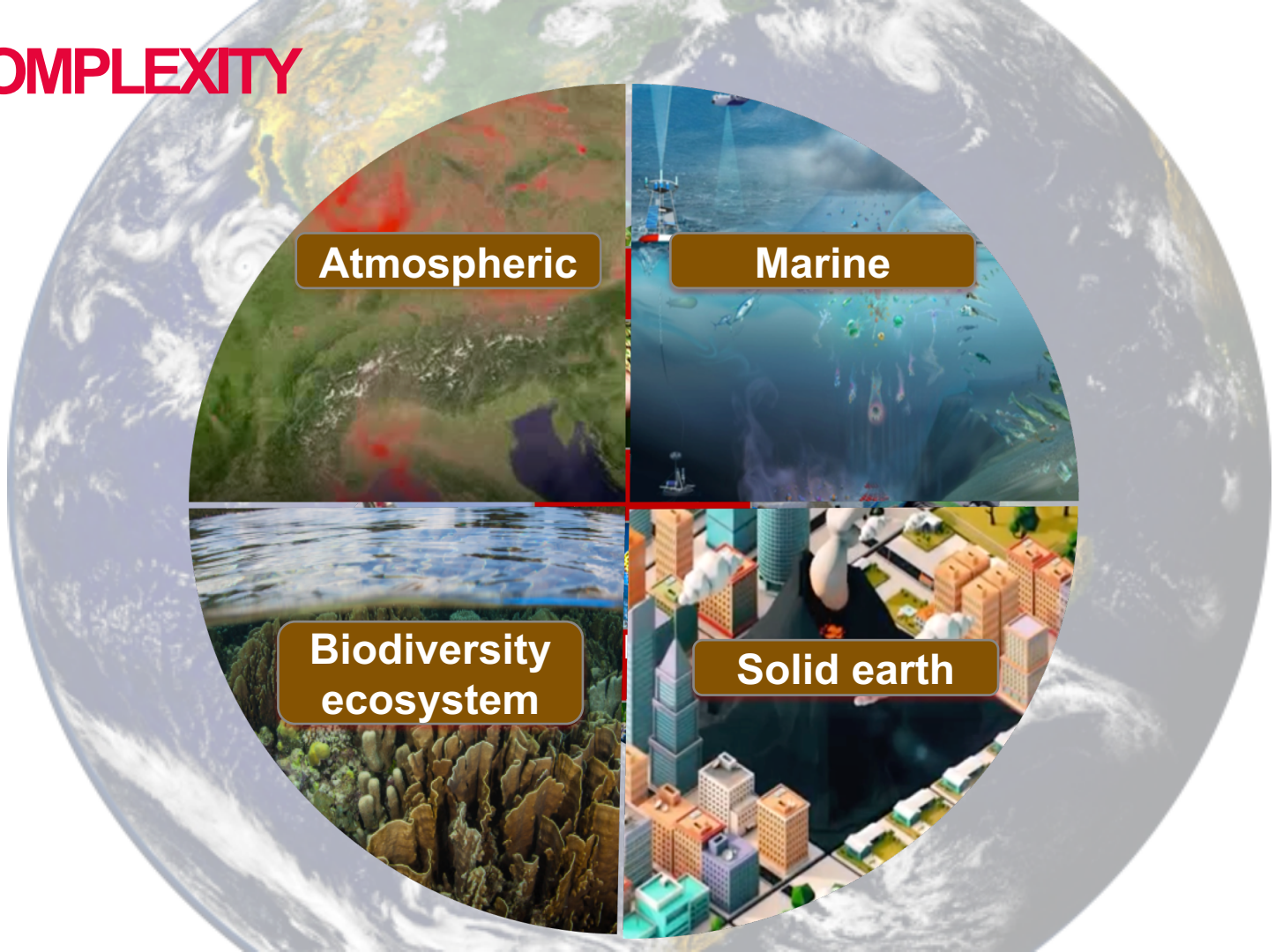
DR. ZHIMING ZHAO

Z.ZHAO@UVA.NL

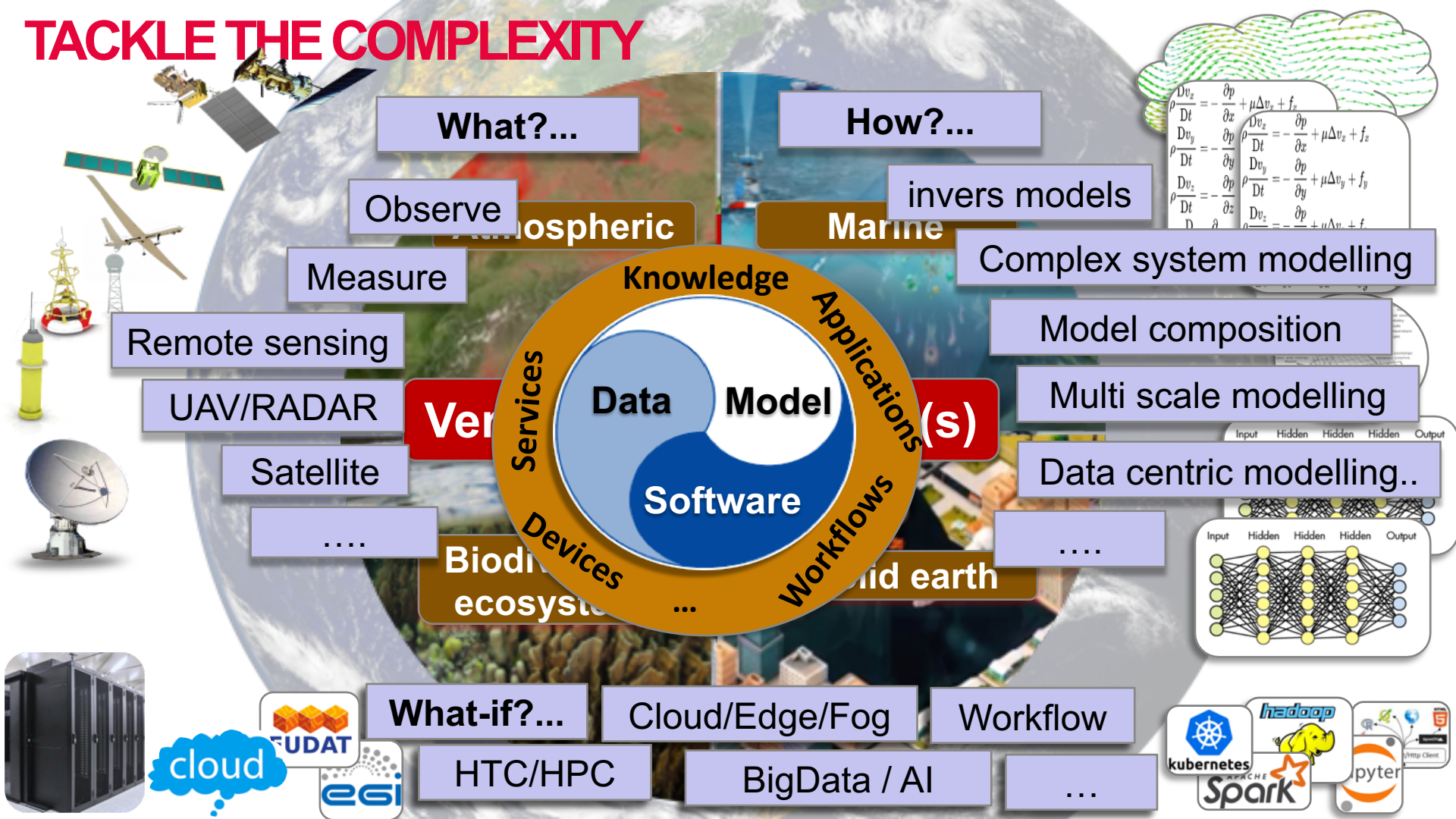
UNIVERSITY OF AMSTERDAM

INTERDISCIPLINARY RESEARCH

COMPLEXITY



TACKLE THE COMPLEXITY



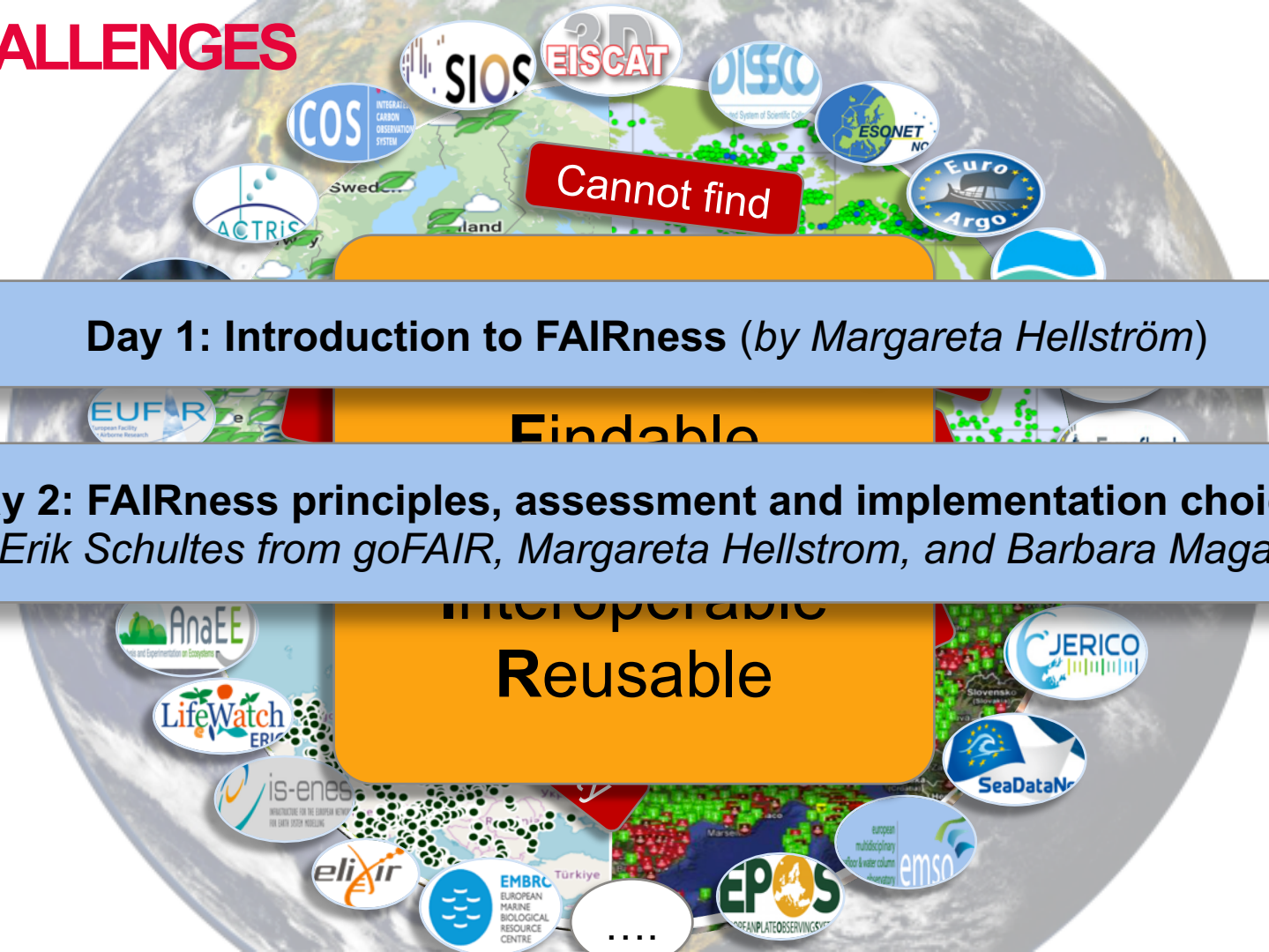
RI CHALLENGES

Research Infrastructure (RI)

- **Facilities, resources and related services** that are used by the scientific community to conduct top-level research
- **knowledge-based resources** such as collections, archives or structures for scientific information;
- enabling **Information and Communications Technology-based infrastructures** to achieve excellence in research.
- Such infrastructures may be "**single-sited**" or "**distributed**" (an organised network of resources),

---- EU ERIC practical guideline [10.2777/79873]

RI CHALLENGES



Day 1: Introduction to FAIRness (*by Margareta Hellström*)

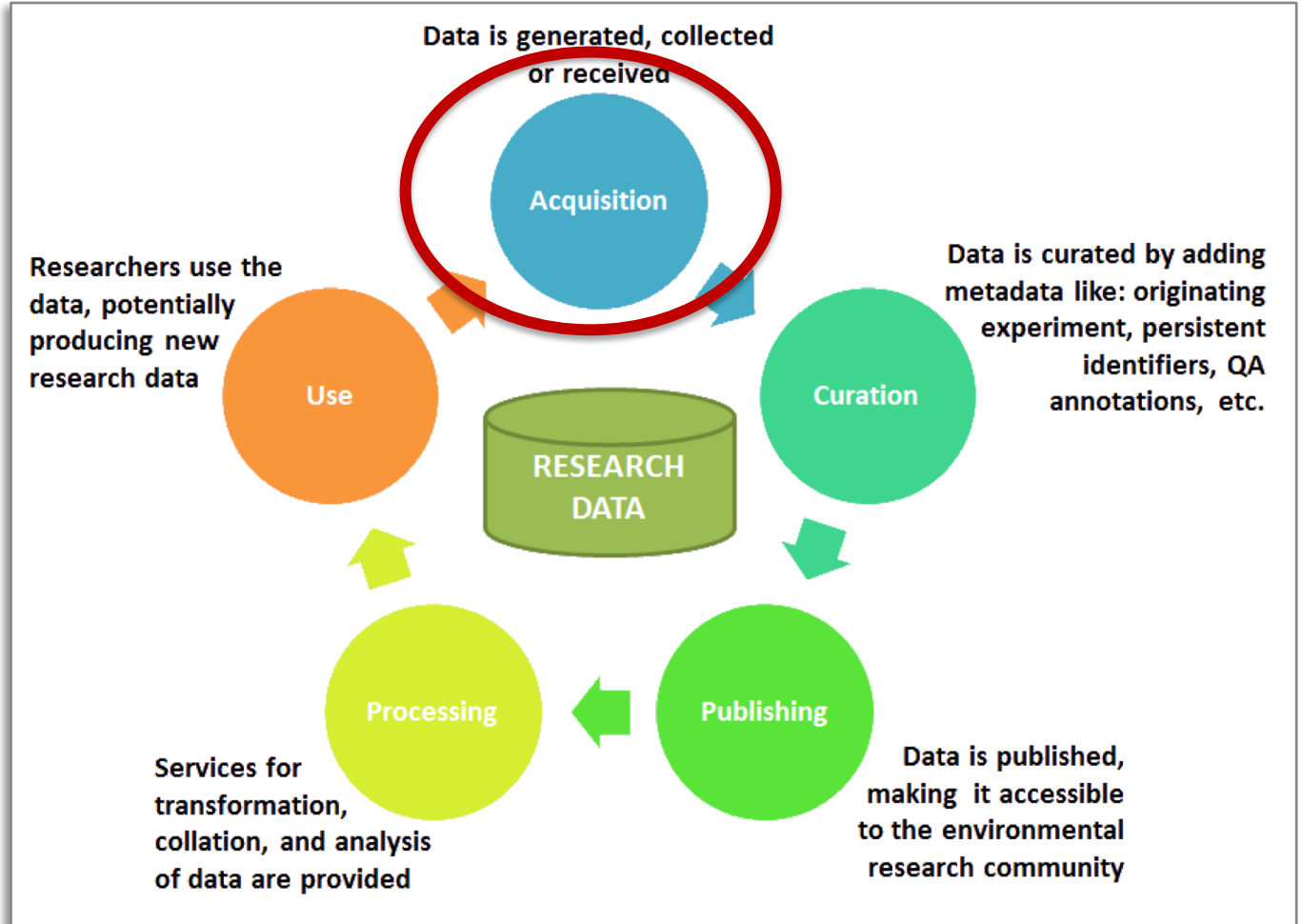
Findable

Day 2: FAIRness principles, assessment and implementation choices
(*by Erik Schultes from goFAIR, Margareta Hellstrom, and Barbara Magagna*)

Interoperable

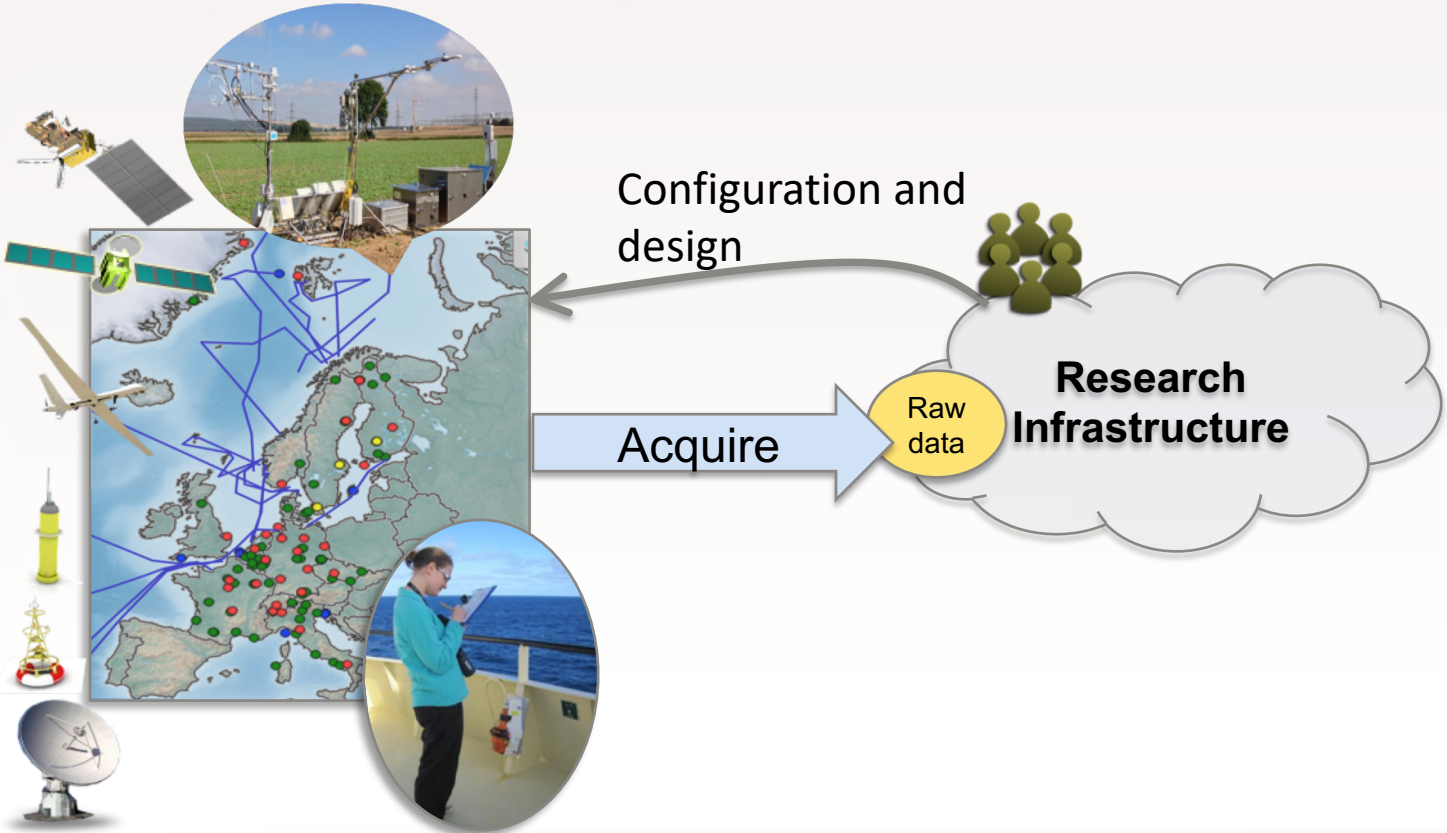
Reusable

DATA MANAGEMENT LIFECYCLE



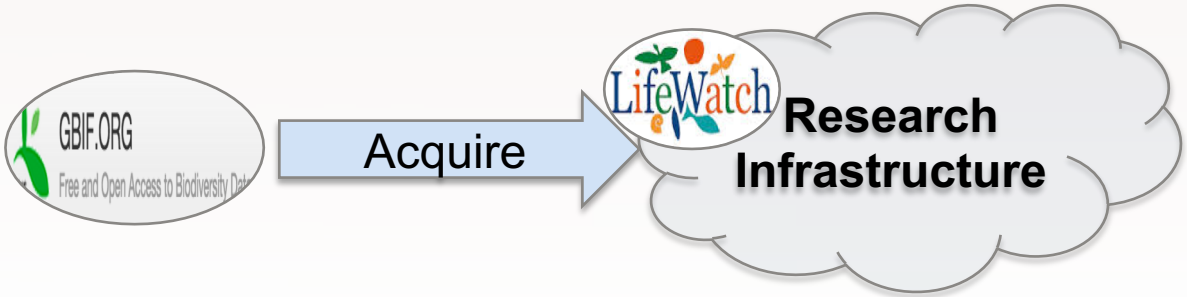
ACQUISITION

- From sensors or observation network

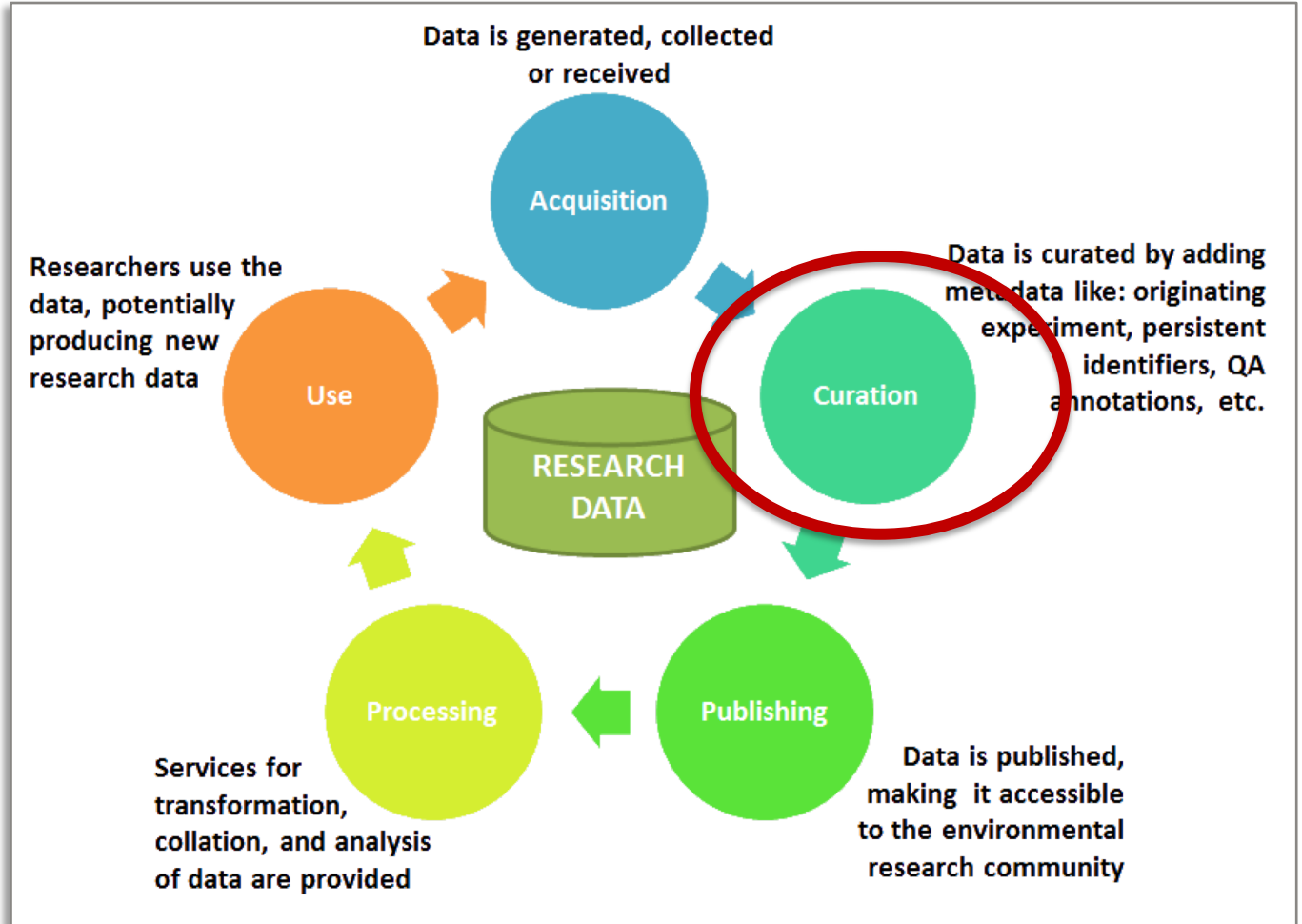


ACQUISITION

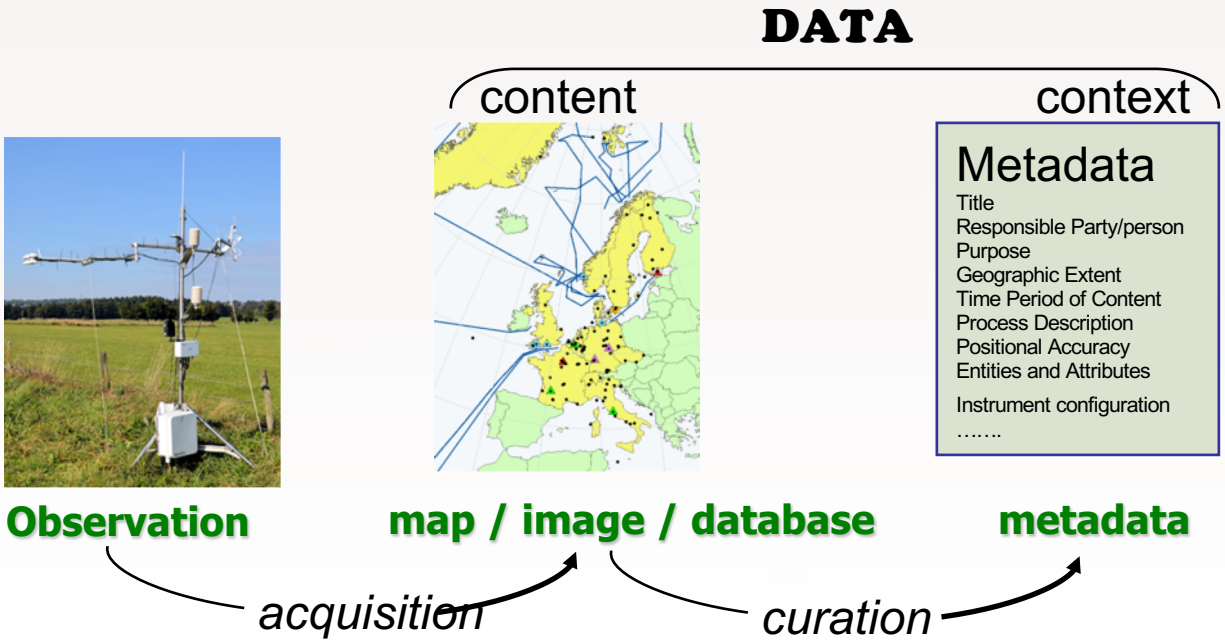
- From sensors or observation network
- From another infrastructure



DATA MANAGEMENT LIFECYCLE



DATA AS A PRODUCT

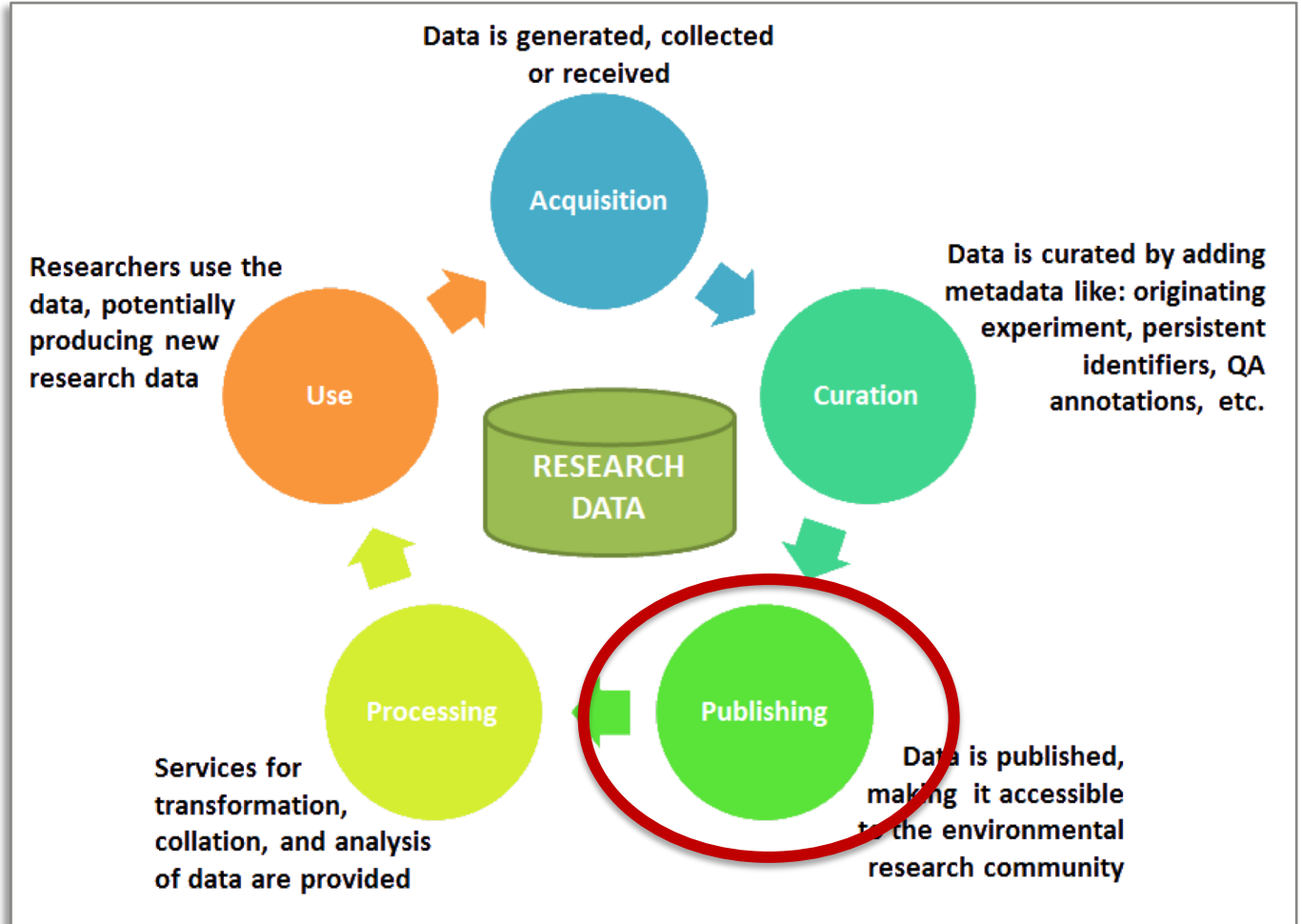


META INFORMATION NEEDED

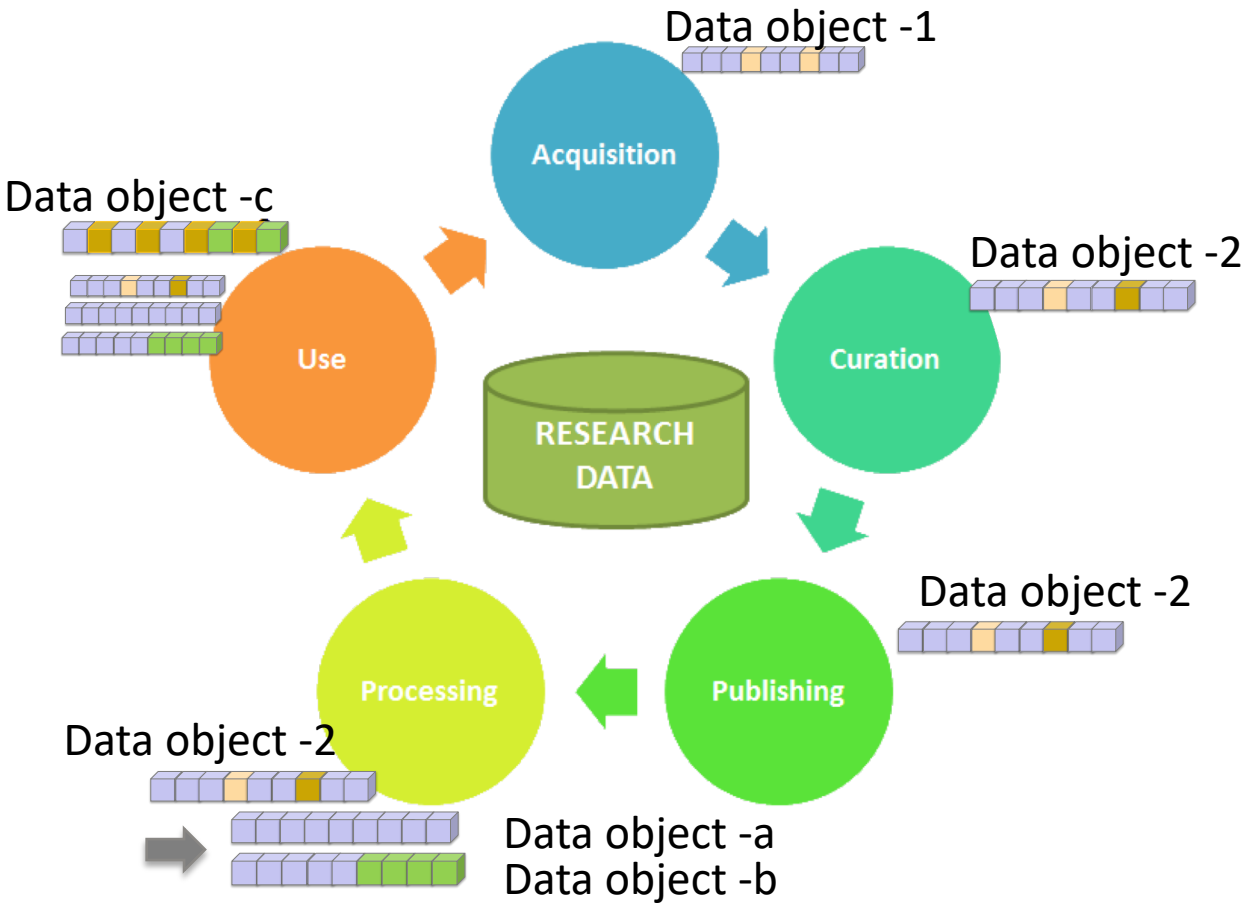
Who	• Is the contact person? • created/collected/measured the data?
Where	• is the data created/collected/measured, e.g., site? • can the data be accessed?
How	• was the data created/collected/measured? • is the instrument configured?
When	• is the time period of the collection/measurement? • was the data created/imported?
Why	• was the data created? • are there missing values?

Day 1: Afternoon
Data curation and preservation (*by José Maria Garcia*)

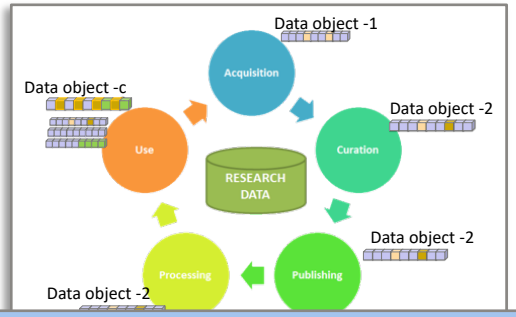
DATA MANAGEMENT LIFECYCLE



DATA IDENTIFICATION AND CITATION



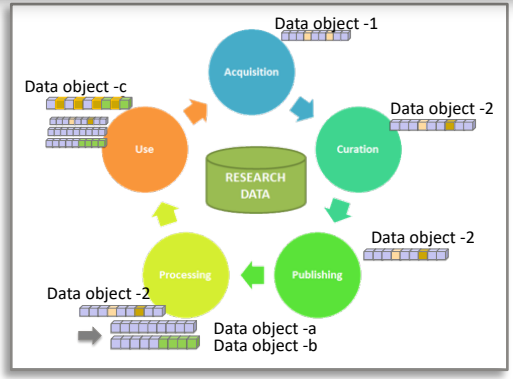
DATA IDENTIFICATION AND CITATION



Data object -1

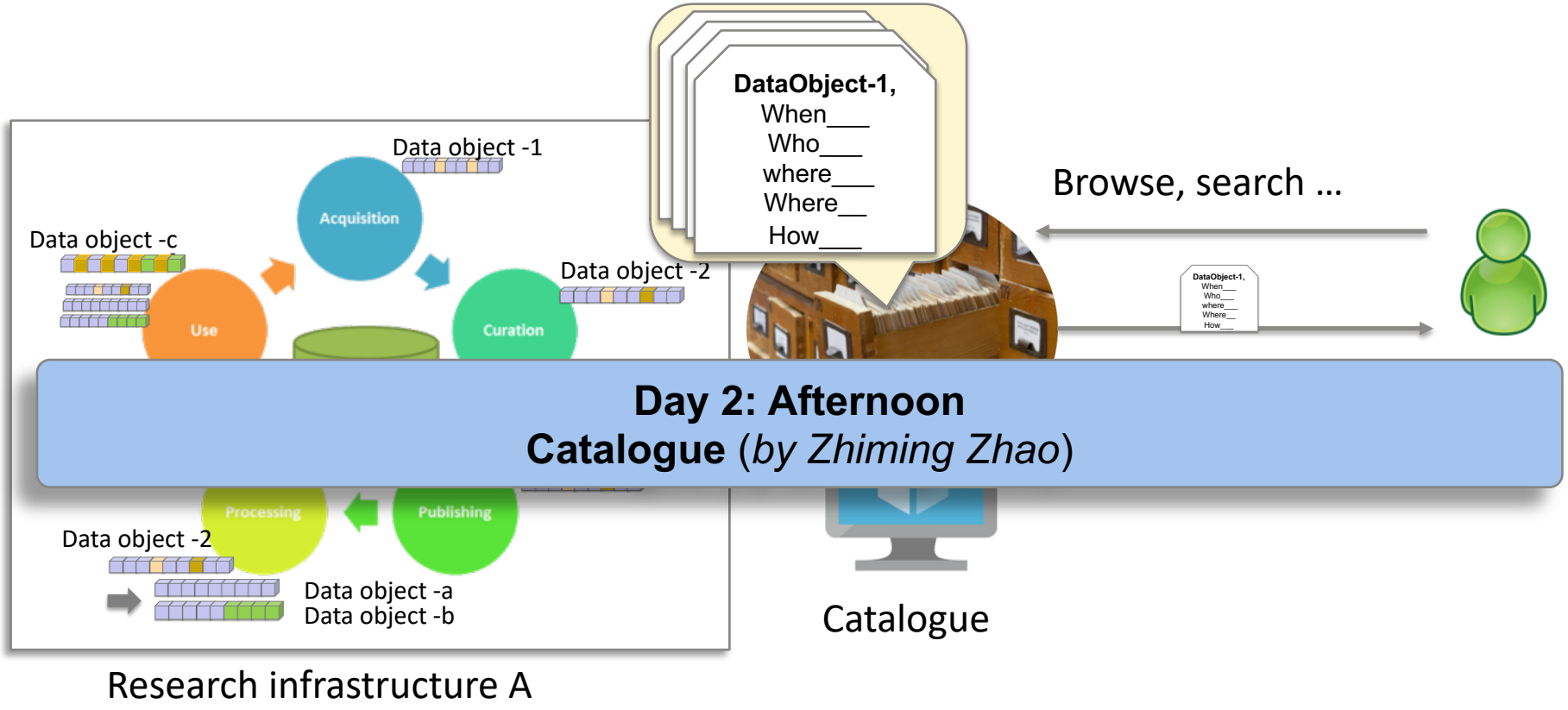
Globally ???

Day 1: Afternoon
Data identification and citation (*by Margareta Hellstrom*)

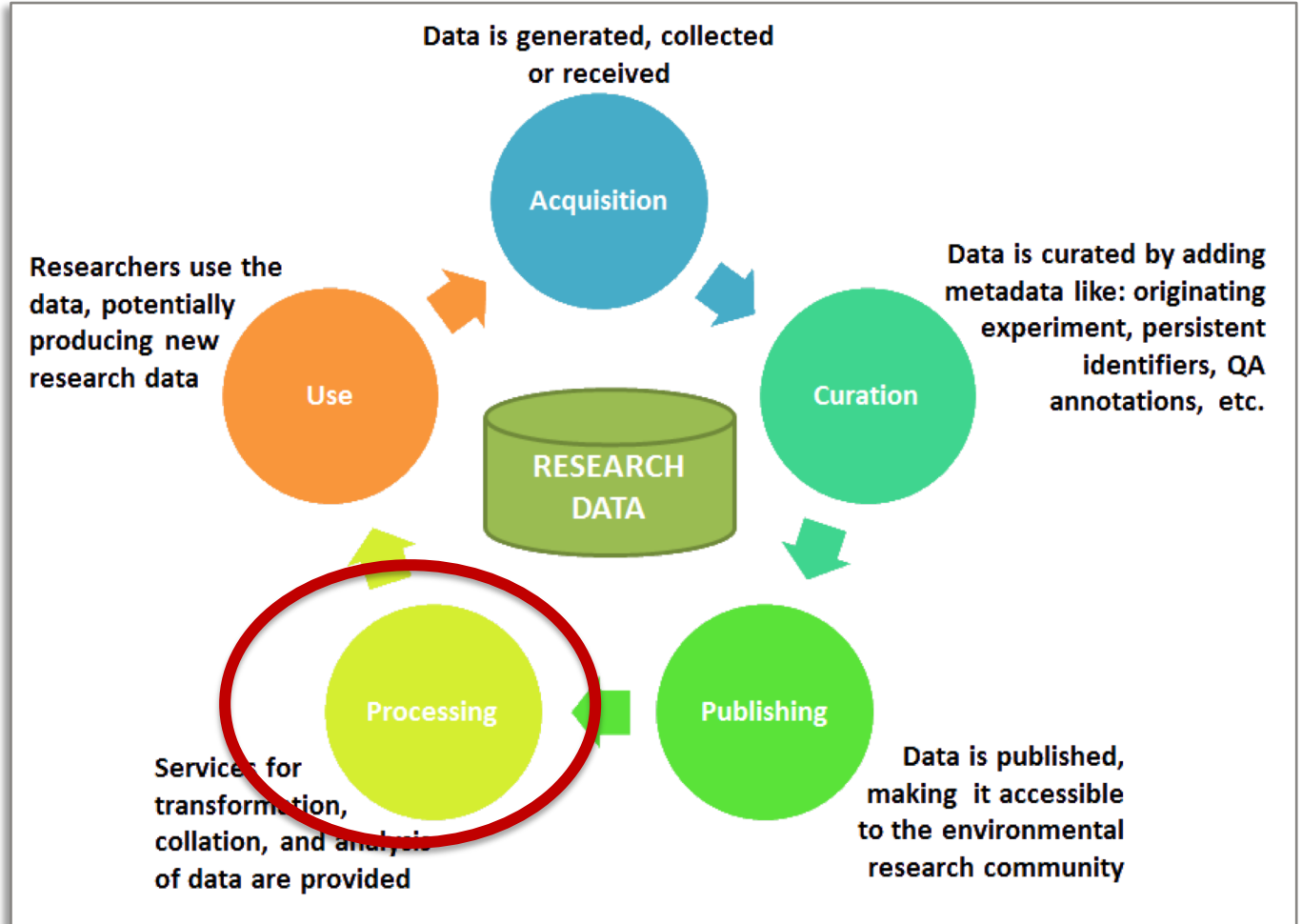


Research infrastructure B

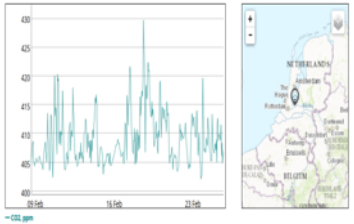
DATA CATALOGUING

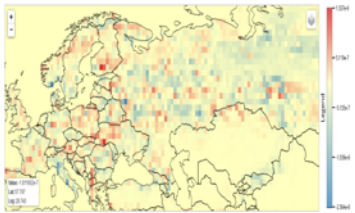


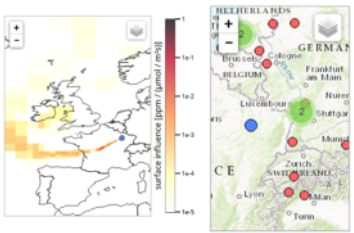
DATA PROCESSING



DATA PRODUCTS

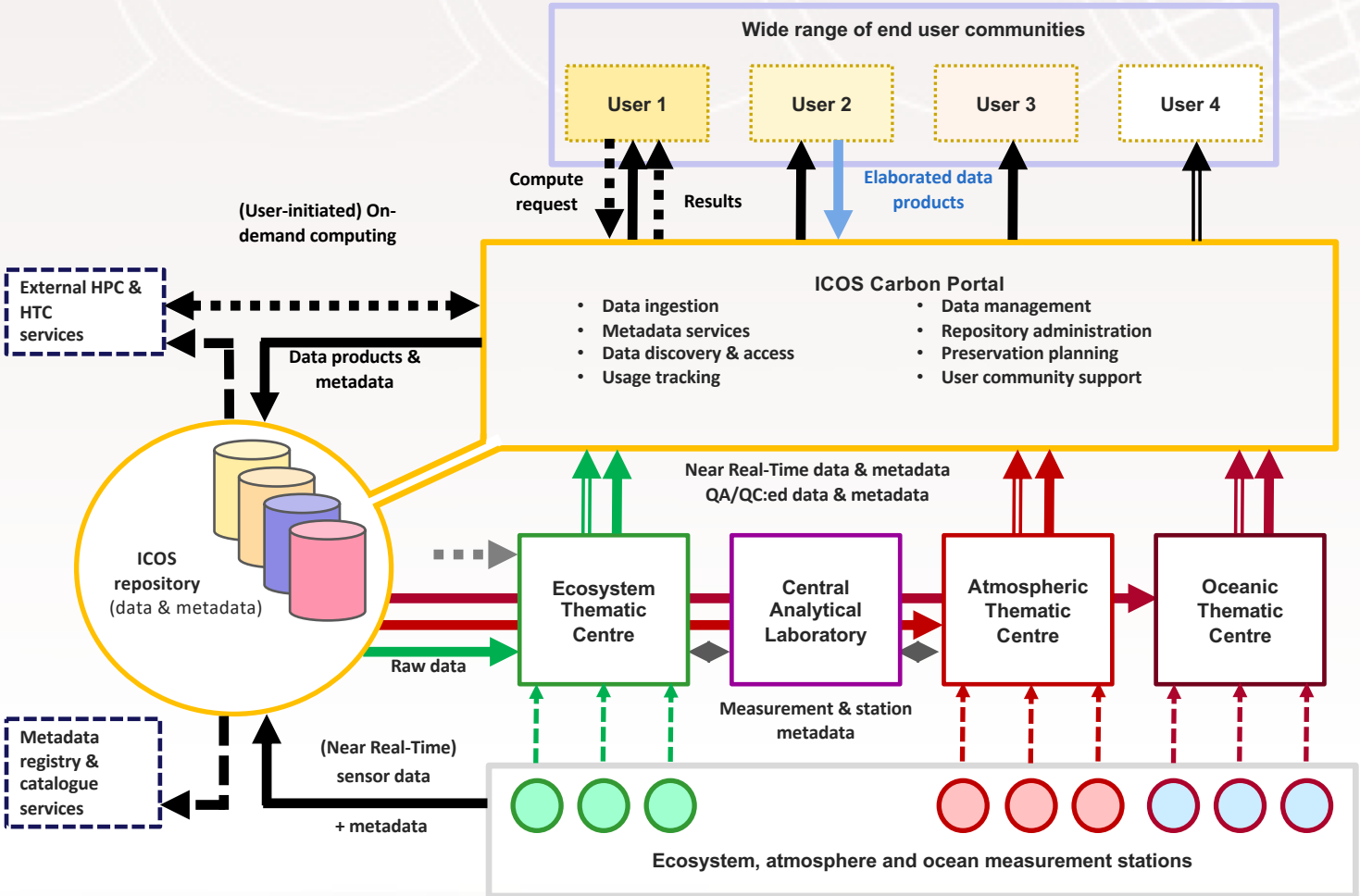




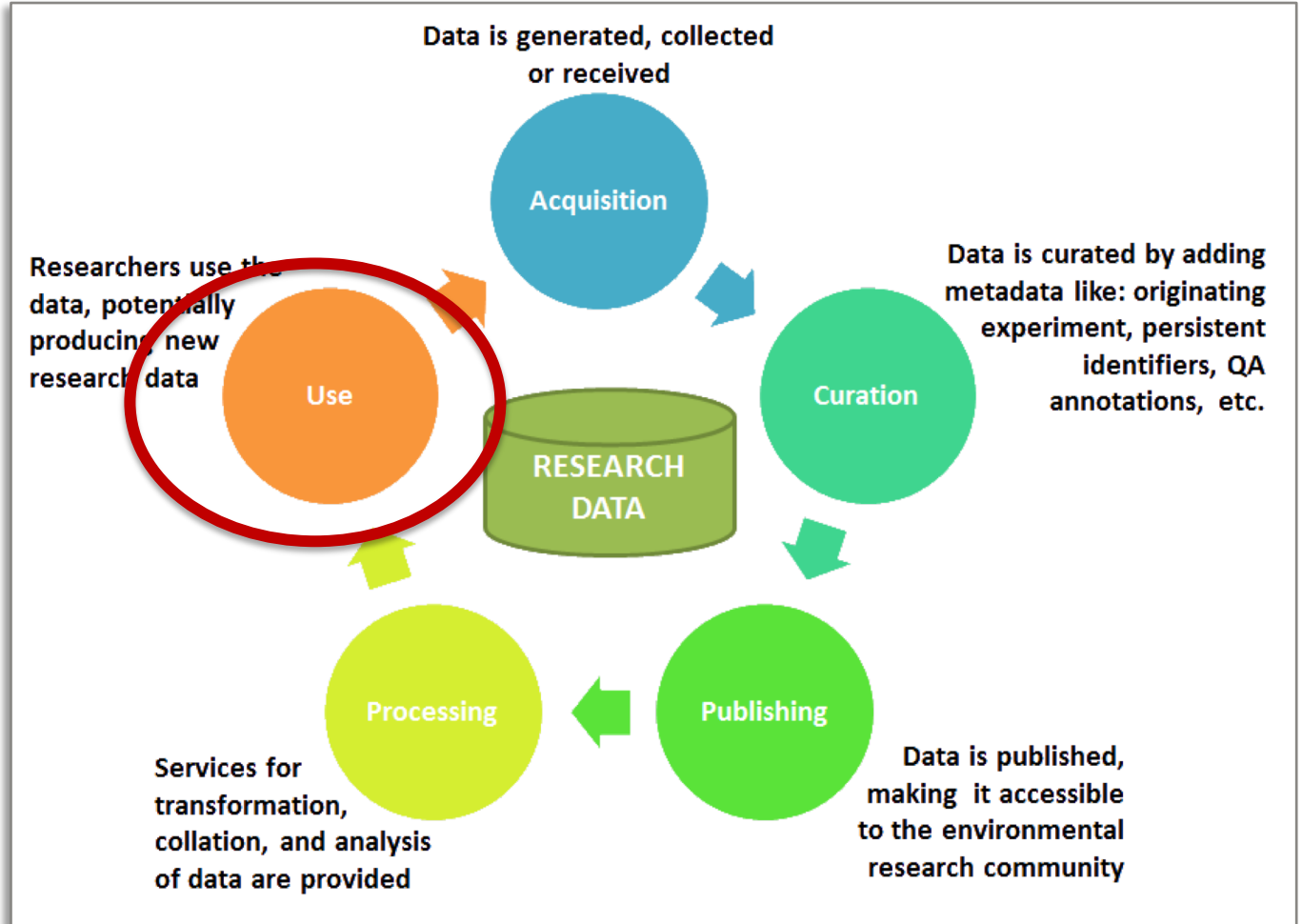


- Raw observational data (**Level 0**)
  - For specialists; available on request
- Near Real-Time data (**Level 1**)
  - For specialists; available on request
  - Mainly automated quality control & processing
  - Time delay & parameter scope varies
- QA/QCed & aggregated data (**Level 2**)
  - Main data product – time series of 30-60 min means
  - Full parameter sets available
- Elaborated data (**Level 3**)
  - Model calculation outputs (atmospheric & ecosystem)
  - Mainly contributed by ICOS data end users

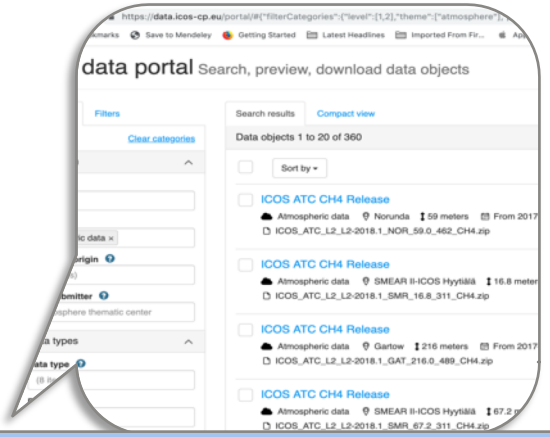
AN EXAMPLE OF ICOS



DATA MANAGEMENT LIFECYCLE



DATA USE



Browse

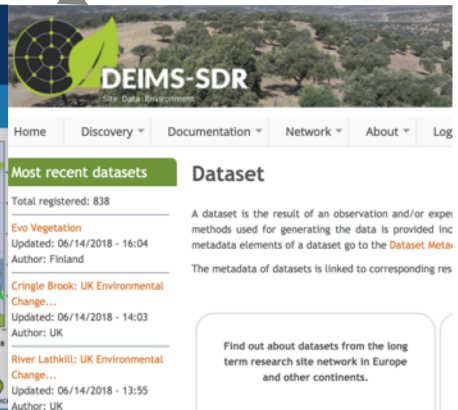
Metadata

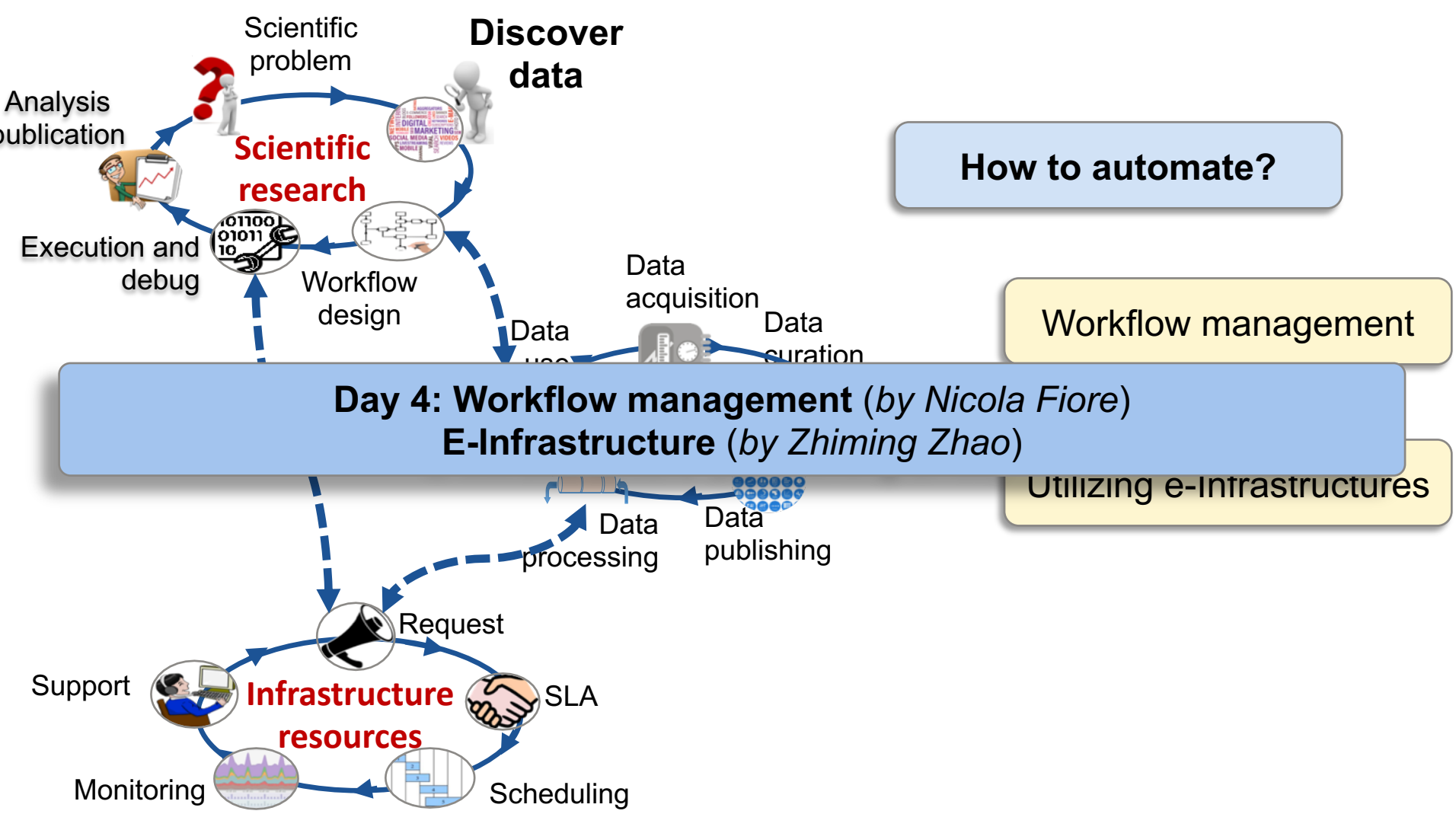
Search…

Ontology

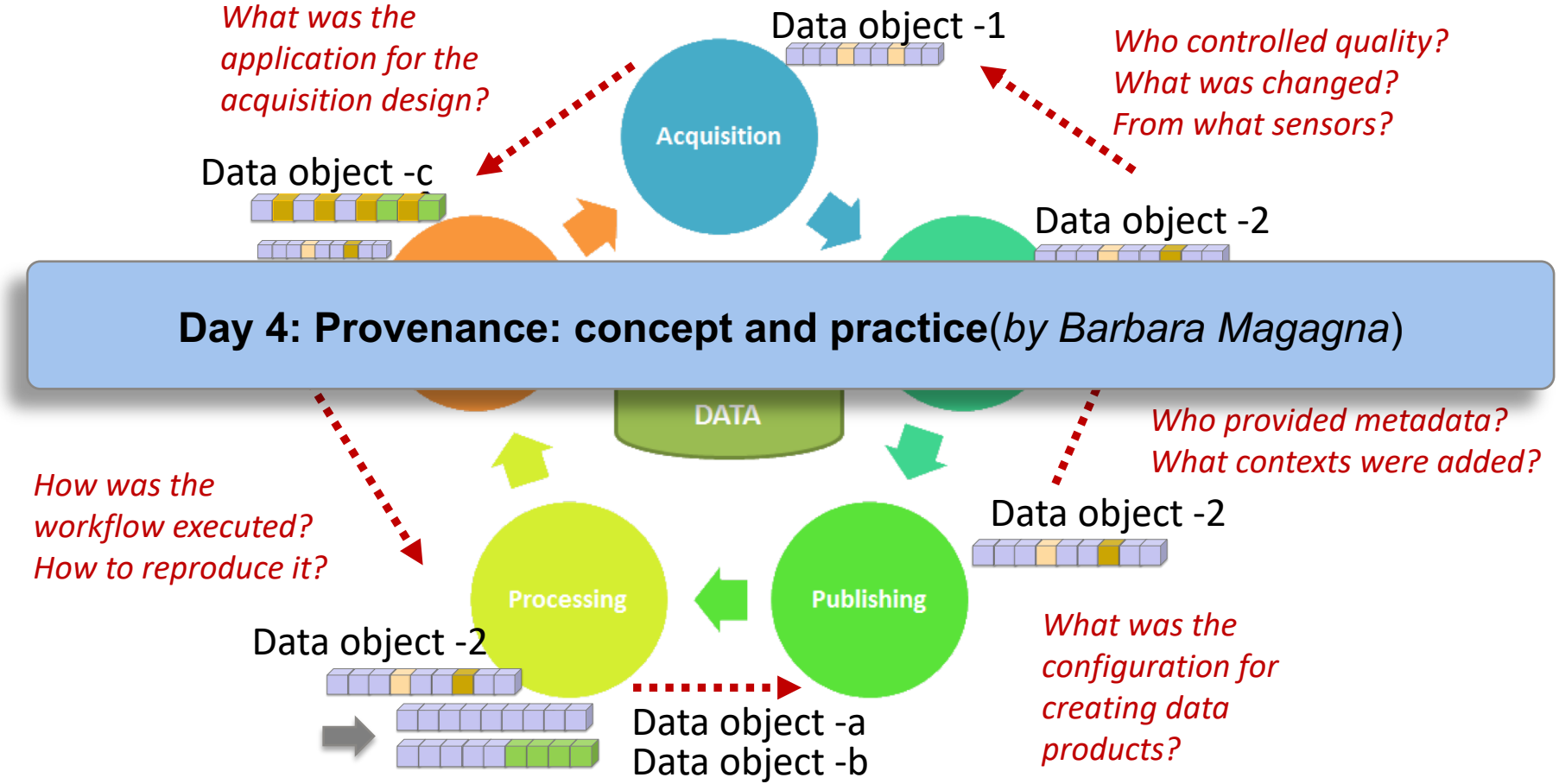
Semantic search…

Day 3: Ontology use (*by Clement Jonquet*)
semantic search (*by Pier Luigi Buttigieg*)

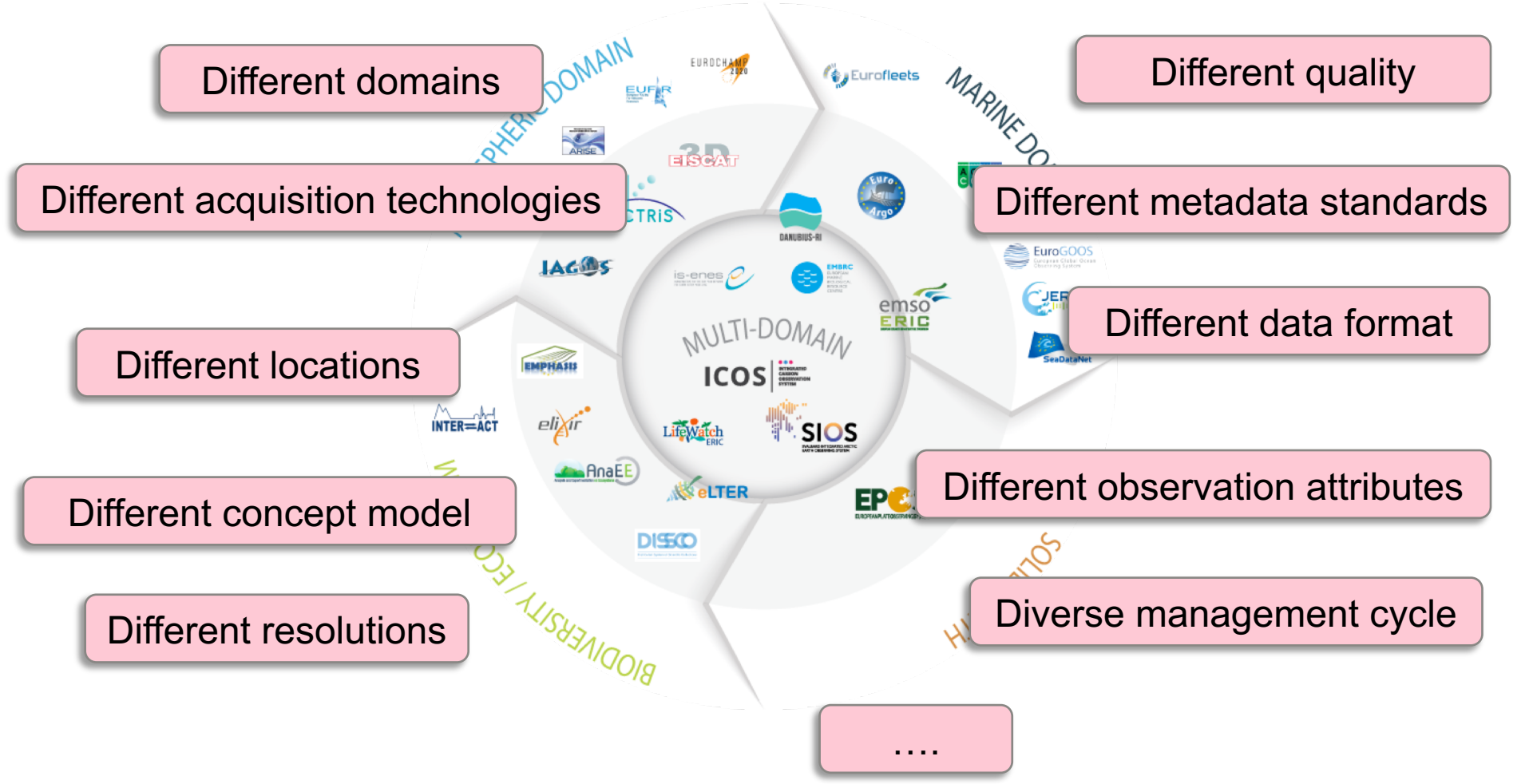



Scientific problem
Discover data
Analysis
ublication
Scientific research
Execution and debug
Workflow design
Data acquisition
Data
Data curation
How to automate?
Workflow management
Day 4: Workflow management (by Nicola Fiore)
E-Infrastructure (by Zhiming Zhao)
Utilizing e-Infrastructures
Data processing
Data publishing
Request
Support
Infrastructure resources
SLA
Monitoring
Scheduling

PROVENANCE

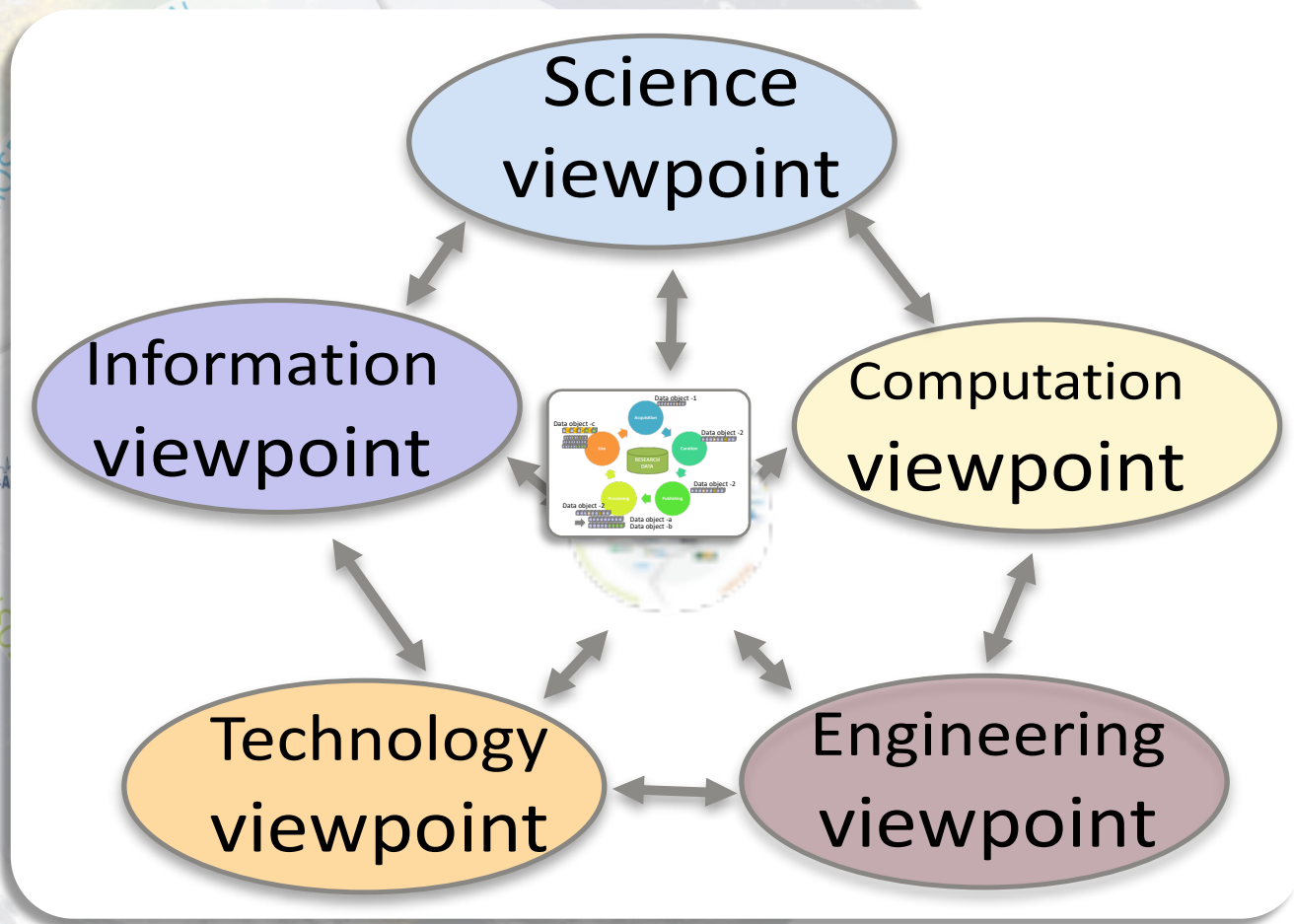

What was the application for the acquisition design?
Data object -1
Who controlled quality? What was changed? From what sensors?
Acquisition
Data object -c
Data object -2
Day 4: Provenance: concept and practice(by Barbara Magagna)
DATA
Who provided metadata? What contexts were added?
How was the workflow executed? How to reproduce it?
Data object -2
Processing
Publishing
Data object -2
What was the configuration for creating data products?
Data object -a
Data object -b

DATA MANAGEMENT IN RESEARCH INFRASTRUCTURE

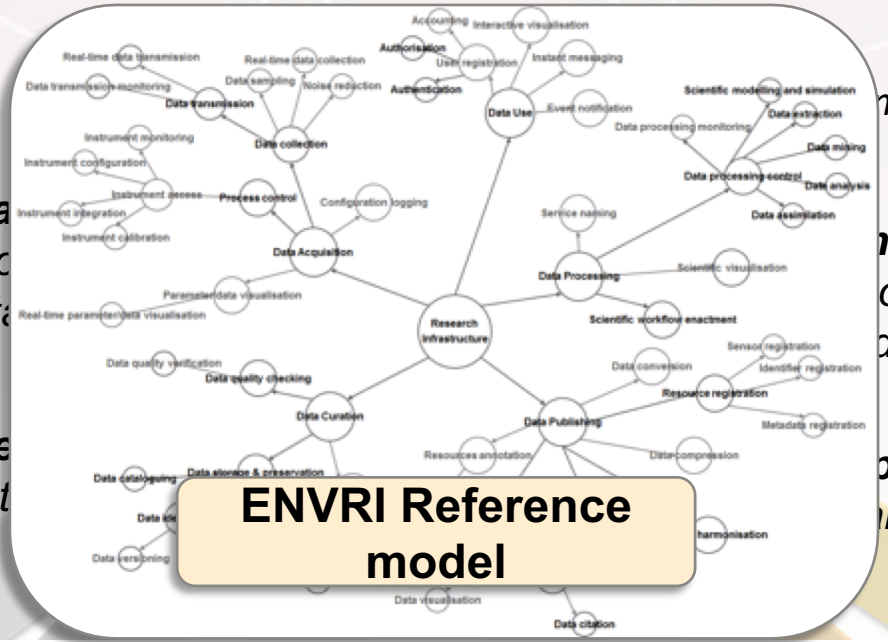


MULTI-VIEW APPROACH

Preparation:
ENVRI –Reference model



ENVRI REFERENCE MODEL (ENVRI.EU/RM)



ENVRI Reference model

nity, Role,

Informationa
data evolutio
data

***nal viewpoint**:*
onents,
d binding.

Enginee
architect

***point**:*
ndards used

SUMMER SCHOOL STRUCTURE

Monday	Tuesday	Wednesday	Thursday	Friday
Introduction - Data management - FAIR	FAIRness principle and assessment	Ontology-1	Workflow E-Infrastructure-1	Wrap up
Student self-introduction	FAIRness Practice	Ontology-2	Workflow E-Infrastructure-2	Assessment
Data curation	Data catalogue	Semantic search-1	Provenance -1	
Data identification and citation	Data management plan	Semantic search-2	Provenance -2	