



ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

ESCAPE towards an open-science virtual research space

Giovanni Lamanna

LAPP/CNRS-IN2P3

NuPECC Meeting , 5th March 2021

<https://projectescape.eu/>



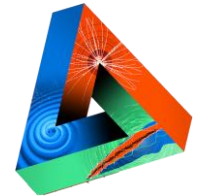
Preamble

G.Lamanna, Director of **LAPP**,
the **CNRS-IN2P3** laboratory in Annecy (FR).

● @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 connecting and at the **interface of Particle, Nuclear and Astrop. domains**



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



● **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



About ESCAPE

- Background context
- Project partnership and work programme
- Highlights, synergies and potential NuPECC interests



Background



- **Open science** for three main ambitions :

- 1) Change the way **citizens** could **perceive** research and **public investments** for research.
- 2) Enable opportunities offered by the **digital revolution** to allow everybody to **participate in the scientific process** by accessing research data.
- ➔ 3) **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.)



ESCAPE proposal in response to H2020-INFRAEOSC-04-2018 call

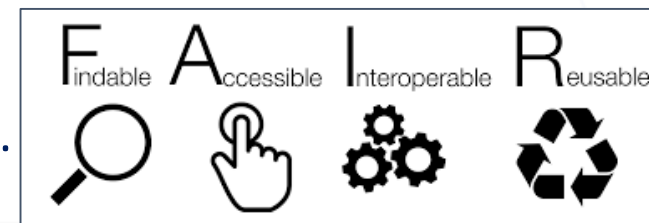
Clusters to ensure the connection of the ESFRI RIs with EOSC (and the construction of EOSC)

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 a cross-border open innovation environment.*
- *Rise the efficiency and productivity of researchers through open data services and infrastructures for discovering, accessing, and reusing data.*
- *Foster the establishment of global standards.*
- *Develop synergies and complementarity between involved research infrastructures.*
- *Adopt common approaches to the data management for economies of scale.*



Working together making data FAIR ...





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.

Indeed (as per the ESFRI vision as well as the current goal) EOSC is more as :

Enable Open (data) Science Commons





ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures



ESCAPE
ESFRI
projects,
landmarks
and a few
more RIs

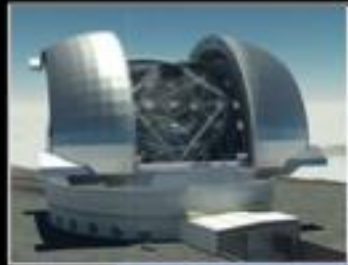
Radio



SKA

**JIVE-
VLBI**

Visible light



ELT

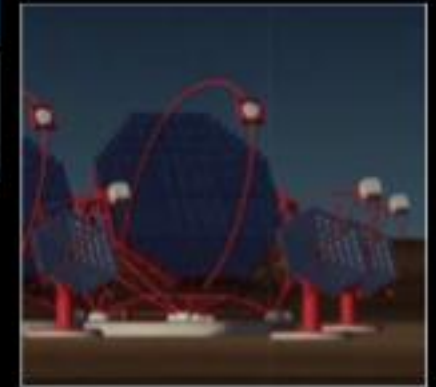


ESO



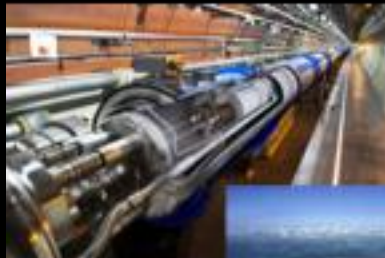
EST

Gamma rays



CTA

**Accelerator-based
Particle Physics**

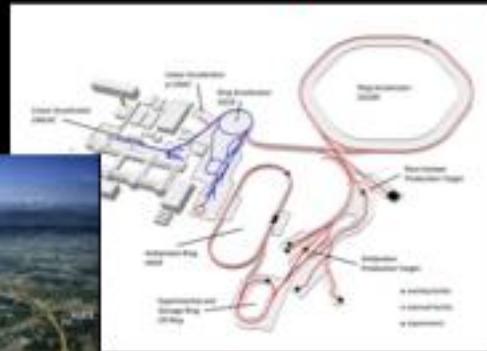


HL-LHC



CERN

**Accelerator-based
Nuclear Physics**



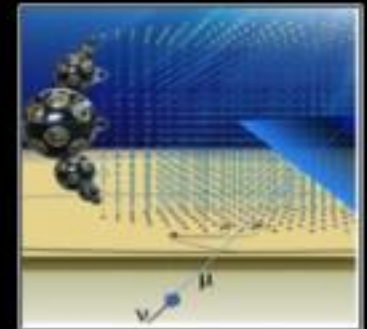
FAIR

**Gravitational
Waves**



EGO-VIRGO

**Cosmic-rays
Neutrinos**



KM3NeT

Astrophysics and Nuclear/Particle Physics: 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;
 - 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

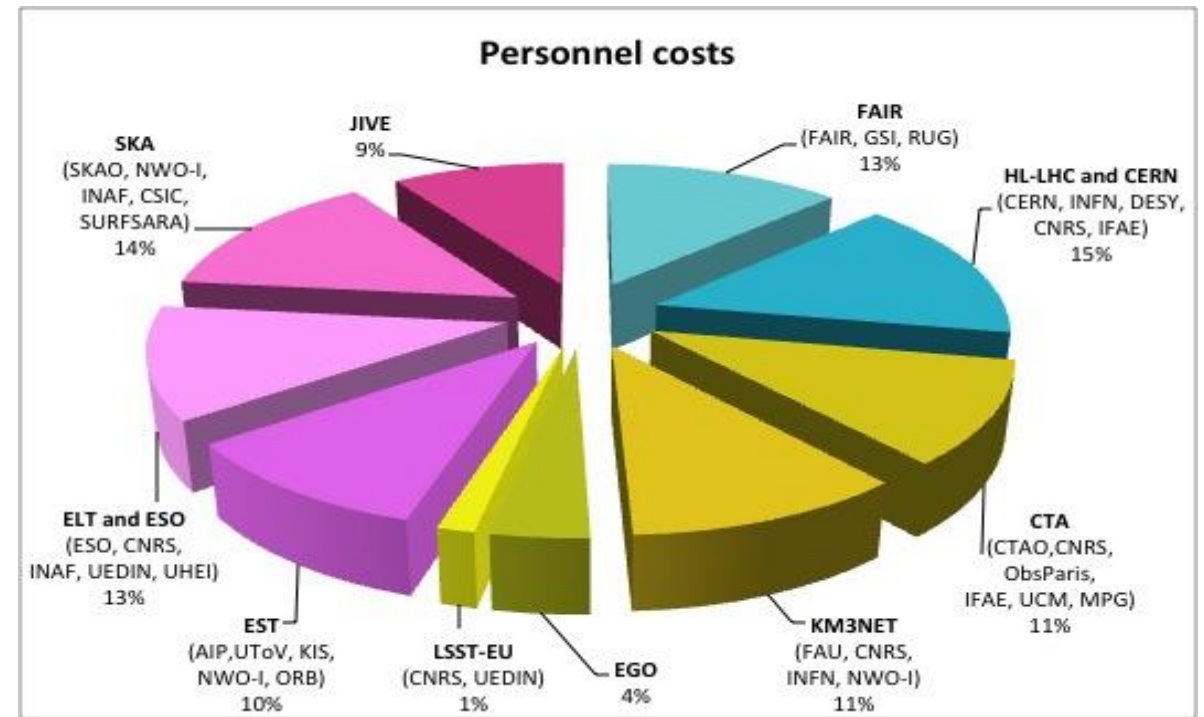


As per H2020 INFRAEOSC-04-2018 call - CLUSTER MEMBERSHIP and PARTNERSHIP:

- 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 Director 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



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



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 co-located** as it is in the GRID.



Software repository as part of the EOSC catalogue



Objectives:

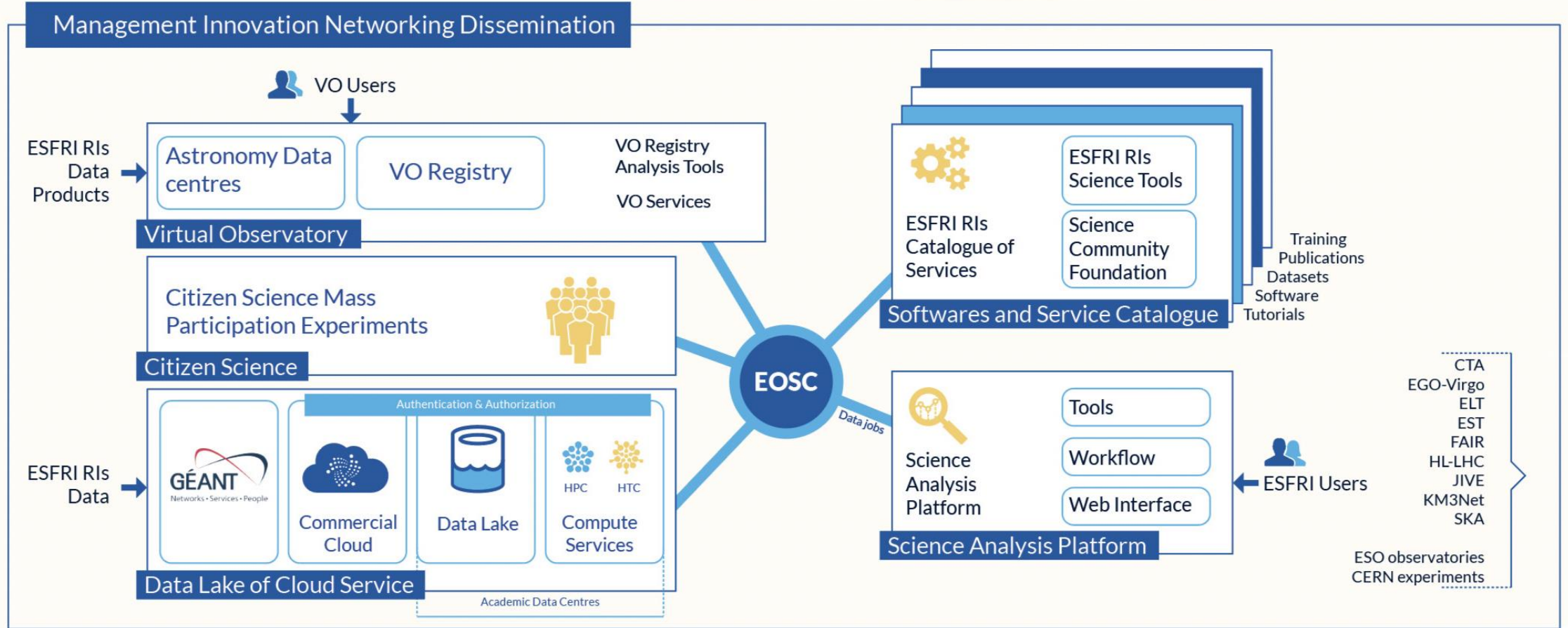
- Each ESCAPE ESFRI RI needs individually to expose and make accessible data & software (each one needs a sort of “start-kit”).
- All together they wish to adopt common solutions and offer a virtual space for interoperability and multi-messenger & multi-probe data research to next generation scientists (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



ESCAPE final goal: building a community-based EOSC cell



- “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.

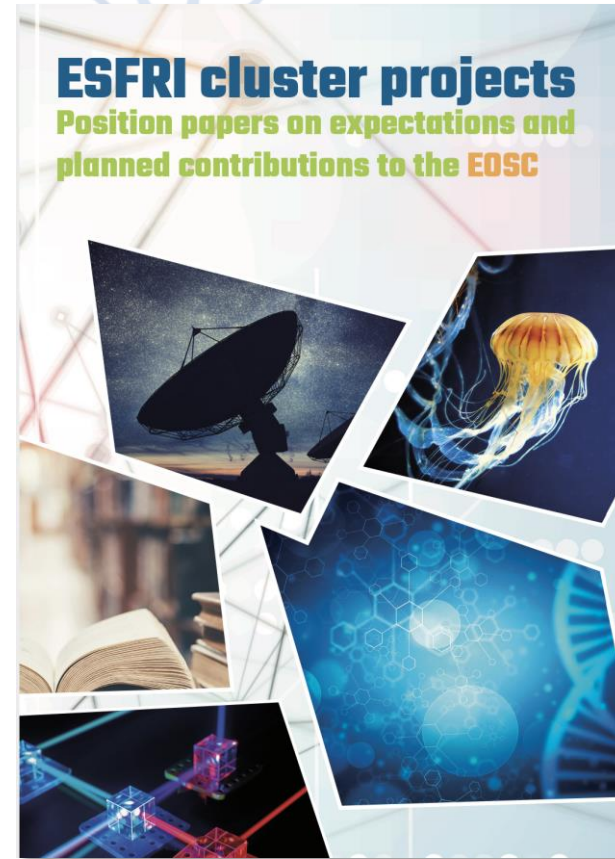
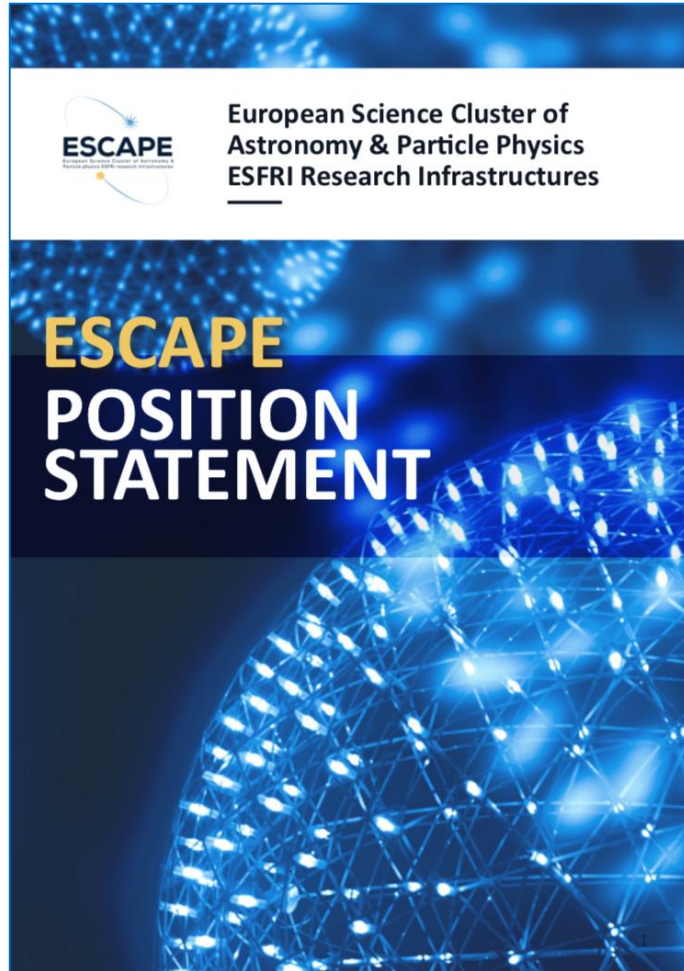


... more about SYNERGY



Broader synergies with other research clusters

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



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>

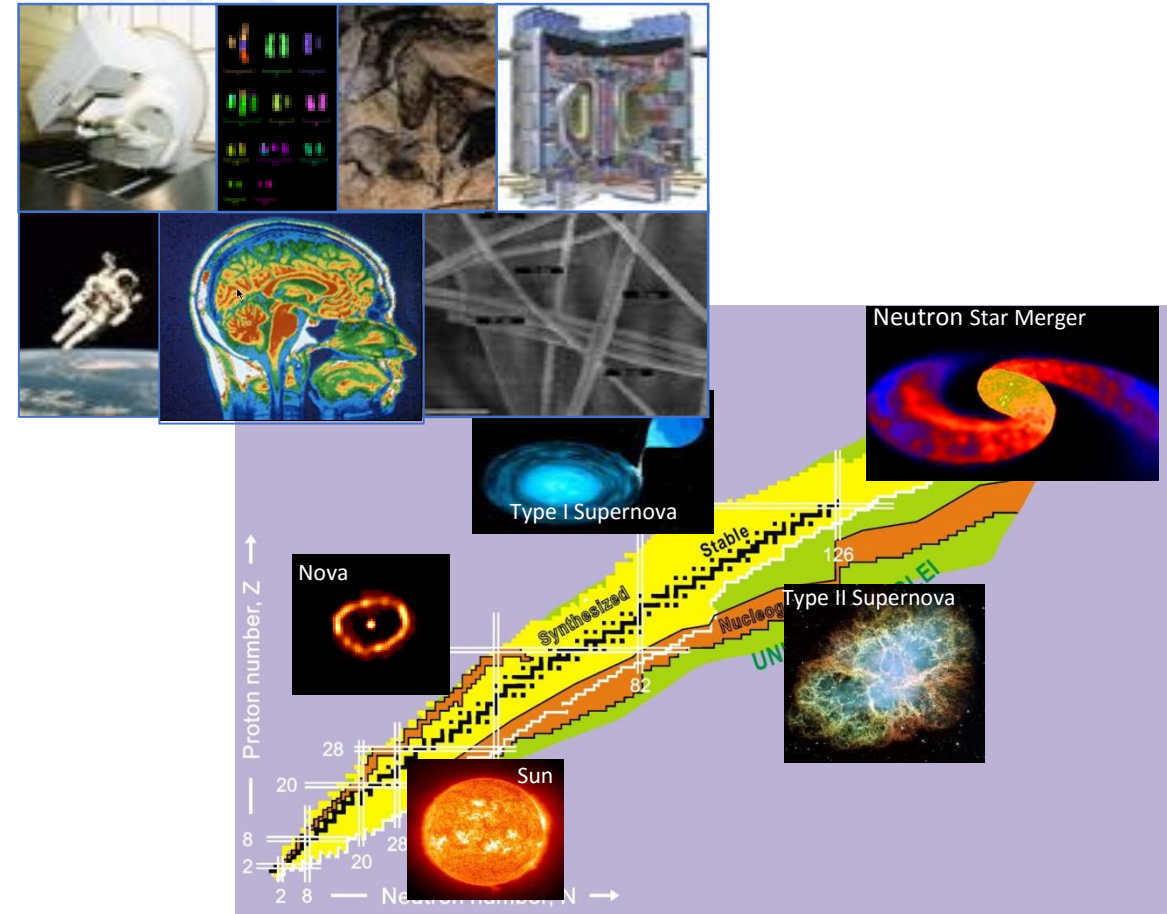
https://www.projectescape.eu/sites/default/files/Escape_position_statement_web.pdf



Open Data Science in Nuclear Physics

Open Science in Nuclear Physics (ESFRI RIs such as **FAIR-GSI**, **SPIRAL2-GANIL**, **ELI-PN**, **NICA** and **MYRRHA**) concerns the “data management policy” issue first and then capacity to leverage data for transdisciplinary research and innovation:

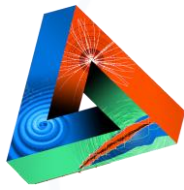
- Nuclear facilities for science: “**open facilities**” exploited for multidisciplinary research (e.g. industrial research, technology, medicine, etc.) imply embargo and proprietary data (but still F.A.I.R. data). Large innovation potential [...] (... *central metadata, standard data format, software tools (for facility exploitation and data access), services (AAI) for scheduling, archive and data embargo monitoring...*)
- Nuclear research data: fundamental research and key data for astrophysics when matter at extreme conditions is studied as the one in giant planets, stars, and during stellar explosions and collisions.
(... *data preservation (as for HEP), data access, science-ready data for “interoperability”, e.g. ApPEC and NuPECC synergies and ... scientific computing are the main topics...*)



Horizon Europe framework opportunities ?

- More NuPECC ESFRI RIs joining ESCAPE;
- Taking over domain-based data-services consolidation and operation;
- Science projects in synergy (within JENA as well as) with other communities (such as neutron sources, environment, health clusters ...);
- Facilitate and increase RIs transnational access and enhance science content in EOSC





ESCAPE for open data: FAIRness towards “VRE”

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

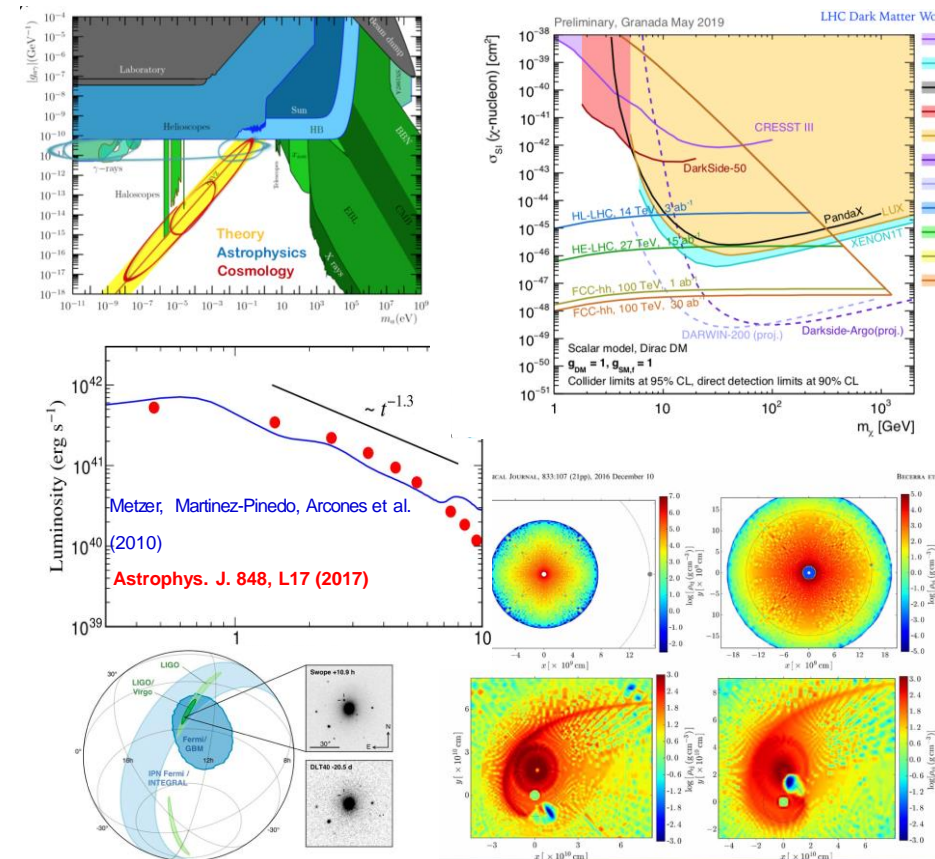
RECALL (from slide 13) 3rd pillar of Open Science :

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.
- It is now in all clusters aiming at enhancing researchers participation in open science and cross-domain scientific research.
- Included by all five Science Clusters in one more H2020 EU funding request



ESCAPE TSPs participating to the JENAA Eols

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:

- do ‘frontier’ multi-messenger science to understand extreme matter and particle processes in strongly curved space-time.
- combine astronomy and e-infrastructure and focus on data organisation
- 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 Eols
(with already about 1000 subscribed scientists)*



JENAS Eol: Initiative for Dark Matter in Europe and beyond: Towards facilitating communication and result sharing in the Dark Matter community (iDMEu)

5 décembre 2019 à 30 décembre 2020
Fuséau horaire Europe/Zurich

Rechercher...

If you would like to endorse this Expression of Interest, please use the menu on the left

Accueil	Following the call for Expressions of Interest by APPEC-ECFA-NuPECC at JENAS 2019 (attached below) for possible projects with interest spanning the high energy physics, astroparticle physics and nuclear physics community, we have drafted an open Eol on dark matter. The text is just below. If you'd like to endorse this initiative and be involved in further activities, please fill the form on the side of this page.
Endorse this Expression of Interest	
Endorsers List	

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

22 septembre 2020
Fuséau horaire Europe/Paris

Accueil	The APPEC-ECFA-NuPECC at JENAS 2019 have recently announced a call for Expressions of Interest (Eol) 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". <i>If you'd like to endorse this initiative and be involved in further activities, please fill the form on the side of this page.</i>
Agenda	
Liste des contributions	
Endorse this Expression of Interest	
List of Endorsers	

Gravitational Wave Probes of Fundamental Physics



An European Union Research Ecosystem in evolution:

- 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).
- “Science Cluster” scheme is a potential model of “coordinating structure”, because it combines the top-down (as for the ESCAPE-SC) and bottom-up (thanks to ApPEC, NuPECC, ECFA /JENAA community consultations) approaches.
- At EC level we can enhance cohesion and coordination (among ApPEC, NuPECC, ECFA), more inclusiveness (**FAIR-GSI + SPIRAL2-GANIL, ELI-PN, NICA, MYRRHA** and more RIs). At the same time we could consider to consolidate the Cluster as a (JENA+AstroNet) “Community Platform RI” to operate O.S.
- The five clusters have already started debating and stimulating discussions for positioning in Horizon Europe framework and envisaging the future of the Cluster consortia. [...]



- ❑ 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
 - 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 longer-term cooperation

- ❑ 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; NuPECC/JENA international community to keep coordinated.



THANK YOU!

ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

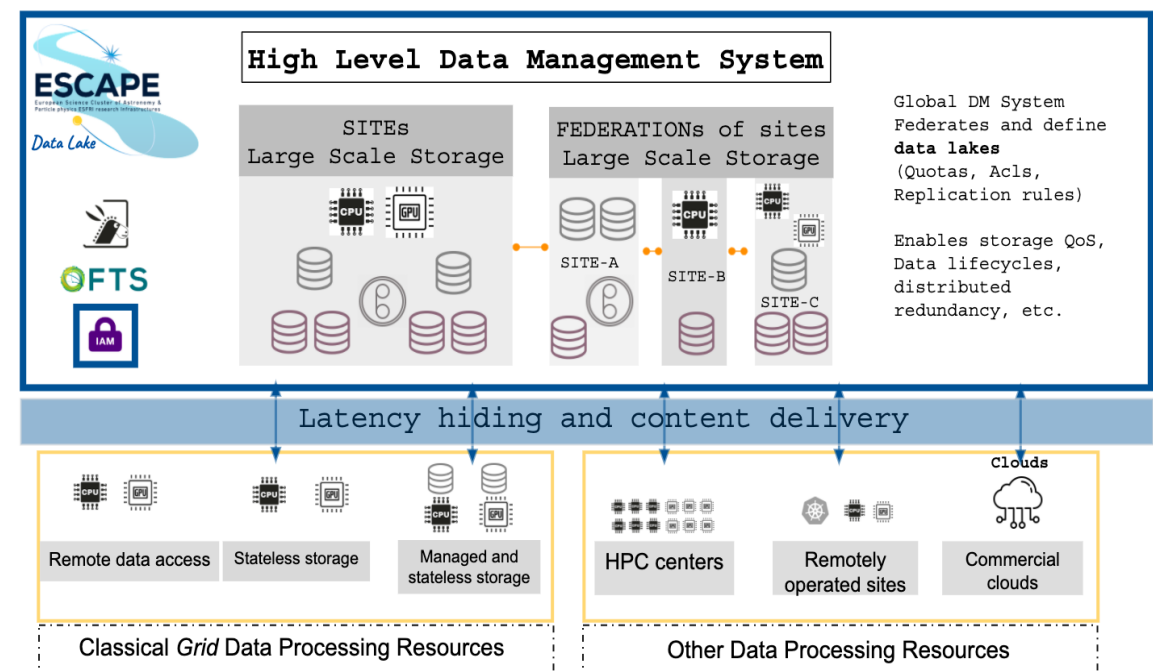


The ESCAPE Data Infrastructure for Open Science

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

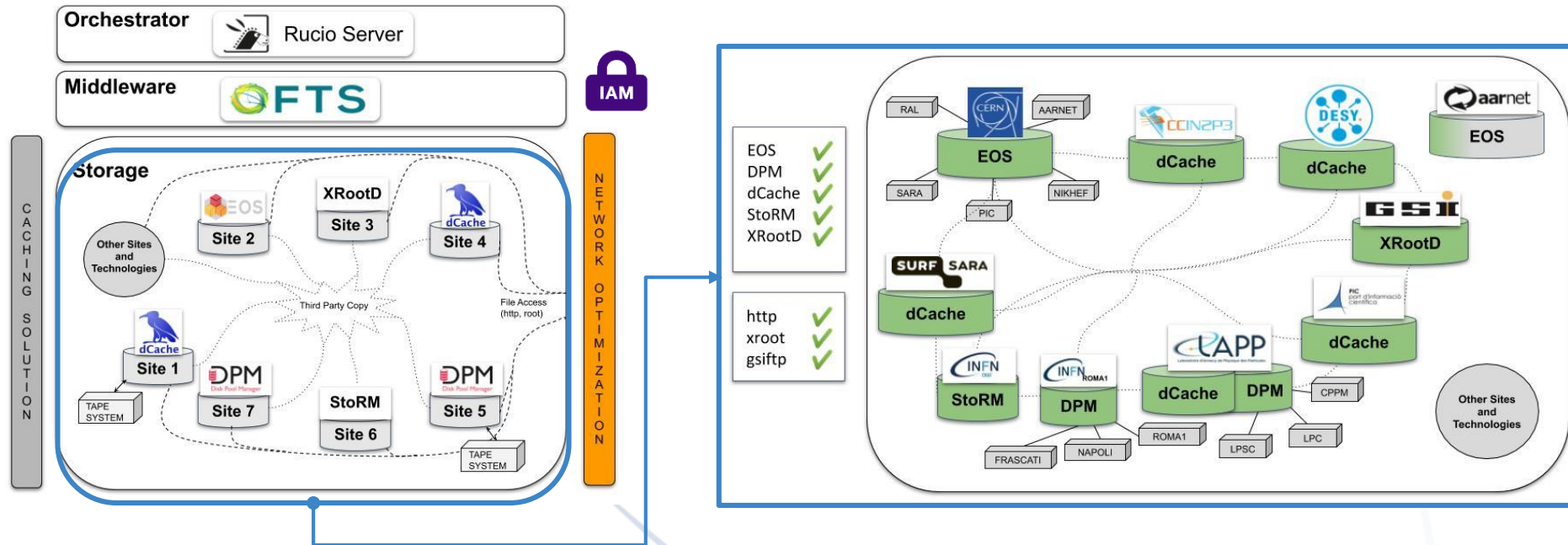


- Provides **global data management** orchestration
- Contribute to deliver **Open Access and FAIR data services**: trustable data repositories; enable data management policies; transparent data access layer.
- Science **projects to drive** the services requirements most suitable to their needs.



First achievements : a functional Data Lake pilot

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



First achievements : Science in the Data Lake

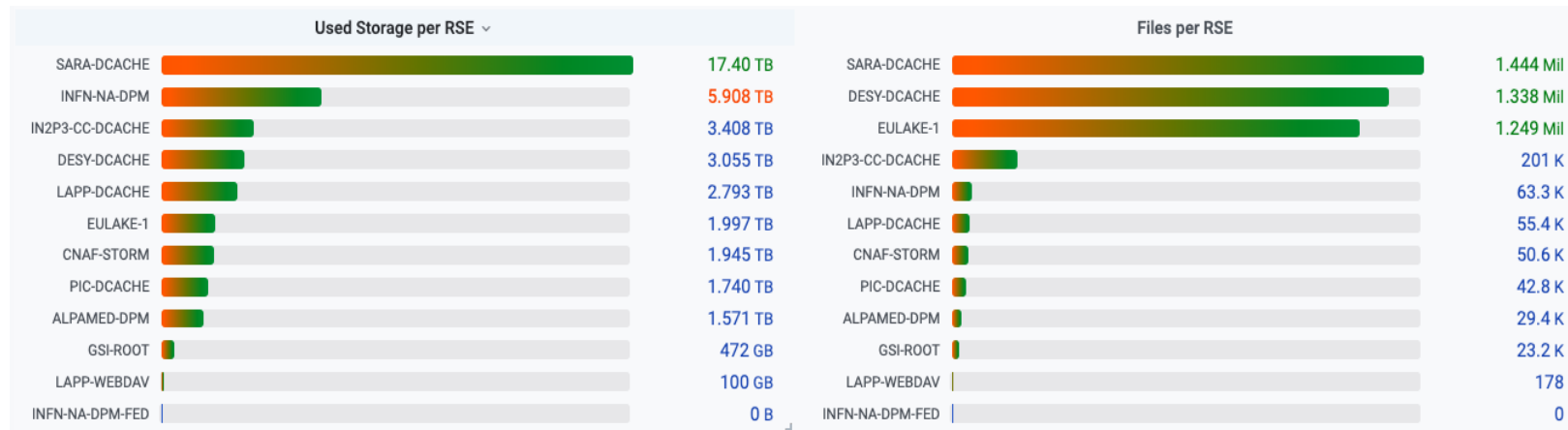
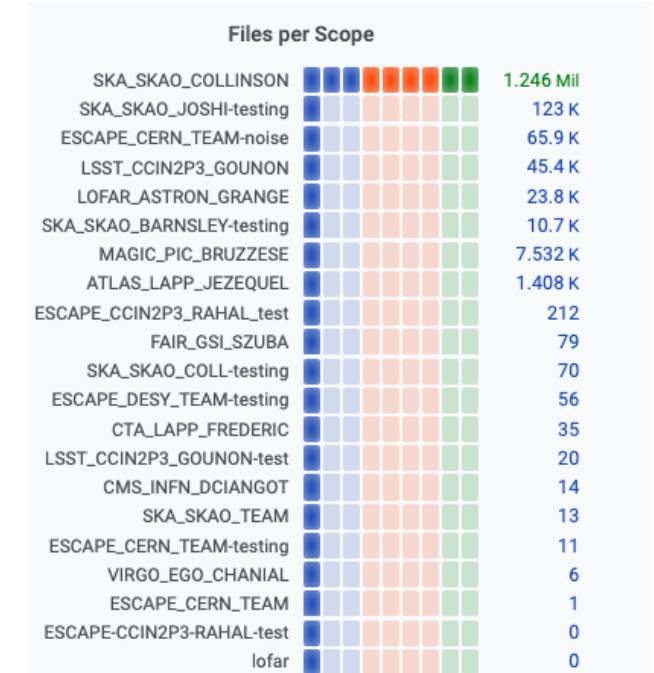
⑩ Strong involvement of ESFRI RIs and other experiments:

- Data injection within the Data Lake by:

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

⑩ Data management demonstrator from **Astroparticle, Radio-astronomy, Gravitational Waves, Cosmology and Particle Physics** communities together on a **common** data management infrastructure

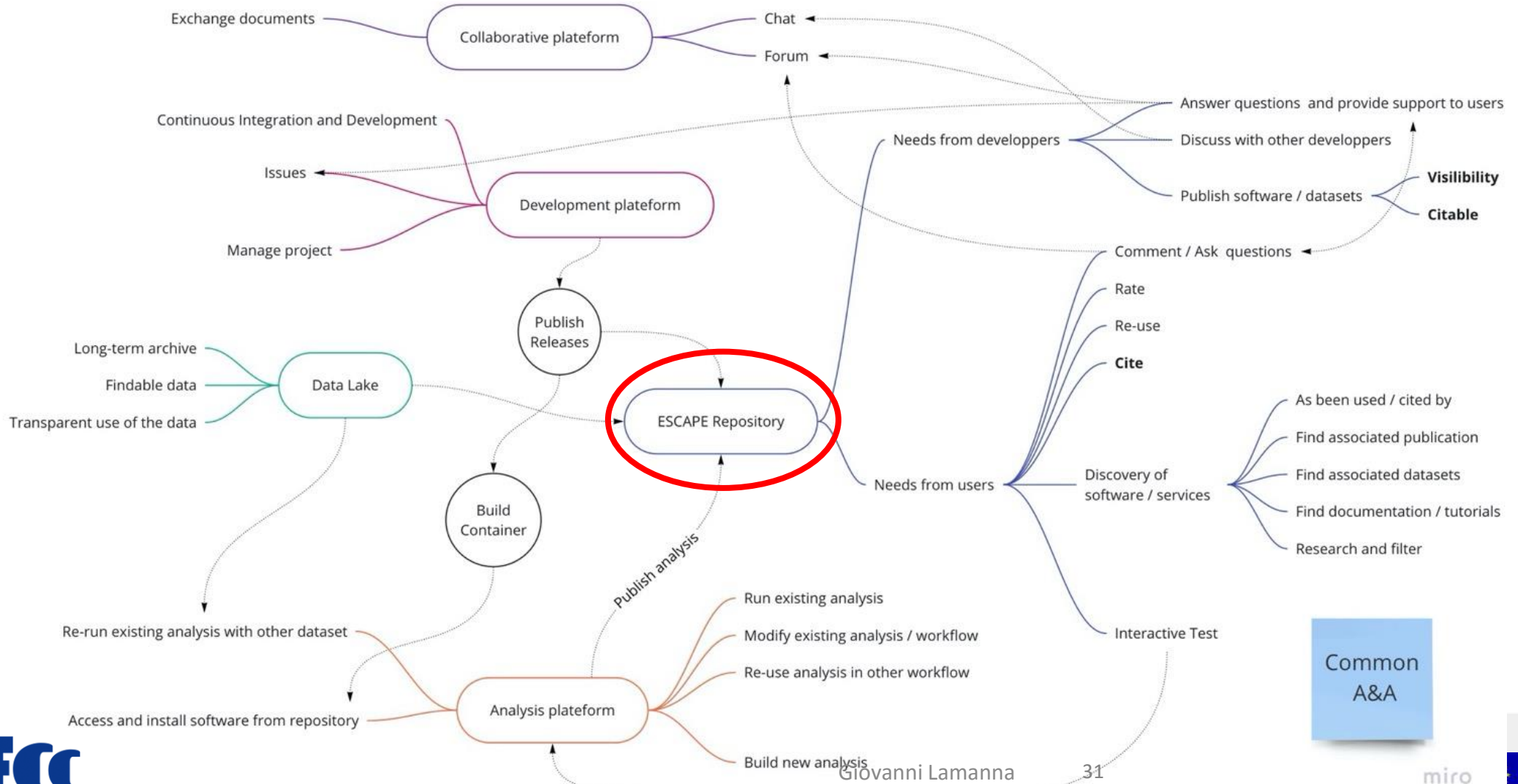
⑩ Pipeline data analysis tests currently in progress



[DEMO](#)



Building up a global Virtual Research Environment



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

From a single click



1. Make a new tag (release)
2. Let the CI do the rest

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



- Long term archived
- Findable
- Citable

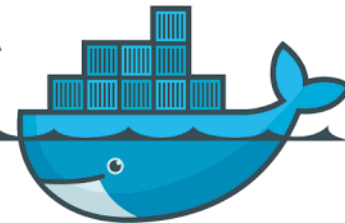


- builds a singularity image

- publishes singularity image

under dev

- builds a docker container



- publishes



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>

The screenshot shows the Zenodo project page for 'The joint-crab bundle'. Annotations with blue arrows point to various features:

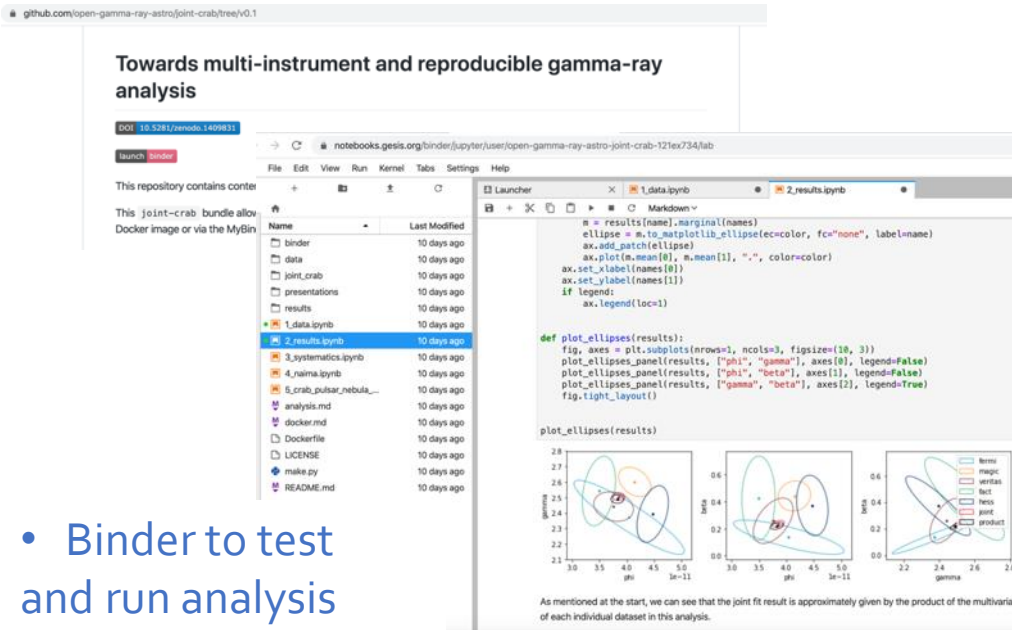
- Source code and data:** Points to the file tree on the left, specifically to the 'data' folder containing FITS files.
- Link to project and article:** Points to the 'DOI: 10.5281/zenodo.2381863' and the 'Cited by' section.
- Cite as:** Points to the 'Cite as' section, which includes the citation text and a text input field for citation styles.
- license:** Points to the 'License (for files): BSD 3-Clause Clear License' section.



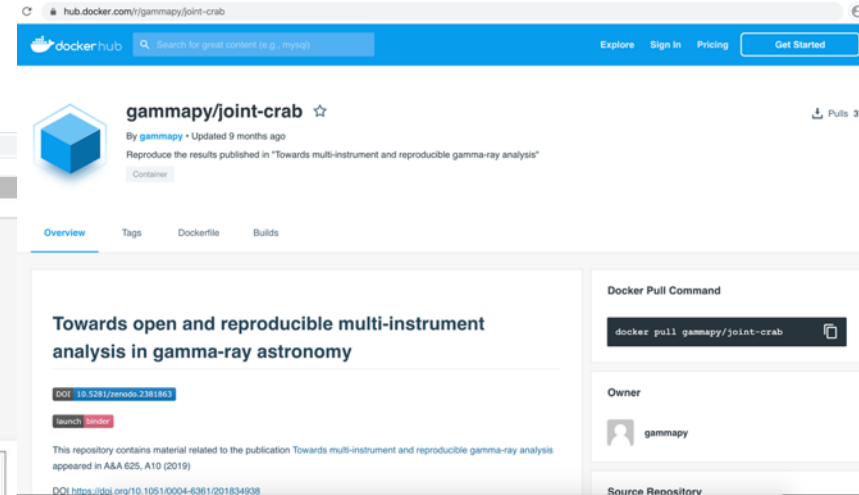
The CRAB multi-instrument gamma-ray analysis with MAGIC, VERITAS, FACT and H.E.S.S.

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

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



- Binder to test and run analysis interactively online



- Docker to ensure reproducibility



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

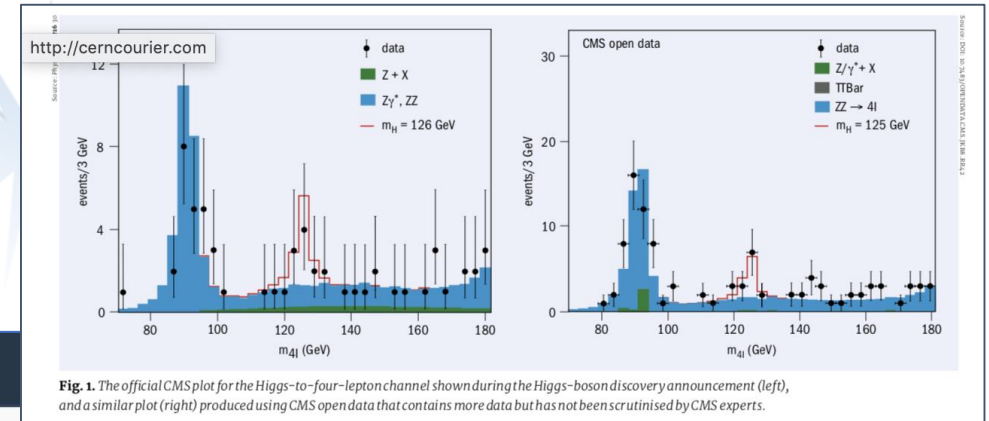


Fig. 1. The official CMS plot for the Higgs-to-four-lepton channel shown during the Higgs-boson discovery announcement (left), and a similar plot (right) produced using CMS open data that contains more data but has not been scrutinised by CMS experts.

opendata
CERN

Explore more than **two petabytes**
of open data from particle physics!

Search

search examples: [collision datasets](#), [keywords:education](#), [energy:7TeV](#)

Explore

- [datasets](#)
- [software](#)
- [environments](#)
- [documentation](#)

Focus on

- [ATLAS](#)
- [ALICE](#)
- [CMS](#)
- [LHCb](#)
- [OPERA](#)
- [Data Science](#)

