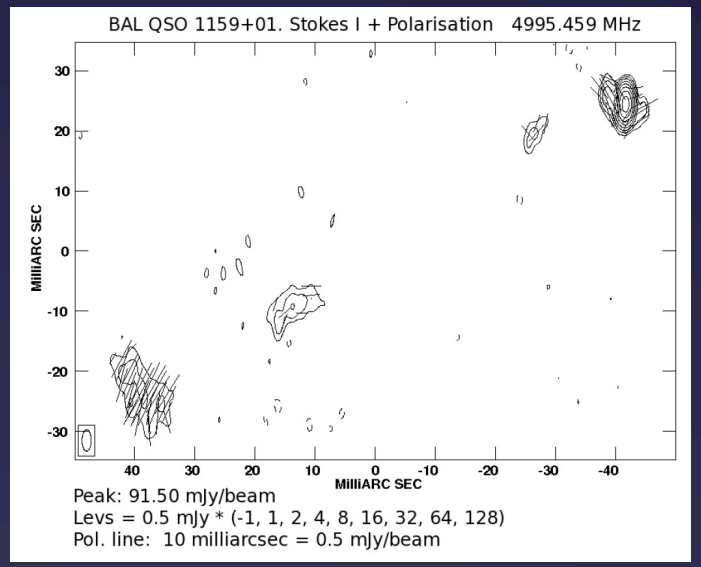
IPARCOS Workshop, 16-17 June 2022

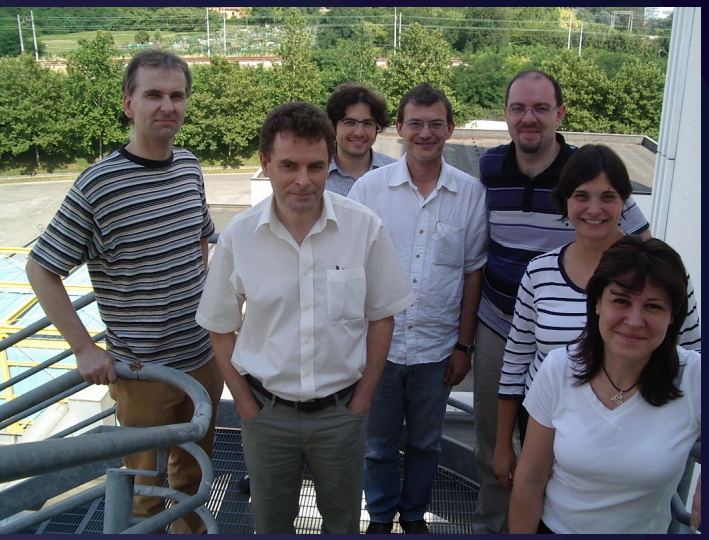
❖ Radio-loud Broad Absorption Line QSOs

Co-directors: K.-H. Mack, J.I. González-Serrano, I. Pérez-Fournón

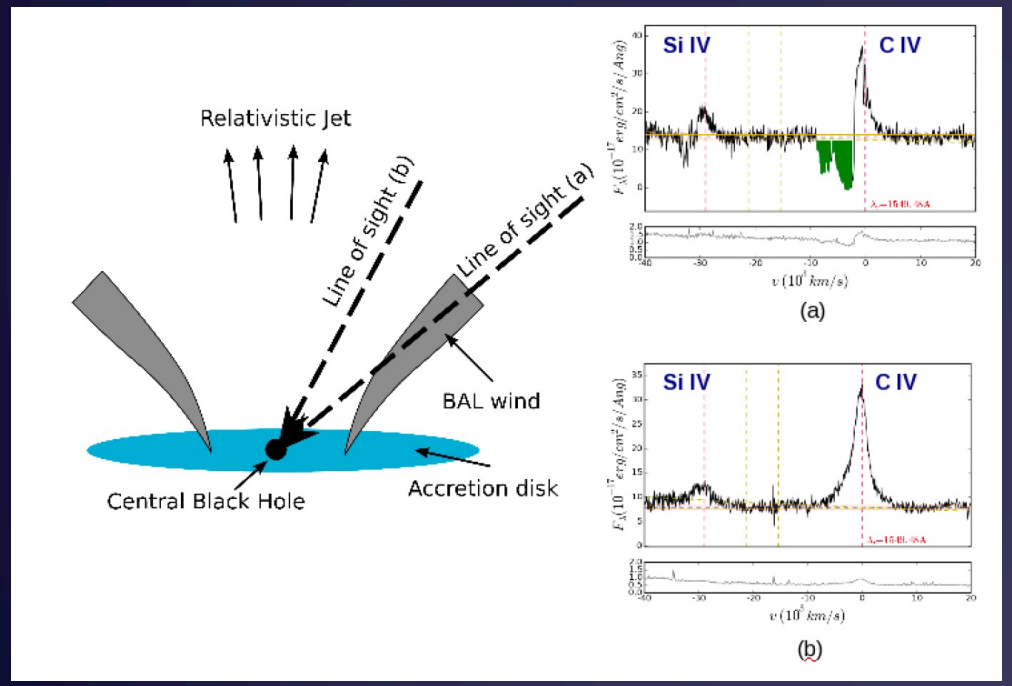


Montenegro-Montes+ 2008, MNRAS, 388, 1853





BAL QSO team, Bologna



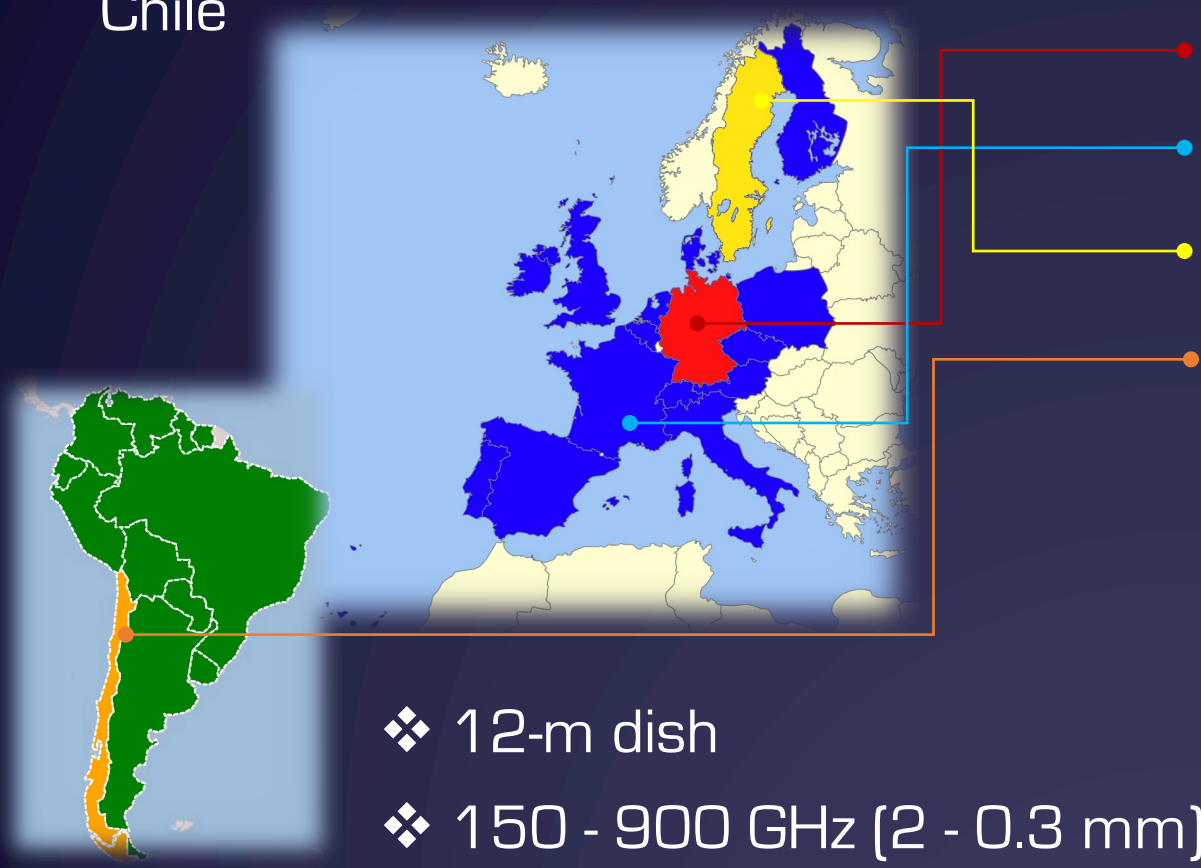
Montenegro-Montes, 2009, PhD T.

- ❖ Radio properties of BAL QSOs to test the evolutionary vs orientation scenarios
  - ❖ Radio morphology at arcsec and mas scale (compactness, orientation)
  - ❖ Synchrotron age, via SED
  - ❖ Polarisation properties (RM)

http://riull.ull.es/xmlui/handle/915/21058

APEX

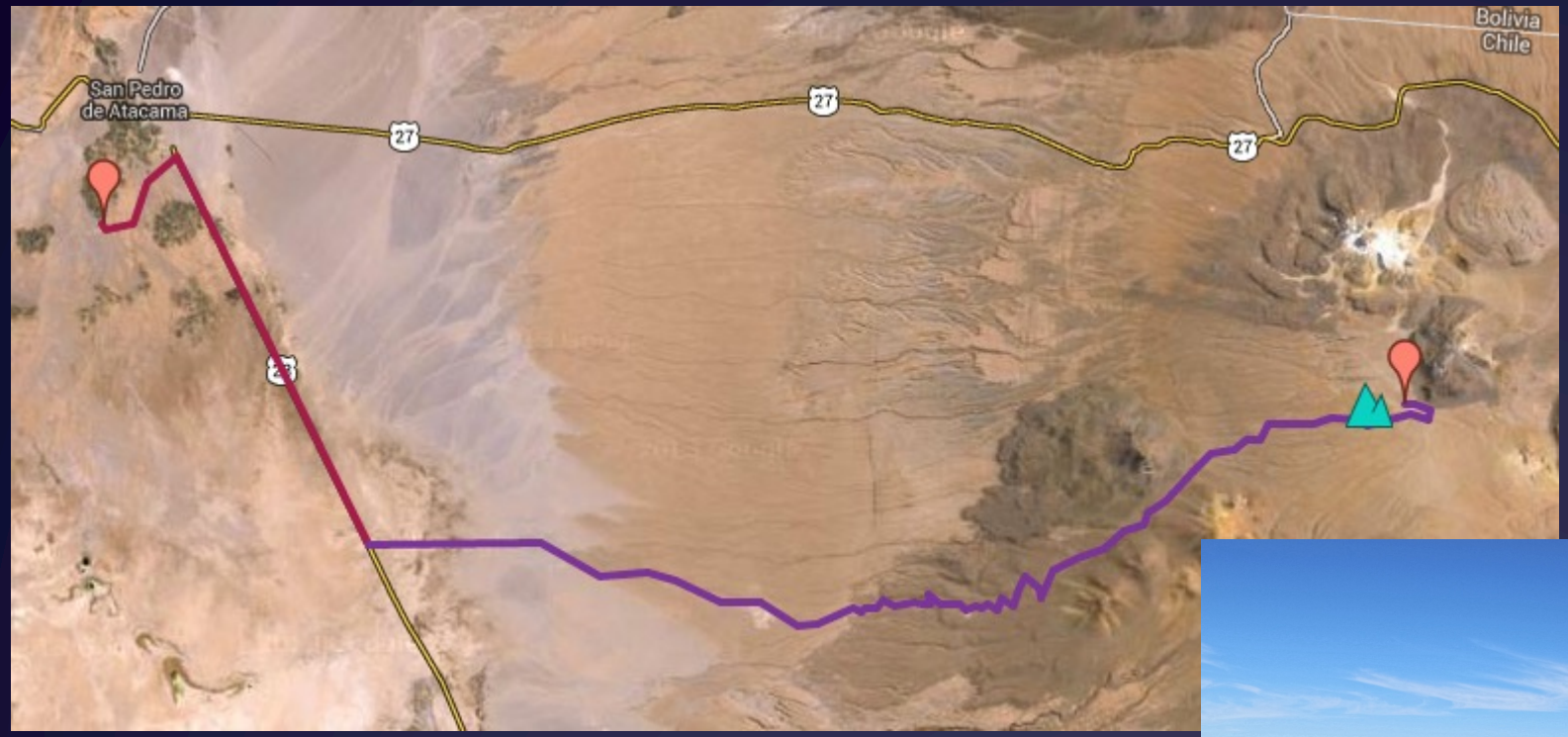
- European partnership in Chile



- 49% MPG and Germany
- 27% ESO member states
- 13% Onsala Space Observatory
- 10% Chile

- 12-m dish
- 150 - 900 GHz (2 - 0.3 mm)
- Operates @5100 m.a.s.l

APEX

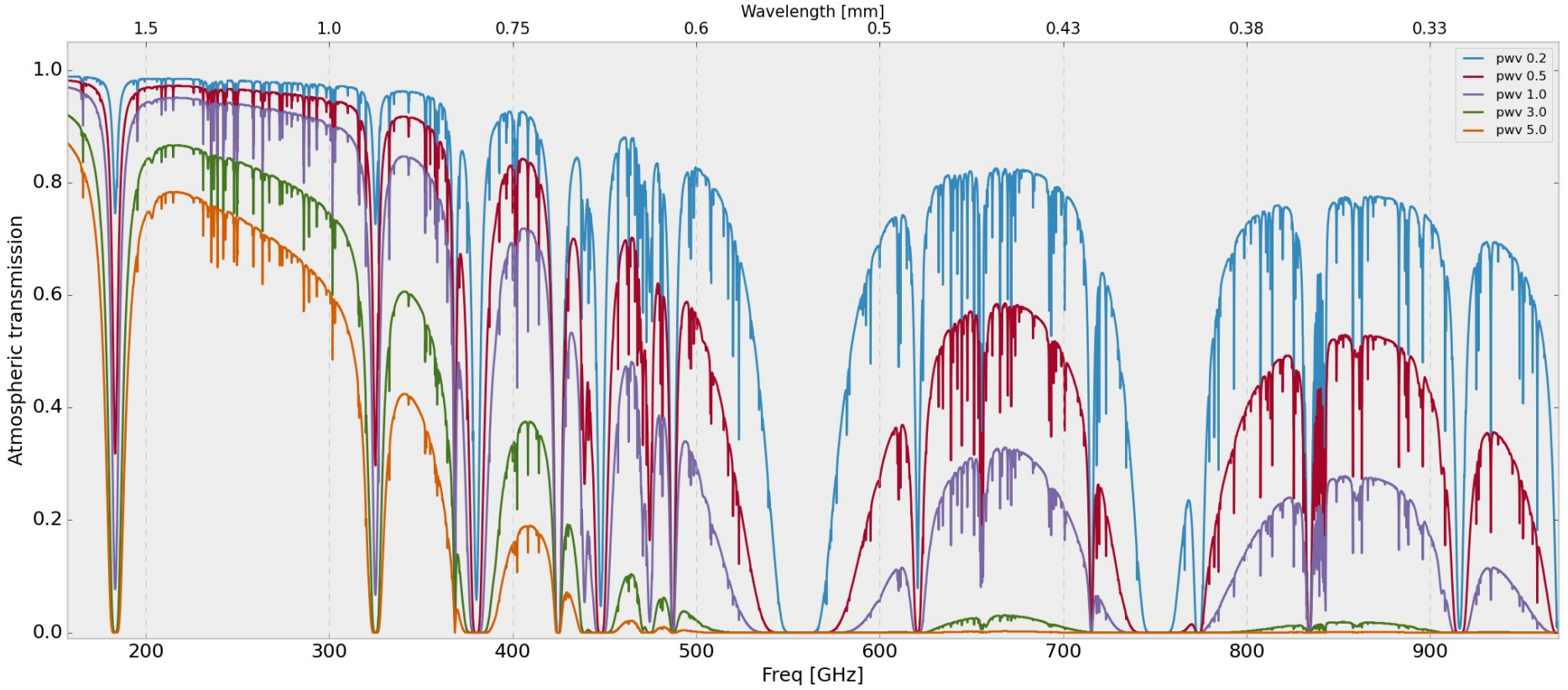

San Pedro
de Atacama
27
27
27
Bolivia
Chile


Credit: ESO


Credit:Clem & Adri Bacri-Normier (wingsforscience.com)/ESO


250
200
150
100
50
0
0
0.750
1.50
2.25
3
3.75
4.50
Mean PWV july 2017

APEX


Wavelength [mm]
1.5
1.0
0.75
0.6
0.5
0.43
0.38
0.33
pwv 0.2
pwv 0.5
pwv 1.0
pwv 3.0
pwv 5.0
Atmospheric transmission
1.0
0.8
0.6
0.4
0.2
0.0
200
300
400
500
600
700
800
900
Freq [GHz]

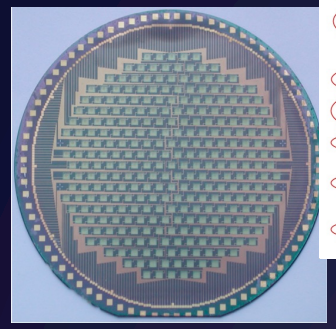

Credit:ESO/H.H.Heyer

APEX

Location	Llano de Chajnantor. 50 km east of San Pedro de Atacama, Northern Chile
Geodetic Coordinates[1]	Latitude: 23º 00’ 20”.8037 South Longitude: 67º 45’ 32”.9035 West Altitude: 5104.47 m
Geocentric coordinates (ITRF2005)[1]	X : 2225039.5297 m Y: -5441197.6292 m Z: -2479303.3597 m
Main reflector	Diameter: 12 m 264 aluminum panels with an average panel surface r.m.s. of 5 µm. Full surface accuracy: < 15 µm r.m.s.
Secondary reflector	Diameter: 0.75 m. Hyperboloidal Aluminum Polished to an average surface r.m.s. of 2 µm.
Manufacturer	Vertex Antennentechnik
Mass	125,000 Kg
Mounting	Alt-azimuthal
Primary focal length	4.8 m
Cassegrain focal ratio f/D	8
Beam width	7.”8 × (800 / f [GHz])
Pointing accuracy (r.m.s.)	< 2” all over the sky Pointing accuracy on track, 0.”6
Receiver cabins	2 Nasmyth (A,B) + 1 Cassegrain (C)

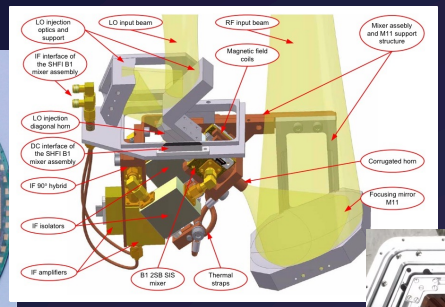
https://www.apex-telescope.org/ns/instruments/

APEX

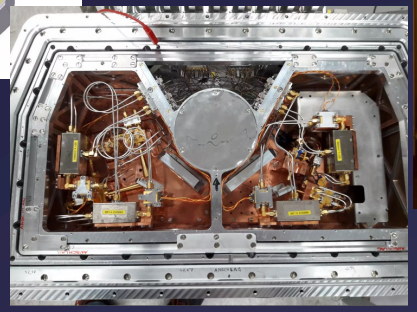


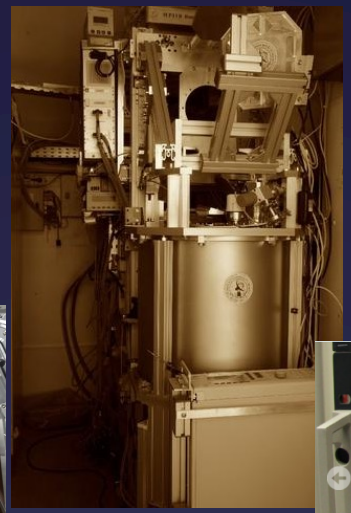
LABOCA

SHeFI



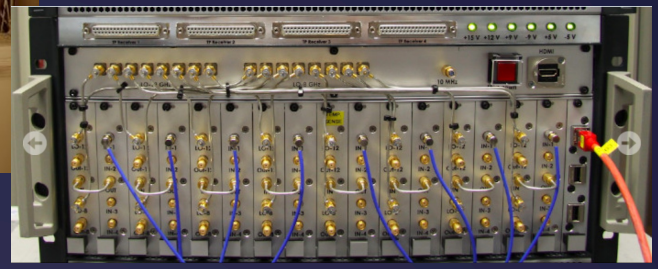
nFLASH





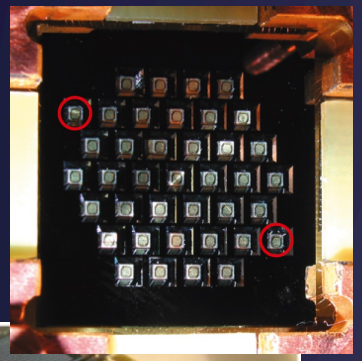
SEPIA

IF/dFFTS4G



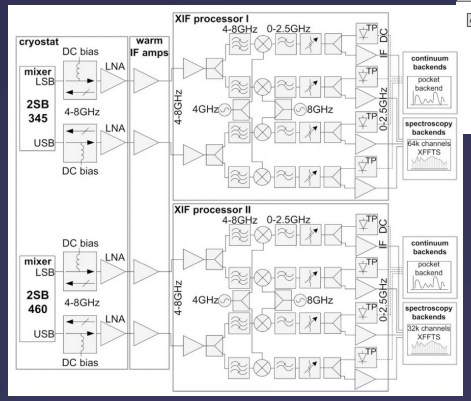
Powerful, varied and versatile permanent instrumentation

SABOCA

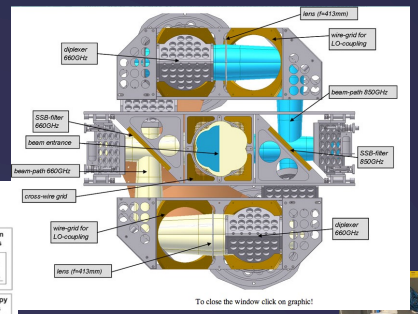


ArTéMiS

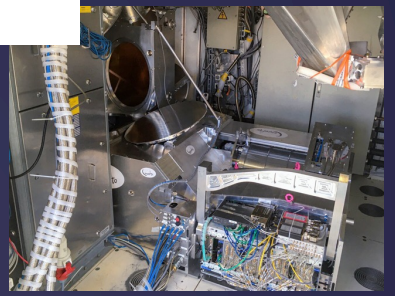
LASMA



FLASH+



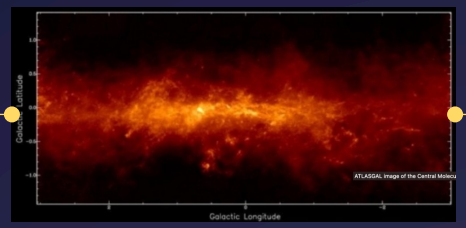
CHAMP+



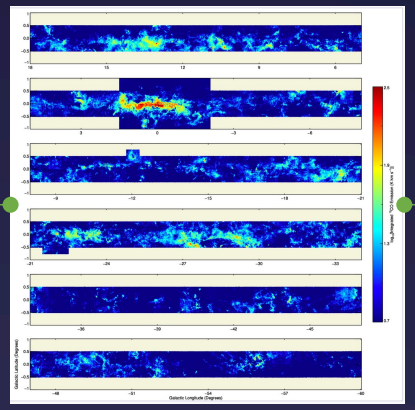
CONCERTO

... and more

APEX

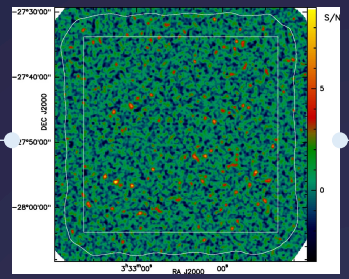


ATLASGAL survey



SEDIGISM survey

Major APEX overhaul: APEX - II



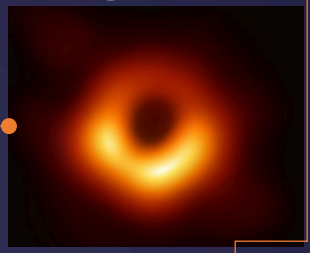
LESS survey

ArTéMiS

APEX External review

nFLASH 230/460

M87 @ EHT



GHReS

LASMA

SEPIA180

CONCERTO

FLASH

LABOCA

SHeFI

SEPIA345

SEPIA660

Inauguration

SABOCA

MoU

Science Operations start

SV

New MoU

2000 2005 2010 2015 2020

Sci ops

APEX

- Night and morning (16h)
- Sequitor / Chajnantor
- Oxygenized container

- All day
- Sequitor (night)
- Chajnantor Morning/afternoon
- Oxygenized container

- All day
- Fully remote (Sequitor)
- SAS system

- Sporadically
- Chajnantor, 5k
- No oxygenized container

We are here :)

Science Operations start

24h Operations

Sciops-R(emote)

COVID: Remote visiting observers

Sciops23+

2000 2005 2010 2015 2020

Sciops

APEX 2018 2022

Principles:

- **Efficiency**

  Minimize downtime
  Automation

- **Data Quality**

  Monitor sub-systems and data

- **Service to partners**

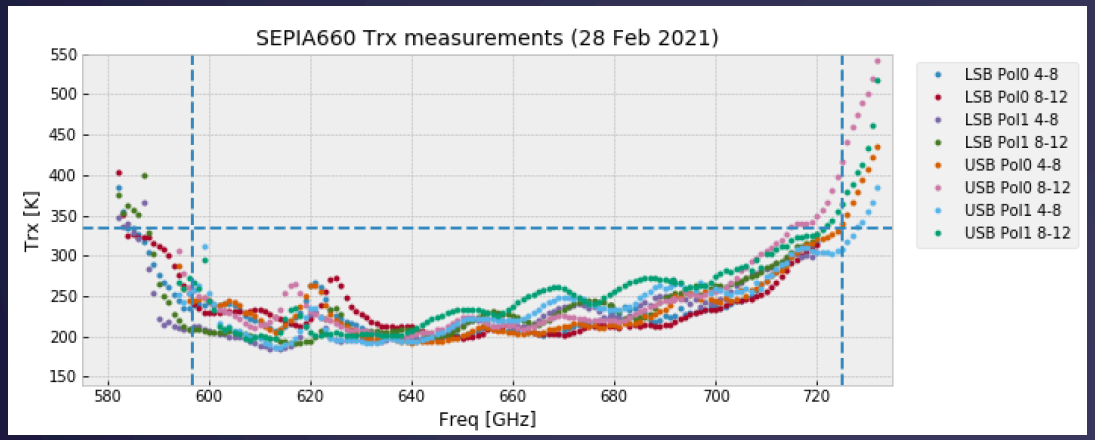
  Tools, documentation, support to instrument teams

Sciops WP-1000				
Observations WP-1100	Data quality & integrity WP-1200	Antenna and Instruments Performance WP-1300	Stakeholder support WP-1400	Internal development and maintenance WP-1500
Observation planning and setup WP-1110	Data quality assurance WP-1210	Antenna system monitoring WP-1310	Commissioning activities WP-1410	Sciops internal and monitoring tools WP-1510
Antenna and instrument operation WP-1120	Scientific goal assessment WP-1220	Instruments performance monitoring WP-1320	Hardware maintenance activities WP-1420	Internal training WP-1520
Observation logging WP-1130	Sientific data backup WP-1230	Troubleshooting WP-1330	Observing training for visiting astronomers WP-1430	Sciops internal documentation WP-1530
Weather and environment monitoring WP-1140	Science Data transfer WP-1240		Documentation for PIs WP-1440	Group coordination activities WP-1540
			Devel. & maint. of tools for PIs/observers WP-1450	Scientific Work (Santiago) WP-1550

Monitoring and Automation efforts

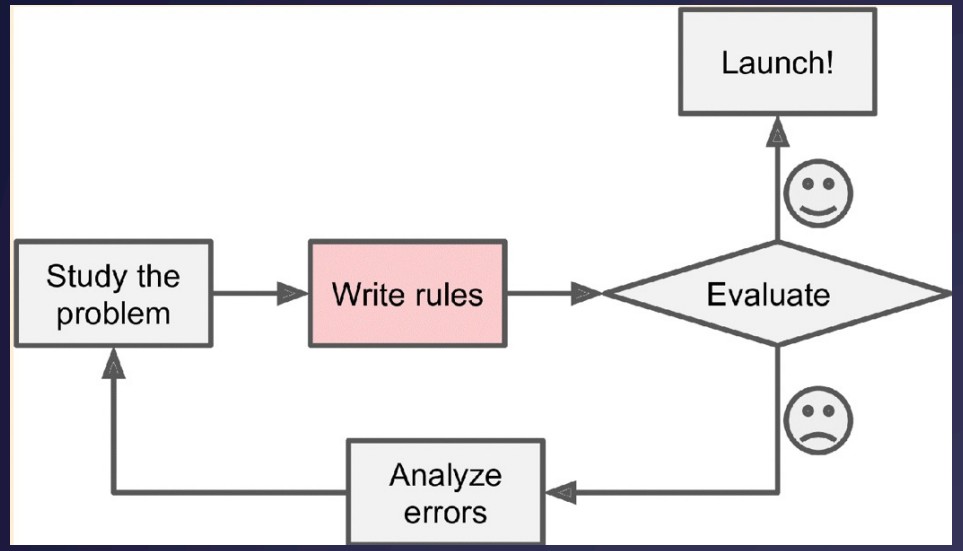
"You cannot manage, what you cannot measure"

- Monitoring points in as many subsystems as possible/useful (telescope, instruments, aux systems)
- Data storage on redundant relational databases (MySQL)
- Graphical dashboards to visualize evolution of KPIs (Grafana)
- Alarms for deviations from expected values
- Python scripting to automate observing macros
- Python/Jupyter to access data and analyse (troubleshooting)
- Anomaly detection (failures, preventive maintenance)

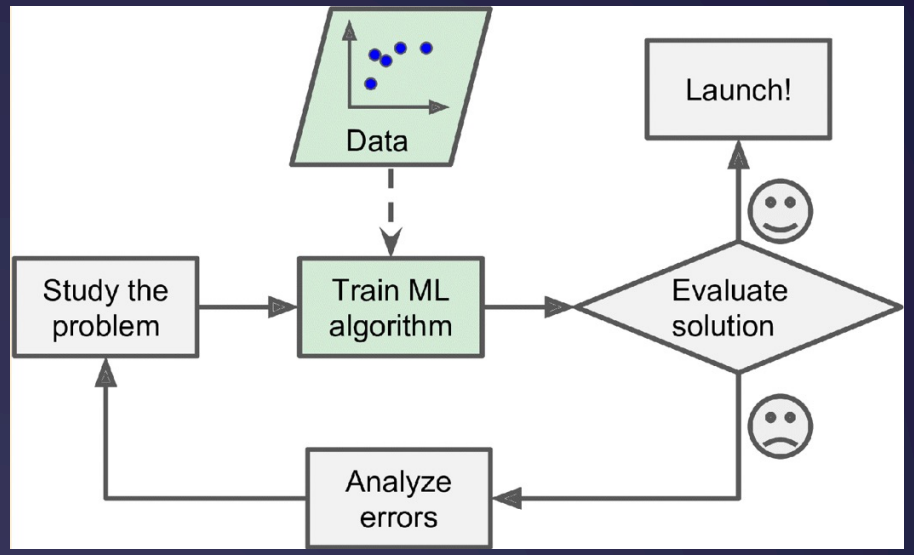


Machine Learning & Artificial Intelligence

- Ability to learn from data [without being explicitly programmed].



Traditional approach

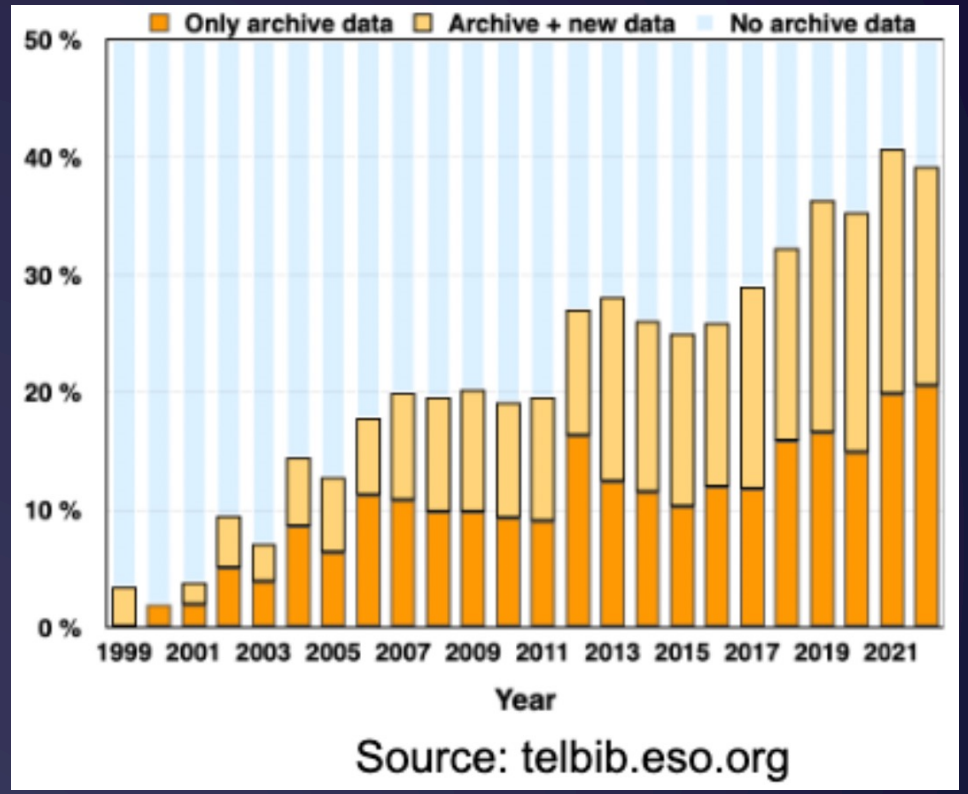


ML approach

Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow , 2nd Edition, by Aurélien Géron (O'Reilly). Copyright 2019 Kiwisoft S.A.S., 978-1-492-03264-9."

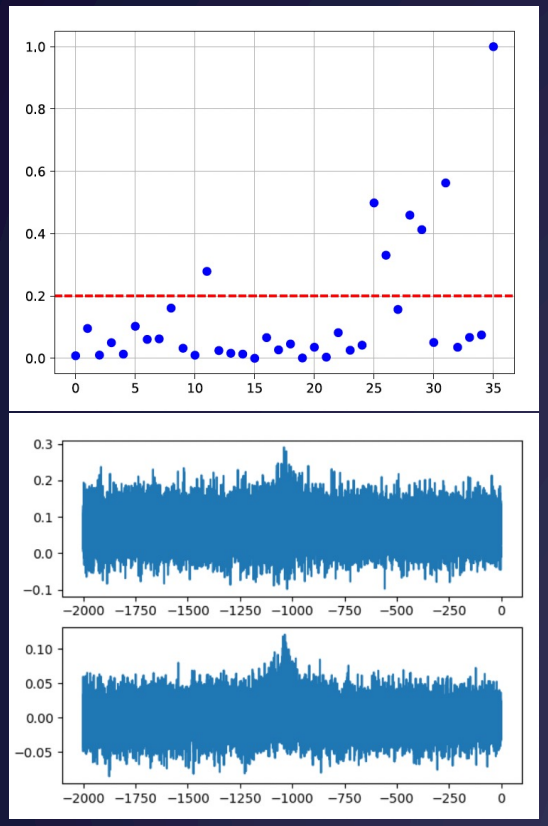
Machine Learning & Artificial Intelligence

- Successfully used in several scientific fields and contexts, including astronomy
- Potential for exploitation of public archives and catalogs
- Era of big data in astronomy (Vera Rubin, former LSST)
- Ethical problems inherent to the use of AI not so much of an issue for astronomical research



Machine Learning & Artificial Intelligence

❖ Successfully used in several scientific fields and contexts, including astronomy



Anomaly detection in APEX het. subscans
(Local Outlier Factor, `scikit-learn`)

F. Montenegro, C. Agurto, P. Venegas, 2019. AIA @ ESO

Science Operations

Token	weight
seqERR_BAD_QUALITY : "<alias>SENSOR20.last({})...	-11.400035
ccsERR_DB_QUALITY : Point <alias>SENSOR20.last...	-11.383230
... "Maintenance" OBS.START "{}" OBS.NTPL "{}"...	-8.649286
... Analysis" TPL.PRESEQ "GRAVITY_gen_tec_Pola...	-8.649286
WAIT -all (blue)	-8.235087
Setting HWP to: {} encs	-7.677115
Forward(b) SETUP to NGCIR3	-7.555012
Last Reply to 'START' from 'NGCIR3' received: ...	-7.555012
Send command 'START' '{},now' to sub-system 'N...	-7.555012
Cropping NGCIR1 failed!	-7.555012
Enable newdata newdataNGCIR3	-7.555012

Dataset Generation	Pre-processing	Event Representation	Technique

- High Level Task
  - Observation
  - Template Exec.
  - PRESET
  - SELINS
- Split in Traces
- Local CSV Files
	- Cleansing: remove bad traces
	- Tokenization (equivalent to stemming, etc, in NLP)
	- Vocabulary generation
		- Time series
		- Tokens
		- Bag of Words (BoW)
		- One Hot BoW
		- Successor Graphs
			- Time series analysis
			- Naive Bayes
			- Cluster detection:
			  - K-Means
			  - SVM
			  - T-SNE
			  - UMAP
			- Topic Modelling
			  - LDA
			- Deep Learning
			  - Transformers
			- Sequence Detection

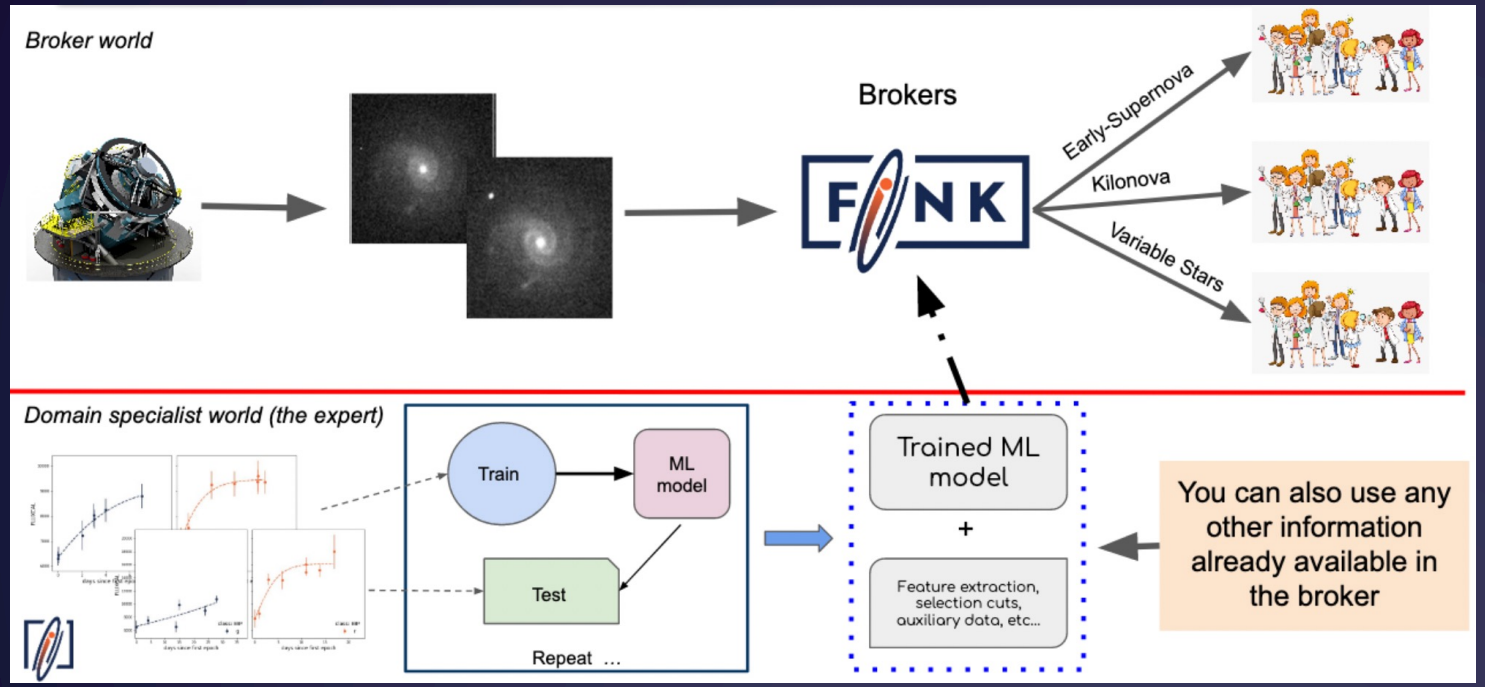
Log analysis in Paranal Observatory (NPL + various techniques)

Juan Pablo Gil, 2022 SCIOPS22 Conference

Machine Learning & Artificial Intelligence

❖ Successfully used in several scientific fields and contexts, including astronomy

Classification and Alert systems for transient objects



The Fink broker, E. Ishiida, 2022. *DOI 10.5281/zenodo.6554459*

Supervised classification of asteroid trails

S. Kruk (ESA).NASA/ESA Hubble Space Telescope

Machine Learning & Artificial Intelligence

- Successfully used in several scientific fields and contexts, including astronomy

Spectroscopic observations of the machine-learning selected anomaly catalogue from the AllWISE Sky Survey★

A. Solarz[1], R. Thomas[1], F. M. Montenegro-Montes[1], M. Gromadzki[2], E. Donoso[3,4], M. Koprowski[5], L. Wyrzykowski[2], C. G. Diaz[3,6], E. Sani[1] and M. Bilicki[7]

Received: 18 May 2020 | **Accepted:** 14 August 2020

Abstract

We present the results of a programme to search and identify the nature of unusual sources within the All-sky Wide-field Infrared Survey Explorer (WISE) that is based on a machine-learning algorithm for anomaly detection, namely one-class support vector machines (OCSVM). Designed to detect sources deviating from a training set composed of known classes, this algorithm was used to create a model for the expected data based on WISE objects with spectroscopic identifications in the Sloan Digital Sky Survey. Subsequently, it marked as anomalous those sources whose WISE photometry was shown to be inconsistent with this model. We report the results from optical and near-infrared spectroscopy follow-up observations of a subset of 36 bright ($g_{AB} < 19.5$) objects marked as "anomalous" by the OCSVM code to verify its performance. Among the observed objects, we identified three main types of sources: (i) low redshift ($z \sim 0.03 - 0.15$) galaxies containing large amounts of hot dust (53%), including three

- Search for anomalous, extrema or "rare" objects in large public datasets

A. Solarz,+ 2020, A&A, 642, A103

Planned contributions to GUAIX/IPARCOS activities

- Contract María Zambrano (atracción de talento) June 2022-2024

UNIVERSIDAD COMPLUTENSE DE MADRID

GOBIERNO DE ESPAÑA MINISTERIO DE UNIVERSIDADES

Plan de Recuperación, Transformación y Resiliencia

Financiado por la Unión Europea NextGenerationEU

- Contribute to planning and management of CATARSIS survey with TARSIS (previous talk by A. Gil de Paz)
- Contribute to software developent for processing, post-processing, visualization and analisis of CATARSIS data.
- Explore synergies with other members of the IPARCOS multidisciplinary team