OBELICS
(**OB**servatory **E**-environments **LI**nked by common **C**hallenge**S**)

APF

Nice, 17/01/2018

Giovanni Lamanna

OBELICS

- Targeting common ESFRI-projects « Data Challenges ».
  - Generate large volumes of data.
  - Need computing resources, intensive simulation and large storage space.

- Scopes:
  - Enable interoperability and software re-use.
  - Enable open standards and software libraries for multi-messenger data.
  - Develop common solutions, share prototypes, exchange experience.

- Objectives:
  - Create an open innovation environment across ESFRI facilities.

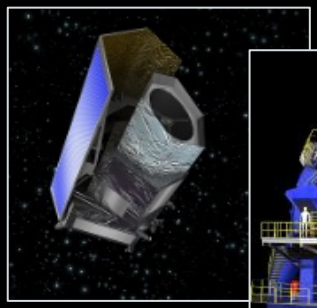
- Expected impact:
  - Economies of scale and saving resources.
  - Contribute to the construction and operation of ESFRI projects.

Radio

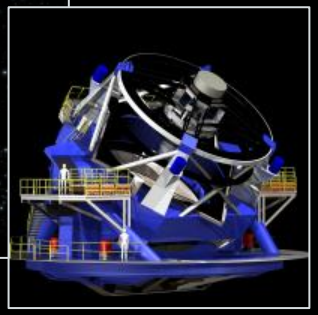
LOFAR

SKA

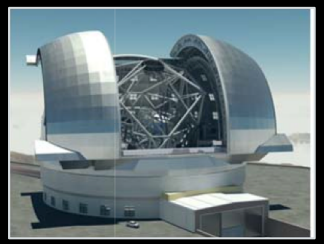
Infrared



EUCLID



LSST



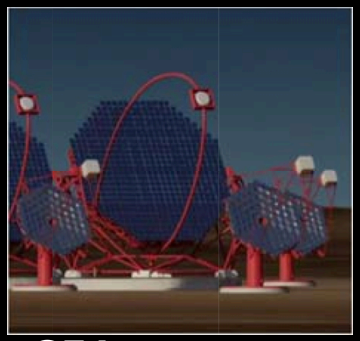
E-ELT

Visible light

X-rays

Gamma rays

HESS

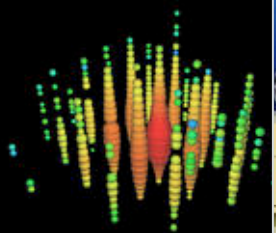


CTA

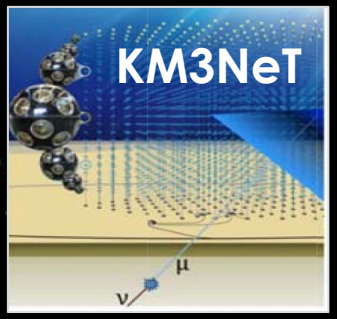
Gravitational Waves

Cosmic-rays **Neutrinos**

LIGO & VIRGO

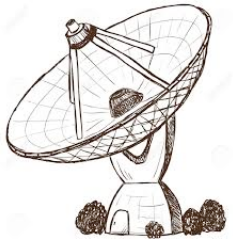


ICECUBE

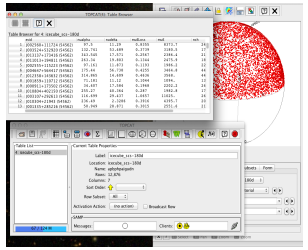


OBELICS

Working on commons along the "data flow":



Data generation → **Data systems integration** → **Data analysis**



Twelve international partners cooperating around three main steps of data pipelines of major ESFRI projects in Astronomy.

Different probes/methods/specifications

Projects	Processing	Main requirements/challenges
EVENT-BASED *(γ-rays, CR, ν)* *CTA, KM3Net …*	Evt-builder, calib. and reconstruction; reduction, real-time science.	Raw big-data (storage & HTC centres). Data formats. Algorithms. On-site operation and reduction. Cooperative science tools. Observatory (A&A). Multi-λ.
IMAGE-BASED *(far-IR, VIS)* *EUCLID, LSST …*	Surveys/deep observation; combining photometer and spectrograph info.; Catalogue of objects.	Big-data products: data base challenges. Graphical processing, Algorithms. Images format. Catalogue preservation and query. HTC centres.
SIGNAL-BASED *(Radio, GW)* *SKA, LIGO-Virgo …*	Noise cleaning; mathematical processing (FT) converting signal in images.	Algorithms. New computing architectures. HPC and HTC combined. Fast soft reduction. Data mining and preservation.

External Coordination

Industrial Cooperation

Task 3.1: Management, User engagement and data Dissemination (MAUD)

Training

Dissemination

Advanced Analysis Software Repository, Low Power Computing Platforms, Open Data format, Large Data Base solutions, A&A services, Workflows management

TASK 3.2 Data Generation & information Extraction (D-GEX)

TASK 3.3 Data systems INTegration (D-INT)

TASK 3.4 Data ANAlysis / interpreta-tion (D-ANA)

SKA | CTA | KM3NeT | LSST | EUCLID | LOFAR | EGO | E-ELT

OBELICS & multi-λ

Four fields where OBELICS scopes can have links with multi−λ policy

1) multi−λ as a potential source of fragmentation and where OBELICS helps proposing shared (digital) solutions.

2) multi−λ as an implication of openness; OBELICS builds a framework for open-source solutions for data archive, access and analysis in a big-data context (multi-RI scheduler; simulations and modelling).

3) multi−λ as a severe time constraint for efficient follow-up analysis. OBELICS produces transversal HPC programming solutions for fast data calibration, compression, transfer and reconstruction + ML ...

4) multi−λ as a motivation for: i) data quality assurance for which RIs are the unique subjects; ii) innovation in digital standards and digital infrastructures (archive); iii) large "community" adoption of programme languages and methods for workflows and visualisation.

Three sub-domains to leverage final
OBELICS expected achievements and to build an
"Open Data multi-lambda and cross-domain research environment".

1) Data sharing

- Solutions for heterogeneous and interoperable regional/ national computing and data resources …
- Open-source software repository. Sharing and co-developing scientific software is critical. Provision of open science tools; services for multi-wavelength quasi-real-time follow-up analysis; modelling and simulations …

2) Data access

- "Open Access and user-support" (for data and software)

Ex.: A "portal" implementing the access to data to combine and to analyse it, visualise the results and implement new workflows.
-> Policy for sharing to be subscribed by the RIs

3) Data re-use

- Preservation solutions (for data, software and framework) and long-term sustainability.

[…]

Conclusions or last considerations:

- New RIs are: Big Data generators, precision measurements, and plethora of results expected in the future.

- New needs and expectations from next generation researchers.

- Policy, community networking, computing infrastructures, scientific software and data services are multi-layers for the implementation of a multi-messenger strategy for the future.