

https://projectescape.eu/

Giovanni Lamanna, LAPP/CNRS 20th November 2020





@LAPP research in HEP (e.g. @CERN: ATLAS, LHCb, HL-LHC, CLIC, FCC and Proto-DUNE) and Astroparticle & Cosmology (e.g. Virgo, CTA, LSST).





 Supporting actions on data science at the interface of both domains and connecting domains among them.



 LAPP-CNRS already involved in co-leading the previous cluster H2020 ASTERICS (astronomy and astroparticle physics)



- LAPP-CNRS coordinating the H2020 ESCAPE science cluster, extended to accelerator-based particle and nuclear Physics
 - A new cluster focus: FAIRness of data for the EOSC (European Open Science Cloud) implementation





Outline

About ESCAPE

- Background context
- Project partnership and work programme
- Highlights (focusing on HEP main interests)
- "Synergy" as key to demonstrate:

20/11/2020

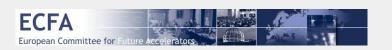
- Achievements
- Consolidation
- Outlook for the future







Background







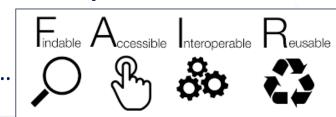
EC science cluster initiative

ESCAPE proposal in response to H2020-INFRAEOSC-04-2018 call Clusters to ensure the connection of the ESFRI RIs with EOSC (and the construction of EOSC)

Expected impact:

- Improve access to data and tools leading to new insights and innovation
- Facilitate access of researchers to data and resources for data driven science.
- Create <u>a cross-border</u> open innovation <u>environment</u>.
- Rise the efficiency and <u>productivity of researchers</u> through open data services and infrastructures for discovering, accessing, and reusing data.
- Foster the establishment of global standards.
- Develop <u>synergies</u> and complementarity <u>between involved research</u> infrastructures.
- Adopt common approaches to the data management for economies of scale.

















EOSC is the European Commission action in response to EU member states' shared policy about the uptake of Open Science:

- EOSC is a cloud for research data in Europe that allows universal access to data
- EOSC will federate existing resources across national data centres, e-infrastructures, and research infrastructures, allowing researchers and citizens to access and re-use data produced by other scientists.







In 2018 the Directors of major pan-European Research Infrastructures (RI) were convened for a shared analysis and decided to leverage the H2020 ASTERICS cluster successful experience towards the next focus ... EOSC!





Visible light Radio Gamma rays JIVE-**EST VLBI** ELT **ESO** CTA SKA Gravitational Cosmic-rays Neutrinos Accelerator-based Accelerator-based Nuclear Physics Waves **Particle Physics** HL-LHC KM3NeT **EGO-VIRGO FAIR**

ESCAPE

ESFRI and

other RIs



Astrophysics and Particle Physics: the results of the pre-analysis

Aligned expectations:

- **Large volumes of data** generators (up to multi-Exabyte scale level)
- "Observatory" and "Facility" type of operation requires global open access and long-term sustainability of research data
- The astrophysics and the accelerator-based particle/nuclear physics ESFRI facilities joined for a multi-probe **approach** towards the understanding of the Universe.
 - Addressing expectations of new generation researchers for a "virtual space" sharing workflows and interoperate data
 - Acknowledge commitment of scientists (on transversal research) .
- Engage with society and citizens

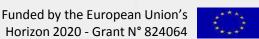
Decide to enhance the coordination:

- leveraging two major complementary excellences in data stewardship:
 - the astronomy **Virtual Observatory** infrastructure;

20/11/2020

- long-standing expertise of the **HEP** community in large-scale distributed computing and big-data management.
- operating a shared open innovation environment, adopting cooperatively the FAIR/Open-Science principles







ESCAPE in a nutshell

- 31 partners (including 2 SMEs)
- 7 ESFRI projects & landmarks: CTA, ELT, EST, FAIR, HL-LHC, KM3NeT, SKA
- 2 pan-European International Organizations: **CERN, ESO** (with their world-class established infrastructures, experiments and observatories).
- 2 European research infrastructures: EGO and JIV-ERIC

Formal commitment of their legal entities and management boards required by EC

- 1 involved initiative/infrastructure: EURO-VO
- 4 supporting European consortia: APPEC, ASTRONET, ECFA and NuPECC.
- Budget: 15.98 M€
- Started: 1/2/2019
- Duration: 48 months (end date 31/1/2023)
- Coordinator: CNRS-LAPP







































































ESCAPE in a nutshell

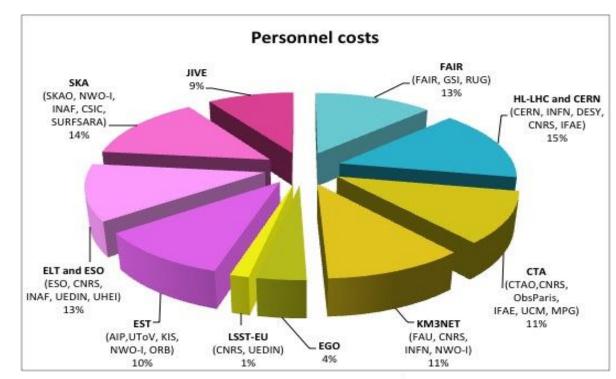
As per H2020 INFRAEOSC-04-2018 call - CLUSTER MEMBERSHIP and PARTNERSHIP:

- The EC funding contributions proportional to the number of pan-European research infrastructures (ESFRI project/landmark) that the science cluster connects to the EOSC.
- Each RI legal entity commits together with a sub-set of associated national stakeholders.



Furthermore:

- The <u>Director</u> of each ESFRI RI is a member of the
 ESCAPE Supervisory Committee (E-SC)
- APPEC, ASTRONET, ECFA, NuPPEC chairs and ESA representative form the ESCAPE External Advisory Board (E-EAB)



Distribution of personnel costs among partners grouped by RI







20/11/2020



The 2020 European Strategy

Quoting some recommendations ... that ESCAPE is addressing (and subjects for discussion in ECFA ..)





ECFA, APPEC (and ASTRONET), NuPECC communities

Organisational issues

20/11/2020

- The particle physics community and the European Commission have a strong record of collaboration. The relationship between the particle physics community and the European Commission should be further strengthened, exploring funding-mechanism opportunities for the realisation of infrastructure projects and R&D programmes in cooperation with other fields of science and industry.
- European science policy is quickly moving towards Open Science, which promotes and accelerates the sharing of scientific knowledge with the community at large. Particle physics has been a pioneer in several aspects of Open Science. The particle physics community should work with the relevant authorities to help shape the emerging consensus on Open Science to be adopted for publiclyfunded research, and should then implement a policy of Open Science for the field.







From Open Science to Open Data

- **Open science** for three main ambitions:
 - Change the way citizens could perceive research and public investments for research.
 - Enable opportunities offered by the digital revolution to allow everybody to participate in the scientific **process** by accessing research data.
- **Accelerate the discoveries** and **increase scientific value** by sharing data and by transferring knowledge within scientific communities.
- **Open data** means 'FAIR Data'
 - support and uptake a set of guiding principles about the way to plan, produce and reproduce scientific data
 - 'Data' are any 'digital objects' (including software, workflows, algorithms, ...)
 - 'Open data' does not mean 'free data'

(depending on what 'data' are being referred to, conditions governing their access and reuse will always be applied.)

20/11/2020

More fundamental motivations for open data in the future:

> "What about more exploratory studies where the theory hasn't vet been invented?

What about engaging undergraduate students?

What about examining old data for signs of new physics?'

(Jesse Thaler)









ESCAPE Work Programme

Data Lake:

Build a scalable, federated, data infrastructure as the basis of open science for the ESFRI projects within ESCAPE. Enable connection to compute and storage resources.



Software Repository:

Repository of "scientific software" as a major component of the "data" to be curated in EOSC. Implementation of a community-based approach for the continuous development of shared software and for training of researchers and data scientists.



Virtual Observatory:

Extend the VO FAIR standards, methods and to a broader scientific context; prepare the VO to interface the large data volumes of next facilities.



Science Platforms:

Flexible science platforms to enable the open data analysis tailored by and for each facility as well as a global one for transversal workflows.

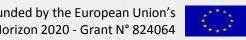


Citizen Science:

Open gateway for citizen science on ESCAPE data archives and ESFRI community

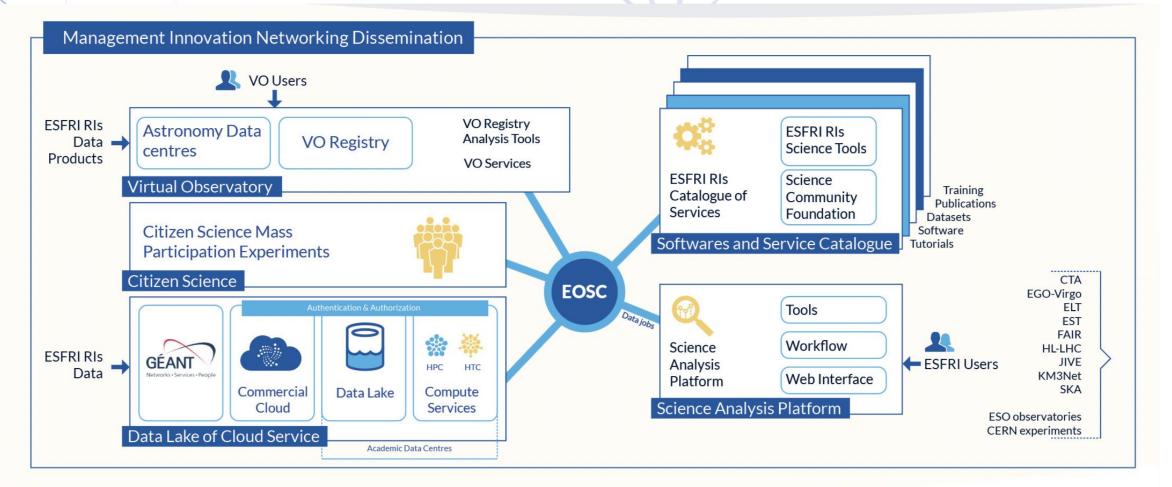








ESCAPE final goal: building a community-based EOSC cell









Highlighting some current results







ESCAPE Work Programme

Data Lake:

Build a scalable, federated, data infrastructure as the basis of open science for the ESFRI projects within ESCAPE. Enable connection to compute and storage resources.



Software Repository:

Repository of "scientific software" as a major component of the "data" to be curated in EOSC. Implementation of a community-based approach for the continuous development of shared software and for training of researchers and data scientists.



Virtual Observatory:

Extend the VO FAIR standards, methods and to a broader scientific context; prepare the VO to interface the large data volumes of next facilities.

Science Platforms:

Flexible science platforms to enable the open data analysis tailored by and for each facility as well as a global one for transversal workflows.



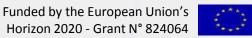
Citizen Science:

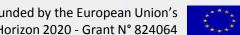
Open gateway for citizen science on ESCAPE data archives and ESFRI community

20/11/2020











ESCAPE, FAIRness and computing

DATA LAKE: A new model for federated computing and storage to face the (overall) **Exabyte scale** of data volumes of next generation ESFRI RIs in ESCAPE.

- The astrop. next RIs (such as SKA et al.) will operate as "observatories". World-wide scientists have open access to science-ready data and potentially lower-level data for further analysis through regional centres.
- The HL-LHC data challenges and the quality of the network implies the evolution from hierarchical (WLCG) to regional distributed data centres.
- FAIR-GSI has similar needs and follows the same path.
- For all: increasing offer of large-scale heterogeneous CPU-only resources (grid, HPC, public and commercial cloud, volunteer computing); **CPU and storage resources are no longer always colocated** as it is in the GRID.





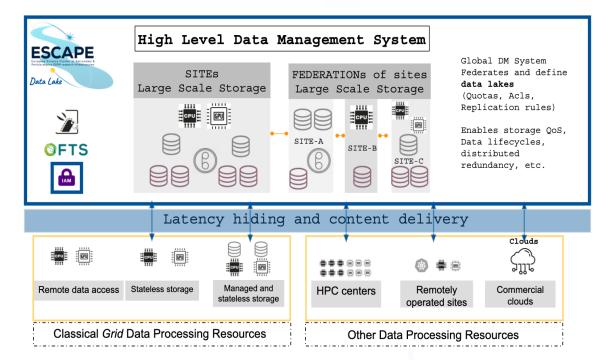


The ESCAPE Data Infrastructure for Open Science

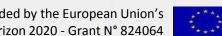
The ESCAPE Data Infrastructure for Open Science (DIOS) aims at delivering a prototype of the Data Lake concept, a common storage infrastructure that:



- Provides **global data management** orchestration
- Contribute to deliver **Open Access and FAIR data** services: trustable data repositories; enable data management policies; transparent data access layer.
- Science **projects to drive** the services requirements most suitable to their needs.



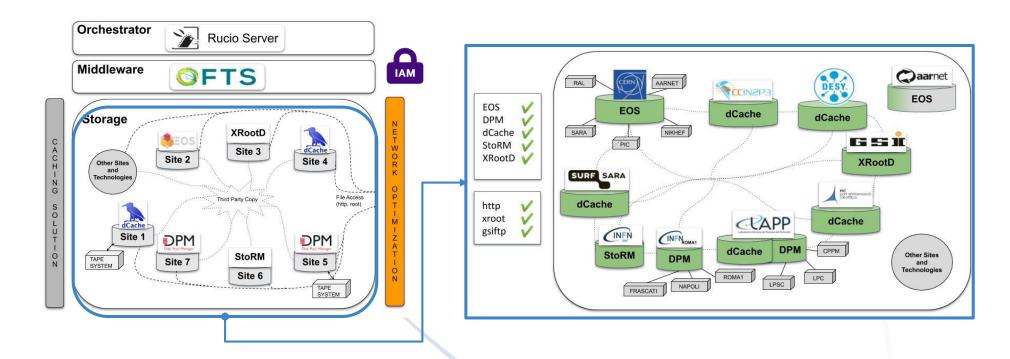






First achievements: a functional Data Lake pilot

- Pilot Data Lake with 10 storage endpoints functional:
 - CERN, DESY, GSI, IFAE-PIC, IN2P3-CC, INFN-CNAF, -ROMA, -Napoli, LAPP-MUST and SURF-SARA
- The high level Data Lake orchestration layer is consolidated









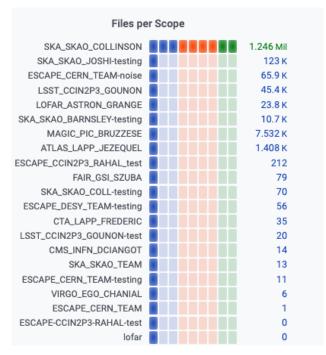
First achievements: Science in the Data Lake

- Strong involvement of ESFRI RIs and other experiments:
 - Data injection within the Data Lake by:

ATLAS, CMS, CTA, FAIR, LOFAR, LSST, MAGIC, SKA, and VIRGO/EGO

- Data management demonstrator from Astroparticle, Radio-astronomy, Gravitational Waves, Cosmology and Particle Physics communities together on a common data management infrastructure
- Pipeline data analysis tests currently in progress





DEMO







Software repository as part of the EOSC catalogue



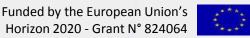
Objectives:

- Each ESCAPE ESFRI RI needs individually to expose and make accessible data & software (each one needs a sort of "start-kit").
- All together they wish to adopt common solutions and offer a virtual space for interoperability and multi-messenger & multi-probe data research to next generation scientists (Astro. & Particles)
- All are willing to co-develop new methods/algorithms, share (novel) software and expose the open science tools under the EOSC catalogue

ESCAPE deliverables:

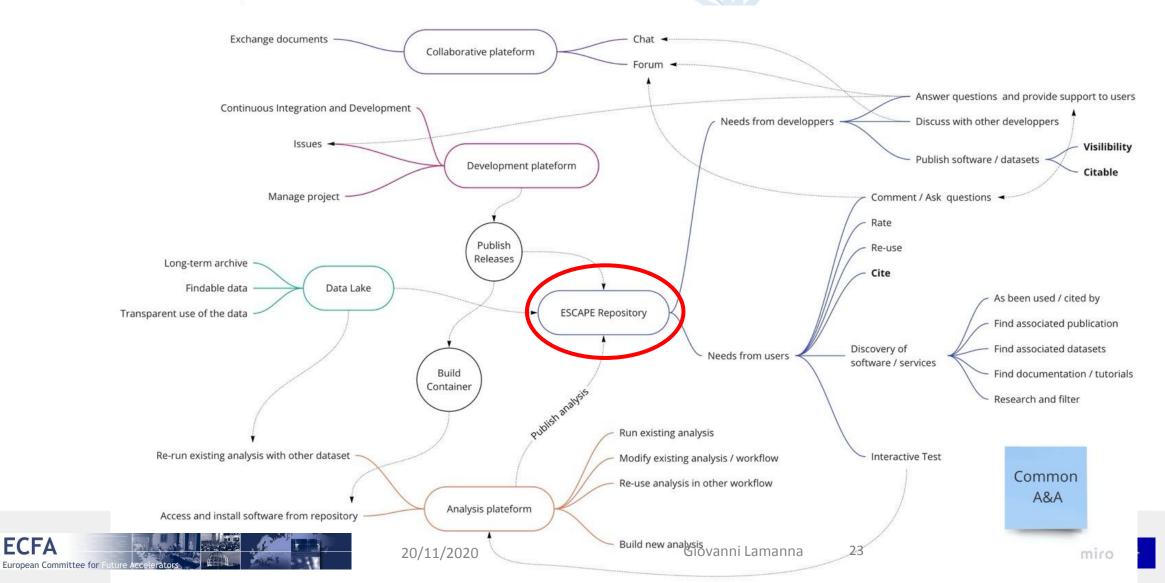
- Establish a community-foundation
- Expose/share software to users via the EOSC catalogue
- Train and guide the scientists/users
- **Provide a scheme to acknowledge and reward scientists** for their commitment







Building up a global Virtual Research Environment





ESCAPE OSSR and Development Platform - how to ease the publication and integration process?

From a single click

Publishes source code (updates your existing record with new versions)



- Long term archived
- Findable
- Citable



- Make a new tag (release)
- Let the CI do the rest





Giovanni Lamanna

publishes singularity image

under dev

builds a docker container













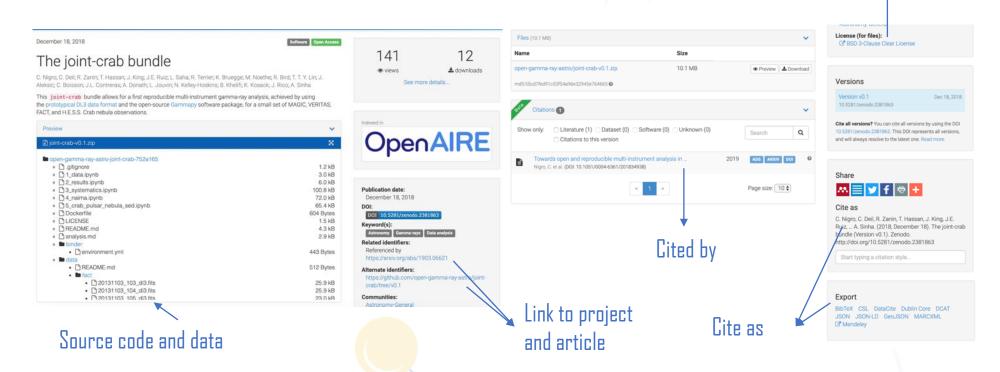
Example project: The CRAB bundle

license

The CRAB multi-instrument gamma-ray analysis with MAGIC, VERITAS, FACT and H.E.S.S.

https://zenodo.org/record/2381863#.XkxcD5NKhhA

https://github.com/open-gamma-ray-astro/joint-crab/tree/v0.1









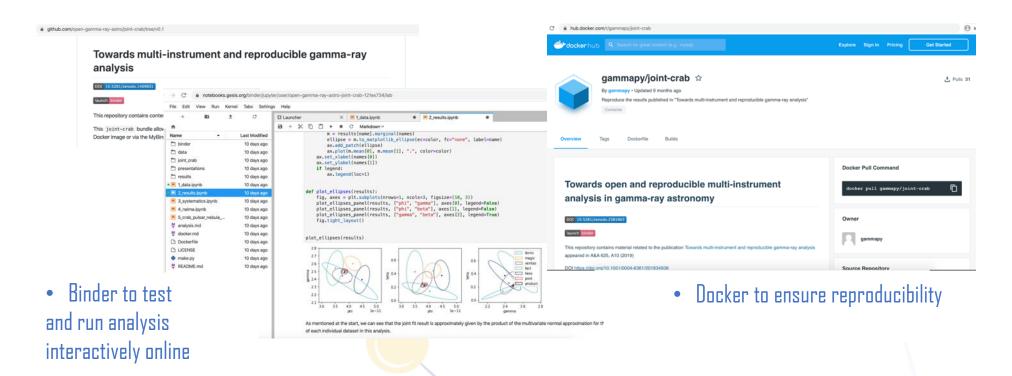
Example project: The CRAB bundle

The CRAB multi-instrument gamma-ray analysis with MAGIC, VERITAS, FACT and H.E.S.S.

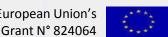
20/11/2020

https://zenodo.org/record/2381863#.XkxcD5NKhhA

https://github.com/open-gamma-ray-astro/joint-crab/tree/v0.1









Interfacing the Particle Physics CERN Open Data portal

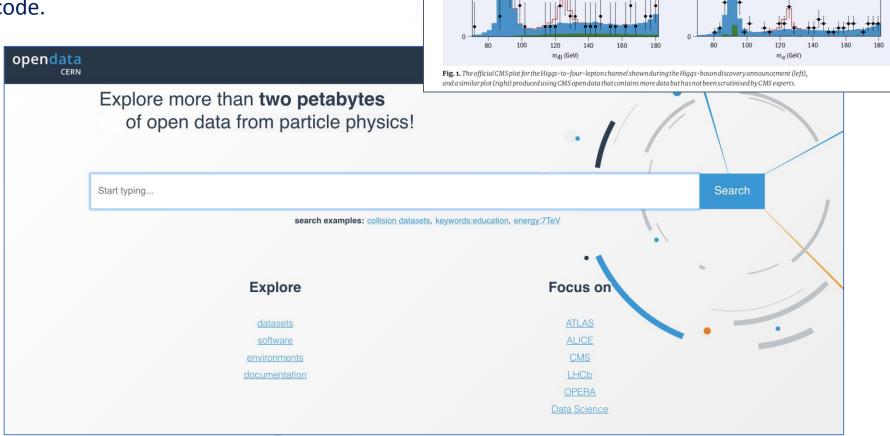
→ data

■ Z + X

- m_H = 126 GeV

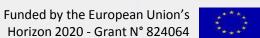
http://cerncourier.com

Today CERN's open-data portal hosts and serves data including many software tools and virtual machines to run the analysis code.





20/11/2020



∳ data

Z/γ*+ X

■ TTBar

- m_H = 125 GeV



The 2020 European Strategy

Quoted not for self-reference, but for an outlook for the future



Other essential scientific activities for particle physics

Large-scale data-intensive software and computing infrastructures are an essential ingredient to particle physics research programmes. The community faces major challenges in this area, notably with a view to the HL-LHC. As a result, the software and computing models used in particle physics research must evolve to meet the future needs of the field. The community must vigorously pursue common, coordinated R&D efforts in collaboration with other fields of science and industry, to develop software and computing infrastructures that exploit recent advances in information technology and data science. Further development of internal policies on open data and data preservation should be encouraged, and an adequate level of resources invested in their implementation.

Deliberation Document on the 2020 update of the European Strategy for Particle Physics

The scientific outcomes of particle physics experiments are made possible by the development of an efficient computing and software infrastructure. Computing and software are profound R&D topics in their own right and are essential to sustain and enhance particle physics research capabilities. There is a need for strong community-wide coordination for computing and software R&D activities, and for the development of common coordinating structures that will promote coherence in these activities, long-term planning and effective means of exploiting synergies with other disciplines and industry. Some recently initiated examples are the HEP Software Foundation addressing the common computing and software challenges related to particle physics, and ESCAPE (European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures) exploring the synergies in the areas of astronomy, astroparticle and acceleratorbased particle physics.



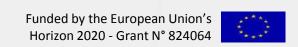




The 2020 European Strategy

- "Cluster" is the most successful (network) tool that the European Commission has ever proposed, thanks to indispensable ingredients: network + funding + focus + high-level commitment + coherence with European policy + multi-disciplines + bottom-up researchers' involvement + training.
 - Science clusters (within the EOSC Association) to build a coordinating structure;
 - Physicists together with data-scientists, researchers in computer science and digital SMEs.
 - Virtual Research Space for open science, R&D and open data uptake.







... more about SYNERGY









ESCAPE for open data: FAIRness towards "VRE"

https://indico.ijclab.in2p3.fr/event/5418/contributions/17542/

RECALL (from slide 13) 3rd pillar of Open Science:

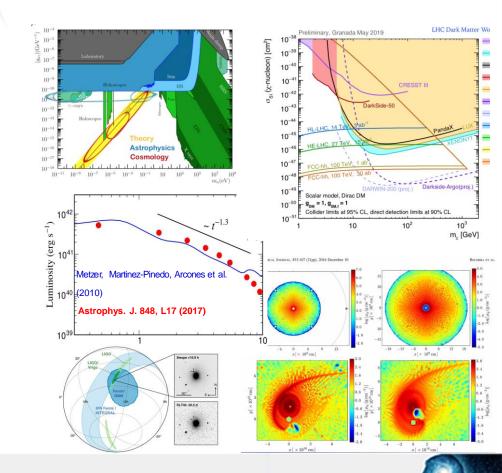
Accelerating discoveries and increasing scientific value by sharing software and knowledge within scientific communities.

"Test Science Projects (TSP)" to validate the ESCAPE towards "virtual research environments":

Some potential examples to pursue:

- Dark Matter: From legacy data of gamma-ray dwarf galaxy observations to ... Data registry and software through multi-messengers data for Indirect Dark Matter search, Direct search, Accelerator results, models and phenomenological studies.
- From high-energy detection of GRBs, FRBs to ... Multi-lambda, nuclear astrophysics, Gravitational Waves, fundamental physics and relativistic astrophysics computational modelling.

A vision proposed by G.Lamanna at JENAS 2019 workshop ... And an important path has been accomplished in one year [...]









Synergy: Test Science Projects

TSPs originally part of the ESCAPE work programme, proposed to validate ESCAPE services for Open Science at the end of the project.

The ESCAPE-TSP concept finds consensus and evolves.

- Thanks to synergies and shared plans with ESFRI-EOSC task force, EC and other Science Clusters.
- Stimulating and/or cooperating with JENAA Eols.

20/11/2020

- It is now in all clusters aiming at enhancing researchers participation in open science and crossdomain scientific research.
- Included by all five Science Clusters in one more H2020 EU funding request (proposal successful but EC grant decision pending ... since more inter-clusters than intra-cluster TSP is now requested.)







ESCAPE TSPs participating to the JENAA EoIs

Dark Matter TSP:

- understand the nature of dark matter by collecting data, analysis pipelines and results from complementary astronomy, particle and nuclear physics sources on a broad platform that will be ultimately be hosted on the EOSC Portal.
- exploit synergies and complementarities across different communities, creating a unique link between dark matter as a fundamental science question and the Open Science ESCAPE services needed to answer it.

Extreme Universe TSP:

- o do 'frontier' multi-messenger science to understand extreme matter and particle processes in strongly curved space-time.
- o combine astronomy and e-infrastructures and focus on data organisation
- o organise data from different wavelengths/messengers and different types of extreme astrophysical transients (SNe, GRBs, FRBs, TDEs) so that they can be easily gathered, analysed and modelled holistically, and not remain fragmented as present.

Linked to two corresponding JENAA EoIs (with already about 1000 subscribed scientists)





"Gravitational Wave Probes of Fundamental Physics" - a cross-cutting initiative

22 septembre 2020

Foundamental Europelhome

Accuell

Agenda

Liste des contributions

Endorse this Expression of Interest

List of Endorsers

The APPEC-ECFA-NuPECC at JENAS 2019 have recently announced a call for Expressions of Interest (Eo) in multidisciplinary projects at the interface between astroparticle, nuclear, and high-energy physics. In response to this call, we have prepared an open Eol on "Gravitational Wave Probes of Fundamental Physics".

If you'd like to endorse this initiative and be involved in further activities, please fill the form on the side of this page.

Gravitational Wave Probes of Fundamental Physics





Part of ESCAPE work programme is to work with PRACE and GÉANT

- Recent agreement is aligned with **ESCAPE** goals
- ESCAPE will collaborate on demonstrators and common aspects
 - e.g. AAI, data delivery to PRACE

ESCAPE synergies

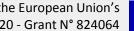
SKAO Signs HPC Agreement With CERN, GÉANT, PRACE



Eckhard Elsen (top left), Director for Research and Computing at CERN; Philip Diamond (top right), SKA Director-General; Erik Huizer (bottom left), Chief Executive Officer of GÉANT; and Philippe Lavocat (bottom right), PRACE Council Vice-Chair, signed the agreement for the new collaboration.

SKA Global HQ, Wednesday 22 July - SKAO has signed a Cooperation Agreement with CERN, the European Organization for Nuclear Research; GÉANT, the pan-European network and services provider for research and education; and PRACE, the Partnership for Advanced Computing in Europe; to overcome challenges related to the use of high-performance computing (HPC) to support large, data-intensive science projects.







ESCAPE synergies

Explore, plan and support potential ESCAPE industrial engagement

- Co-developments with digital SMEs, e.g.
 - Wavefier: real-time Machine Learning Classifier for transient signals in Gravitational Waves
 - **Gamma-Learn**: real-time Machine Learning pipeline for Gamma-ray astronomy
 - Combining ESCAPE with European Regional Development Fund programme
 - Leveraging industrial ICT cooperation schemes (within ESCAPE) ESFRI RIs)

20/11/2020



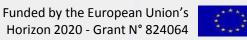


















High-priority future initiatives



The 2020 European Strategy

- A. An electron-positron Higgs factory is the highest-priority next collider. For the longer term, the European particle physics community has the ambition to operate a proton-proton collider at the highest achievable energy. Accomplishing these compelling goals will require innovation and cutting-edge technology:
- the particle physics community should ramp up its R&D effort focused on advanced accelerator technologies, in particular that for high-field superconducting magnets, including high-temperature superconductors;
- Europe, together with its international partners, should investigate the technical
 and financial feasibility of a future hadron collider at CERN with a centre-of-mass
 energy of at least 100 TeV and with an electron-positron Higgs and electroweak
 factory as a possible first stage. Such a feasibility study of the colliders and
 related infrastructure should be established as a global endeavour and be
 completed on the timescale of the next Strategy update.
- A. The energy efficiency of present and future accelerators, and of computing facilities, is and should remain an area requiring constant attention. Travel also represents an environmental challenge, due to the international nature of the field. The environmental impact of particle physics activities should continue to be carefully studied and minimised. A detailed plan for the minimisation of environmental impact and for the saving and re-use of energy should be part of the approval process for any major project. Alternatives to travel should be explored and encouraged.
- B. Particle physics, with its fundamental questions and technological innovations, attracts bright young minds. Their education and training are crucial for the needs of the field and of society at large. For early-career researchers to thrive, the particle physics community should place strong emphasis on their supervision and training. Additional measures should be taken in large collaborations to increase the recognition of individuals developing and maintaining experiments, computing and software. The particle physics community commits to placing the principles of equality, diversity and inclusion at the heart of all its activities.







The 2020 European Strategy and an outlook for the future

FCC-IS, R&D detector and accelerator and ESCAPE

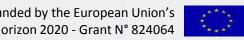
- CERN and the H2020 FCC-IS consortium explore the possibilities to leverage the ESCAPE OSSR platform to build the EOSC for the detector development activities (in line with the expected Horizon Europe work programme).
- Any digital object from R&D work should be FAIR and accessed from a single catalogue (ex.: Solid-state detector legacy data; co-creation with industries data; shared dev. with light and neutron sources community; potentially any NIM paper is concerned [...])



Environment and **Diversity/recognition**

- Attracting and respecting young brilliant minds
- -> support **credit systems**, **acknowledgement**, **reward and apply certification** for the results and contributions of all **researchers engaged** in co-operative work that allow all to reap the benefits of open science.
- A certified open archive to be exploited by any new research facility (e.g. FCC) to share innovation, practices and methods about energy/water/heat management, environmental protection, etc.

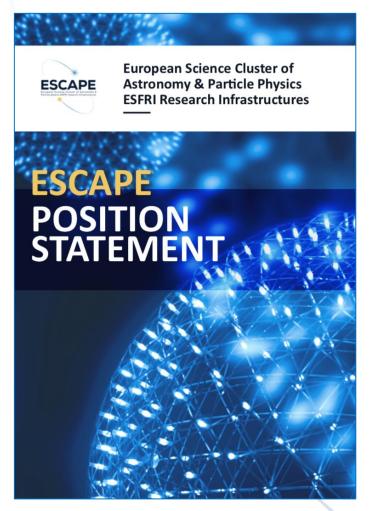






Broader synergies with other research clusters

Gathering the contributions from all RIs Directors (E-SC)



20/11/2020





Five thematic Science Clusters founded under INFRAEOSC-04-2018 (80% of ESFRI RIS)

https://zenodo.org/record/4044010#.X2oaYtaxVcs https://zenodo.org/record/3675081#.X2R2PJNLhTY

https://www.projectescape.eu/sites/default/files/Escape_position_statement_web.pdf









Some final (personal) considerations

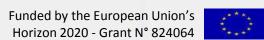
An European Union Research Ecosystem in evolution:

20/11/2020

- At present definitely more cohesion and cooperation (e.g. the five science clusters) among disciplines (also in view of the next EU strategy in research).
- Other disciplines are getting organised (often leveraging the corresponding "Science Cluster" towards an ERIC, a League or building a thematic "Community platform RI").
- "Science Cluster" scheme is a potential model of "coordinating structure", because it combines the topdown (as for the ESCAPE-SC) and bottom-up (thanks to ECFA/JENAA community consultations) approaches.
- At EC level we should underline CERN role (as our reference legal entity) for HEP and not exclusively as the reference Laboratory. At the same time we could consider a sort of "Community Platform RI" (JENAA RI?).
- Convening the Directors of our (astro. & part.) ESFRI RIs is key (e.g. ESCAPE-SC) to establish a shared view, focus on next challenges together in EU, also bridging with other "clusters" (e.g. photon and neutrons).
- The five clusters have already started debating and stimulating discussions for positioning in Horizon Europe framework.

[...]







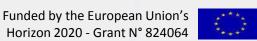
Summary

- ☐ ESCAPE brings together Astronomy, Astrophysics, Astro-Particle, High Energy and Nuclear Physics communities
 - Common interests in Exabyte-scale FAIR data management and open science
 - Objectives are science-driven (multi-messenger/multi-probe key approach) as well as commonality and synergies across infrastructure, services and tools.
- ☐ Broader synergies within a large scientific community and for innovation/society
 - Facilitate or follow up high-level cooperative agreement among flagship RIs

20/11/2020

- Test Science Projects (TSP) to enhance researchers commitment in Open Science and building EOSC by focusing on transdisciplinary scientific objectives (and supporting JENAA)
- Committing in and leveraging ESCAPE for cooperation with FCCIS (and detector/accelerator R&D data).
- ☐ Broader synergies with the other ESFRI science cluster projects towards new focus
 - All acting in concert towards the EOSC aligned goals and common interests across a broad range of European Research actors
 - New challenges, opportunities and potential focus in Horizon Europe. "Cluster tool" to be leveraged; HEP and ECFA/JENA international community to keep coordinated.







THANK YOU

