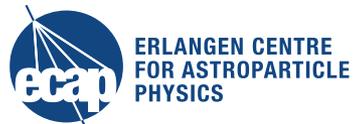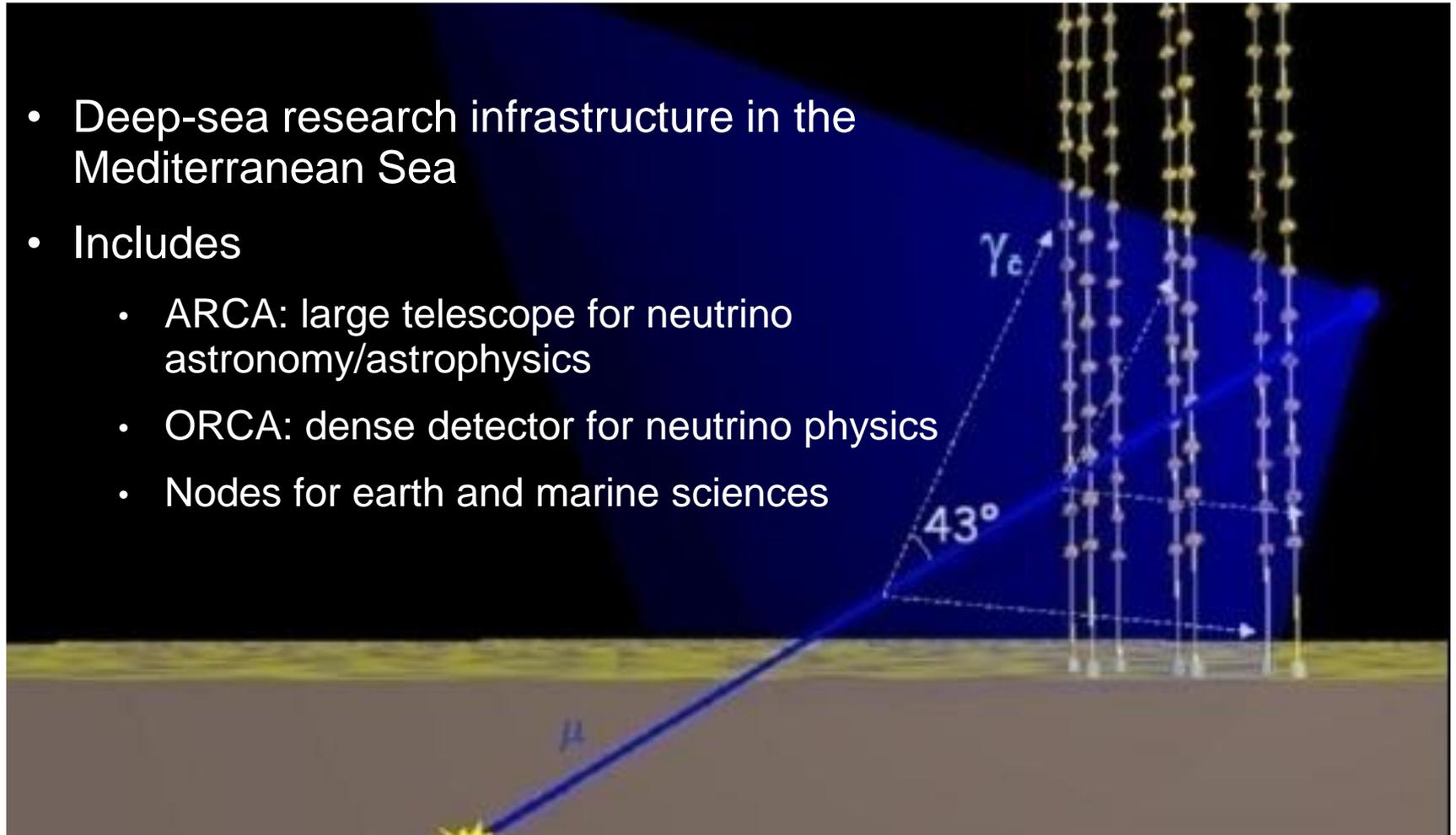# KM3NeT e-Needs

Uli Katz
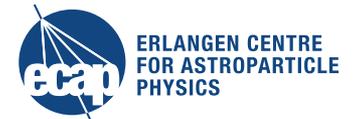KAPPA Workshop, Karlsruhe
2 November 2017

What is KM3NeT?

- Deep-sea research infrastructure in the Mediterranean Sea

- Includes

  - ARCA: large telescope for neutrino astronomy/astrophysics

  - ORCA: dense detector for neutrino physics

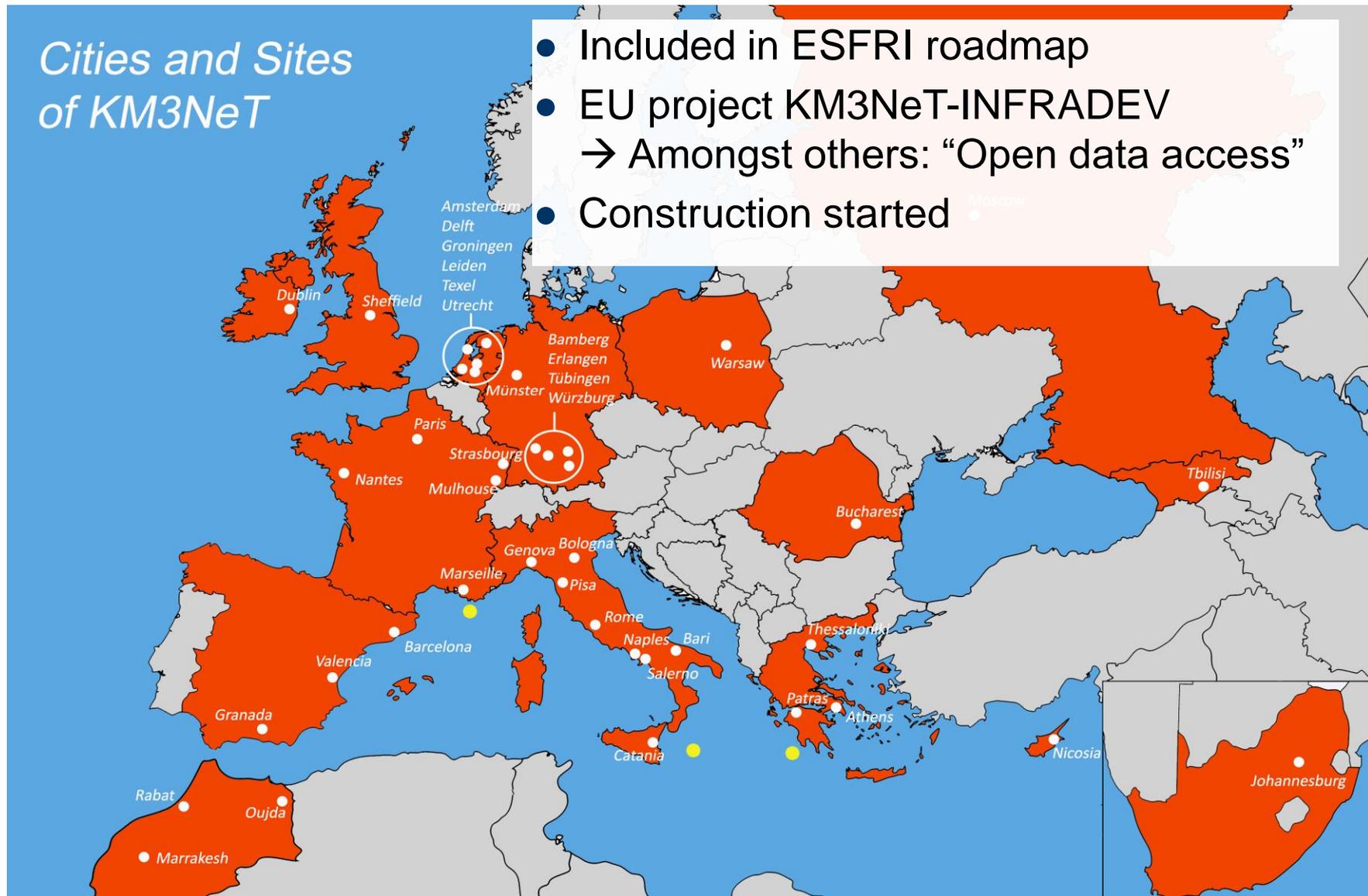 - Nodes for earth and marine sciences

The KM3NeT Collaboration

Cities and Sites of KM3NeT

- Included in ESFRI roadmap
- EU project KM3NeT-INFRADEV
  → Amongst others: "Open data access"
- Construction started

KM3NeT detector and data

Detector

- 31 3-inch PMTS   = 1 digital optical module (DOM)
- 18 DOMs   = 1 detection unit (DU)
- 115 DUs   = 1 building block (BB)
- 3 BBs   = KM3NeT Phase 2

Altogether: 192500 PMTs, 5-10 kHz single-photon rate each,
all data to shore ➔ some GB/s to filter ➔ ~100 MB/s to storage

Data processing

- Online data filter (on shore)
- Calibration and reconstruction
- Data analysis

+ Monte Carlo
+ Online analysis for alerts etc.
+ Earth and sea sciences
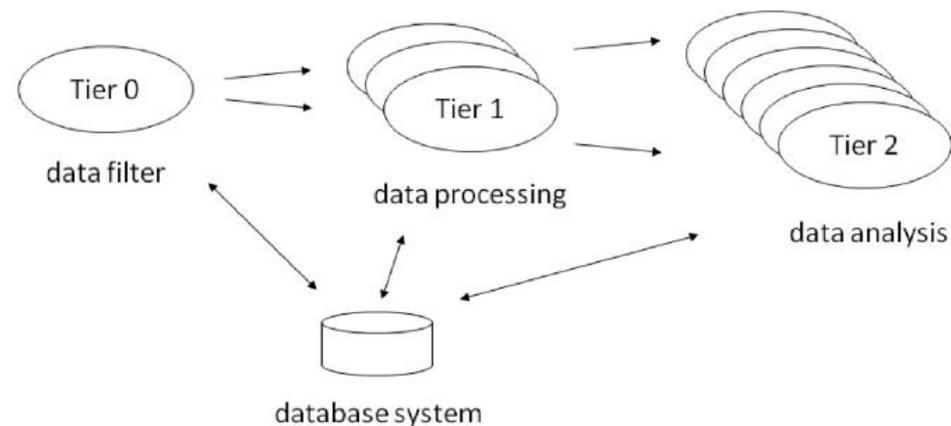
Requires advanced
data management

The data management plan

- Computing model

- FAIR data:
  - findable
  - accessible
  - interoperable
  - re-usable

KM3NeT INFRADEV – H2020 – 739560

KM3NeT Data Management Plan

KM3NeT-INFRADEV GA DELIVERABLE: D4.1

Document identifier:	KM3NeT-InfraDev-WP4-D4.1_v1.4
Date:	22 June 2017
Work package:	WP4
Lead partner:	FAU
Document status:	Endorsed by PMB and KM3NeT IB
Dissemination level:	Public

Computing Model: General Scheme

- Tier-like structure, mixed access: GRID + direct (batch)

Tier	Computing Facility	Processing steps	Access
Tier-0	at detector site	triggering, online-calibration, quasi-online reconstruction	direct access, direct processing
Tier-1	computing centres	calibration and reconstruction, simulation	direct access, batch processing and/or grid access
Tier-2	local computing clusters	simulation and analysis	varying

- Data transfer between the computing centers based on GRID access tools (where applicable)
- Central services funded through CNRS and INFN, additional services by the collaboration institutes

Computing model
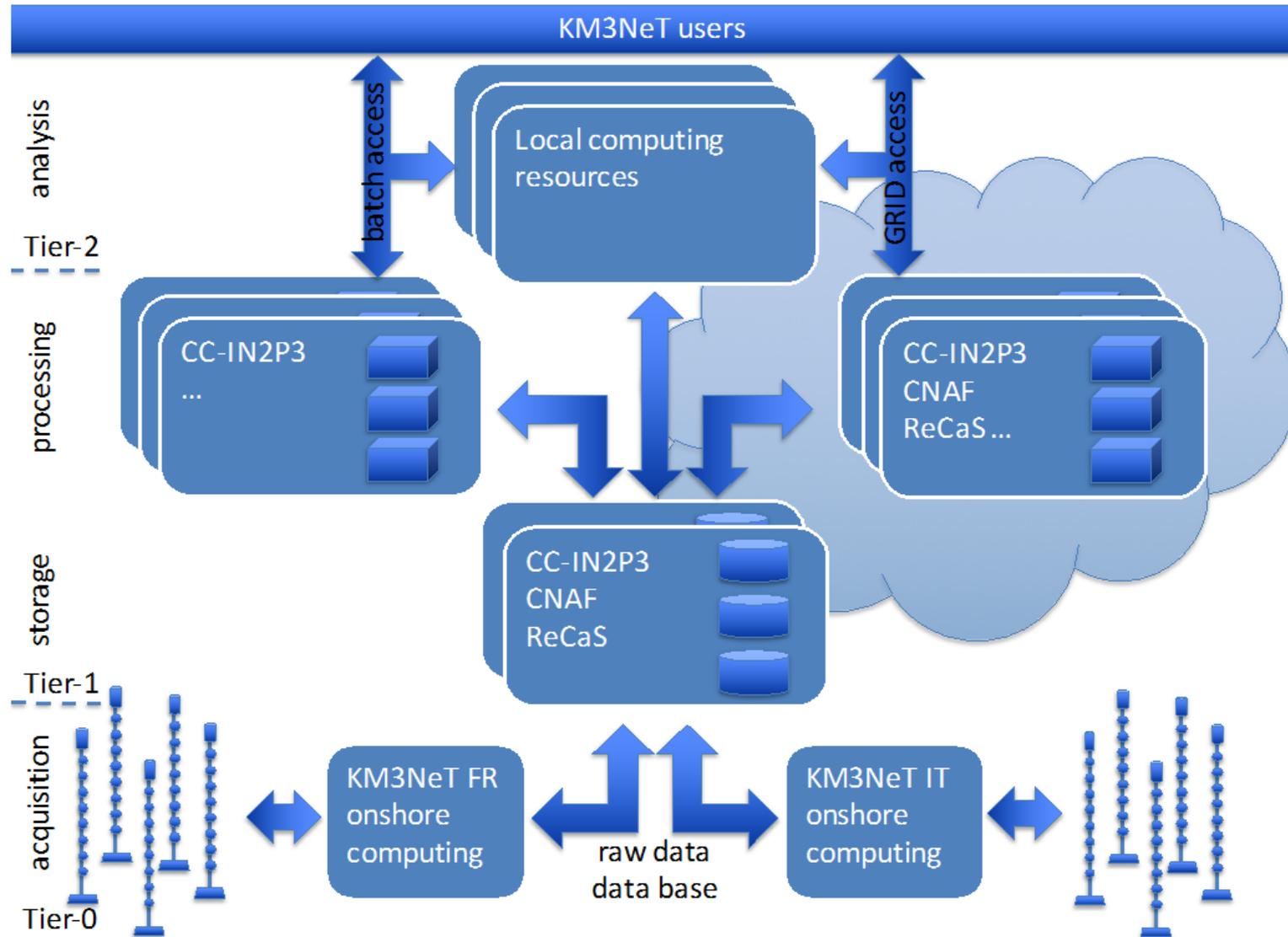
Computing Needs (per year)

- Includes:
  - Data processing (calibration and reconstruction)
  - Simulation

per year	size (TB)	computing time (HS06.h)	computing resources (HS06)
One Building Block	1000	350 M	40 k
Phase 1 (constr. 2016-2019)	300	60 M	7 k
- first year of operation	100	25 M	3 k
- second year of operation	150	40 M	5 k
Phase 2 (constr. 2019-2023)	2500	1 G	125 k

Data Management Plan

- Data formats and meta-data following common practice (root, xml, ascii) → allow for integration in eCommons

- Use of existing eCommons (e.g. from ANTARES) and specific development via ASTERICS

  - contacts with e.g. EGI, GÉANT, GARR, RENATER, surfNET

- Data storage:

  - essential: reproducibility of all scientific results and data usability over full time of experiment (+ at least 10 years after end of operation)

  - parallel storage of low- and high-level data at CC-Lyon and CNAF (long-term commitments, pledged resources)

  - central services like software repository, central software builds

- Data processing by specialized service group

- Data access: via WAN/GRID access tools (xrootd, iRODS and gridFTP)

Open-access data policy

- Public access to data
  - Summary data (event information plus quality information) after fixed latency (typically 2 years)
  - Web-based downloads of data and softwares
  - Also: Simulation data

- On request: More (detailed) data, earlier releases, etc.

- Harmonise with wider community ($\rightarrow$ GNN and ASTERICS), use common tools/platforms ($\rightarrow$ Virtual Observatory, VO), link to other eCommons

GNN = Global Neutrino Network;
ASTERICS = Astronomy ESFRI and Research Infrastructure Cluster;

KM3NeT needs in GADCAP[1]

- Centralised Tier-2 services (data access, data analysis, …)

- Possibly further Tier-1 centre (data processing)?

- Platform for DeepLearning approaches

- Support for public data access and data conservation
  (services under development, implementation not yet settled)

- Integrating APP activities:

  - Platform for common tools (VO) and (meta)data in common formats

  - Training

  - Outreach

[1] German Analysis and Data Centre for Astroparticle Physics

Summary

- KM3NeT construction started, detector with 3 blocks by ~2023

- Data management plan and computing model established

- KM3NeT adopts tier-model for data management
  (large data volumes, significant CPU requirements)

- Open data access in preparation (includes data archiving)

- KM3NeT could profit strongly from a German
  Analysis and Data Centre in Astroparticle Physics

Beware: Seek synergies with
EOSC = European Open Science Cloud