

European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

# The ESCAPE Cluster for Open Science in the EOSC

Cristiano Bozza - University of Salerno and INFN

COST CA18108 Workshop on future challenges and opportunities in QGMM - Napoli (Italy)





ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.



# The FAIR approach and Open Science

All latest breakthroughs in science are the result of expensive, longlasting, personpower-intensive efforts

#### AND

The scientific method is all about reproducibility of experiments, analyses and results!

- Share
- Cooperate
- Cross-check
- Cross-fertilize





# The FAIR approach and Open Science

The FAIR approach (see <u>https://www.go-fair.org/fair-principles/</u>):

- Findable: make data <u>automatically</u> findable, by means of metadata that are easy to understand both to humans and machines
- Accessible: once found, data should be easy to access, both for human and machines, with proper authorization/authentication
- Interoperable: data should be stored and retrieved in ways that allow them to be integrated with other data and used in different analysis workflows
- **R**eusable: *metadata and data should be well-described so that they* can be reproduced or combined in different settings







۲

Access portal:

# FAIR data and Open Science: the EOSC

The European Open Science Cloud aims "to provide European researchers, innovators, companies and ٠ citizens with a federated and open multi-disciplinary environment where they can publish, find and reuse data, tools and services for research, innovation and educational purposes." (from the EC website)







# FAIR data and Open Science: the EOSC

• Marketplace and catalogue in the portal (<u>https://eosc-portal.eu/</u>)

Facility		Portal
Institut Laue Langevin	2 Institut Laue-Langevin Data Portal	
Istituto Nazionale di Astrofisica	2 Organisation: Institut Laue Langevin	
Virtual Atomic and Molecular	2 Scientific domain: Physical Sciences, Chemical Sciences, Biological Sciences	
Data Centre	Add to comparison Add to favourites	
European Spallation Source ERIC	1	
Show 228 more	Jupyter Notebook can be used to create and share	
Dedicated for	<ul> <li>documents that contain live code, equations, visualizations,</li> </ul>	A1177
Researchers	17 and text.	
Research Groups	13 JNcbutcasdtclcevat	
Research Communities	9 Data analysis with Jupyter Notebook	
Students	7 Organization – European Supplementary Dudiction Excilibut	ESKF
Research Projects	6 Scientific domain: Physical Sciences	The European Synchrotron
Show 13 more	Add to comparison Add to favourites	
Related Infrastructures and blatforms	•	也 OTHER ORDER TYPE
Find or choose from the list helow	LOFAR Science Processing	
Find of choose from the use below	LOFAR Science Processing	ASTRON
GARR Container Platform	2 Providing a service to generate science-ready LOFAR data, enabling discoveries in astronomy to	
FCI		
EGI	<sup>1</sup> happen faster and more easily.	
EGI-ACE	happen faster and more easily.     Organisation: NWO-I Netherlands Institute for Radio Astronomy (ASTRON)	
EGI-ACE SURF	1     happen faster and more easily.       1     Organisation:       1     Provided by:       1     SURF	
EGFACE SURF The Dataverse Project	1       happen faster and more easily.         1       Organisation:       NWO-I Netherlands Institute for Radio Astronomy (ASTRON)         1       Provided by:       SURF         1       Scientific domain:       Physical Sciences	
EGFACE SURF The Dataverse Project Show 69 more	1     happen faster and more easily.       1     Organisation:       1     Provided by:       2     SURF       1     Scientific domain:       2     Add to comparison   Add to favourites	
EGI-ACE SURF The Dataverse Project Show 69 more	1       happen faster and more easily.         1       Organisation:       NWO-I Netherlands Institute for Radio Astronomy (ASTRON)         1       Provided by:       SURF         1       Scientific domain:       Physical Sciences         1       Add to comparison       Add to favourites	Ů OTHER ORDER TYPE

					Contact us 🐱 He
	Object name *			_	
	1E 1740.7-2942		R	esolve	
	RA *		Dec *		
	203.57043033		-29.74510007		
	Start time *	End tim	e* T	ime unit	
	2017-03-06113:2	2017	03-06115:32:27.	ISO/ V	
NTEGRAL ISGRI INTEGRAL JEM-X INTEGRA	L SPI-ACS Polar	Antares GW Lega	scySurvey		
Instrument query parameters :	<b>^</b> 3	- 20 keV			[2022.07.05T00:44:18] 🔺 🗙
OSA Version Radius		JS9 Download <b>6</b>	Catalog Query paramete	rs Log Share 🕄	API code <b>0</b>
OSA11.2 ~ 4	deg√	🔆 View on Renku 🖯			
		Sig. Range: -5.08 65.26		Toggle Log. Norm	
Windows (ScWs) ③ Select for time range ○ Custom list ○ Custom list in file Maximum number of ScWs	3	82			∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞ ∞
50	3	2.0			40 -
50 INTEGRAL data access priviledge © Public O All Private EIXX module number	3	776	6 <sup>6X 5-1</sup>		20
50 INTEGRAL data access priviledge © Public O All Private FENX module number © JEMX-1 O JEMX-2	3	78	O <sup>GX 5-1</sup>		20 20







# ESCAPE overview

#### https://www.projectescape.eu/



ABOUT US = SERVICES = ESFRI PROJECTS OPEN SCIENCE NEWS = EVENTS LIBRARY INTERNAL

6

#### European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures







# ESCAPE overview

ESCAPE partners: Astrophysics & Particle Physics Research Infrastructures













#### FAIR data management DIOS: a data-lake infrastructure in ESCAPE







DIOS: a "data lake" prototype infrastructure to support FAIRness of data at Exabyte scale

- Storage services
  - PB-scale capacity available deployed at DESY, SURF-SARA, IN2P3-CC, CERN, IFAE-PIC, LAPP, GSI and INFN(CNAF, ROMA and Napoli)
- Data management and orchestration service: RUCIO (from ATLAS/LHC)
  - Instances running now at CERN, SKAO, IN2P3 and PIC
- File transfer and data movement: FTS (from WLCG/CERN)
  - Instances at CERN and PIC and using various protocols: http, gridftp and xrootd (and swift for HPCs)
- Content delivery and latency hiding: XCache (from XrootD/SLAC)
  - Deployments at GSI, CERN, LAPP and IN2P3
- Global Information System: CRIC (from ATLAS/LHC) deployed at CERN
- Auth/Authz/IM (AAI): Indigo IAM (from INDIGO-DATACLOUD, AARC/AARC2 and FIM4R projects). The ESCAPE IAM instance is deployed and operated in INFN/CNAF







• DIOS Resource allocation







• DIOS usage example



**LOFAR**: astronomical radio source 3C196 made using LOFAR data. The raw visibility data was downloaded via RUCIO from the EULAKE-1 and processed on Open Nebula at SurfSara using the container based LOFAR software



**EGO/VIRGO**: Upload 4h of Virgo public data sampled at 4kHz from an EGO server to the datalake. Download them to CNAF. The data is split into 1s samples. Making available the real-time strain data to pipelines and tools assessing the data quality.





• DIOS usage example





1. Data injected to the DL from three radio source observations in external locations2. User in external location download the data, process and store results back to the DL	3. User interested in combining results stored with other public data to cover also visible spectrum	4. Combined optical data from the Hubble located via the VO (WP4)	5. Optical and radio data aggregate in via the <b>ESAP (WP5), combined</b> <b>analysis done.</b> Results uploaded back to the DL.
---	--	--	---



From left to right: Radio image, Optical image and the Combined image (LOFAR with optical contours)



Cristiano Bozza - University of Salerno and INFN







#### FAIR and Open software OSSR: Open Source Scientific software Repository







The goal: populate the EOSC with a (sustainably) growing library of open source software for scientific research

#### https://escape2020.pages.in2p3.fr/wp3/ossr-pages/



zenodo	Search ESCAPE 2020 Q	Upload	Communities	<b>⊕</b> Log	in 🕼 Sign up				
ESCAPE 202	.0								
P All versions Access Right	Found 18 results.		< 1 >	Mos	Sort by: t recent v asc. v				
🗌 Open (31)	May 28, 2022 (v1.4.51) Software Ope FairRootGroup/FairMQ: v1	May 28, 2022 (v1.4.51) Software Open Access View View							
File Type	Alexey Rybalchenko; Dennis Klein; Nicolas Winckler, fuhlig1; Mikolaj Krzewicki; Giulio Eulisse; Mohammad Al- Turany; Matthias Richter; Gvozden Neskovic; Sebastien Binet; Dr. Christian Tacke; Radoslaw Karabowicz; AndreyLebedev; David Rohr; Matthias Kretz; Sandro Wenzel; Teo Mrnjavac; TKollegger; Laurent Abhecetche:								
☐ Zip (15) ☐ Pdf (12) ☐ Gz (3) ☐ Json (3) ☐ Md (1)	Changelog: v1.4.51 - 2022-05-28 Bugfix fix(shm): Add nullptr check on segmentInfo to Monitor::ResetContent() (#429) fix(shm): Monitor::ResetContent(): Reset user data after recreating the management segment, to avoid bumping into locked interprocess locks. #429) fix(shm): Monitor::Cleanup(): open m Uploaded on May 28, 2022 4 more version(s) exist for this record								
□ Simg (1) □ Tar (1)	May 13, 2022 (v0.20) Software Open Gammapy: Python toolbox	Access for gamm	na-ray astronor	ny	View				
Keywords	Donath, Axel; Deil, Christoph; Donath, Axel; Deil, Christoph; C	Terrier, Régis io; Wood, Ma llis; 💿 Nöthe, Harris Jopat	; 🔟 Ruiz, José Enri tthew; Olivera, Laur Maximilian; Vorokł han: Fidalgo, Davic	que; Kıng, Johannes; 💿 Rem a; Paz Arribas, Manuel; Giunt 1, Olga; 💿 Sipőcz, Brigitta; Let : Mohrmann, Lars: Lennarz F	y, Quentin; Jouvin, i, Luca; 💿 Khélifi, faucheur, Julien; )irk: Hailaoui Jalel;				
ESCAPE (7) Jupyter-notebook (4) Astronomy (2)	de Almeida Coutinho, Alexis, Wege Michael; Watson, Jason; Voruganti Aguasca Cabot, Arnau; Emery, Gab Arpit; Tibaldo, Luigi; Gallant, Yves; Jean-Philippe; Nakashima, Kaori, d	nmat, Matthi , Arjun; Poon, iriel; Tiziani, D Spir-Jacob, M e Bony, Math	ani, Fidaigo, David as; Papadopoulos, Helen; Tollerud, Er iomenico; Minaya, I arion; Chen, Andrev ieu; Craig, Matthew	, monmann, Cars, Leffilarz, L Dimitri; Chakraborty, Nachike k; Armstrong, Thomas; Joshi gnacio; Brügge, Kai; Siejkows v; 🕜 Zanin, Roberta; Bradley, ; Saha, Lab; Lemière, Anne; V	nr, najiaou, Jalel, ta; Droettboom, , Vikas; Bray, Erik; ski, Hubert; Gogia, Larry; Lenain, inicius, Zé; Barbary,				
Analysis (2)	Kyle; Vuillaume, Thomas; Ginsburg Bigas, Oscar; Cahill, Tyler; Zabalza,	, Adam; Morc Víctor; Kerze	uende, Daniel; Con ndorf, Wolfgang; D	treras, José Luis; Vega Garcia eiml, Peter; Panny, Sebastian;	a, Laura; Blanch Manconi, Silvia;				





OSSR development platform: not only repository but also a working tool ۰







- OSSR Example: ConCORDIA Containers for CORSIKA on DIRAC
  - A set of turnkey containers for cosmic ray simulation

	Setup GL		
PARAMETERS:	•		
Energy Hadronic Interaction Model			*
Low Energy Hadronic Interaction Mo	del		~
Detector Geometry		Se	etup GUI
ADDITIONAL		INPUT:	
OFTIONS.		RUNNR - RUN NUMBER:	1
		EVTNR - NUMBER OF FIRST SHOWER EVENT:	1
1a - Cherenkov version:	1 - Photons counted only in the ste	NSHOW - NUMBER OF SHOWERS TO GENERAT:	1
	2 - Photons counted in every step (	PRMPAR - PARTICLE TYPE OF PRIM. PARTICLE:	14
	<ul> <li>3 - No Cherenkov light distribution</li> <li>1 Emission apple is wavelength in</li> </ul>	ESLOPE - SLOPE OF PRIMARY ENERGY SPECTRUM:	-2.7
	<ul> <li>2 - Emission angle depending on w</li> </ul>	ERANGE_MIN - ENERGY RANGE OF PRIMARY PARTICLE (MIN):	1.e+5
1b - Cherenkov version using	1 - Particles at detector level not st	ERANGE_MAX - ENERGY RANGE OF PRIMARY PARTICLE (MAX):	1.e+5
Bernlohr IACT routines (for telescopes):	2 - Particles at detector level are st	THETAP_A - RANGE OF ZENITH ANGLE (DEGREE):	20
1c - apply atm. absorption, mirror	1c - apply atm absorption mirror	THETAP_B - RANGE OF ZENITH ANGLE (DEGREE):	20
reflectivity & quantum eff.:		PHIP_A - RANGE OF AZIMUTH ANGLE (DEGREE):	-180
1d - Auger Cherenkov longitudinal distribution:	1d - Auger Cherenkov longitudinal	PHIP_B - RANGE OF AZIMUTH ANGLE (DEGREE):	180
		SEED1_A - SEED FOR 1. RANDOM NUMBER SEQUENCE:	1
<b>.</b>		SEED1_B - SEED FOR 1. RANDOM NUMBER SEQUENCE:	0
GUI to cust	omize	SEED1_C - SEED FOR 1. RANDOM NUMBER SEQUENCE:	0
		SEED2_A - SEED FOR 2. RANDOM NUMBER SEQUENCE:	2
containers	and	SEED2_B - SEED FOR 2. RANDOM NUMBER SEQUENCE:	0
containers		SEED2_C - SEED FOR 2. RANDOM NUMBER SEQUENCE:	0
manage ich		OBSLEV - OBSERVATION LEVEL (IN CM):	100.e+2







- OSSR Training
  - Software as first class EOSC citizen
  - Enable software custodians



O7 June 2021 to18 June 2021
 ESCAPE Summer School 2021



12/07/2022







# Connecting ESFRI projects to EOSC through the Virtual Observatory



Funded by the European Union's Horizon 2020 - Grant N° 824064





# ESCAPE CEVO

- Aim make Astronomy data and services FAIR to support Open Science
- Objectives :
  - Assess and implement the connection of ESFRI and other astronomy research infrastructures to the EOSC by the Virtual Observatory.
  - Refine and pursue implementation of FAIR principles for astronomy data via common interoperability standards - extending the VO to new communities.
  - Establish data stewardship practices for adding value to scientific content of ESFRI data archives.





۰

# ESCAPE CEVO

- The Virtual Observatory is an operational framework for interoperable access to astronomical data and services
- A pioneer of FAIR data sharing an existing global framework populated by major data providers (space and ground based) that is heavily used by the community (e.g. ESA Gaia mission data access is fully VO)
- Built on International Virtual Observatory Alliance (IVOA) standards
  - Recognised in the ESFRI roadmap (2018), supported in Europe by Euro-VO (VO Partners + EC projects since ~2001), recognised in ASTRONET roadmaps (2008, 2014, 2021)













# ESCAPE CEVO

• CEVO Examples – Aladin evolution, integration with analysis tools









### **ESFRI Science Analysis Platform**

12/07/2022 Cristiano Bozza – University of Salerno and INFN 25

Funded by the European Union's Horizon 2020 - Grant N° 824064





- ESAP: a focal point to integrate services from other providers
- Web UI and API gateway
- "Centralized ESAP": access to a spectrum of ESCAPE services 000
- "ESFRI ESAP": a way for individual infrastructures, projects, etc. to integrate diverse capabilities







- ESAP Workflow example: query
  - User identifies relevant catalog services configured in this instance of ESAP
  - User submits search terms to multiple catalogs using consistent ESAP UI
  - Search results returned to user and displayed in unified form







- ESAP Workflow example: compute
  - User identifies convenient
     Jupyter system configured in
     this instance of ESAP.
  - User sends retrieved catalog data to notebook
  - Interactive analysis session serves to prepare bulk processing







- ESAP Workflow example: compute
  - User sends instructions for a batch job to a DIRAC cluster
  - Compute cluster retrieves data from the Data Lake
  - Batch processing
  - Results to Data Lake and user notified of completion









- ESAP Workflow example: visualize/publish
  - User identifies a visualization service
  - User initiates a visualization session, passing location of compute results
  - A DOI can be created and the results can be shared to a wider community







#### **ESCAPE offspring: Test Science Projects**







 ESCAPE Test Science Projects: use the EOSC and ESCAPE-developed tools to set up multimessenger analysis platforms

Wavefier: a prototype for realtime analysis





- ESCAPE Test Science Projects: use the EOSC and ESCAPE-developed tools to set up multimessenger analysis platforms
- Example: Extreme Universe TSP

Main Research Area	Objects/sources	Messengers	ESF/RI involved	ESCAPE services EOSC- Future integrations	Data Analysis tools (AI,ML)	Pilot project(s)	Computing resources required	Partner PM involved
Compact objects	Pulsars, FRBs, Off-nuclear AGN	radio, optical, X- ray,	LOFAR	Multiwavelength platform/Software catalogue,VO tools	Data science, Machine Learning	1) Radio astronomy: FRBs, pulsars, plerions, off- nuclear AGN	Compute cluster, Jupyter hub, Rucio Data lake	42 PM Astron/UvA
High energy Astrophysics	GRBs, jets, AGN, BNS, CCSN	neutrinos, gamma-ray, radio,X-ray, GW,	CTA, Virgo, KM3NET, SKA,LSST	Multimessenger platform/Software catalogue, Virtual Observatory tools	Model comparison, Machine Learning	1)GRB/ neutrino/GW analysis, 2) Blazar MWL/neutrino	GPU cluster Jupyter hub	12 PM UvA, 6 PM FAU. 4 PM CNRS , 24 PM SNS
Fundamental physics	Dark matter, GR, Primordial Universe	G <mark>W</mark> ,	Virgo, Einstein Telescope	Template banks, generation software,	Machine learning approach	1) DM template bank and ML analysis pipeline	GPU cluster Jupyter hub	10 PM INFN,12 PM UvA, 12 PM SNS, 2 PM FAU





• Example: Extreme Universe TSP/High Energy Astrophysics







# Example: Extreme Universe TSP/Detecting DM with the Einstein Telescope

- Presence of DM 'spikes' around BHs can alter inspiral dynamics
- GW waveform gradually goes out of phase with the corresponding vacuum-only waveform
- Possibility to detect and constrain dense DM 'spikes' with just a few cycles of GW 'dephasing' → but these subtle differences can be hard to detect!









- Example: Dark Matter
- Implement end-to-end dark matter analysis workflows on the EOSC, taking into account diversity in datasets (from kilobytes to petabytes), workflows (Jupyter notebook vs Big Data analysis), experimental communities (from individual theorists to large international labs)



s on the EOS es to petabyt perimental pational labs)	C, Tess), $10^{-34}$ $10^{-34}$ $10^{-35}$ $10^{-36}$ $10^{-37}$ $10^{-38}$ $10^{-38}$ $10^{-39}$ $10^{-44}$ $10^{-44}$ $10^{-44}$ $10^{-44}$ $10^{-44}$ $10^{-45}$ Direct de $10^{-45}$ $10^{-45}$ Direct de	Direct Detection plane       Direction plane       SM	ype, DM Type $X, g_{DM} = 0.Y$ ) intion experiments exclusion 90% CL. D Experiment 1 D Experiment 2 D Experiment 3 rojection 2 $10^3$ $m_{DM}$ [GeV] Credit: ESA/Gala/DPAC	$\begin{array}{c} \text{Illustrative ex} \\ 10^{-20} \\ 10^{-21} \\ 10^{-22} \\ 10^{-23} \\ 10^{-24} \\ 10^{-25} \\ 10^{-26} \\ 10^{-27} \\ 10^{-28} \\ 10^{-29} \\ 10^{-30} \\ 10^{-31} \\ 10^{-32} \\ 10 \end{array}$	Mample     Indirect detection plane       Indirect detection plane       ID Experiment 1       ID Experiment 2, projection       ID Experiment 3, projection       Collider search       Vpe, DM Type       8 - 0.Y       II02     103       I02     103       I04       MDM [GeV]
	Direct Detection	Colliders	Astrophysics	Theory	Indirect Detection
RELEVANCE	DM that interacts inside the detector (WIMPs, axions)	produce DM and probe the dark interaction	necessary for all	necessary for all	detect annihilating/ decaying DM through its decays (i.e. neutrino searches, gamma rays)
EXPERIMENTS INVOLVED	Darkside	ATLAS			KM3Net, CTA







#### • Example: Dark Matter

- Implement end-to-end dark matter analysis workflows on the EOSC, taking into account diversity in datasets (from kilobytes to petabytes), workflows (Jupyter notebook vs Big Data analysis), experimental communities (from individual theorists to large international labs)
- Onboarded onto the VRE (Virtual Research Environment)
  - Reinterpretation of search for particles mediating dark matter at ATLAS (LHC)
  - Analysis of Fermi-LAT data for dark matter constraints from dwarf galaxies
  - Analysis of KM3NeT experiment data for DM indirect detection
- To be onboarded onto the VRE
  - Two new full ATLAS/LHC searches for dark matter mediators
  - Real-time analysis
  - Leptons + missing transverse momentum
  - A reinterpretation of CMS search for dark matter production at the LHC
  - Compression algorithms using Machine Learning
  - Simulation pipeline for DarkSide experiment for direct detection of dark matter

The VRE server







- Example: IRF from KM3NeT for Dark Matter and Extreme Universe TSPs
- Instrument Response Function of neutrino telescope provides a quantitative estimation of the event rate
- It allows avoiding extensive MC simulations each time for a new configuration of neutrino source
- Different configurations of neutrino sources:
  - Diffuse source
  - Point source with power law E<sup>-a</sup>
  - Extended source
- Using gammapy interface will give an easily combination with other gamma experiments like CTA









#### **ESCAPE** Collaboration Agreement

12/07/2022 Cristiano Bozza – University of Salerno and INFN 39

Funded by the European Union's Horizon 2020 - Grant N° 824064





# ESCAPE Collaboration Agreement

 ESCAPE will become a sustained "Community Platform": (ESF)RIs core partners as Parties in a new Collaboration Agreement to operate Open Science and cooperate to address new topics in ERA







# ESCAPE Collaboration Agreement







