



# Open Science in KM3NeT

Jutta Schnabel

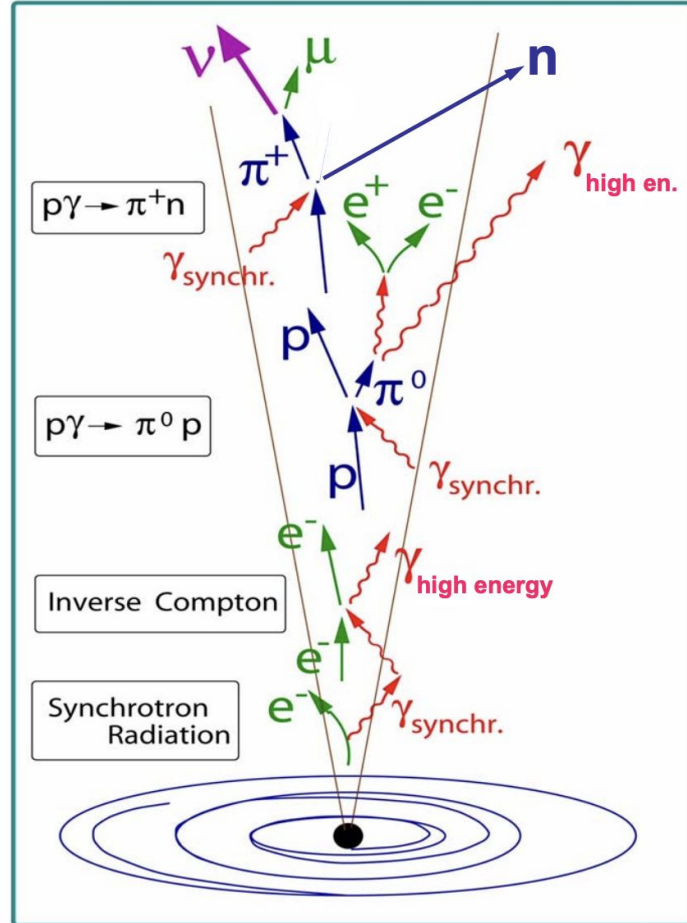
EPAP Seminar, King's College London

Erlangen & London, 04/11/2022

# KM3NeT - Building the instrument

# Neutrino generation and sources

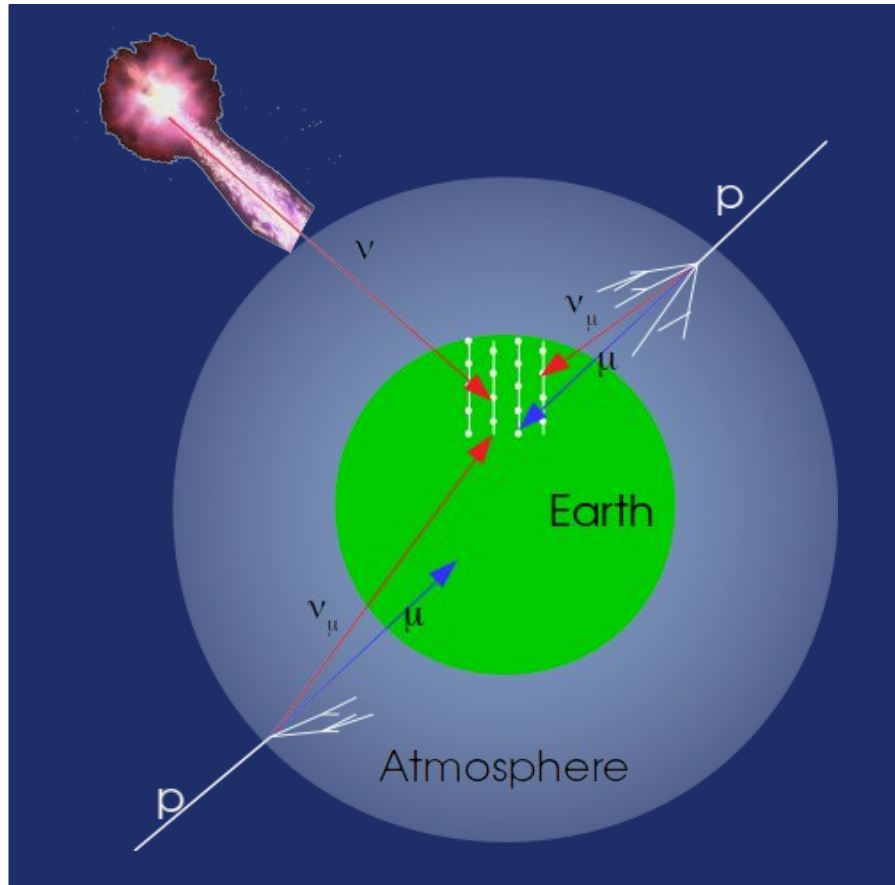
Cosmic, atmospheric or anything else



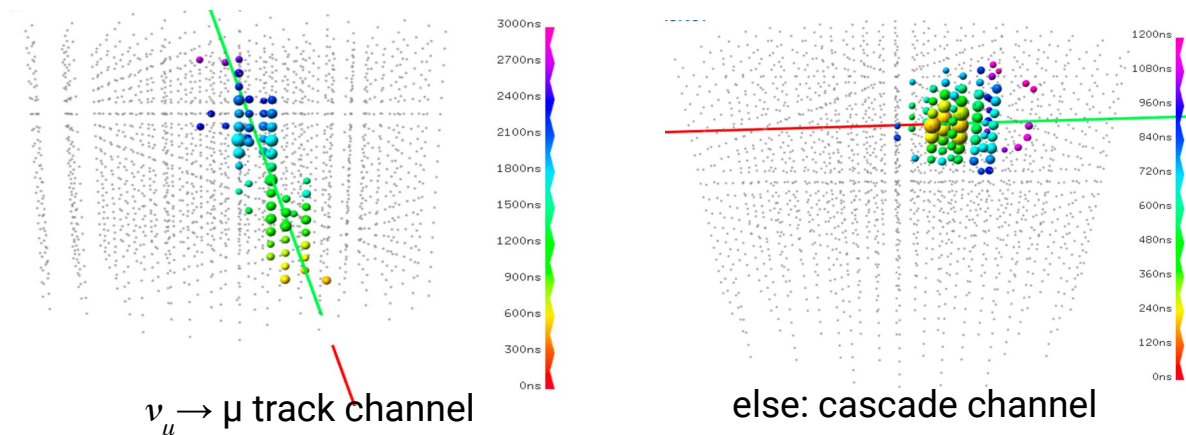
- in cosmic sources
  - classical  $\pi^{+/-} \rightarrow e^{+/-} + \nu_e \nu_\mu \bar{\nu}_\mu$
  - acceleration mechanisms for high energies
  - flavour composition 1:1:1 at detection
- in atmosphere
  - from cosmic ray interaction
  - background for cosmic search, signal for neutrino property study
- other interactions, e.g.
  - dark matter: WIMP annihilation
  - ...

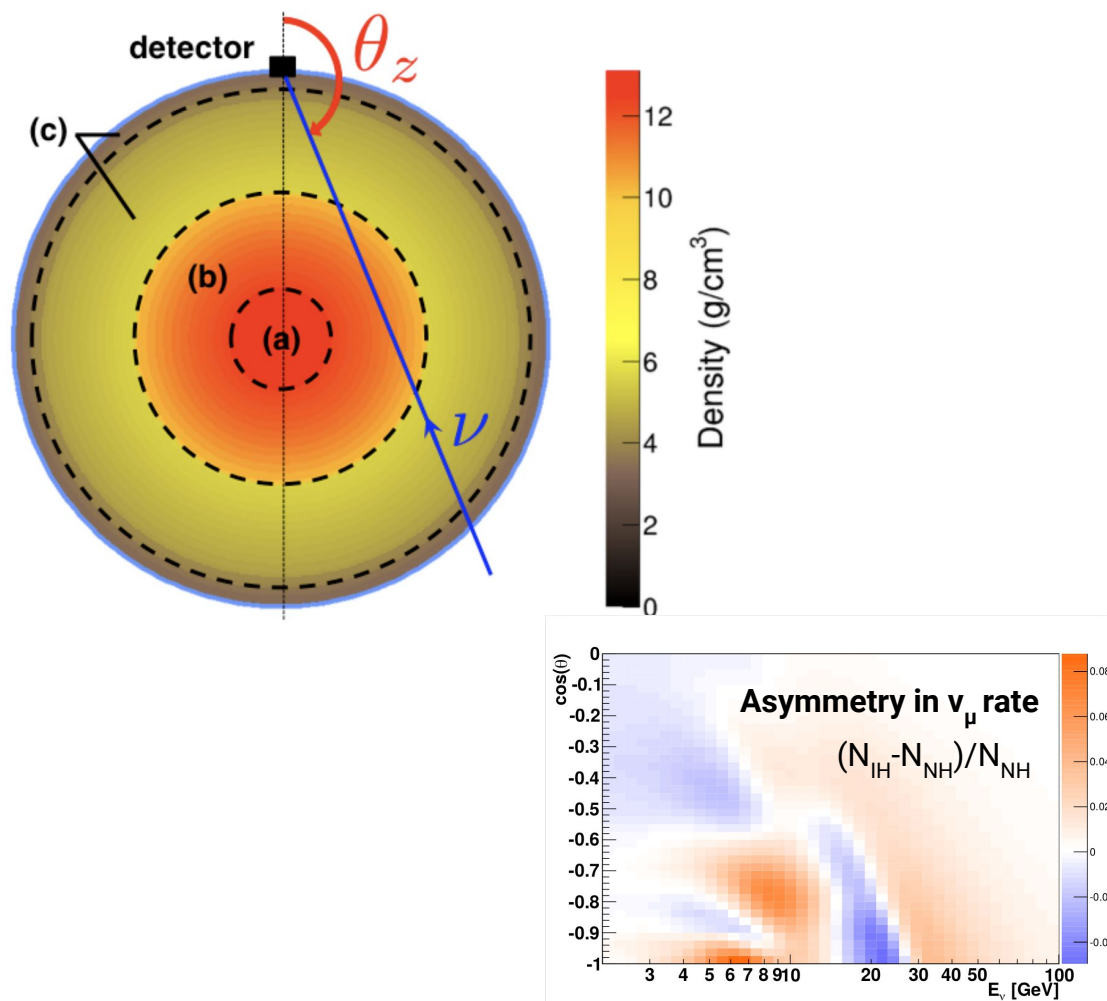
# Detection principle

Cherenkov radiation from neutrino interaction secondaries



- neutrino interaction producing high-energy charged leptons
  - Cherenkov radiation allows directional reconstruction
  - radiative processes allows energy reco
- Earth used for shielding of atmospheric muons
- resolution
  - track channel: better than  $0.1^\circ$  for  $E > 100 \text{ TeV}$
  - cascade channel: better than  $2^\circ$

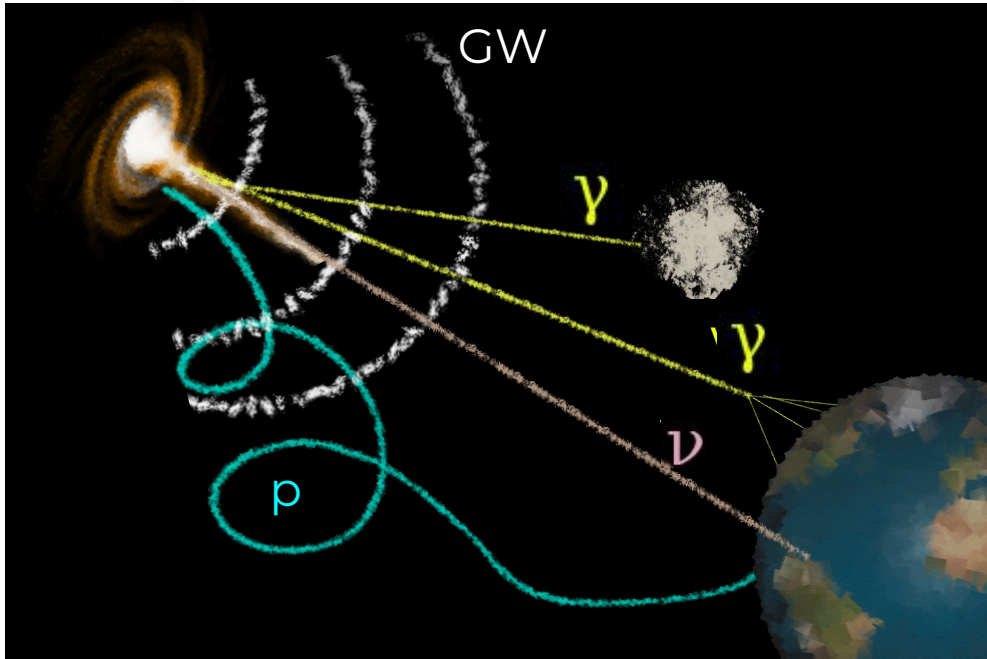




## Oscillation Research with Cosmics in the Abyss

- **signal:** neutrinos produced by cosmic rays in atmosphere
- few-GeV energy range
- neutrino properties through oscillation studies
  - mass hierarchy
  - sterile neutrinos & other
- characteristic patterns of neutrino appearance/disappearance at different energies/pathlength

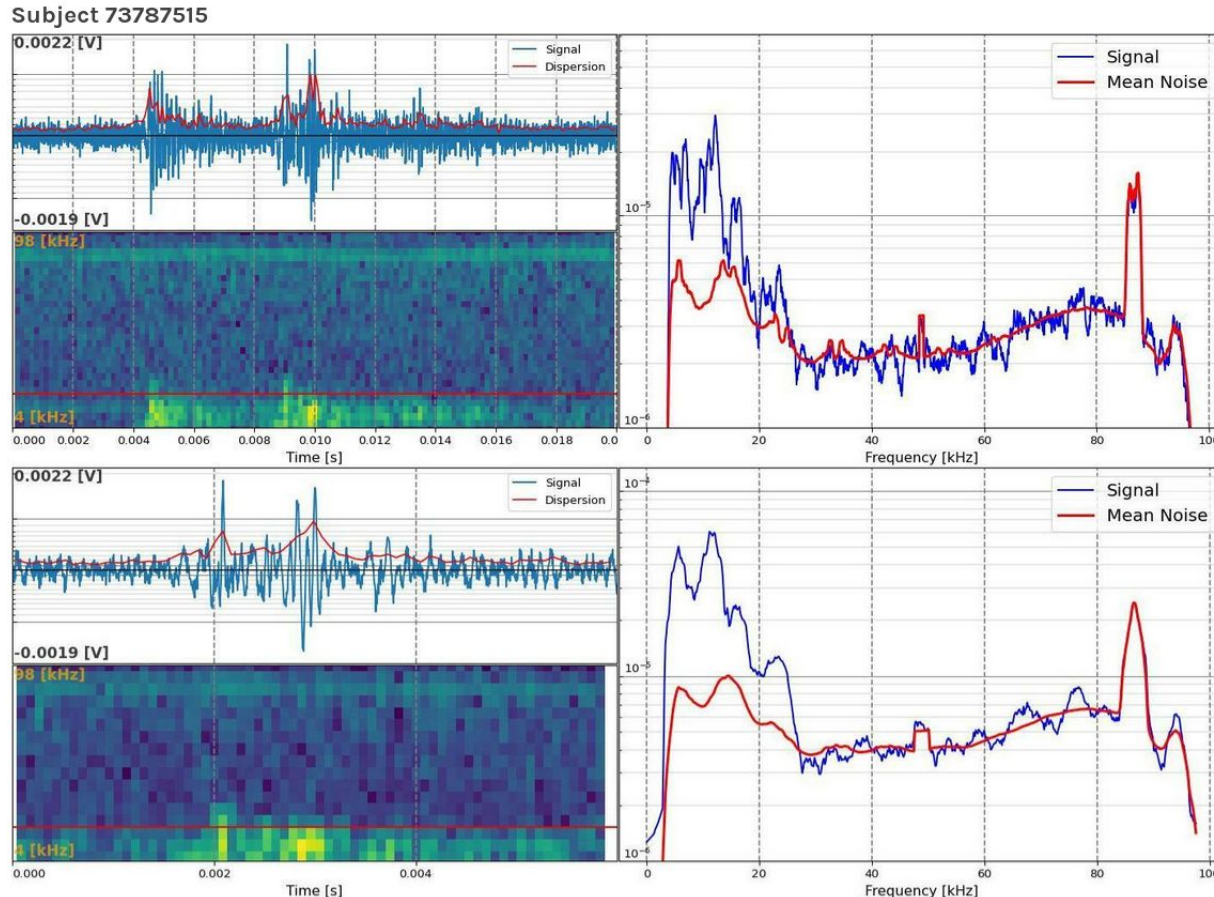
## Astroparticle Research with Cosmics in the Abyss



- superior angular resolution
- decides origin, energy spectrum, flavour composition
- 87% of the sky mapped including the Galactic Center
- energies between  $O(10 \text{ GeV})$  -  $O(\text{PeV})$

# Extra: Earth and sea science

Using the infrastructure for a wide science program



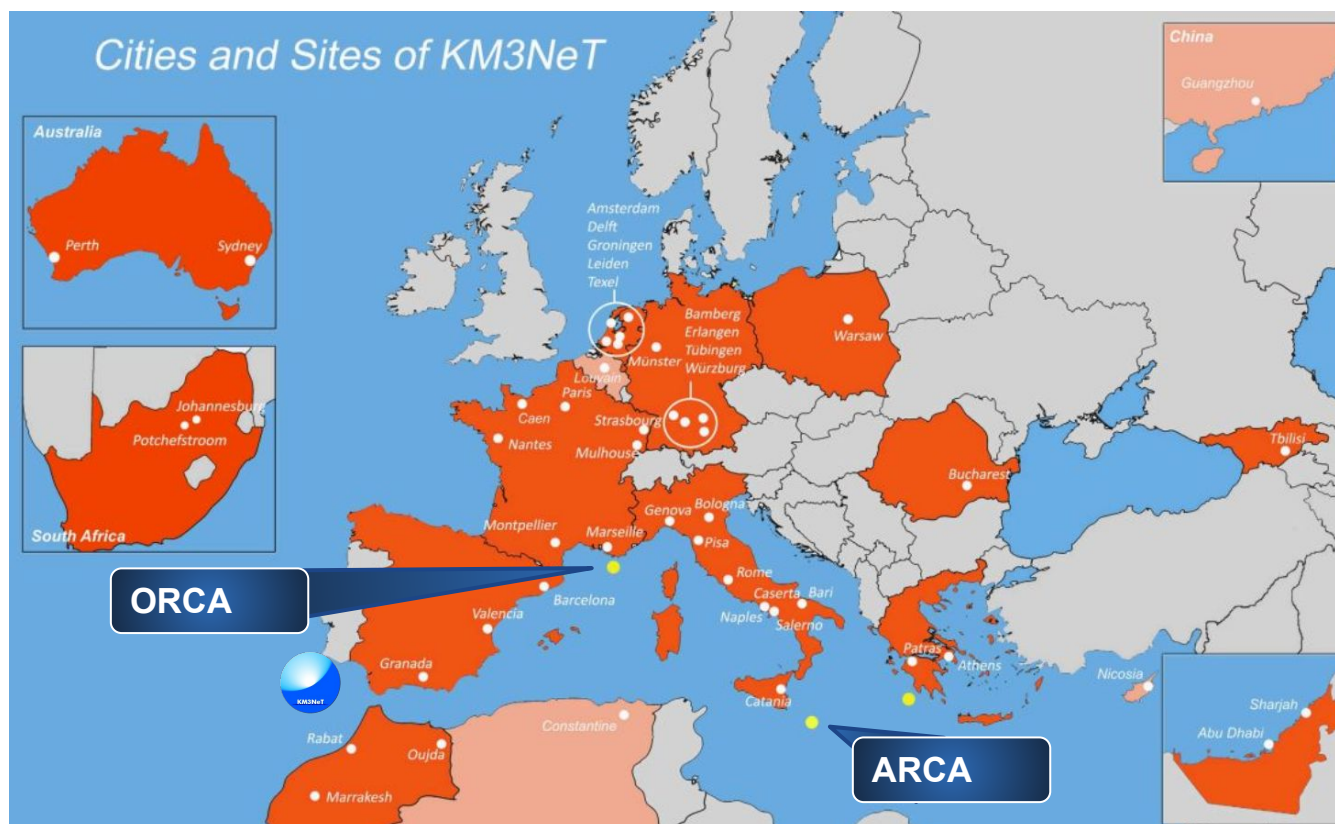
[zooniverse project](#)

- access for Earth and Sea science community
  - dedicated instrumentation installable on Detection Units
  - or use of standard equipment of detectors
- standard equipment
  - water optical and oceanographic properties
  - behaviour of bioluminescent organisms
  - measurement of sea currents
  - identification of acoustic noise sources
- collaboration with the European Multidisciplinary Seafloor Observatory (EMSO)

# Who is KM3NeT?

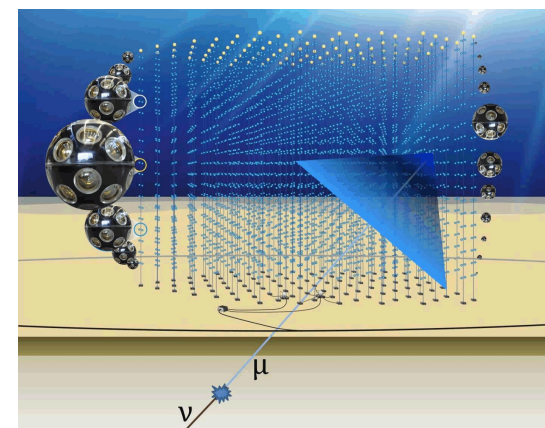
The detectors and collaboration

LoI: [arXiv:1601.07459](https://arxiv.org/abs/1601.07459) [astro-ph.IM]



## KM3NeT

are Water Cherenkov detectors for high-energy neutrinos in the Mediterranean Sea, under construction

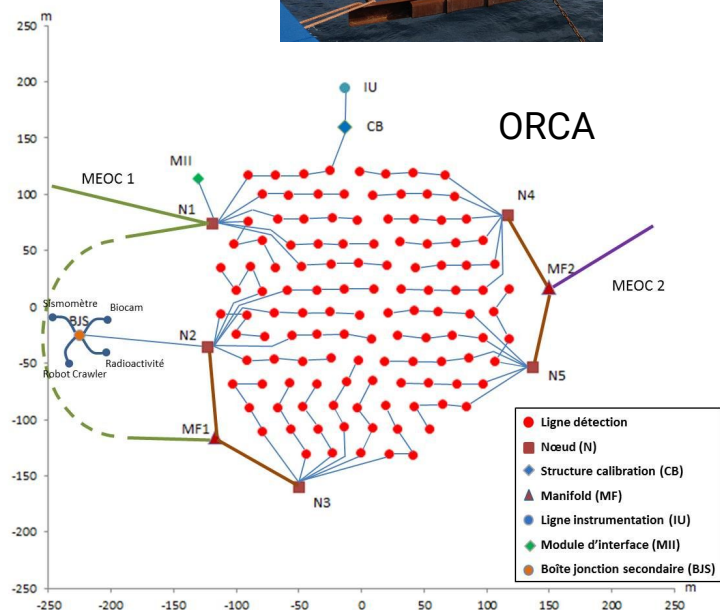
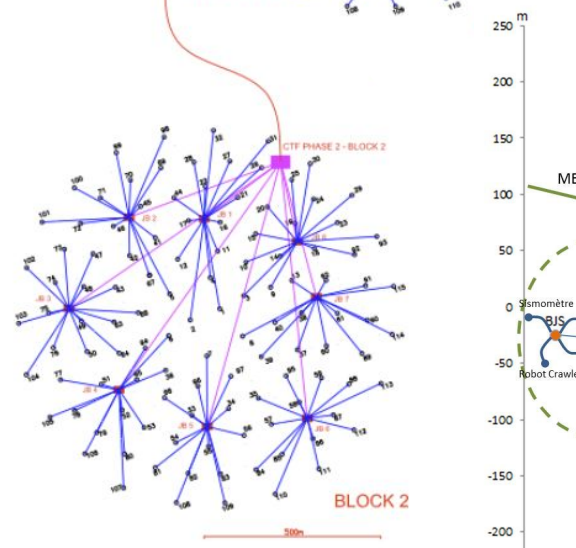
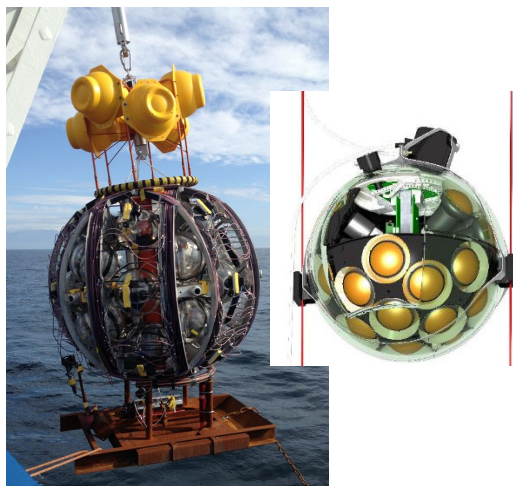
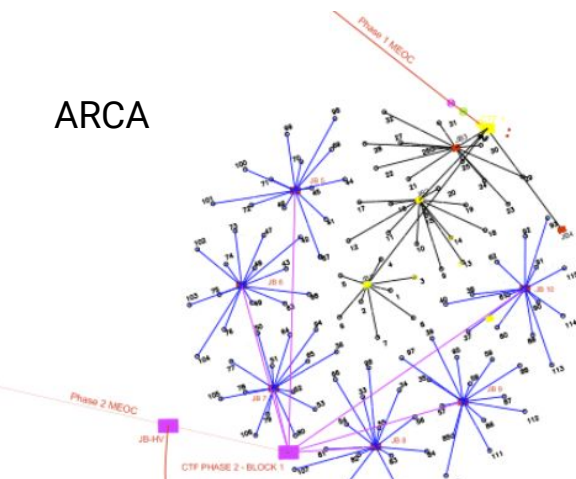




# Deployment and installation

Same implementation, different detectors

ARCA

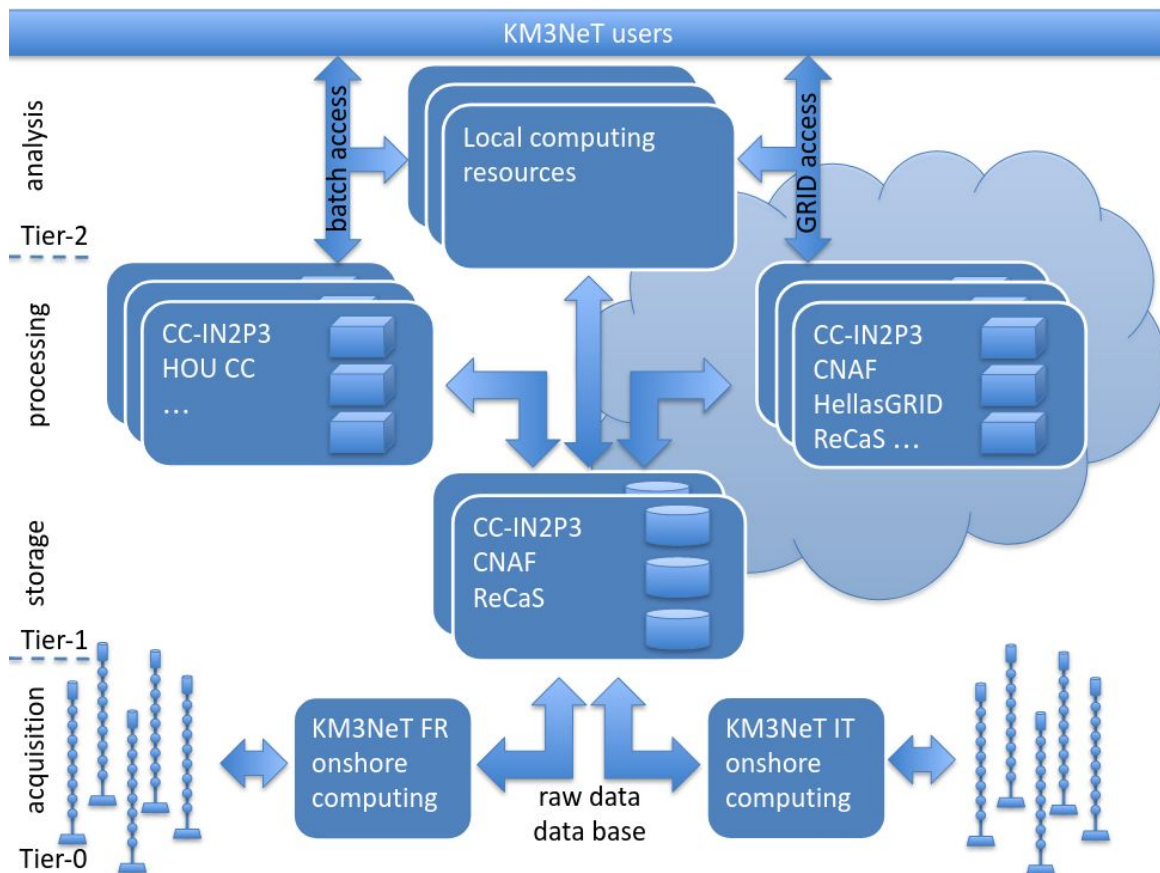


- Detection Units (DUs) containing multiple Digital Optical Modules (DOMs)
- anchored to the sea bed, uncoiled using a launcher vehicle to place DU within 1m accuracy on floor [S. Aiello et al 2020 JINST 15 P11027](#)
- several DU deployments achievable during a sea operation
- connection to floor network using ROV
- readout of all data to shore

	String Spacing (m)	DOM Spacing (m)	Depth (m)	Instrumented mass (Mton)
<b>ORCA</b>	20-23	9	2470	5-8
<b>ARCA</b>	90	36	3400	500*2

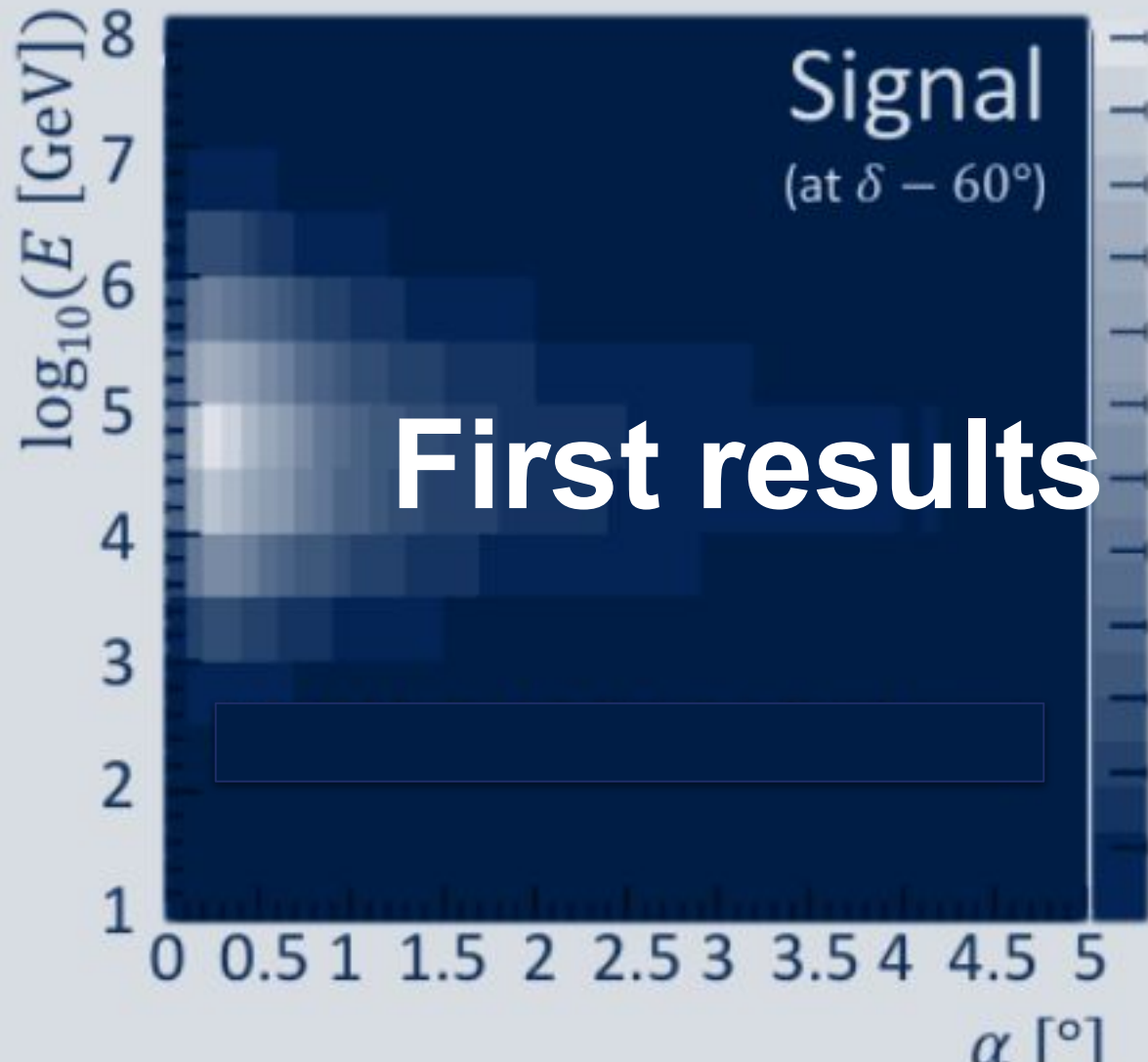
# Data acquisition and computing

Big data: high volume, parallelized, distributed

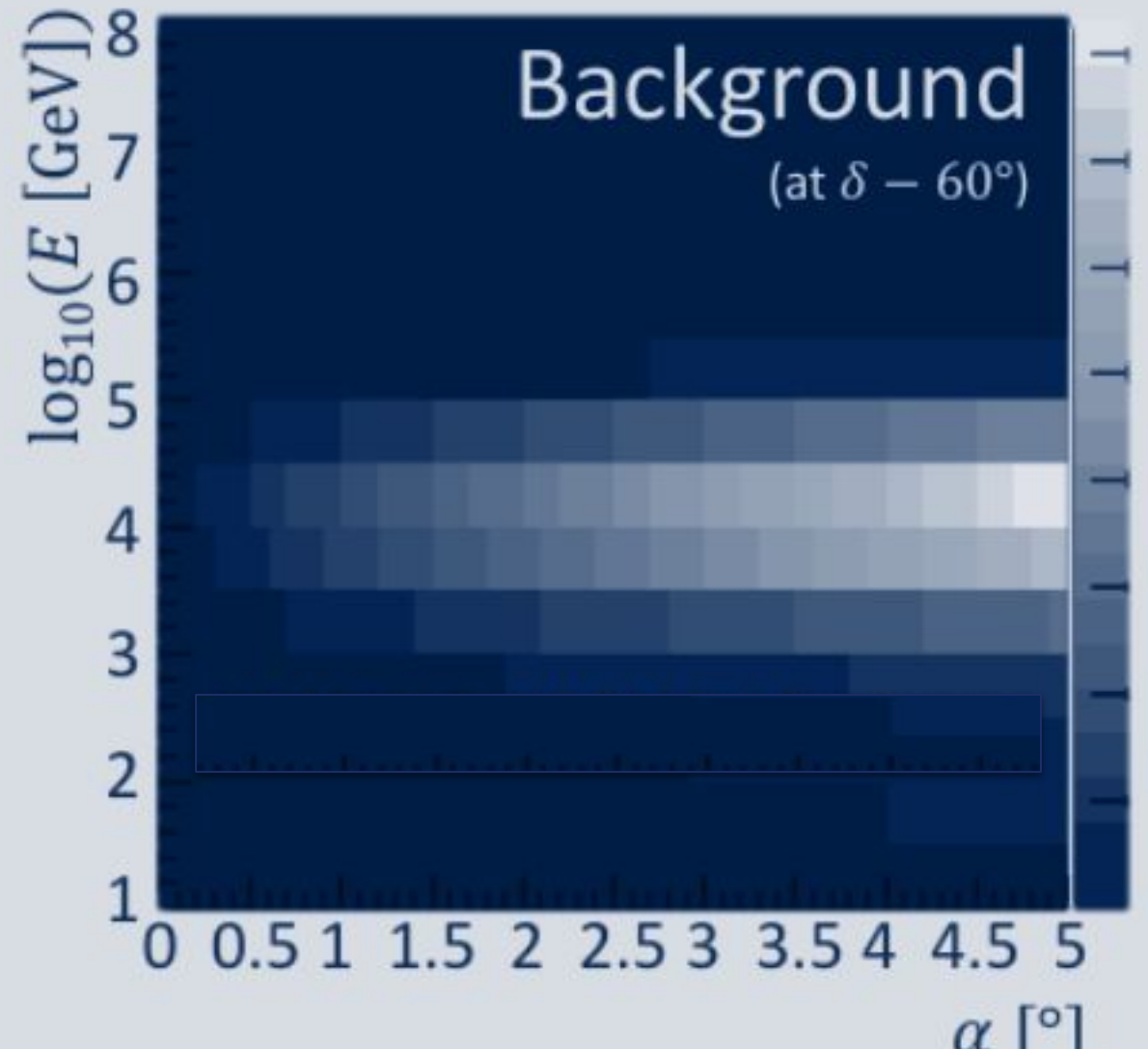


- all digital data from the PMTs sent to shore
  - processing on computing cluster in real time
  - full MC simulation of data processing starting from neutrino generation
- physics events filtered through triggering software
  - different filters can be applied to the data
  - for full building block about 25 Gb/s reduced to 1/100
  - storage for further processing
- model based on the LHC computing model
  - hierarchical data processing system
  - distributed & parallel computing

[KM3NeT-InfraDev-WP4-D4.1](#)

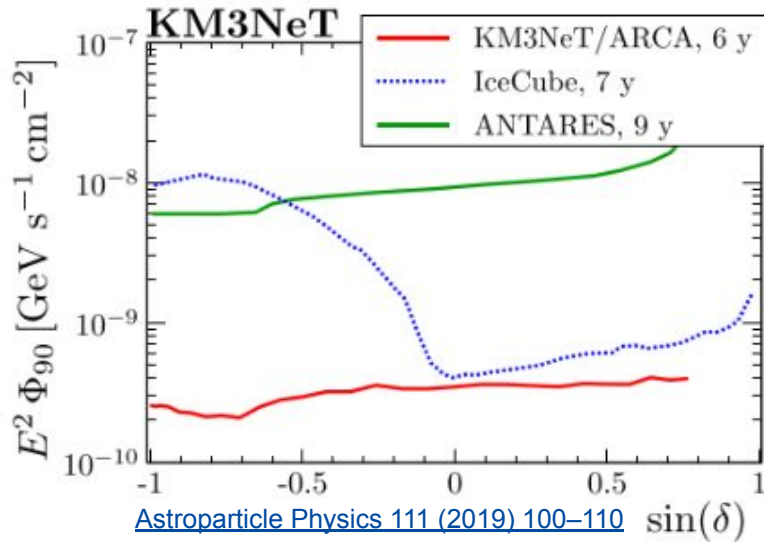


First results

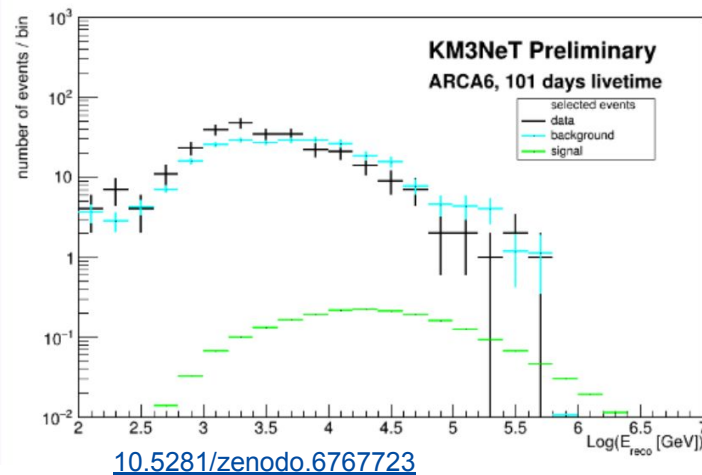


# All-sky neutrino flux and point sources

Cosmic neutrino searches from point-like to very extended sources

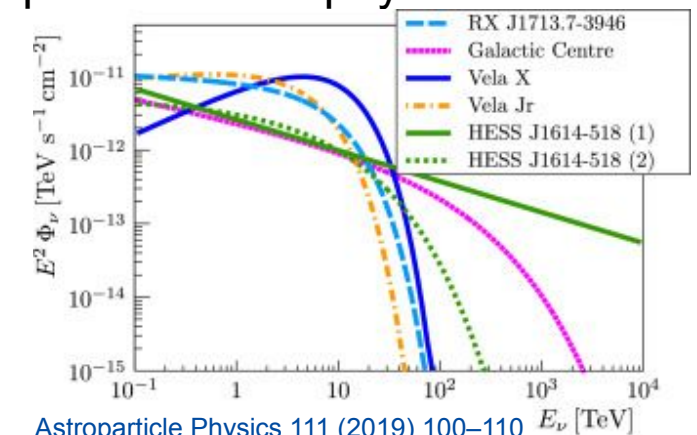


energy spectrum of all-sky cosmic neutrino event selection



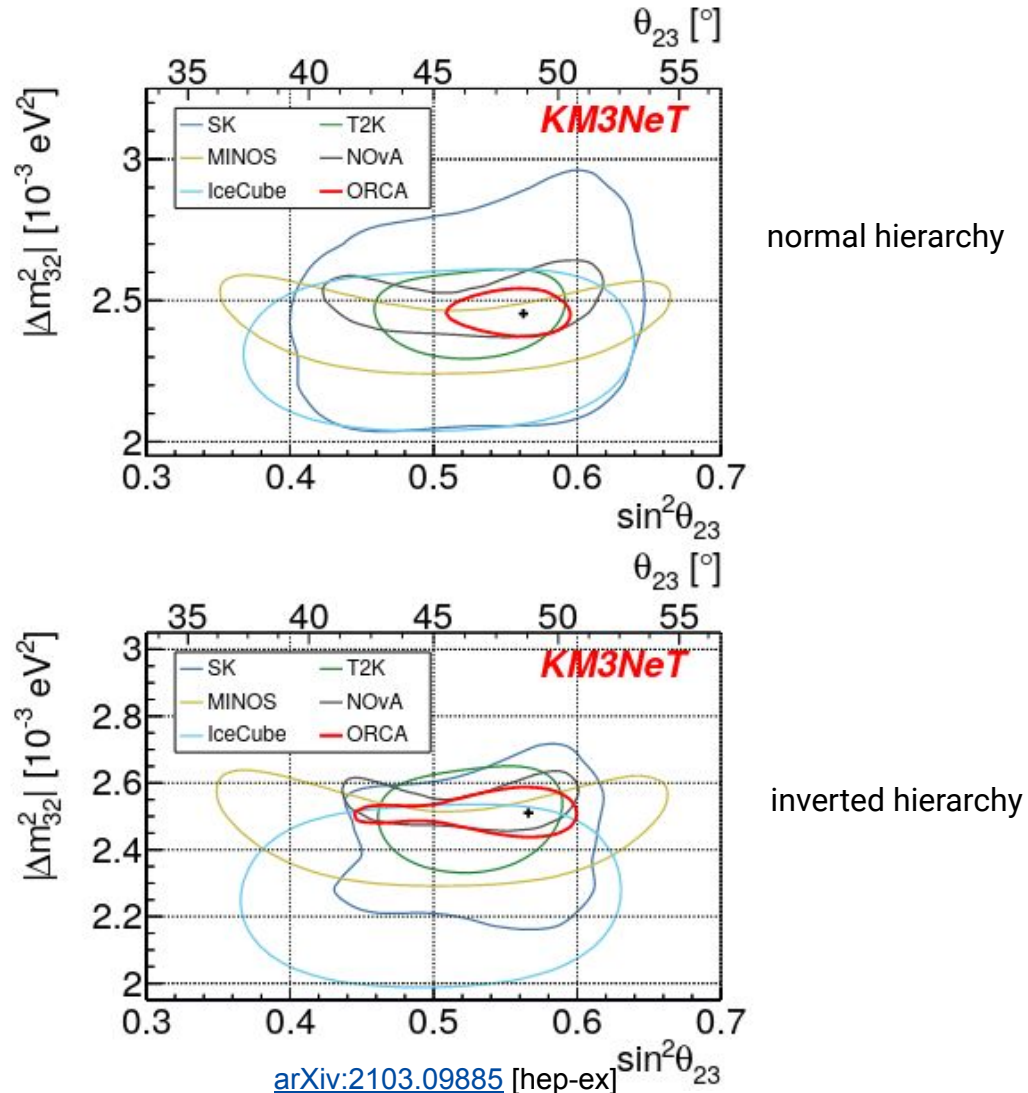
energy spectrum of all-sky cosmic neutrino event selection

- all-sky neutrino flux (ARCA 101 days)
  - $\Phi_{90\%CL} = 17.3 \times 10^{-18} \text{ GeV}^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$
  - $\Phi_{\text{test}} = 1.44 \times 10^{-18} (E/100\text{TeV})^{-2.28} \text{ GeV}^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$
  - observation with  $3\sigma$  significance possible in about six years of operation for most intense sources
- extended source, e.g. galactic center
  - on/off-zone analysis [10.5281/zenodo.6767723](#)
- neutrino flux from potential astrophysical neutrino sources



# Neutrino oscillations

Determine basic neutrino properties



- after 3 years of data taking
  - sensitivity for neutrino mass ordering at  $4.4\sigma$  (normal) and  $2.3\sigma$  (inverted)
  - precision to measure  $\Delta m_{32}^2$  and  $\theta_{23}$ :  $85 \cdot 10^{-6} \text{ eV}^2$  (normal),  $75 \cdot 10^{-6} \text{ eV}^2$  (inverted) ordering
  - unitarity test of the leptonic mixing matrix: exclude  $\nu_\tau$  event rate variations larger than 20% at  $3\sigma$  level

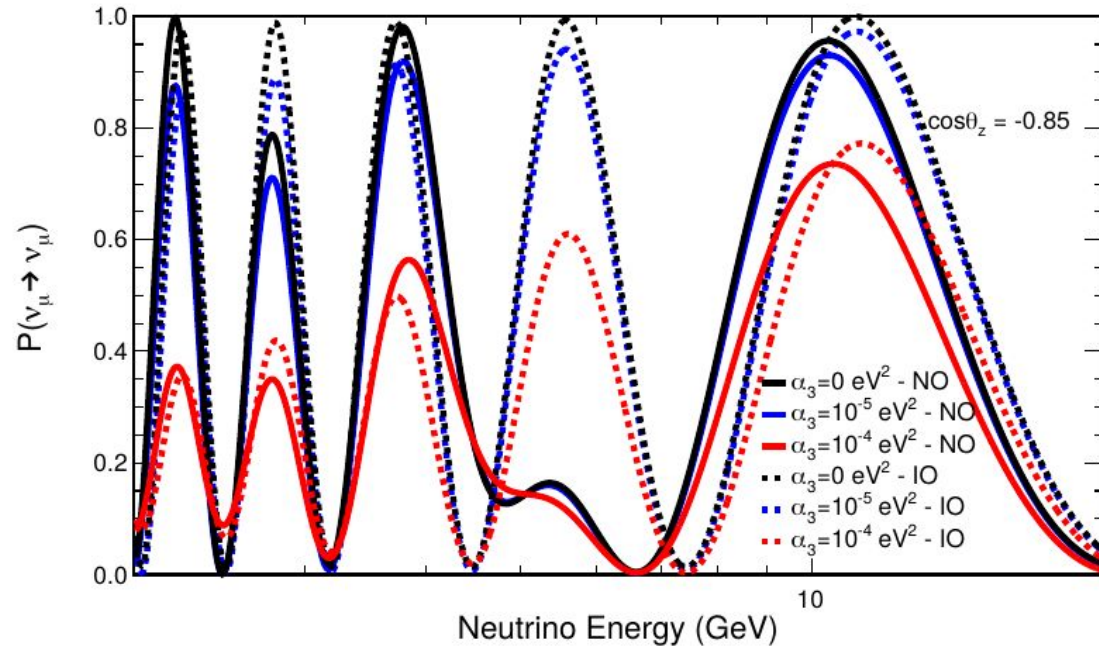
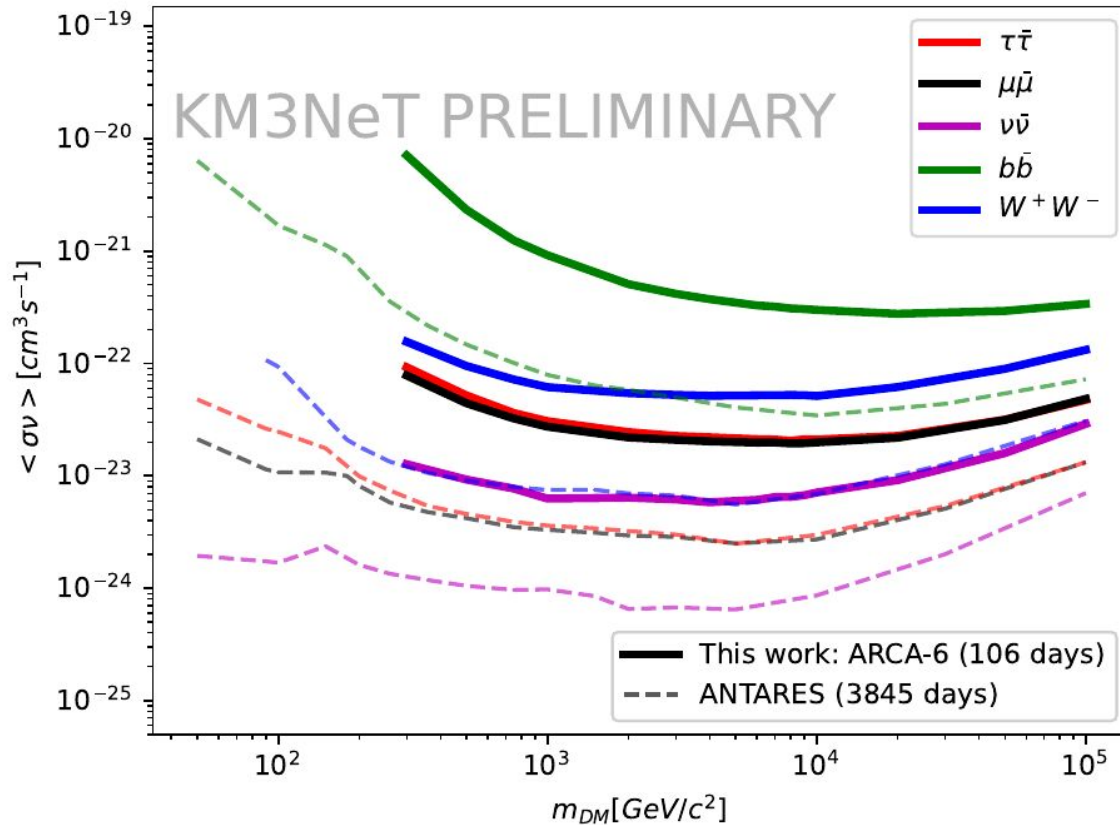


Fig. 1: Muon neutrino survival probabilities in the presence of decay

- probing neutrino invisible decay (neutrino mass state  $\nu_3$  decays into a sterile neutrino) ([zenodo.org/record/6758959](https://zenodo.org/record/6758959))
- sterile neutrino searches: active-sterile mixing with mass squared differences  $\Delta m_{41}^2$  between  $10^{-5}$  and  $10$   $\text{eV}^2$  ([zenodo.org/record/6804567](https://zenodo.org/record/6804567))
- non-standard interactions: sub-dominant effects in the oscillation patterns ([zenodo.org/record/6785232](https://zenodo.org/record/6785232))
- quantum decoherence from quantum gravity ([zenodo.org/record/6781033](https://zenodo.org/record/6781033))

# Dark matter

## Neutrinos from WIMPs



WIMP from the galactic center

### WIMPs decaying

- with masses from  $300 \text{ GeV}/c^2$  to  $100 \text{ TeV}/c^2$ ,
- five annihilation channels:  $\mu^+\mu^-$ ,  $\tau^+\tau^-$ ,  $\nu_\mu\nu_\mu$ ,  $b\bar{b}$ ,  $W^+W^-$

### sources with WIMP accumulation

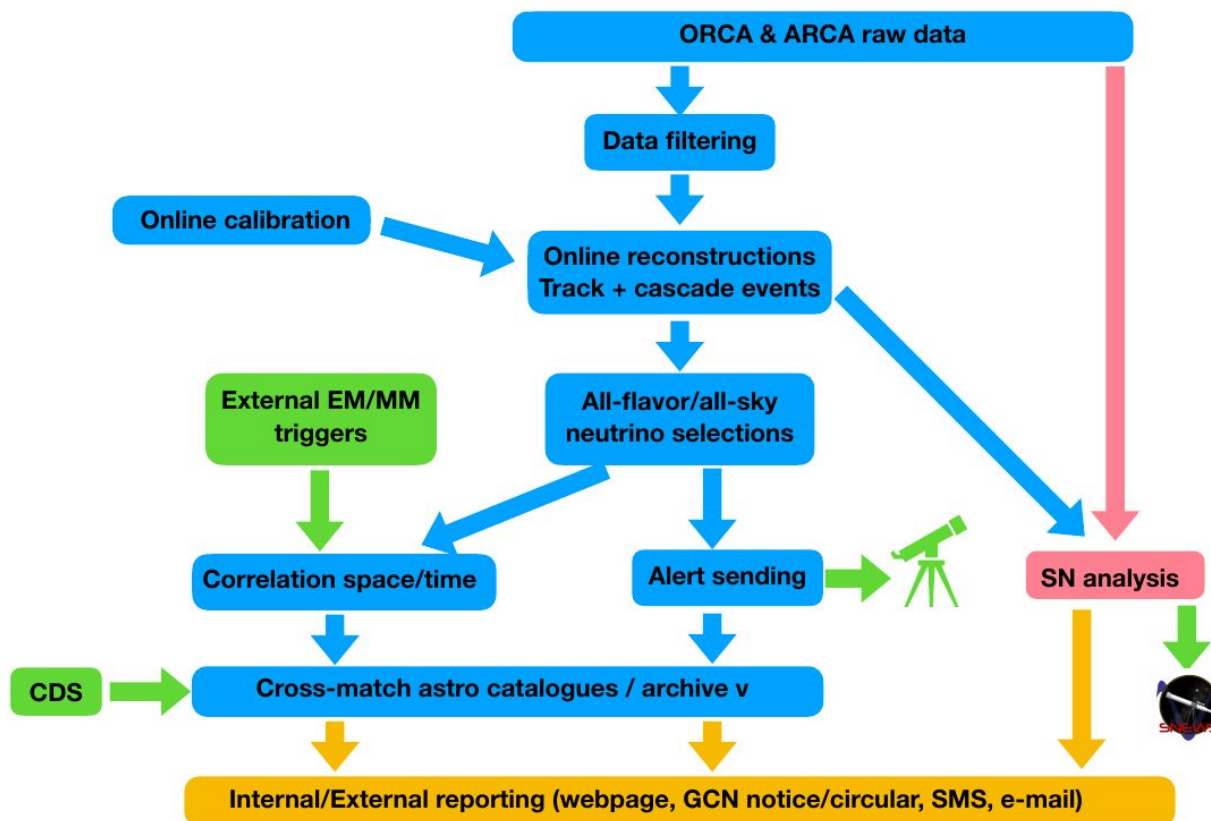
- galactic center ([zenodo.org/record/6785348](https://zenodo.org/record/6785348))
- solar core ([zenodo.org/record/6775092](https://zenodo.org/record/6775092))

### 106 days of ARCA data available

- expecting competitive limits with increased data taking

# Multimessenger alerts

Sending out alerts on neutrino events



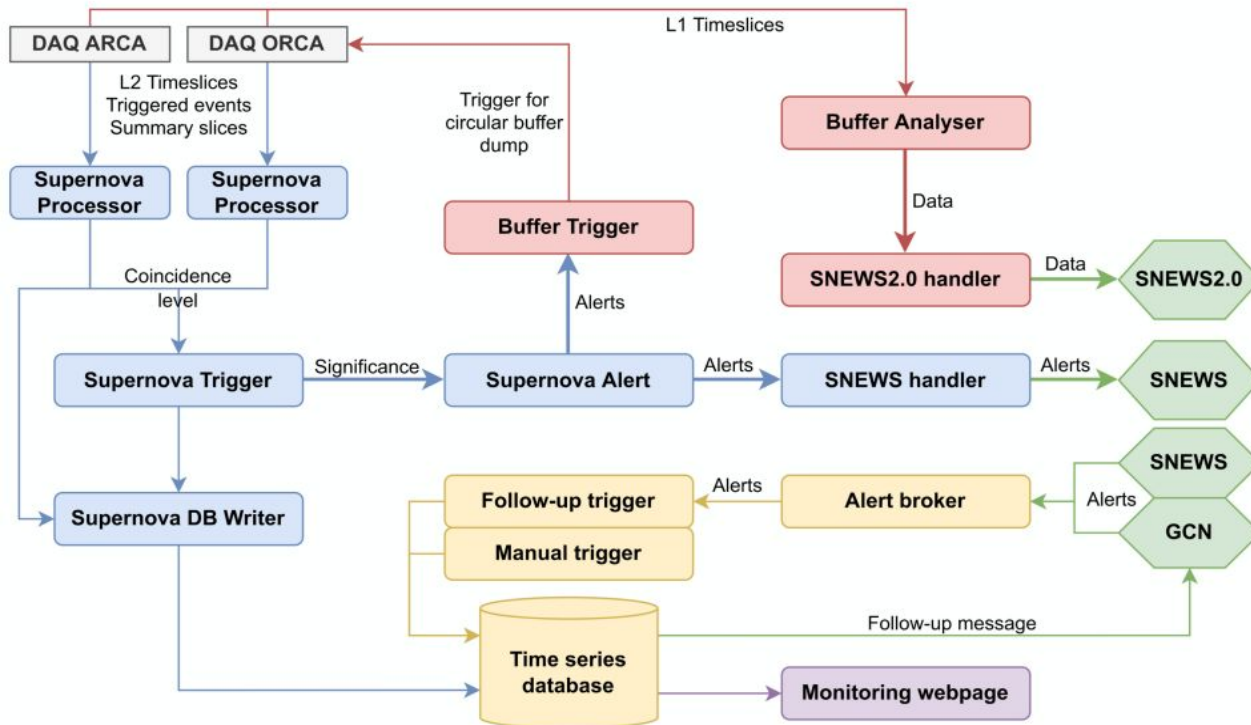
[zenodo.org/record/6805417](https://zenodo.org/record/6805417)

- various alert types
  - supernova monitoring for prompt alerts, generation latency < 20 s
  - receive external EM/GW/v alerts
  - send all flavor, all-sky v alerts, multiplets & HE (GeV - PeV) neutrino alerts
- two pipelines
  - MeV supernova: alert
- planned alert types
  - neutrino triggers (~1/month): high energy and Multiplet neutrino alerts
  - physics triggers (~1-2/month): correlated neutrinos based on astrophysical properties (AGN/TDE/CCSN/GRB/Sun...)



# Supernova monitoring

Different observation mode: high flux of neutrinos

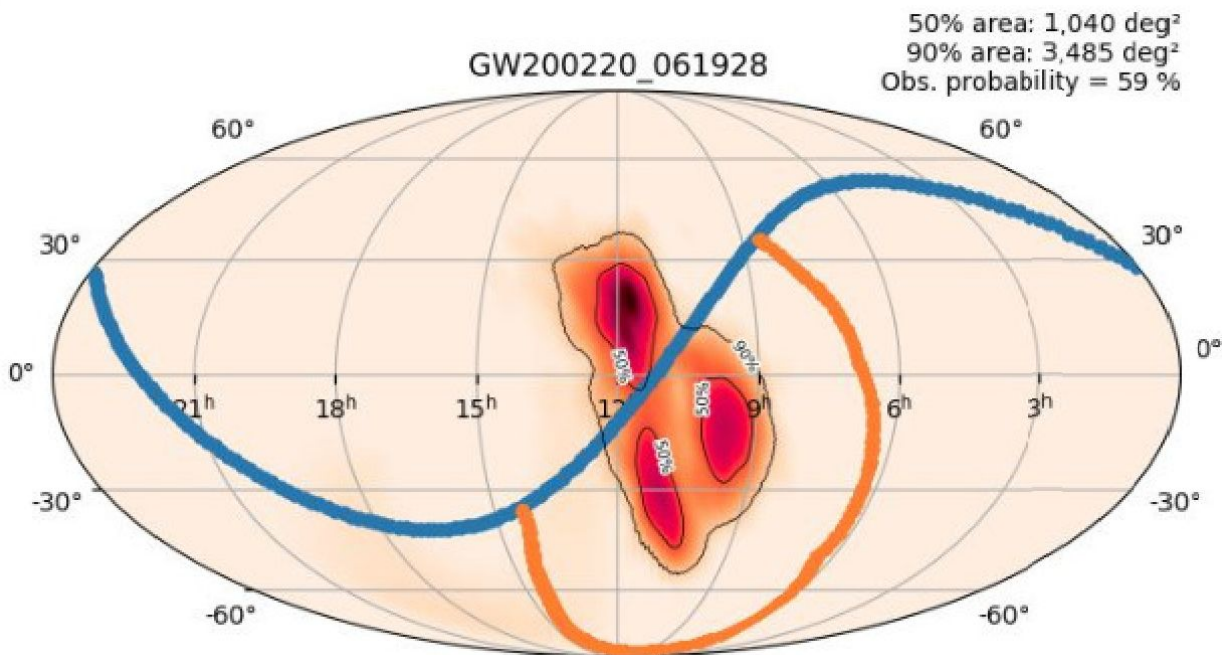


<https://zenodo.org/record/6785410>

- high flux of low-energy neutrinos, generating positrons producing increase in photon rate
- CCSN detectable as excess of coincidences above the optical background KM3NeT
- building on both detectors for event identification
- injected to Supernova Early Warning System (SNEWS)
- $5\sigma$  discovery potential for Galactic and near-Galactic events

# Multimessenger analyses: Gravitational waves

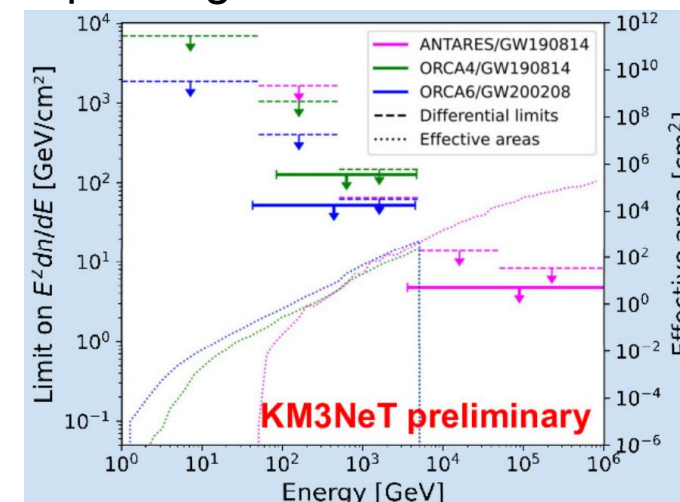
Timing coincidences with GW alerts



[zenodo.org/record/6805229](https://zenodo.org/record/6805229)

Limits on $\phi = E^2 dn/dE$		
(all-flavour, $E^{-2}$ spectrum)	ORCA4	ORCA6
<b>Limits [GeV/cm<sup>2</sup>]</b>	100-500	50-200
<b>5-95% neutrino energy range</b>	70 GeV - 5 TeV	40 GeV - 5 TeV

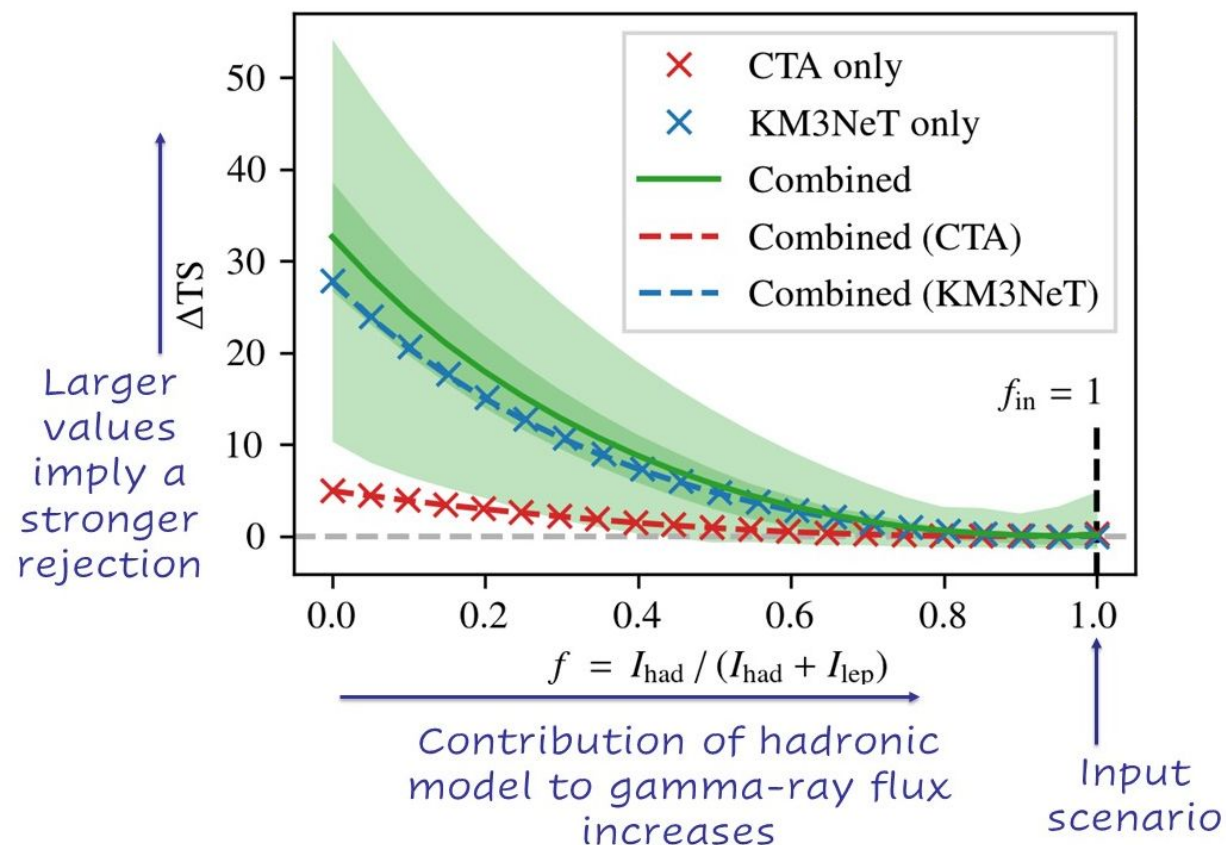
- search for MeV-TeV neutrinos from compact binary mergers (LIGO and Virgo, 3rd observation period)
  - independent search focusing on the MeV and GeV-TeV ranges
  - stacking analysis for neutrino emission in different populations of mergers
- ORCA already surpassing ANTARES in GeV range



# More multimessenger analyses

Coincidences with other neutrino detections & gamma rays

Vela X



[zenodo.org/record/6785224](https://zenodo.org/record/6785224)

- combine with gamma ray observations
  - joint binned likelihood analysis
  - for 4 different sources (VelaX et al)
  - pure leptonic / hadronic scenarios
  - gain from combined CTA/KM3NeT analysis
- follow-up IceCube neutrino alerts
  - looking for a signal excess around the best Active Galactic Nuclei (AGN) counterpart (4 alerts in 2021/22)
  - estimating background from off region
  - time window of  $\pm 1$  day
  - extended search time window: 1 month

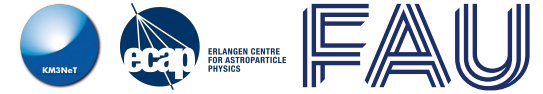
[zenodo.org/record/6805372](https://zenodo.org/record/6805372)



# How to get the best out of our science?

# The dilemma of modern science

How do we optimize our work?



Incentives for individual success are focused on getting it published, not getting it right

Nosek, Spies, & Motyl, 2012

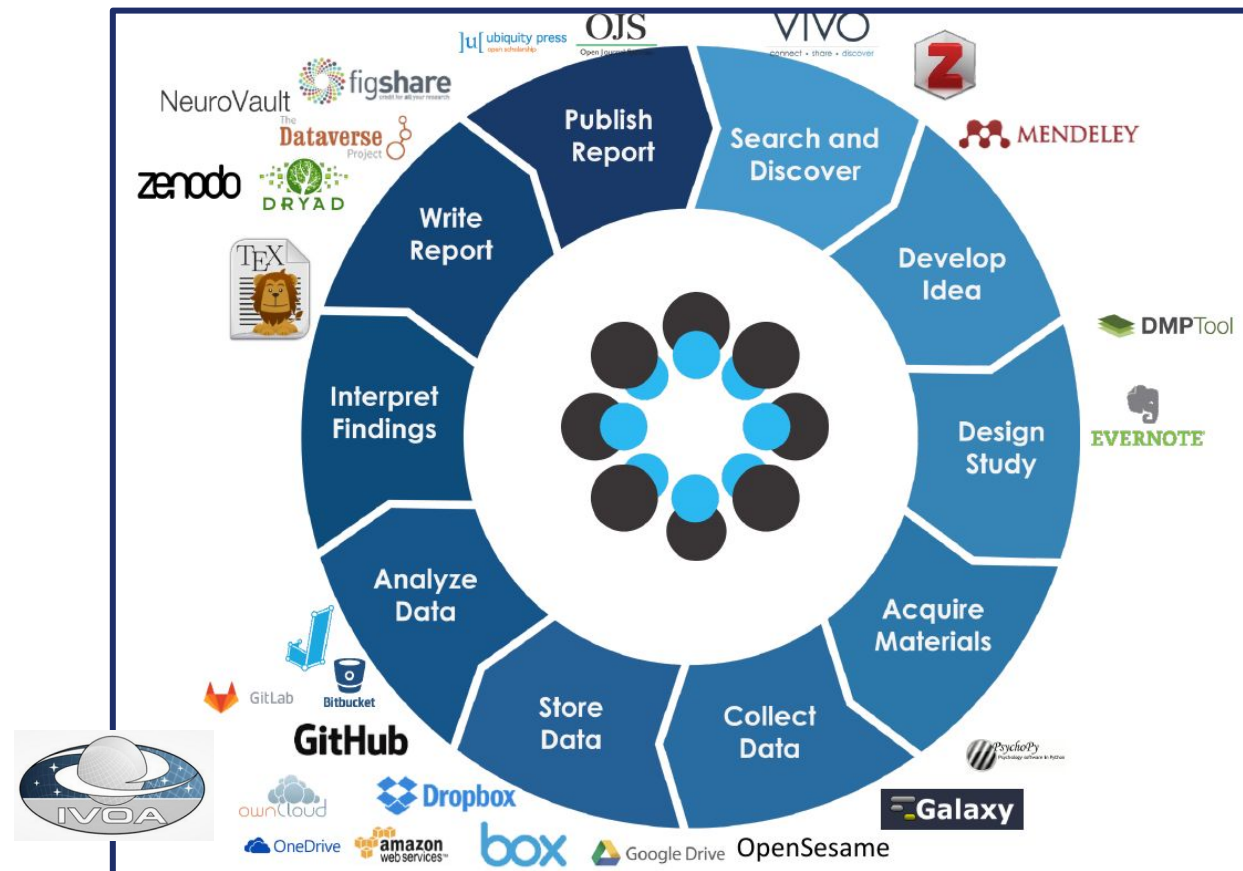
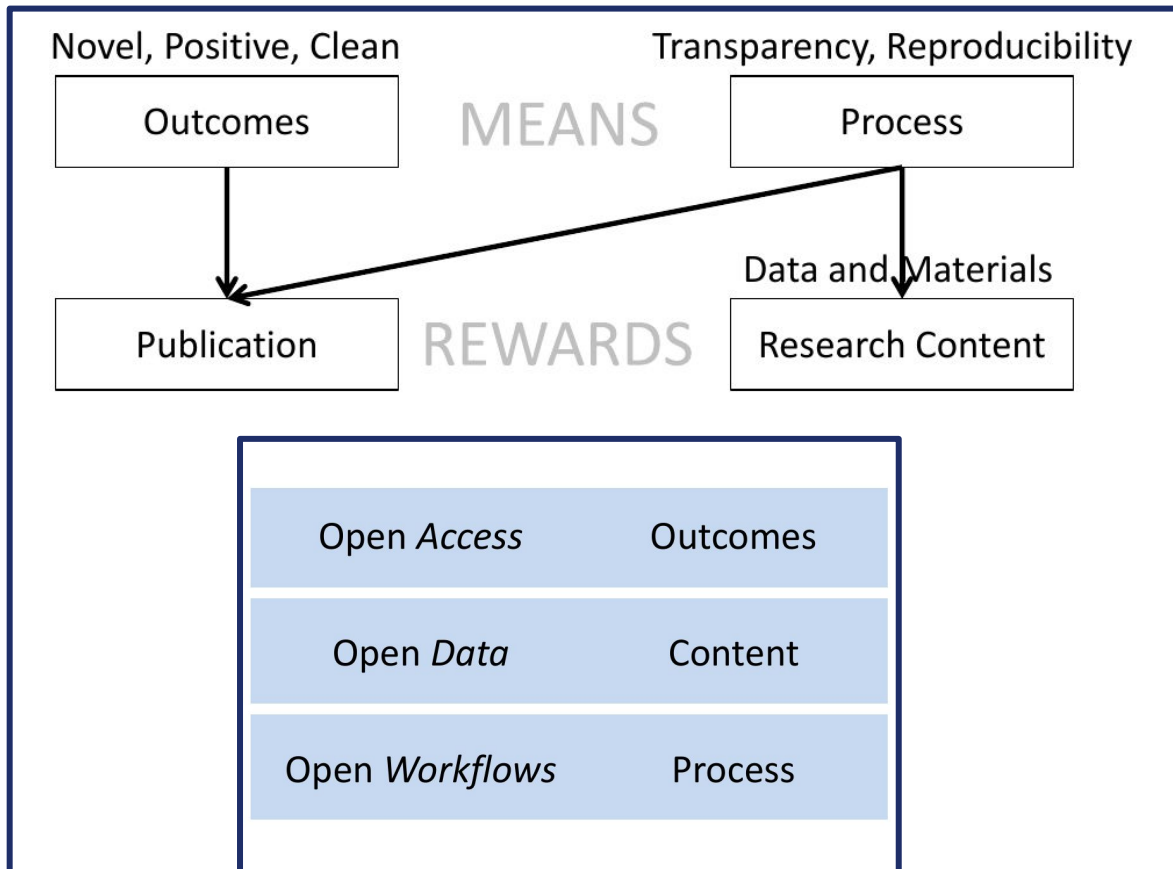
## Barriers

1. Perceived norms (*Anderson, Martinson, & DeVries, 2007*)
2. Motivated reasoning (*Kunda, 1990*)
3. Minimal accountability (*Lerner & Tetlock, 1999*)
4. I am busy (*Me & You, 2016*)

B. Nosek, “[Improving openness and reproducibility in research](#)”, in [BD2K Guide to the Fundamentals of Data Science](#)

# Changing the focus

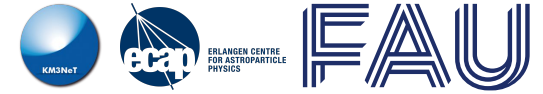
Publish research, not papers



B. Nosek, “[Improving openness and reproducibility in research](#)”, in [BD2K Guide to the Fundamentals of Data Science](#)

# The Open Science Paradigm

What does it actually mean?



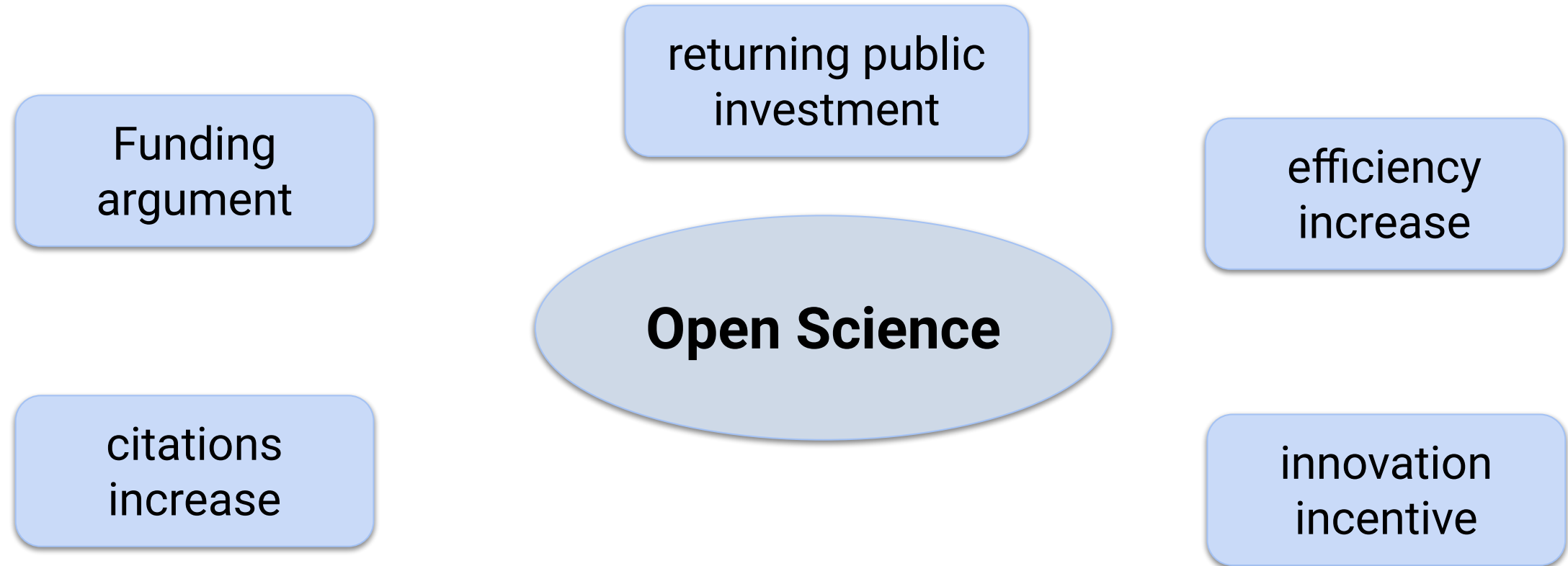
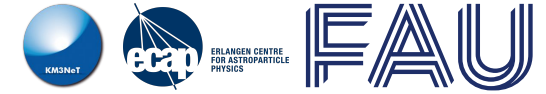
**Open science** is the movement to make scientific research (including publications, data, physical samples, and software) and its dissemination accessible to all levels of society, amateur or professional.

Open science is transparent and accessible knowledge that is shared and developed through collaborative networks.

*Wikipedia (Open Science)*

# Why care about Open Science?

Beyond the moral high ground

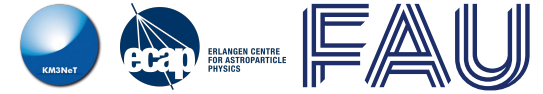


OECD (2015-10-15), "Making Open Science a Reality", OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing, Paris. <http://dx.doi.org/10.1787/5jrs2f963zs1-en>



# Spelling it out: FAIRness

Making research findable, accessible, interoperable, re-usable



**FAIR** initiative to make scientific data available

This requires:

- Metadata to all science products
- Long-term storage and legal backing (licenses)
- Infrastructure to find and access
- Interfaces to integrate data and workflows for new use
- Resources to develop, integrate and maintain

→ Applicable not only to data, but also **software** and **workflows**!

**FAIR Guiding Principles for scientific data management and stewardship**

<https://www.go-fair.org>





Metadata, Identifiers, Archiving



Data

Publication

Software



arXiv.org



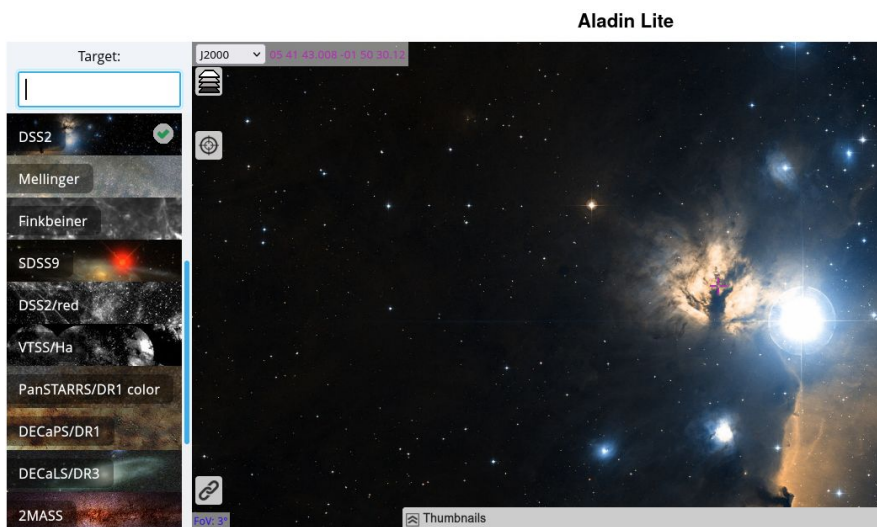
# How to make it ...

Integration in different environments

# Accessible

Data

Software



analysis environments

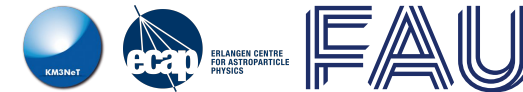


containerization  
notebooks

# How to make it ...

Standardizing data formats and use

# Interoperability



Data

Software

## Documents & Standards

DOCUMENTS XML SCHEMA VOCABULARIES DOC SUBMISSION

- *Technical Specifications*
- *Notes*
- *Promotion process*
- *IVOA Technical Assessment and Roadmap Documents*
- *Submission Log*



## Technical Specifications

>>

Group	Title	Most stable	In progress	Version history
App	SAMP - Simple Application Messaging Protocol	1.3		1.3 1.3 1.3 1.3 1.3 1.2 1.2 1.2 1.11 1.11 1.10 1.00
	VOTable - VOTable Format Definition	1.4		1.4 1.4 1.4 1.4 1.4 1.4 1.3 1.3 1.3 1.2 1.2 1.2 1.20 1.20 1.10 1.00
	MOC - HEALPix Multi-Order Coverage Map	1.1	RFC	2.0 2.0 2.0 2.0 2.0 2.0 1.1 1.1 1.1 1.1 1.1 1.0 1.0 1.0 1.0 1.0



Defining schemas  
providing converters

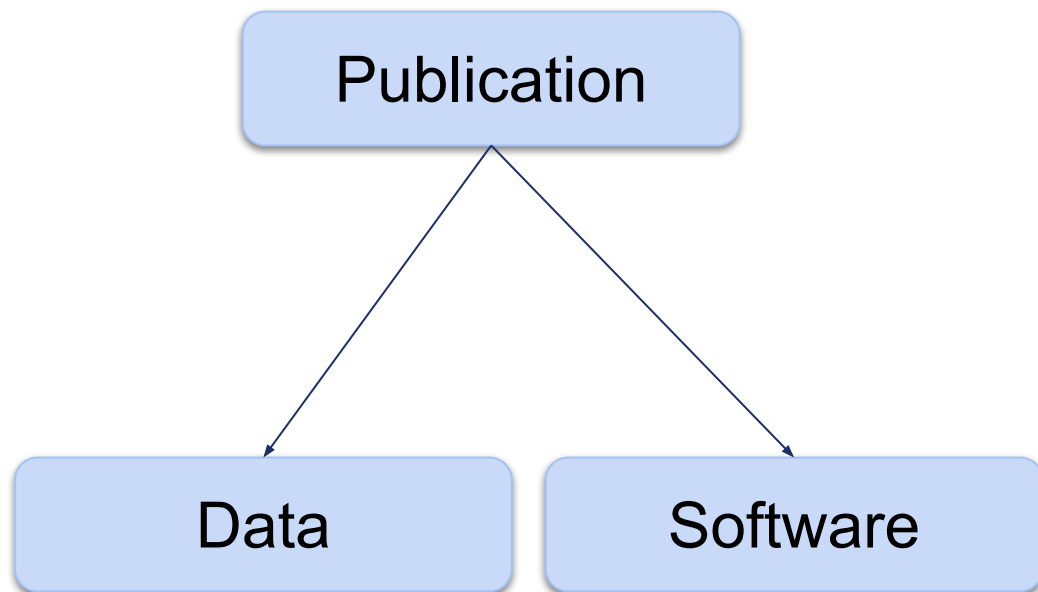
# How to make it ...

Packing it all together & make it understandable

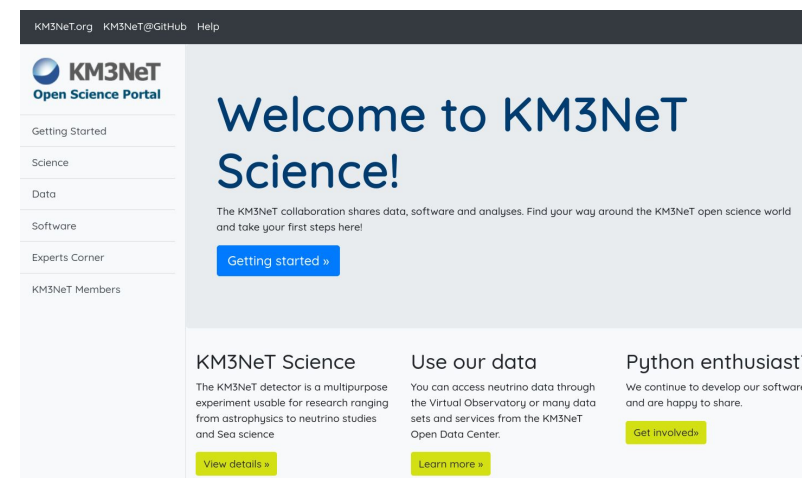
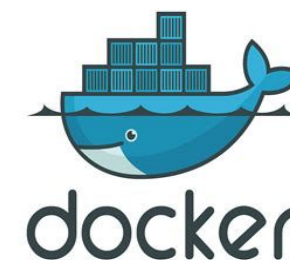
Reusable



ERLANGEN CENTRE  
FOR ASTROPARTICLE  
PHYSICS



Analysis repositories  
+ documentation!



DSS2

Mellinger

Finkbeiner

SDSS9

DSS2/red

VTSS/Ha

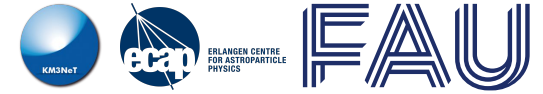
PanSTARRS/DR1 color

DECaPS/DR1

# The Open Science Program in KM3NeT

# First step: Put the topic on the agenda

Decision by the Institute Board of KM3NeT

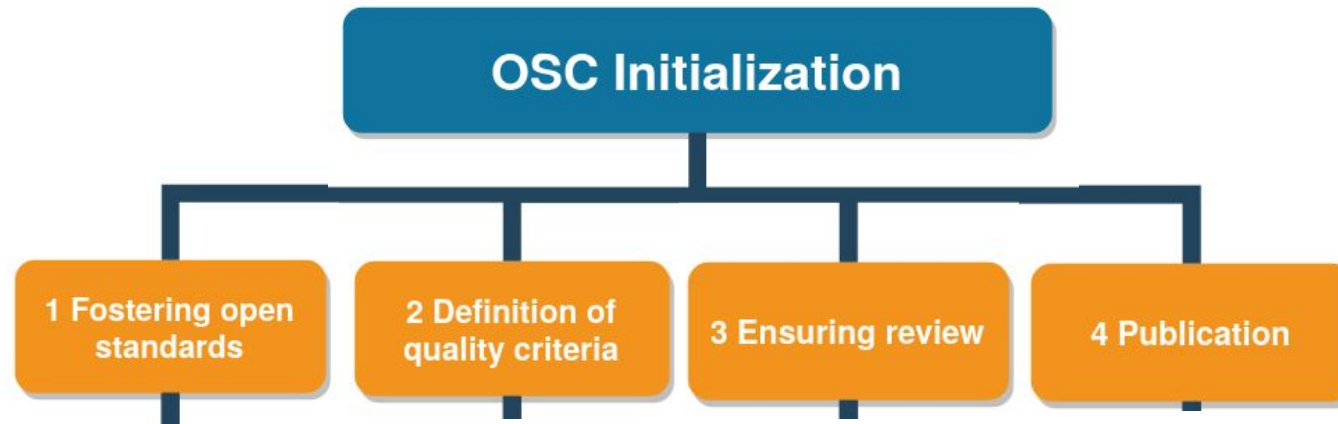


## Open Science Policy

**KM3NeT supports the aims of open data and open science and commits to implement the necessary steps wherever possible. This includes open access data supporting publications, open source software and open data in general including the information needed to appropriately use the data.**

## Implementation in the collaboration needs

- Discussions
- Development
- Standardization & Guidelines
- Training



### Open Science Committee

**KM3NeT installs an Open Science Committee that works in parallel to the OC and CC and sets, maintains and further develops the procedures for KM3NeT open science.**



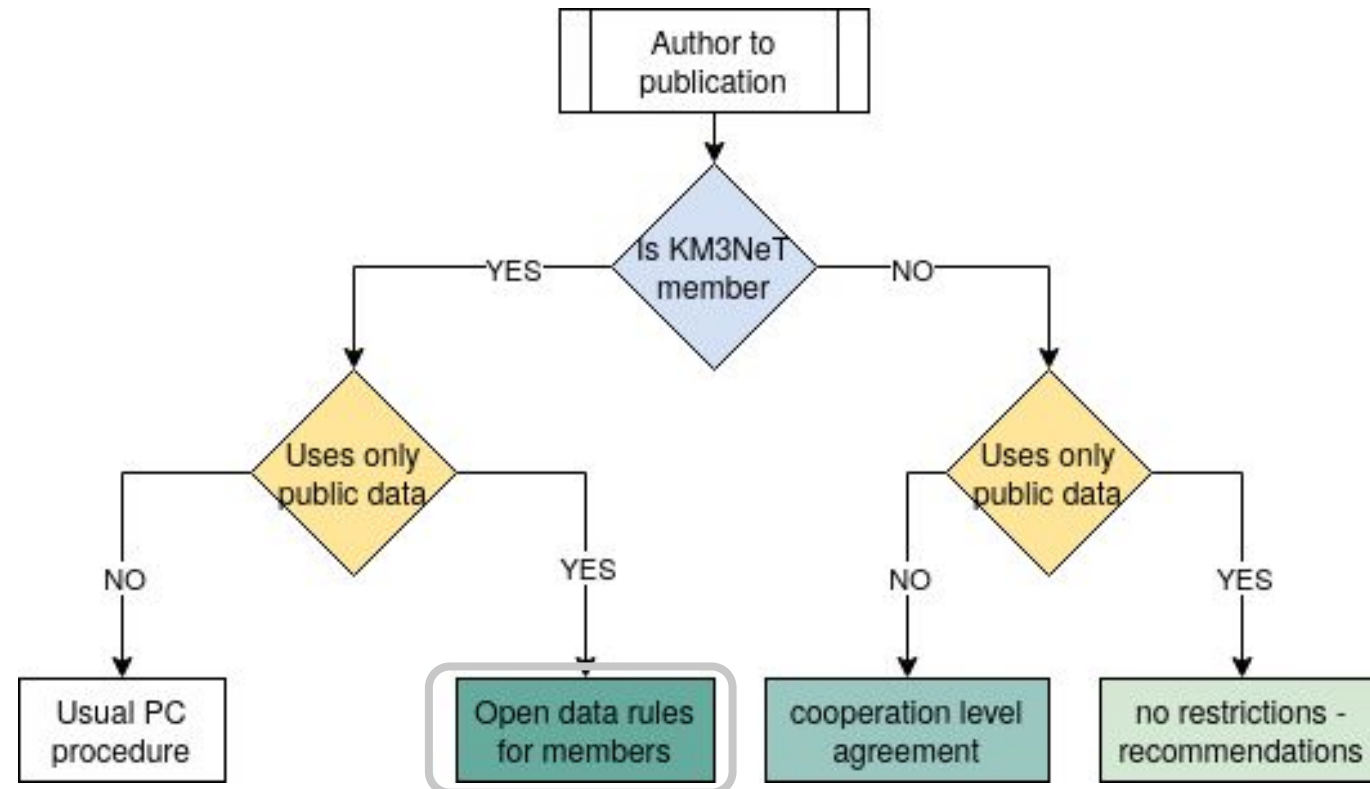
# Addressing issues of policy

How to deal with authorship?

Discussion with PC on rules for use of public data for KM3NeT members  
Cause for tension and problems in other collaborations.

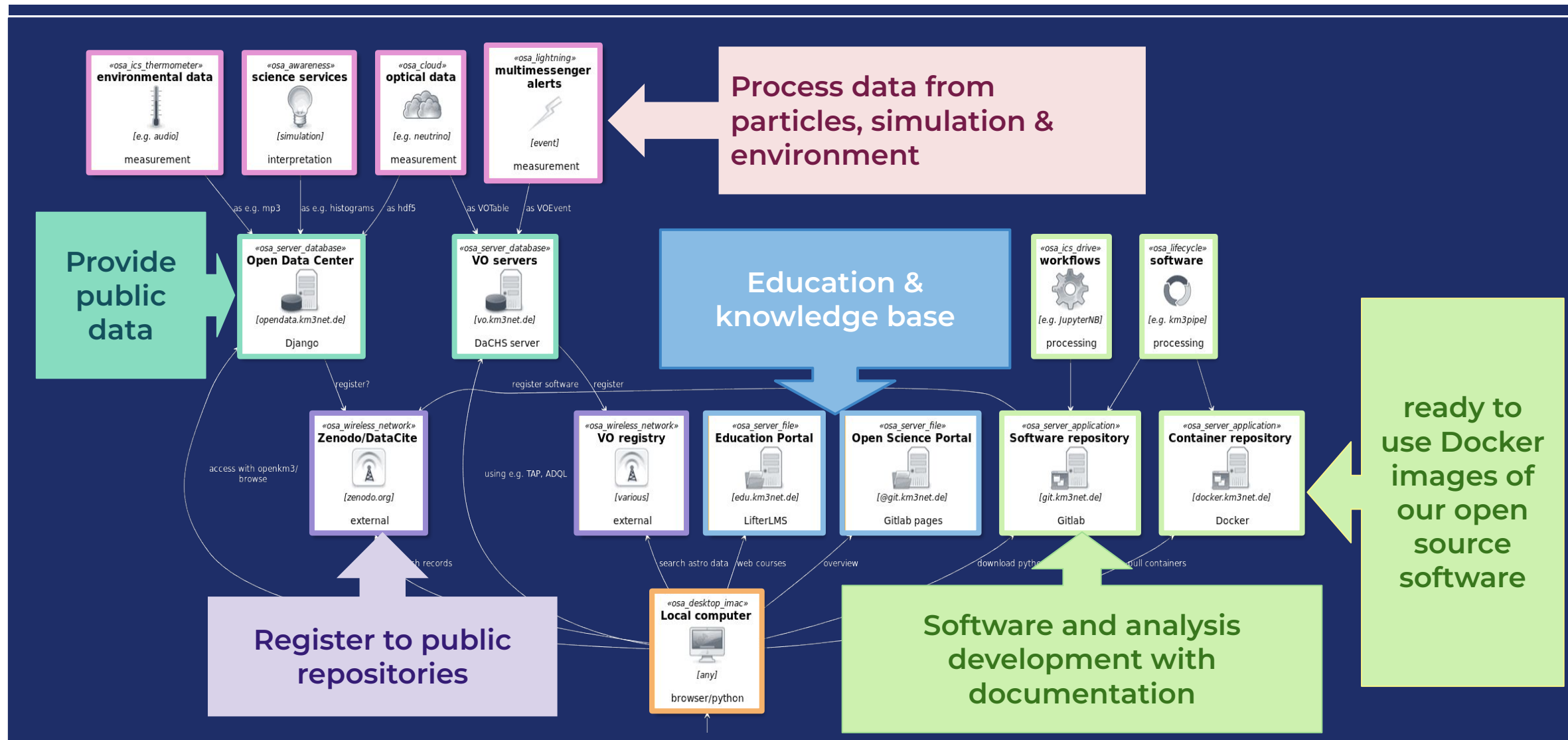
## Main question:

How to find a balance between “right to first discovery” and fostering broad use of data?

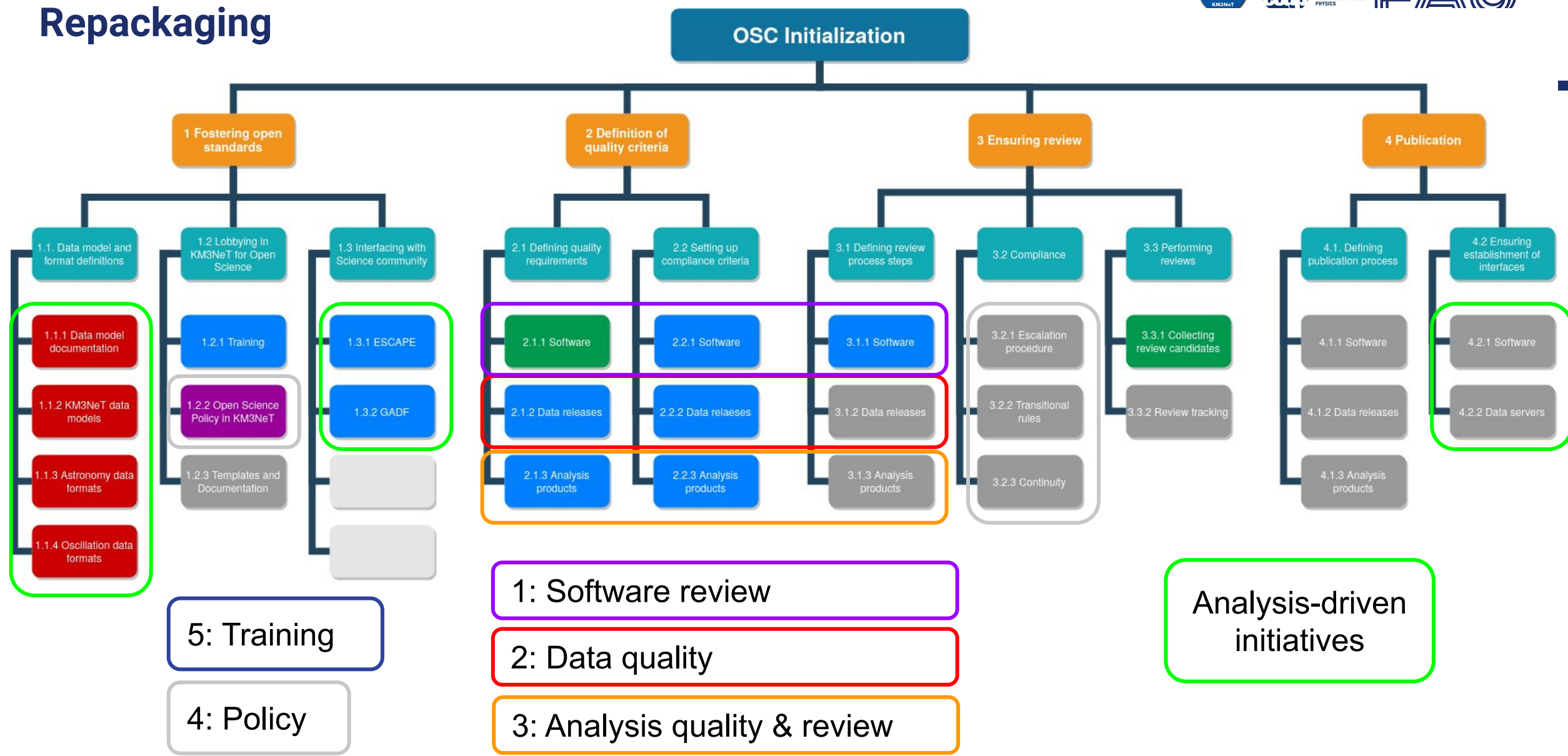


# System development: Starting with prototyping

In data, software, services with the KM3NeT Open Science environment



# Repackaging

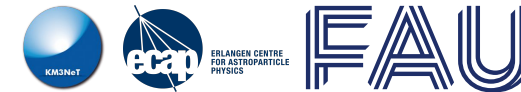


# Hands-On Open Science



# Projects towards Open Science

Give and take - and make new plans



## Offer data products

Neutrino events & instrument response

Software based on ROOT & python

Analysis pipelines in astro- & neutrino physics



## Use best practices

Data management based on HE physics

Open Software & containerization



## Cooperate

Virtual Observatory for astronomy

VODF initiative for VHE astroparticle physics

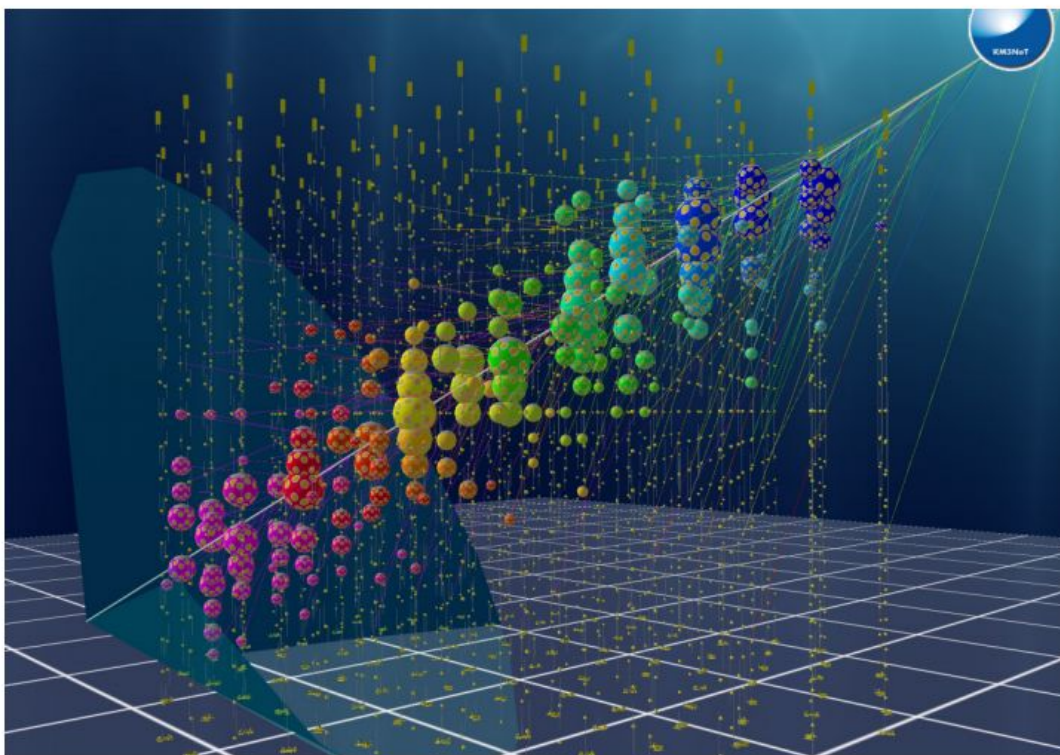
EOSC & ESCAPE for HTC & science platform

# Focus data: neutrino events

The main output of KM3NeT

## „Full“ event (i.e. particle detection!)

event identification	detector status	<photon detections $\bar{x}$ , t, A>
----------------------	-----------------	--------------------------------------



## „Reduced“ event

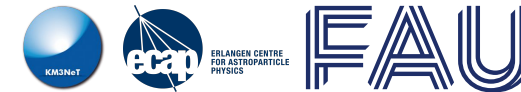
reconstructed particle properties	direction time energy, resolution ...
-----------------------------------	---

Decl» [deg]»	RA» [deg]»	Nhit» [deg]»	Beta»	MJD [days]
19.5»	68.2»	21»	1.0»	54138.3105
-60.0»	26.5»	33»	0.8»	54138.5830
-29.8»	82.1»	34»	0.3»	54140.2299
-8.6»	271.8»	41»	0.3»	54140.6394
-32.3»	261.4»	45»	0.5»	54142.7042
-66.7»	149.9»	52»	0.8»	54159.4158
-13.0»	93.6»	25»	0.7»	54160.4830
-26.2»	266.7»	28»	0.8»	54160.6180
23.5»	121.7»	41»	0.5»	54161.4361
-70.7»	47.1»	30»	0.9»	54165.5838
-55.0»	284.4»	36»	0.5»	54169.0685

Example files of KM3NeT ROOT files:  
<https://github.com/KM3NeT/km3net-testdata>

# Neutrino events in astrophysics

Sharing in the Virtual Observatory

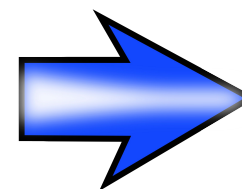


- Running server with DaCHS software
- Registered as data provider to the VO ([ivo://km3net.org](http://ivo://km3net.org))
- Can publish data sets to the VO registry - done with Antares 2007-2017 data
- Data accessible through widely used tools in **astrophysics** (Aladin, Topcat ...)

**Virtual observatory server**

<http://vo.km3net.de/>

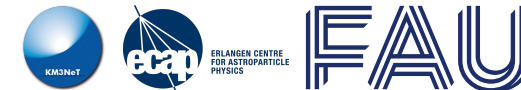
The screenshot shows a web interface for a Virtual Observatory resource. On the left is a sidebar with a KM3NeT logo and buttons for 'Help' and 'Service info'. Below these are 'Metadata' sections for 'Description', 'Keywords', and 'Creator'. The main content area is titled 'Information on resource 'ANTARES 2007-2017'' and contains the following text: 'Neutrino candidates from full-sky search from 2007-2017', 'Services defined within this resource descriptor' with a bullet point for '[ANTARES 2007-2017](#)', 'Tables defined within this resource descriptor' with a bullet point for '[ant20\\_01.main](#) - queryable through [TAP](#) and [ADQL](#)', and a '[Manage RD]' link. At the bottom, it says 'Please report errors and problems to the [site operators](#). Thanks.'



Integration of further (and legacy) ANTARES data in KM3NeT environment planned

# Events outside of astrophysics

Dealing with less standardization in particle physics



## Introducing: The KM3NeT Open Data Center

- For all data not publishable through the IVOA, serving as interface and/or server to the data
- Including also link to data sets on VO server
- Based on Django REST API
- Usable for event data sets (hdf5-files with standardized metadata), plots or services, environmental data ...
- Data accessible through webpage, through REST-API or python based package (openkm3)

**Open Data Center**

<http://opendata.km3net.de>



**KM3NeT**

## Open Data Center

We make our data available for you!

You can find a description of how to use our data at the

### Current uploads

**KM3NeT test data**

ORCA 4-line events

[More Info](#)





# How to add relevant probability functions?

Adding instrument response, background and other auxiliary simulated information

No VO standard for neutrino regime - starting our own

- extended header with content identifier
- tabulated data (e.g. bin content, function values ...)
- ready for use in python environment

Example: Effective area for ANTARES 2007-17 Point Source search

```
ks.print_index()

annotated_aeff0
=====
header:
  name: ANTARES 2007-2017 effective area
  description: effective area for E-2 source sp
  contact: antares.spokesperson@in2p3.fr
  instrument: ANTARES
  license: Creative Commons 4.0 Internation
  reference: https://antares.in2p3.fr/public
```

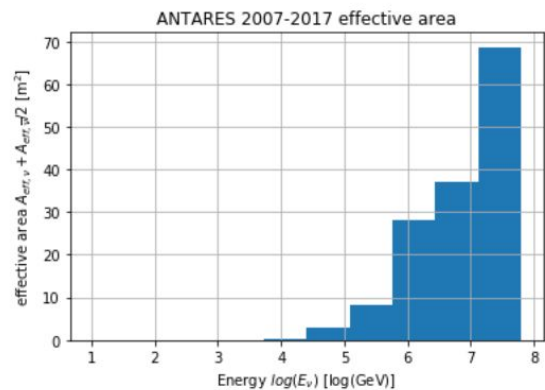
```
plain = ks.get("annotated_aeff0") # not so nice, make i
plain.data
```

```
{'Log(E1/GeV)': [1.0,
1.2,
1.4,
1.6,
1.8,
2.0,
2.2,
2.4,
2.6,
2.8,
3.0,
3.2,
3.4,
3.6,
...]}
```

```
table = ks.get("annotated_aeff0", loadoption = "pandas")
table.data
```

	Log(E1/GeV)	Log(E2/GeV)	AEFF/m2
0	1.0	1.2	5.153940e-12
1	1.2	1.4	1.614220e-10
2	1.4	1.6	4.891220e-09
3	1.6	1.8	1.245090e-07
4	1.8	2.0	9.157530e-07
5	2.0	2.2	4.372520e-06

```
nice = ks.get("annotated_aeff0", loadoption= "plot")
```



```
table.get_origin()
```

```
{'range': {'time': {'measurement_start': datetime.datetime(2007, 1, ...
'measurement_stop': datetime.datetime(2018, 1, 1, 0, 0),
'ktype': 'tbd'},
'zenith': {'min': -45, 'max': 0, 'ktype': 'tbd'}}}
```

```
table.get_paraminfo()
```

```
{'lower_edge': {'columnname': 'Log(E1/GeV)',
'name': 'Energy',
'description': 'logarithmic reconstructed energy of the neutrino',
'unit': 'log(GeV)',
'symbol': 'log(E_{\nu})'},
'upper_edge': {'columnname': 'Log(E2/GeV)',
'name': 'Energy',
'description': 'logarithmic reconstructed energy of the neutrino',
'unit': 'log(GeV)',
'symbol': 'log(E_{\nu})'},
```

get as table

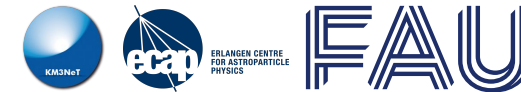
get as dataframe

get as plot

get metadata

# Environment integration

Making it all python accessible



pip install git+https://git.km3net.de/open-data/openkm3

- dependency: pyvo for VO interface
- “KM3Store” as access point for all data + services

```
from openkm3.openio import KM3Store
ks = KM3Store()
```

```
table = ks.get("one_week_orca", ["events"], "pandas")
table.data
```

	angular_error	azimuth	dirz	energy	internalID
0	0.004341	5.108108	-0.968124	22.017775	km3net.44.615
1	0.001181	3.358250	-0.990520	128.639694	km3net.44.615
2	0.003534	6.146142	-0.744035	24.362326	km3net.44.615
3	0.008247	2.954967	-0.784588	15.714084	km3net.44.615
4	0.006967	5.618736	-0.529338	137.616933	km3net.44.615

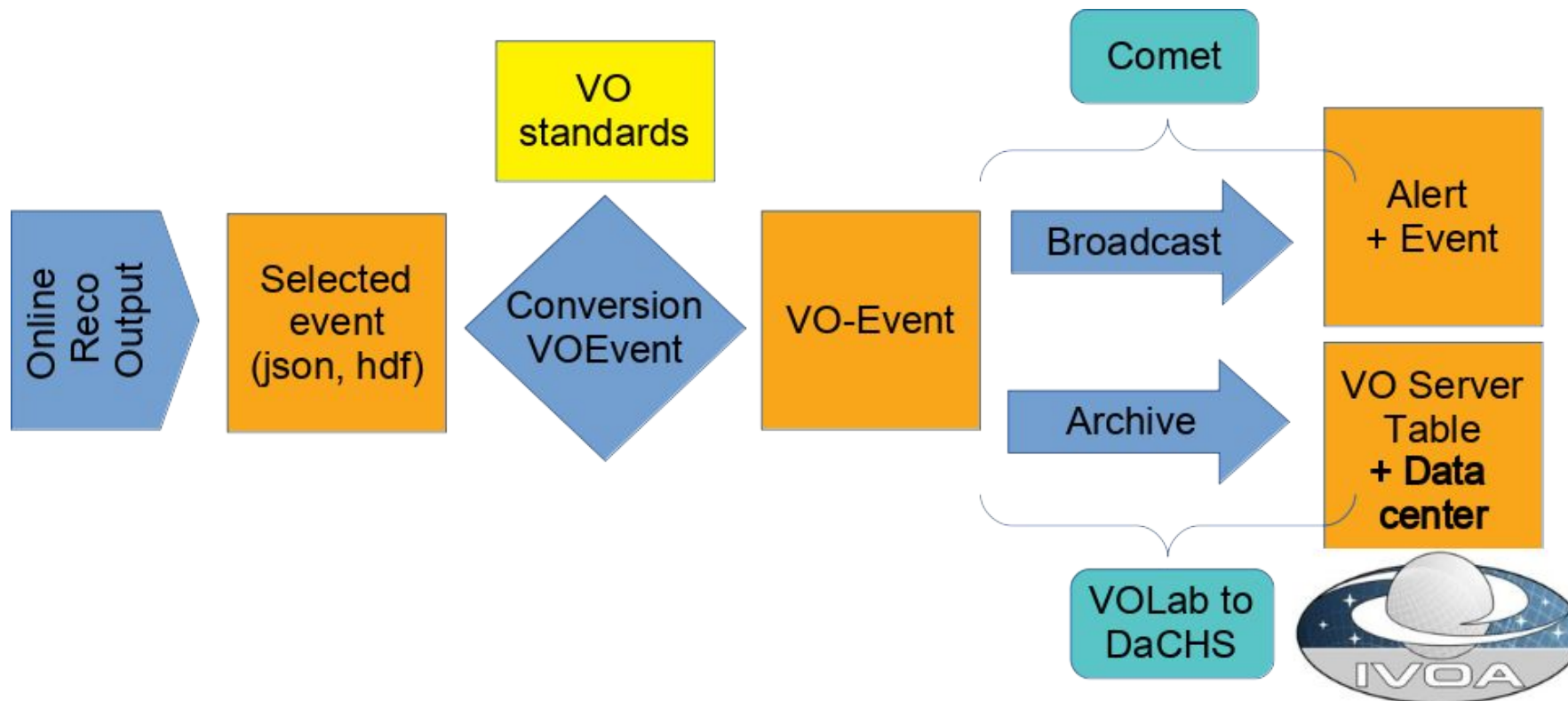
- entries are data sets, single files or services
- access logic coded in package

```
ks.print_index(include_technical=True)
```

```
one_week_orca
=====
tables: ['events', 'group_info', 'header']
header:
    author:          b'The KM3Net collaboration'
    contact:         b'opendata@km3net.de'
    instrument:      b'ORCA'
    license:         b'Creative Commons 4.0 International'
    measurement_start: b'2019-10-03T06:00:00.544000000'
    measurement_stop: b'2019-09-17T06:00:00.202000000'
    reference:       b'http://www.km3net.org/'
url:    http://vo.km3net.de:82/storage/one_week_orca.h5
type:   application/x-hdf5
local:  /home/jutta/Desktop/openkm3/examples/orca_data/.openkm3/one_week
```

# Single alert events

Different form and pipeline for VOEvents



- Converting alert info (json) to VOEvent
- Access additionally provided through **data center**

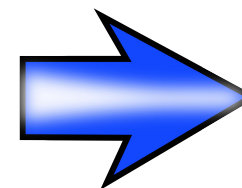
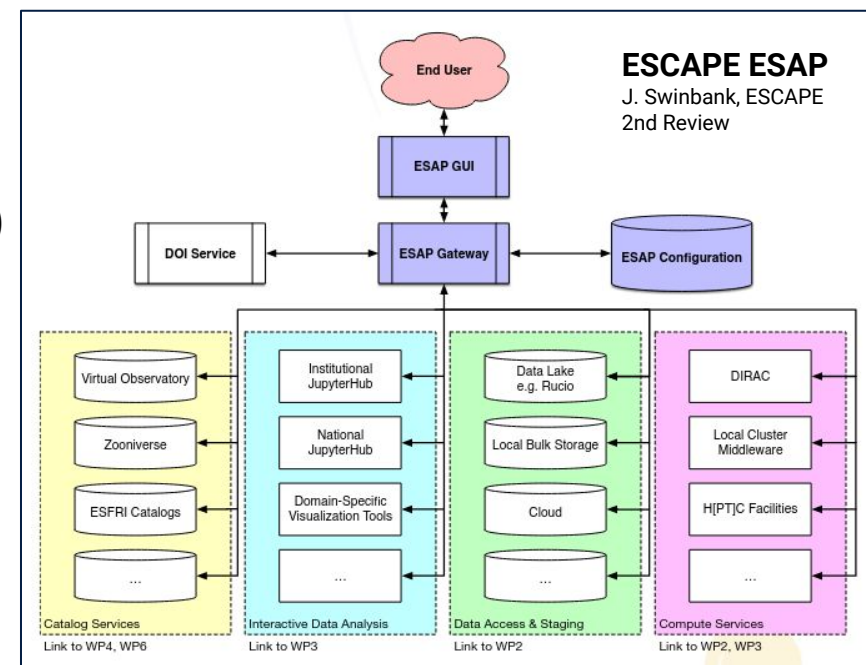
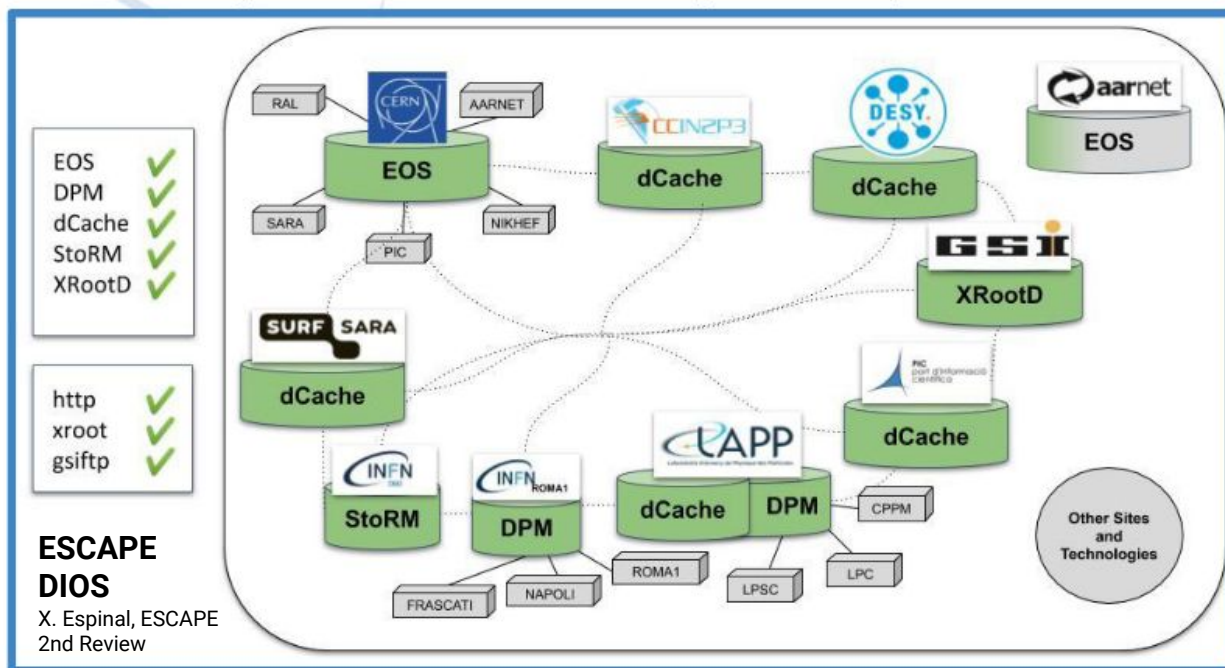
## MM VO Event

[https://git.km3net.de/j\\_schnabel/kmeta/-/tree/master/examples/applications/VOEvent](https://git.km3net.de/j_schnabel/kmeta/-/tree/master/examples/applications/VOEvent)

# Learning from HE physics: Data processing

Using the strength of particle physics community - sharing technology

- CERN-developed software dedicated to high-volume data processing
  - [Rucio](#) for scientific data management
  - exploring the [Dirac interware](#) for job management
  - exploring workflow management software ([snakemake](#), nextflow)
- projects for development & integration: ESCAPE



using HEP software for data management

# Cooperation: VHE Open Data Forum

Started from the Gamma-ray Astronomy Data Formats Initiative (GADF)

- Steered by committee with representation from experiments (including neutrinos & wide-field gamma instruments)
- Group of three lead editors to manage format development

## References

- Github repository: <https://github.com/VODF>

## Participation from KM3NeT side

Member to the Steering Committee: Kay Graf

Lead Editor for neutrino-related developments: Jutta Schnabel

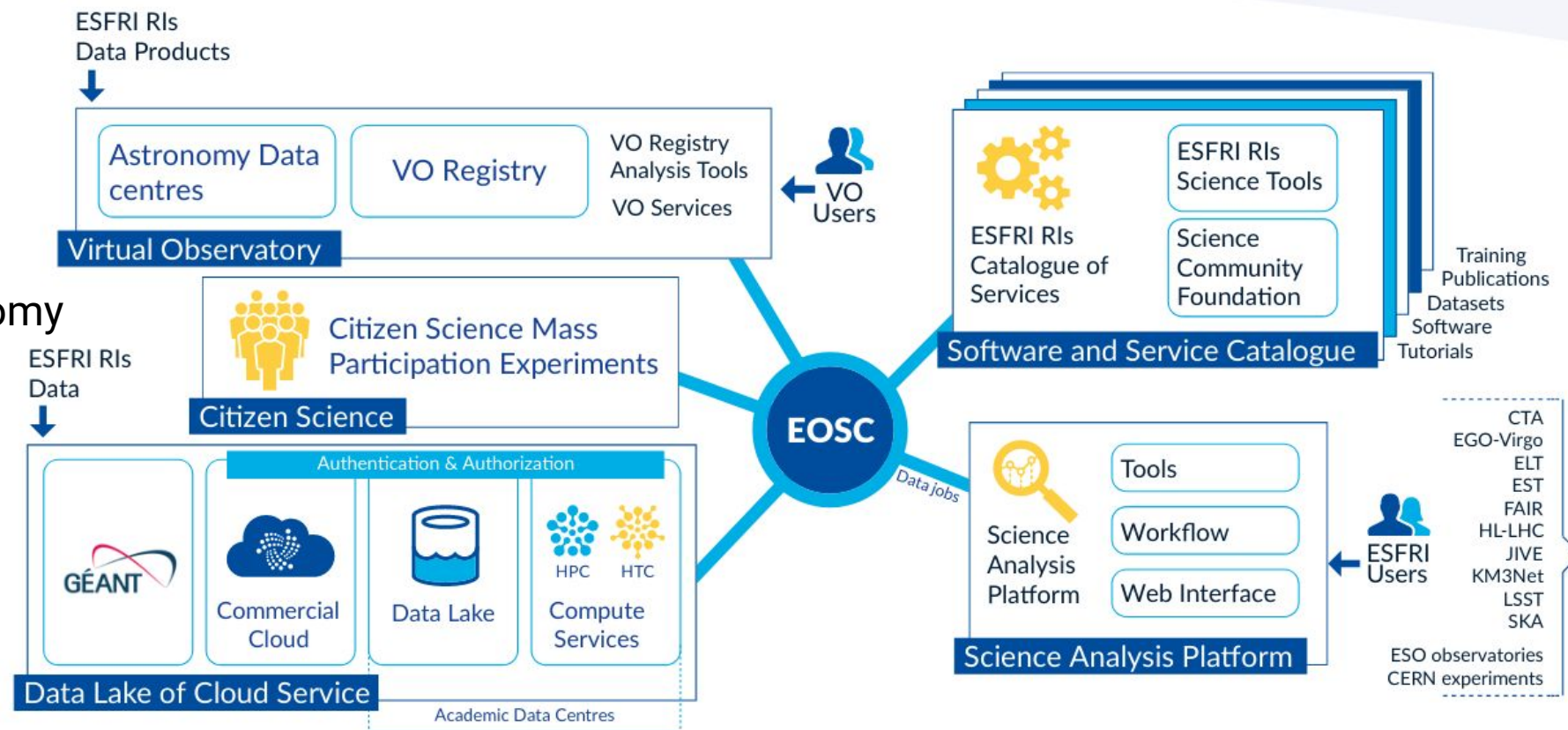


# Share & combine data, software and workflows

**ESCAPE** project as part of the European Open Science Cloud (EOSC):  
Science platform to integrate data and software from various ESFRIs

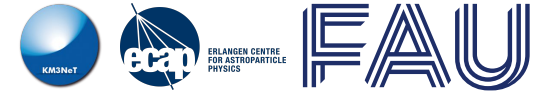
→ Services under development

→ Followed up in the EOSC future project on dark matter & HE astronomy



**Project ESCAPE**  
<https://projectescape.eu>

# So, what does “Open Science” mean?



- Understanding FAIRness of data and research
- Adopting the goal of Open Science as a collaboration
- Creating an Open Science environment for reproducible science
- Integrating with the wider science community
- Co-developing common standards and platforms



**Let's do science together  
&  
Thank you for your attention!**