

# INTEROPERABLE DATA MANAGEMENT AND SERVICES IN ENVRI RIS

DR. ZHIMING ZHAO

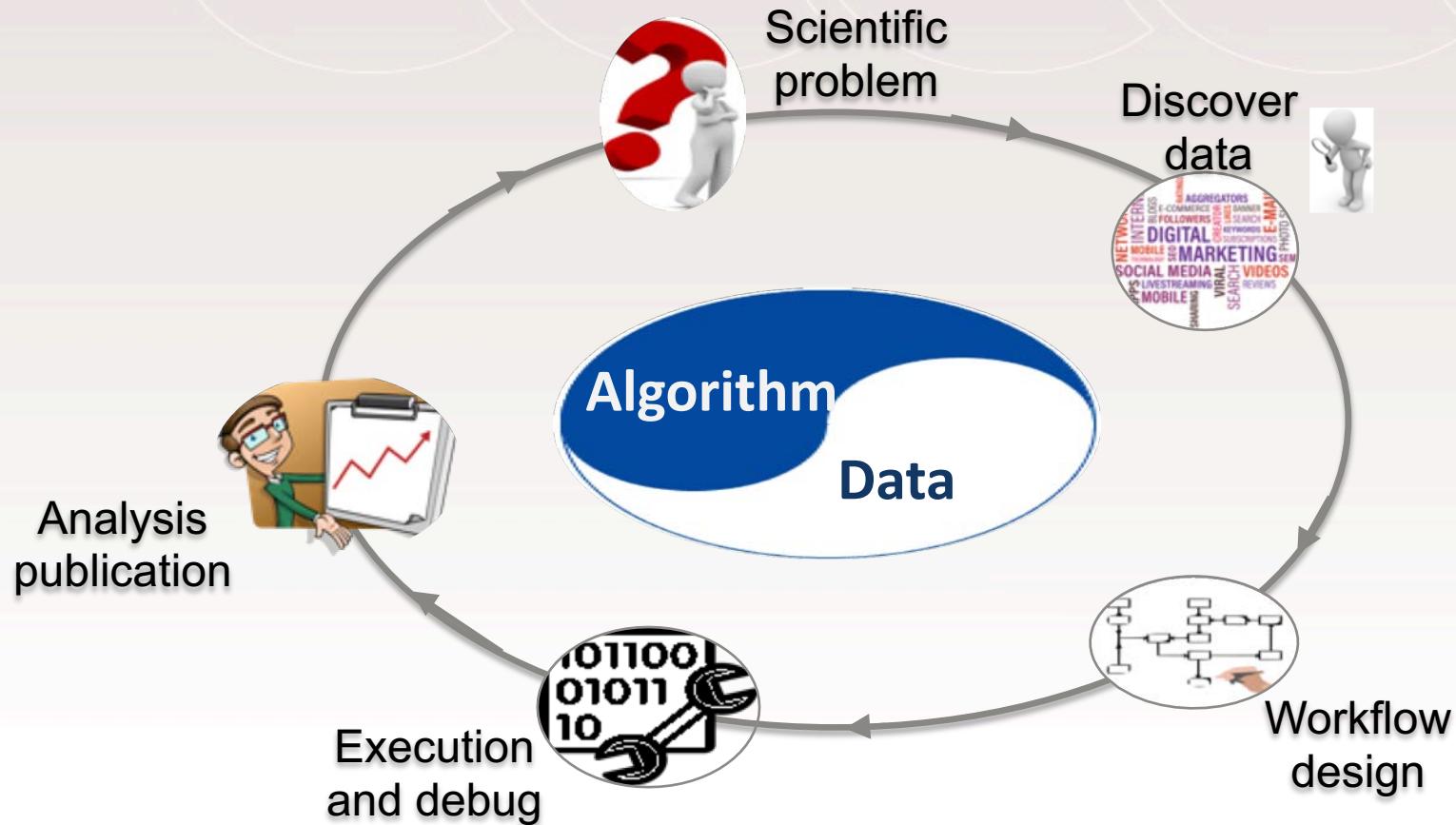
UNIVERSITY OF AMSTERDAM

ON BEHALF OF ALL PARTNERS IN THEME2

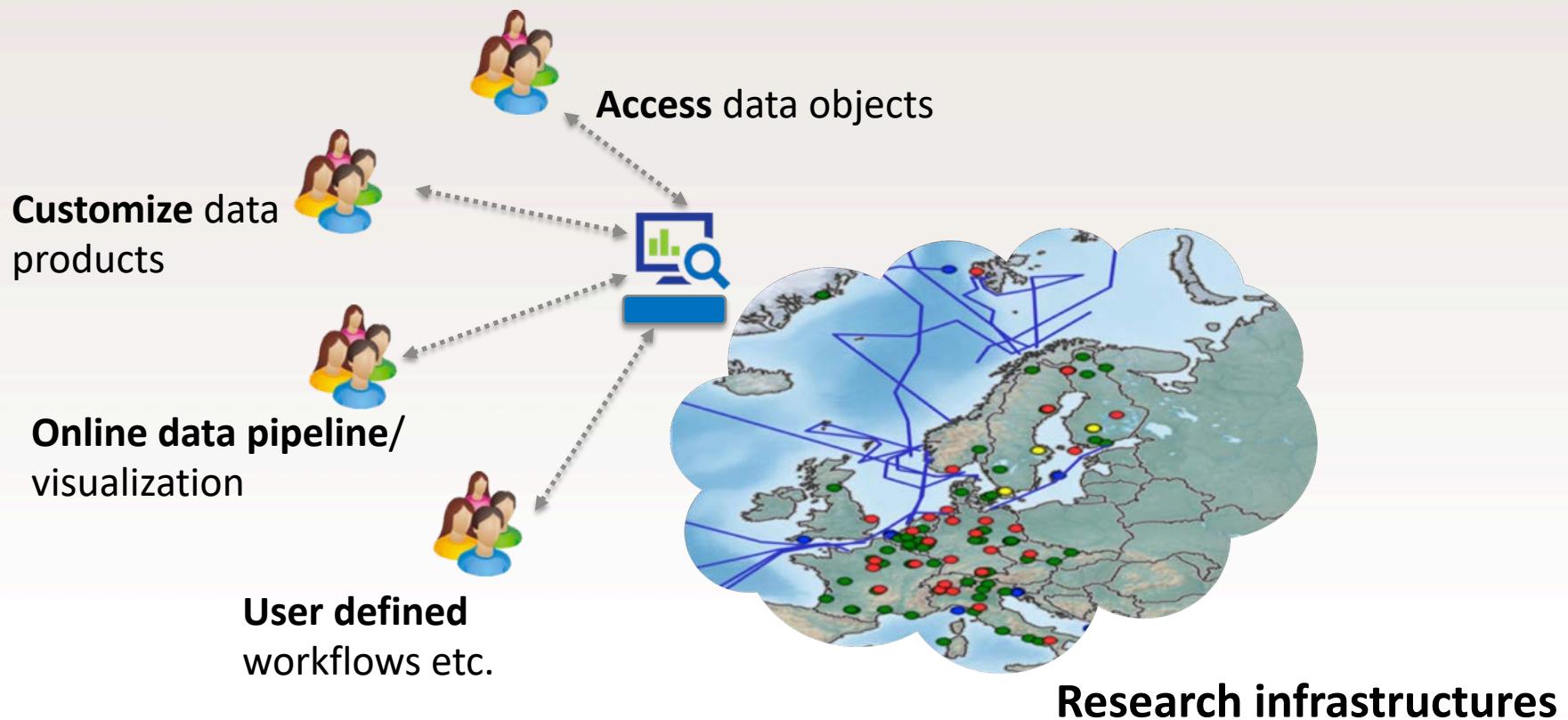


Supporting environmental research  
with integrated solutions  
**- the Earth is our lab**

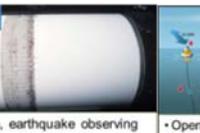
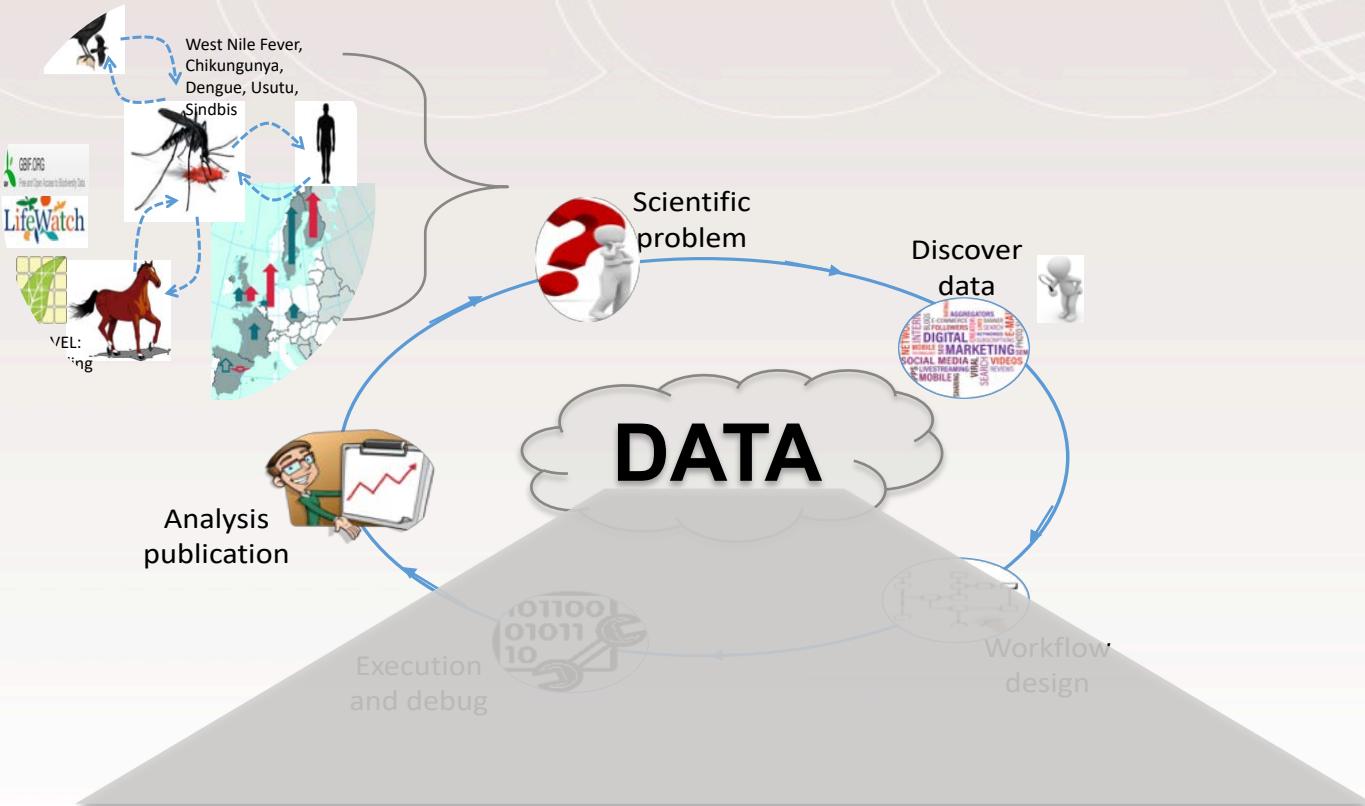
# DATA PLAYS A CENTRAL ROLE IN SCIENTIFIC RESEARCH



# RESEARCH INFRASTRUCTURES AIM TO BREAK BOUNDARY OF ISOLATED DATA SOURCES



# DATA CENTRIC SCIENCES ARE GETTING COMPLEX



• Upper space observing

• Deep sea observing

• Volcano, earthquake observing

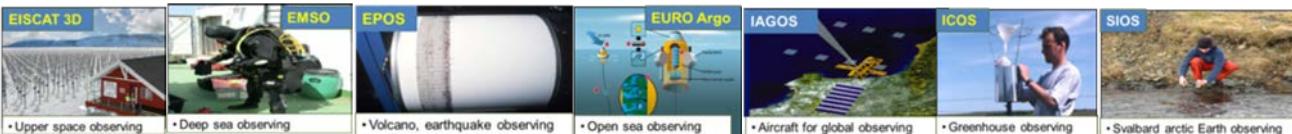
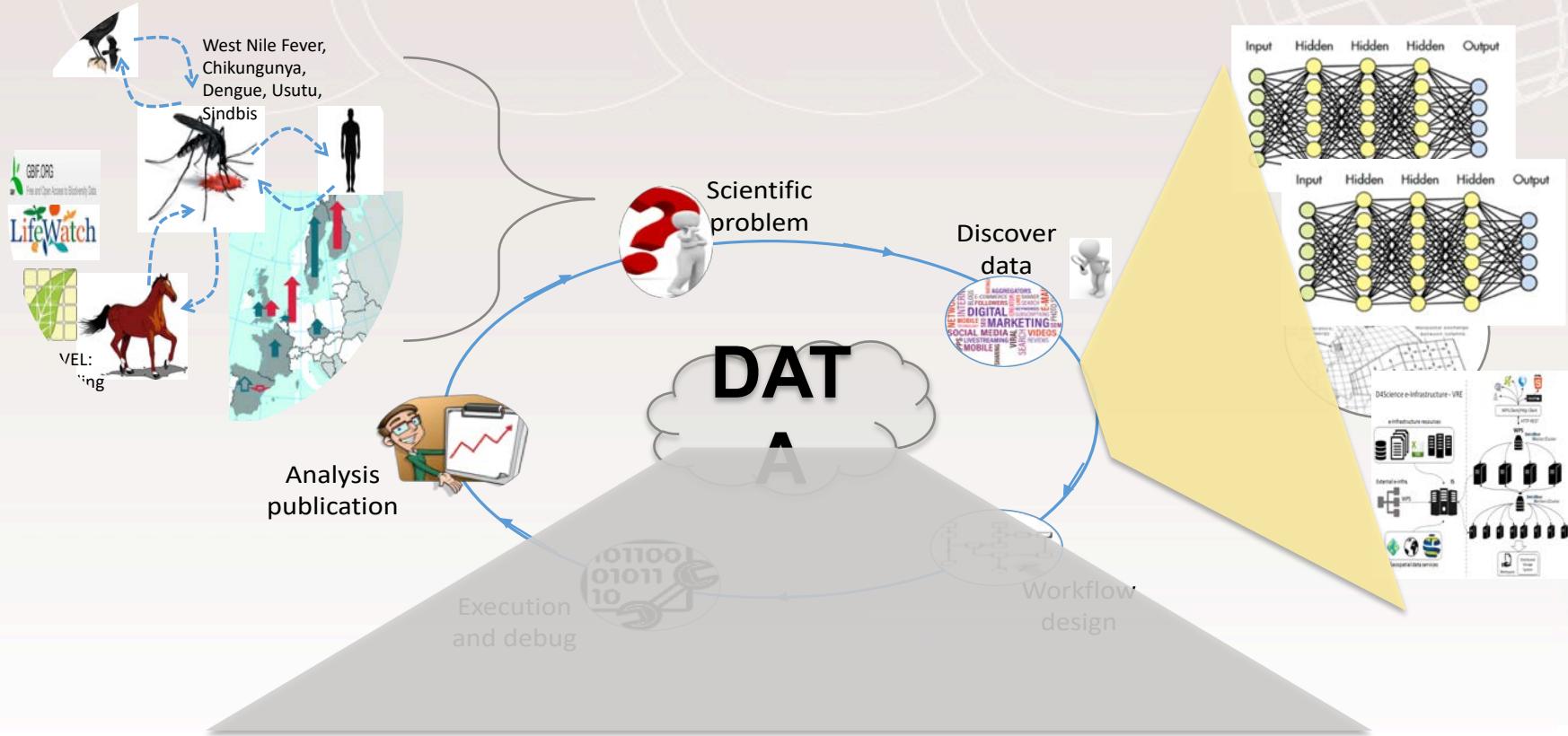
• Open sea observing

• Aircraft for global observing

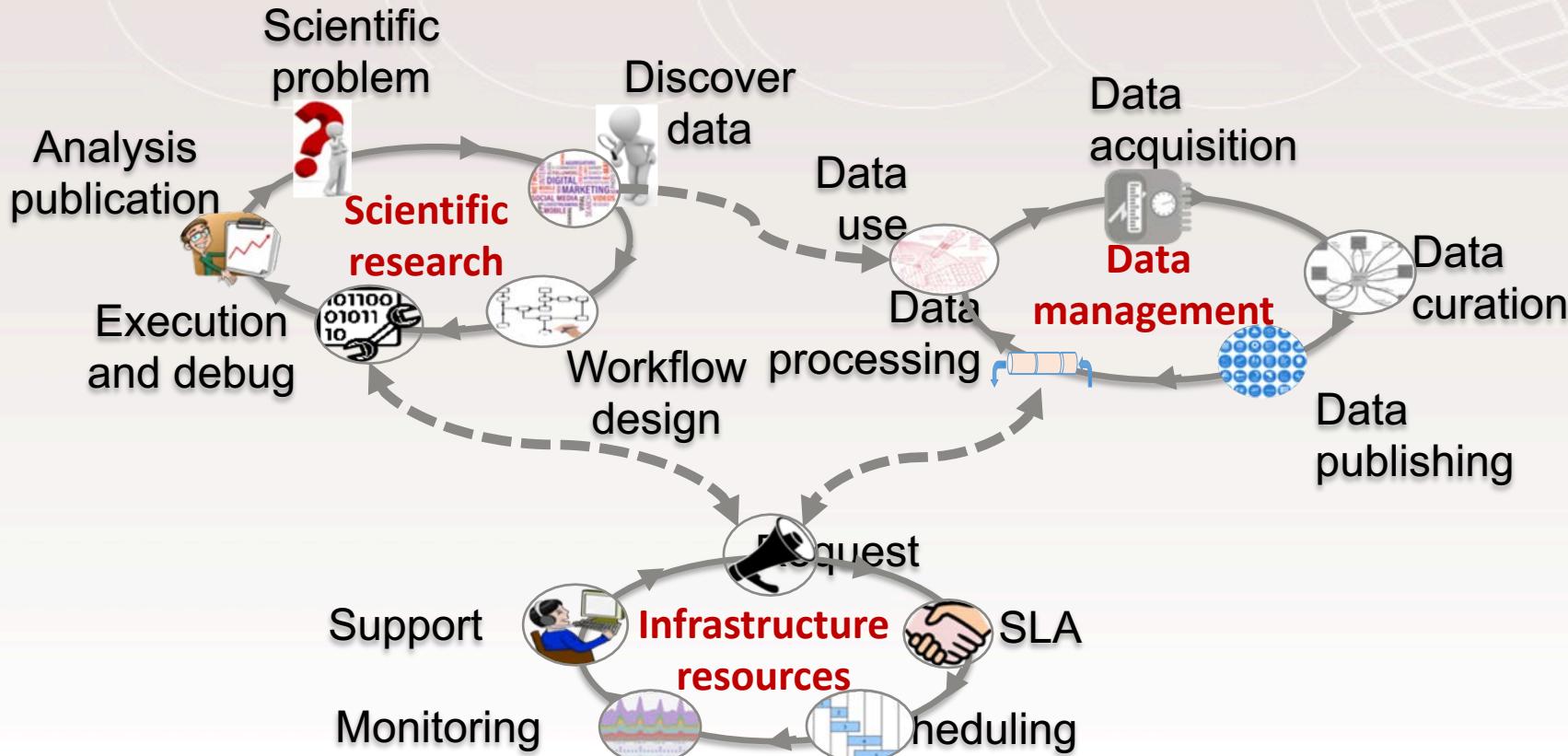
• Greenhouse observing

• Svalbard arctic Earth observing

# DATA CENTRIC SCIENCES ARE GETTING COMPLEX



# DATA CENTRIC SCIENCES ARE GETTING COMPLICATED



# THEME2 SERVICE PORTFOLIO

## A. Reference model related

- A1: Reference model training service - CU
- A2: Open information linking for ENV-RIs - UvA
- A3: ENVRI knowledge base - UvA
- A4: RI architecture design – NERC

## B. Theme2 service pillar

- B1: Linked open data ingestion and metadata service– ICOS/LU
- B2: D4science data analytics - CNR
- B3: Dynamic real-time infrastructure planner - UvA
- B4: Curation - NERC
- B5: Flagship cataloguing - IFREMER
- B6: Provenance - EAA

## C. Reusable solution from use cases/RIs

- C1: Data subscription service - EUDAT
- C2: Pipeline for semantic annotation of relational DB – ANAEE/INRA
- C3: Data / metadata generation from semantic annotations- ANAEE/INRA
- C4: Dynamic ecological information management system (DEIMS)- LTER/EAA
- C5: Biodiversity Community Portal (LifeWatch/LTER)- EAA

## D. Software quality check and testbed

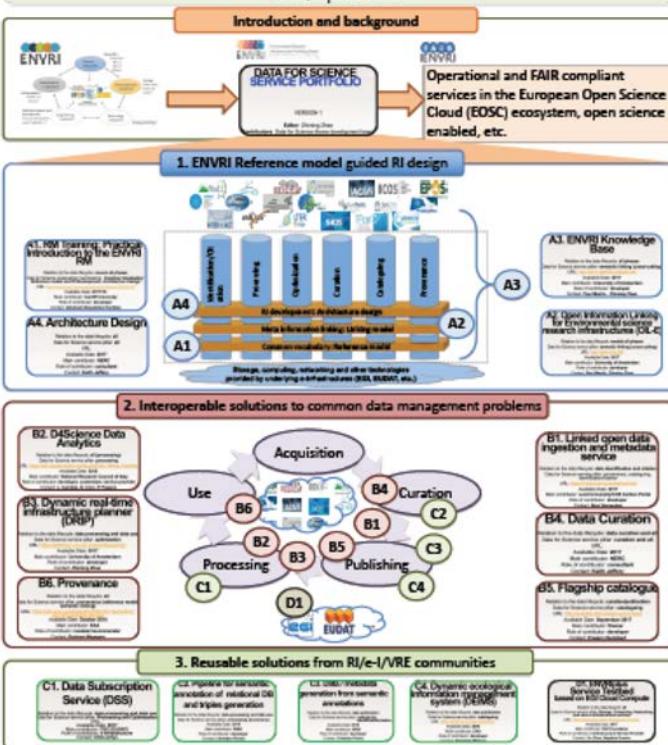
- D1: Envriplus service test bed - EGI

### Software and services towards interoperable data management

— A Service portfolio from the ENVRI<sup>PLUS</sup> Data for Science theme

Zhiming Zhao, Paul Martin, Malcolm Atkinson, Abraham Nierva De La Hidalga, Alex Hardisty, Keith Jeffery, Alex Vermeulen, Margareta Hellström, Leonardo Candela, Erwan Quimbert, Barbara Magagna, Chris Ariyo, Christian Pichot, Thierry Carval, YinThe ENVRI<sup>PLUS</sup> Data for Science Theme technical team

Chen, Baptiste Grenier



EU Horizon 2020 research and innovation programme under grant agreement No 654142



Environmental Research  
Infrastructures Providing Shared  
Solutions for Science and Society

Contact: Dr. Zhiming Zhao ([z.zhao@uva.nl](mailto:z.zhao@uva.nl)).  
ENVRI IRI: [www.envri.eu/](http://www.envri.eu/)  
Portfolio:  
<https://services.science.org/europa/envri/udata-science/portfolio/>



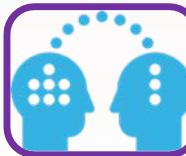
# SOME EXAMPLE SOLUTIONS IN THE PORTFOLIO



A1. Reference model training service



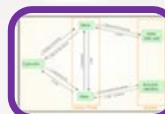
A2. Open information linking for environmental RIs



A3. ENVRI knowledge base



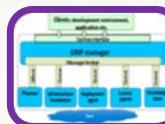
A4. Architecture guideline



B1: Linked open data ingestion and metadata service



B2: D4science data analytics



B3: Dynamic real-time infrastructure planner



B4: Flagship cataloguing

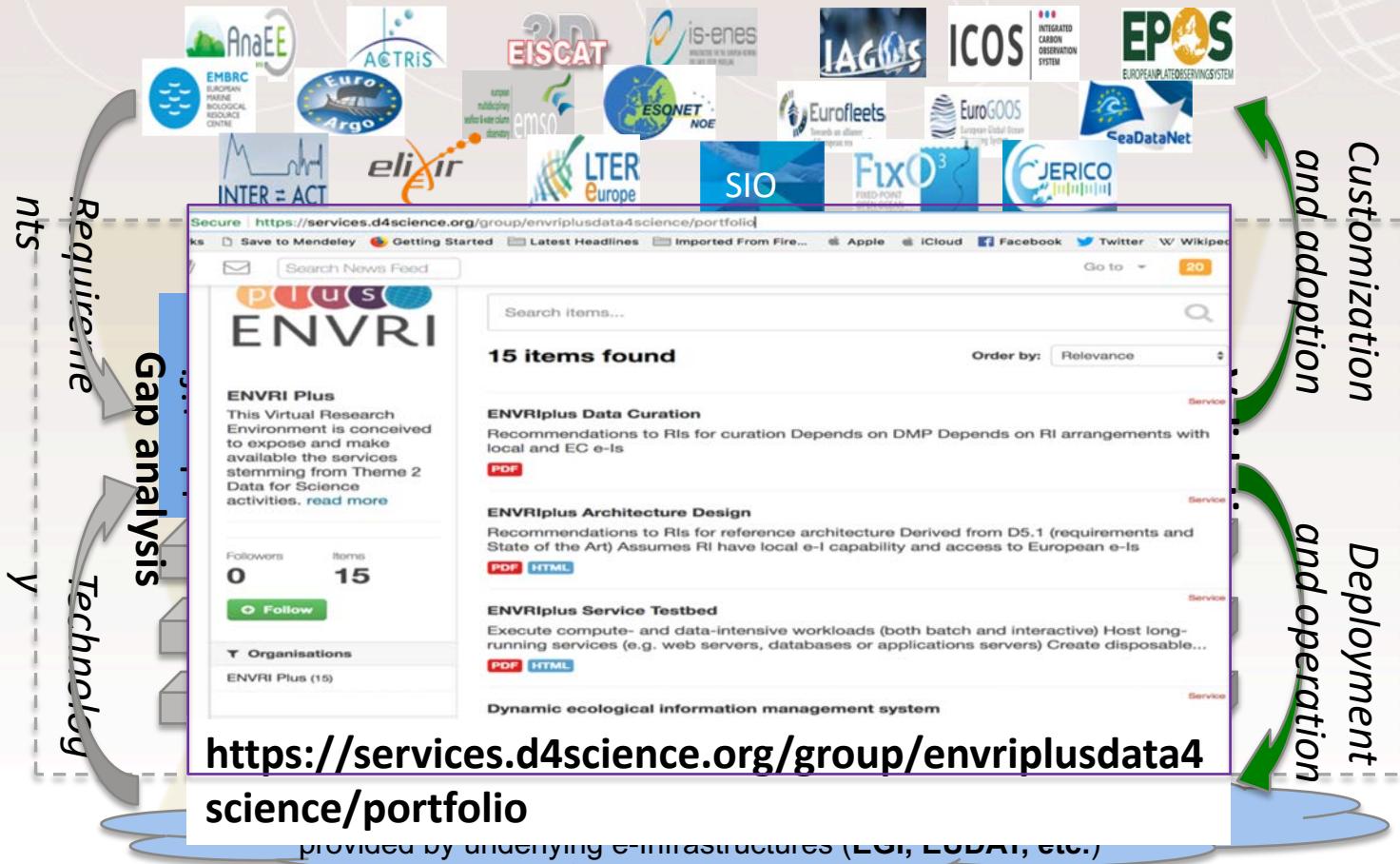


B5: Curation



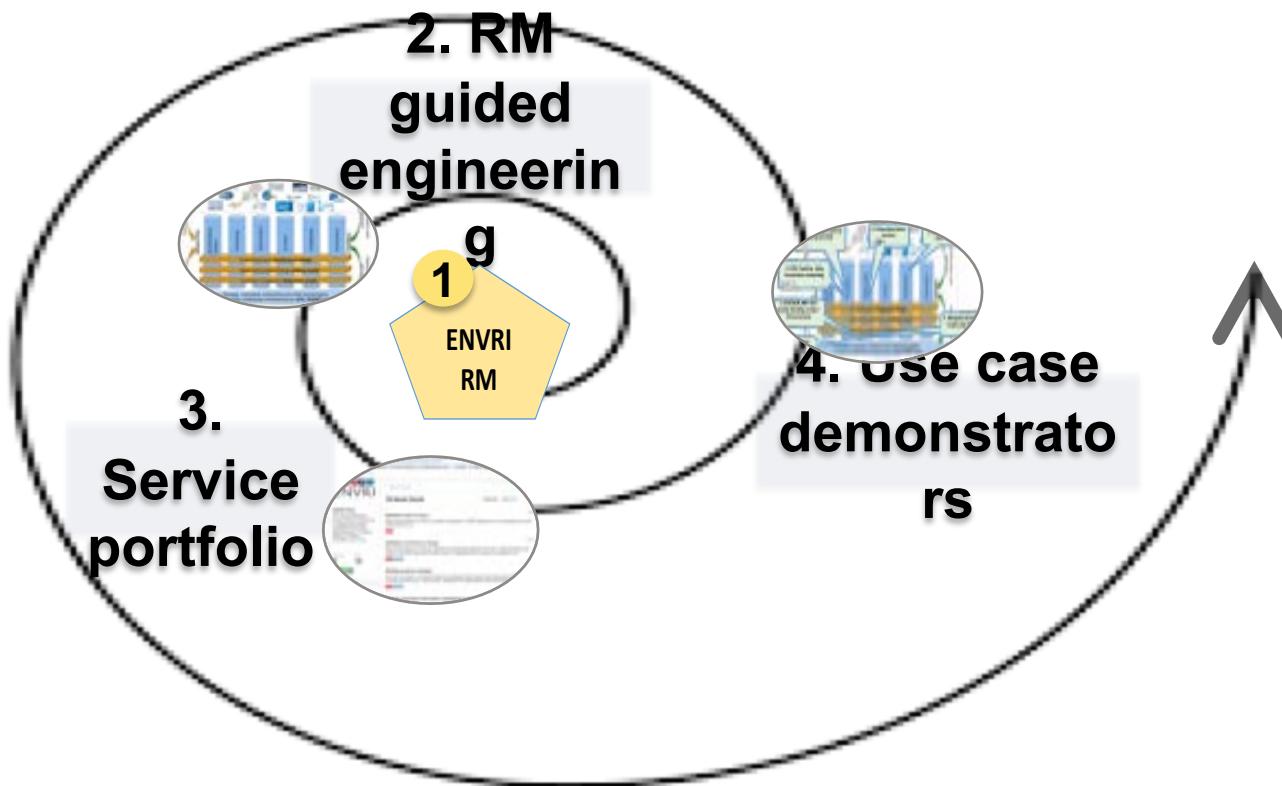
B6: Provenance

# SERVICE PORTFOLIO: REUSABLE SOLUTIONS

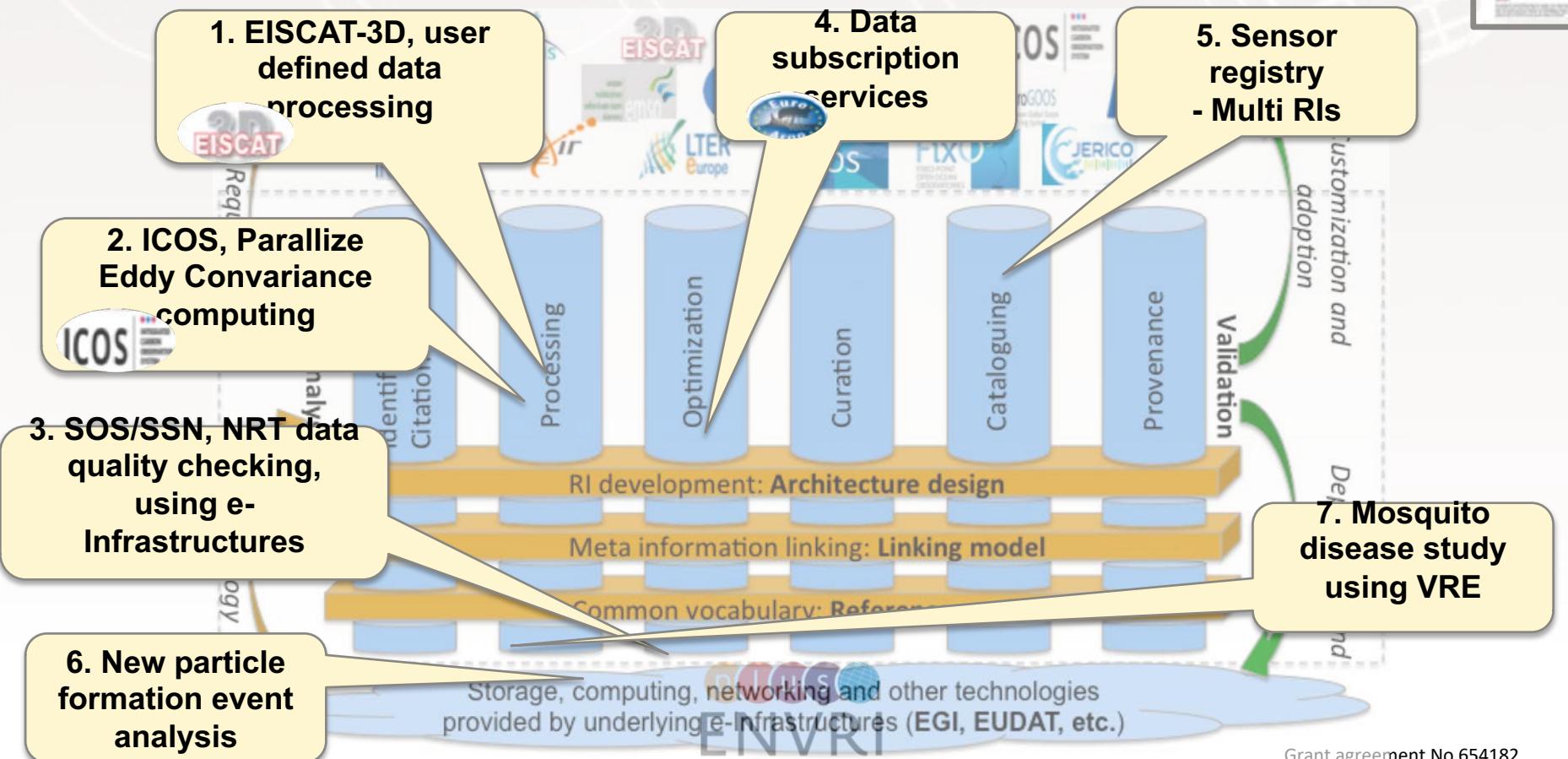


# HIGHLIGHTS 4: USE CASE DEMONSTRATOR

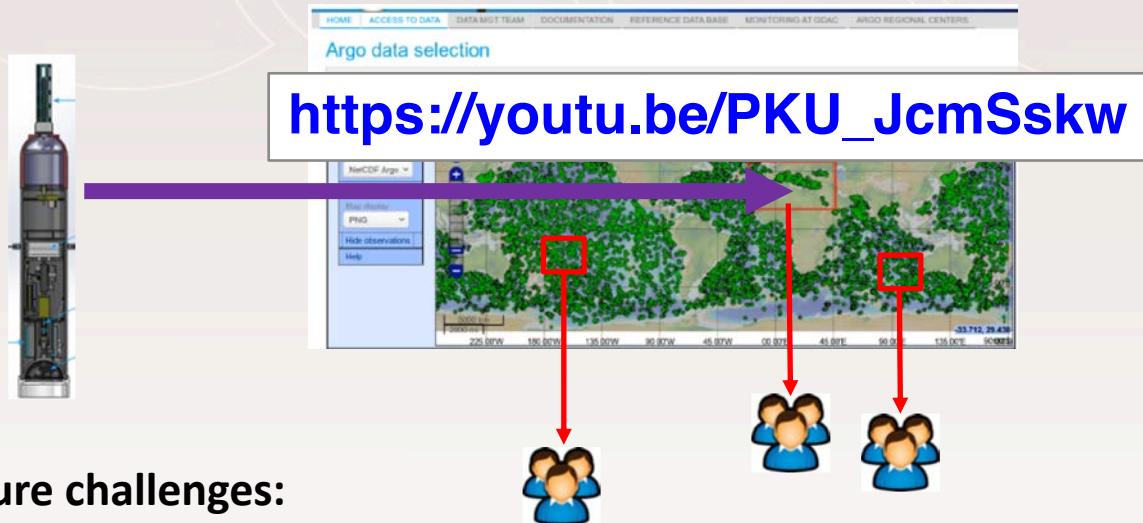
Theme 2 ICT specialists



## **HIGHLIGHT 4: USE CASE DEMONSTRATORS**



# DEMONSTRATOR 1: SUBSCRIBING DATA FROM RI



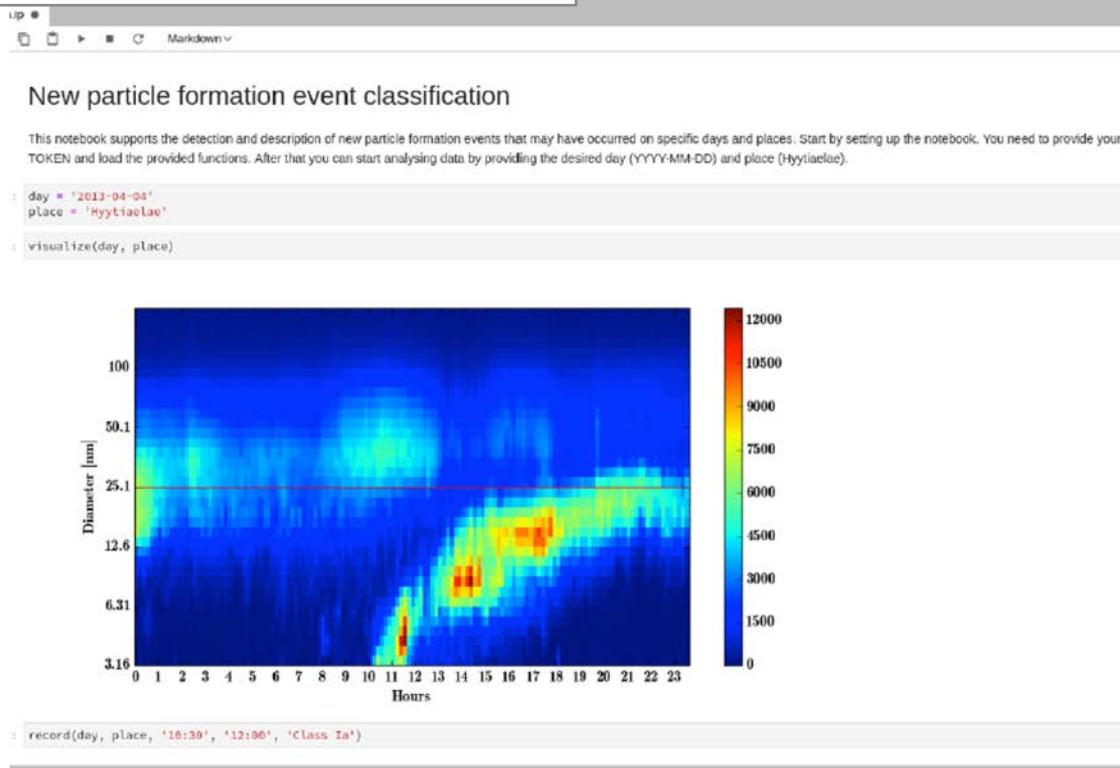
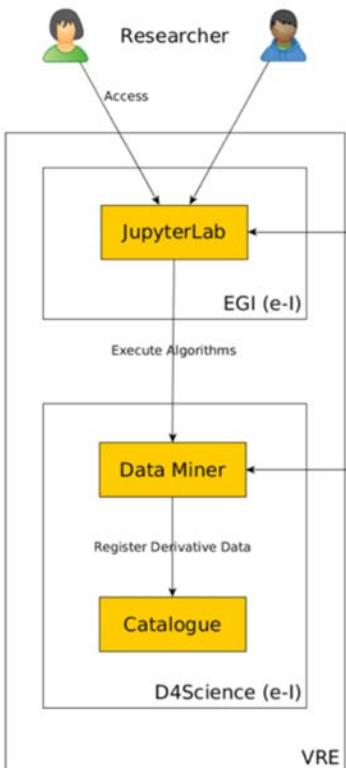
## Infrastructure challenges:

1. Subscriptions are diverse and changing
2. Subscriptions require different data region, time duration and computation
3. Subscriptions are for different purposes: simulation models, real-time decisions, or research etc.
4. Subscriptions have different requirements: e.g., time critical or high QoS

Koulouzis S. et al., (2019), Time-critical data management in clouds: challenges and a Dynamic Real-time Infrastructure Planner (DRIP) solution. *Concurrency Computat Pract*

# DEMONSTRATOR 2: NEW PARTICLE FORMATION EVENT ANALYSIS

<https://youtu.be/ra9W7b5DbgI>

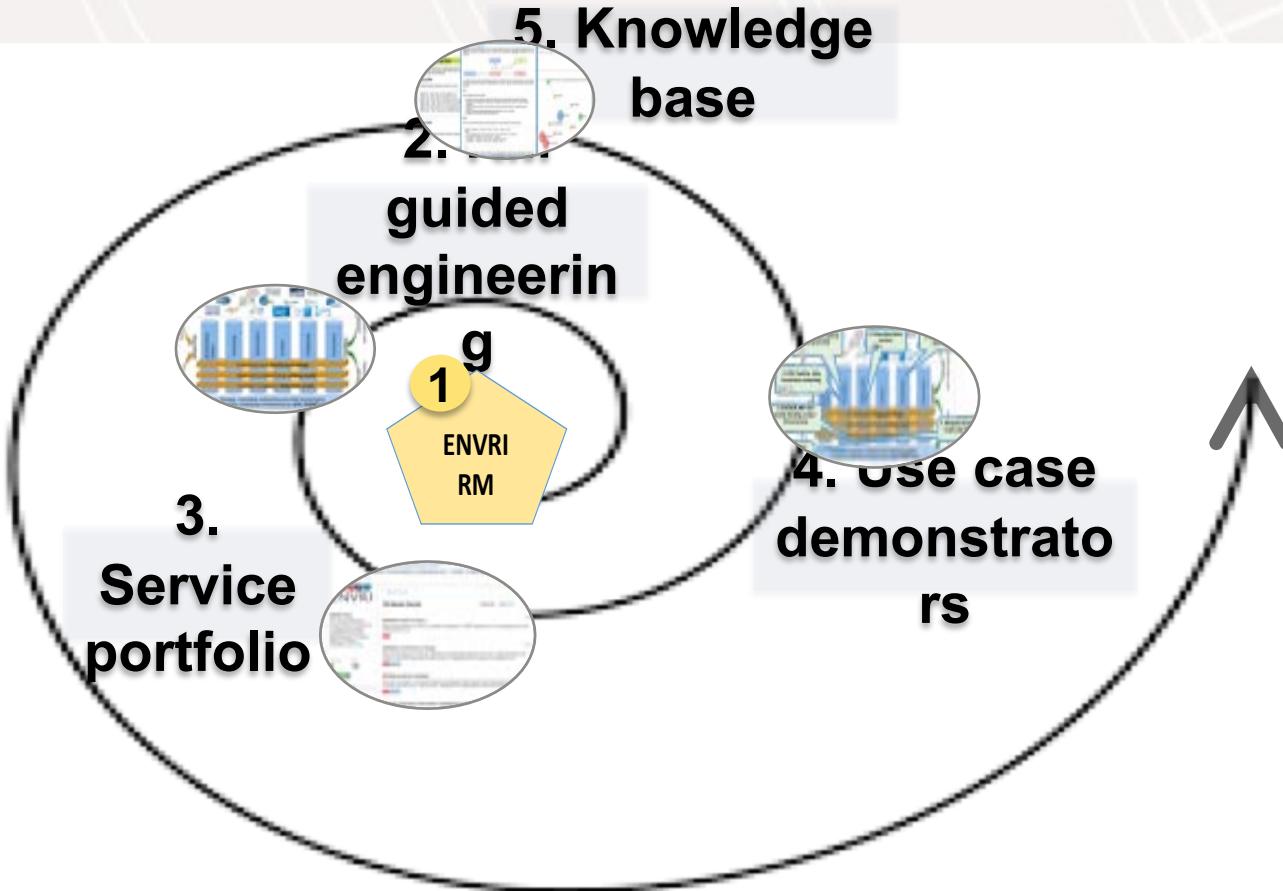


## MORE INFORMATION

**D9.1 and D9.2** Service deployment in computing and data e-Infrastructures (Version1 and 2)

# HIGHLIGHTS 5: KNOWLEDGE BASE

Theme 2 ICT specialists



# THEME2 KNOWLEDGE BASE



Architecture, communities, roles, data management services, quality control, tools, standards etc.

The screenshot displays three main sections:

- ENVRI+ Knowledge Base:** A "notebook" for queries. It shows a code snippet in SPARQL:

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/1999/02/22-rdf-rdf#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-rdf#>
PREFIX base: <http://www.oil-e.net/ontology/oil-e>
PREFIX rm: <http://www.oil-e.net/ontology/envri-rm>
PREFIX data: <http://www.envri.eu/community/data/>
```

- Scenario analysis:** A section titled "We want to be able to identify the research infrastructure component scenarios." It includes a code snippet:

```
SELECT ?property_label ?object_label
WHERE {
  ?addMetadata rdfs:subClassOf ?newProperty ?property
  ?newProperty rdfs:subPropertyOf ?hasValue
  ?hasValue rdfs:label ?property_label
  ?newProperty rdfs:label ?object_label
}
```

- Graph visualization:** A network diagram showing nodes like "envriplus", "EnvCommunity", "EnvProject", "EnvRI", "EnvRIplus", and "EnvRIplus". Edges represent relationships such as "base\_supersetProject", "base\_prefLabel", "base\_homepage", and "base\_isMemberOf".



Prototype: <http://oil-e.vlan400.uvalight.net>

plus  
ENVRI

10.05.2020

Requirement: <https://bit.ly/KB-REQ>

Grant agreement No 654182

# KEY HIGHLIGHTS

1. ENVRI reference model (*common vocabularies*)
2. Reference model guided engineering (*methodologies*)
3. Theme2 service portfolio (*key development results*)
4. Use case demonstrators (*validation*)
5. Theme2 knowledge base (*knowledge curation*)
6. Theme2 summer school (*knowledge transfer*)
7. Theme2 book (*knowledge transfer*)
8. Towards ENVRI-FAIR (*towards FAIR and EOSC*)



Horizon 2020



Grant agreement No 654182

# WRAP UP

