



**Sharing scientific data and products
in solid Earth sciences:
the EPOS (European Plate Observing System) research
infrastructure experience
1 – Open Science & Access Rules**

Massimo Cocco
EPOS ERIC Executive Director

www.epos-eu.org
massimo.cocco@ingv.it
massimo.cocco@epos-eric.eu

Solid Earth Science

- Different communities involved
- Multidisciplinary data & services
- Community building required
- Services to society operational
- Geo-Hazards
- Geo-Resources
- Anthropogenic hazards



EARTHQUAKES

VOLCANIC ERUPTIONS

TSUNAMIS

TECTONICS

GEODETTIC DATA

LABORATORIES

- Earth scientists have a long lasting tradition in data **acquisition**, collection, **quality-control** and **standardization**
- They have skills and experiences in feeding and implementing **metadata** and services for **qualification**, **storage** and **accessibility**
- **Research infrastructures** represent the facilities and resources to strengthen **data management** and **interoperability** through **e-science innovation**

EARTHQUAKES

VOLCANIC ERUPTIONS

TSUNAMIS

TECTONICS

GEODETIC DATA

LABORATORIES

Open Science

Principles:

- **Open Access to Literature** from Funded Research
- **Data from Funded Research in the Public Domain**
- **Access to Research Tools** from Funded Research
- **Invest in Open Research & Cyber infrastructures**

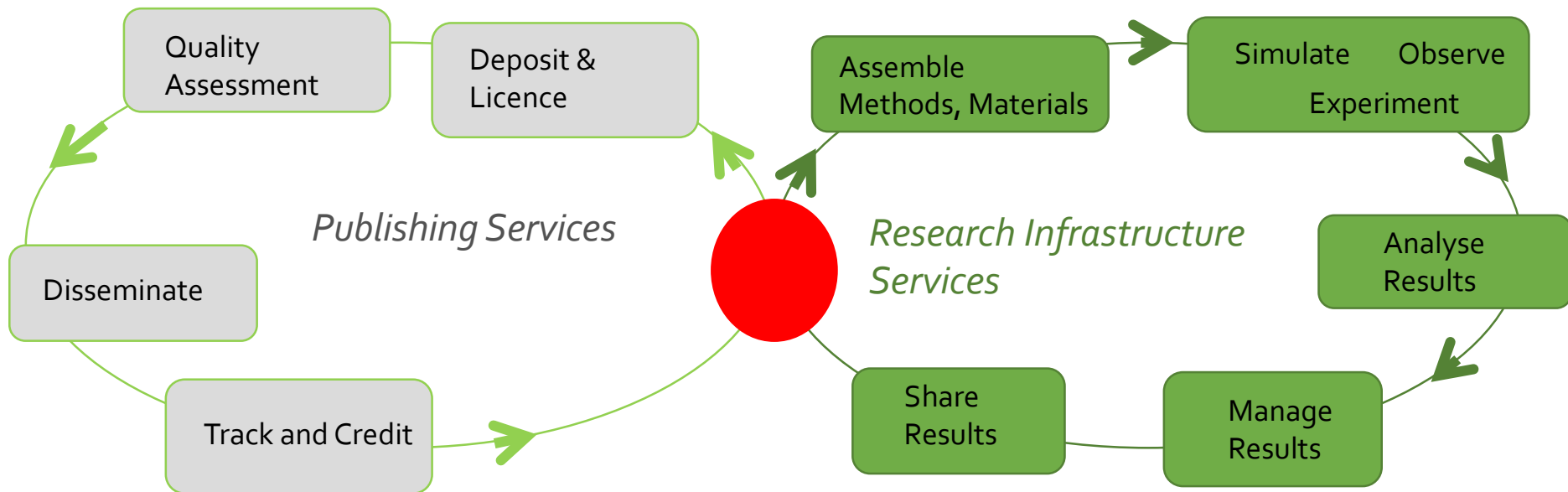
Open Data

Principles:

- Open data is data that **anyone can access, use and share**
- Open data becomes usable when **made available in a common, machine-readable format**
- Open data must be **licensed**. Its license must permit people to use the data in any way they want (...)



Why does this matter?



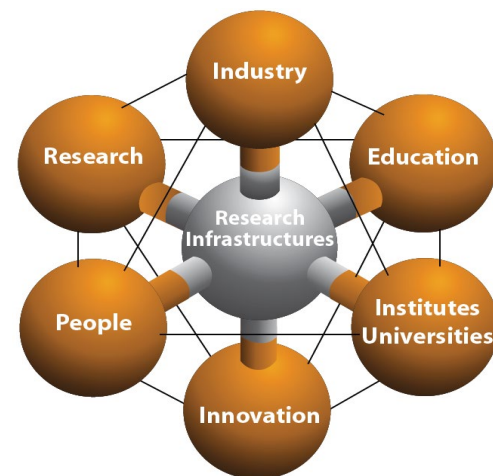
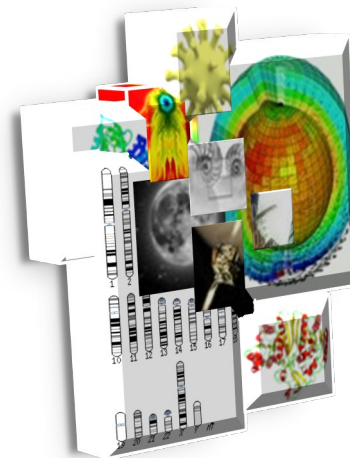
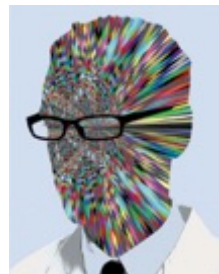
Moving between different environments

Recreating / accessing common environments

Fragmented, multivarious and complicated ...

Access to data: a global challenge

- Grand Challenges facing the world demand open access to data making them FAIR:
 - Findable
 - Accessible (Understandable)
 - Interoperable
 - Re-usable (Reproducible)
- More value from reuse of data
- More and more researchers are seeing the value of sharing data
- Many countries developing open research **data policies**
- Central Role of **Research Infrastructures**



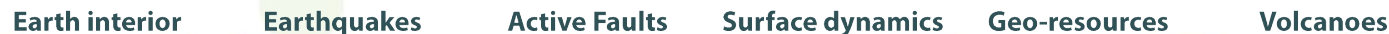
- provide **virtual access** to data and **physical access** to facilities
- allow **communication** with different stakeholders
- tackle **sustainability** challenge from a technical, legal, governance and financial point of view
- strengthen **data management** and **interoperability** through e-science innovation



EPOS is a **long-term plan for the integration** of research infrastructures for solid Earth Science in Europe

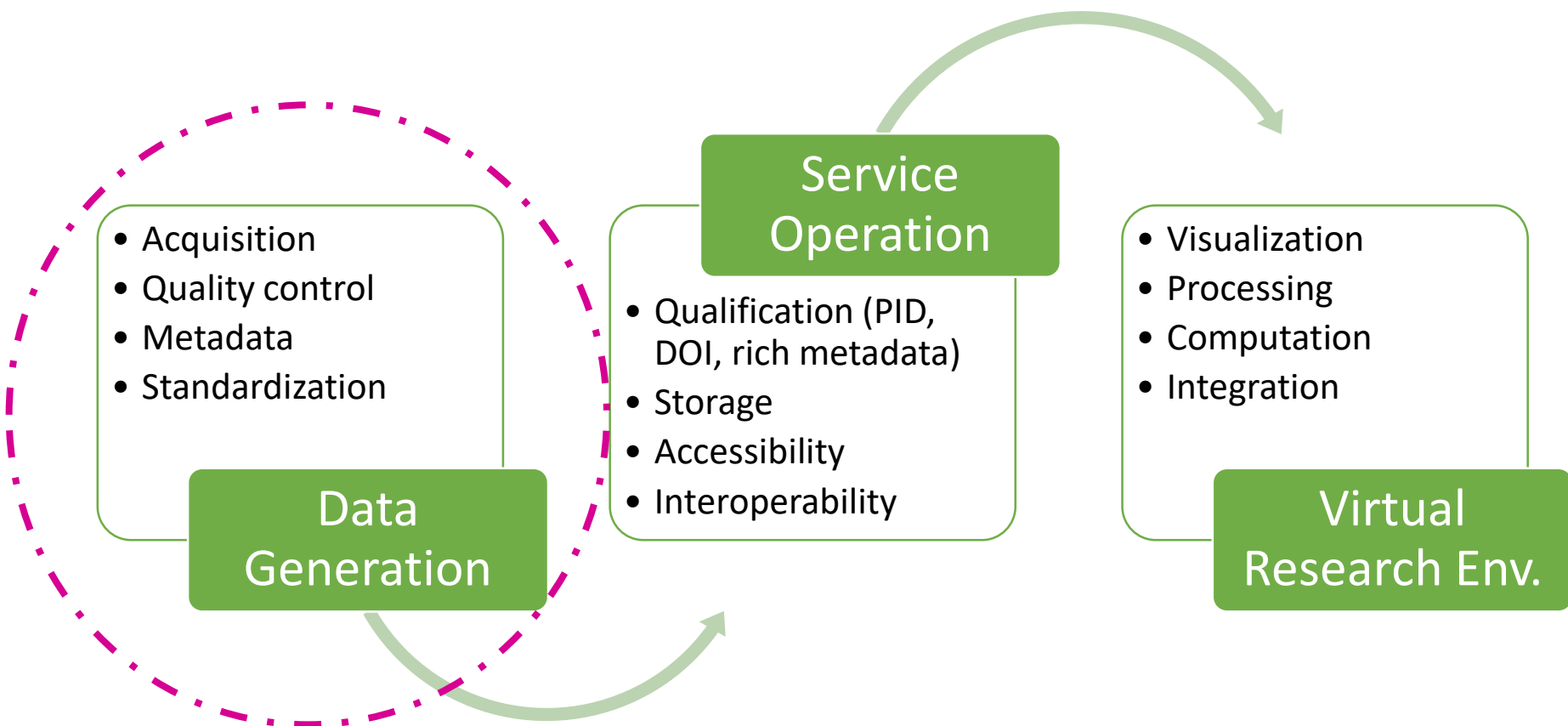
EPOS integrates the **existing (and future)** advanced European facilities and infrastructures into **a single, distributed, sustainable research infrastructure** taking full advantage of new **e-science opportunities**





PROCESSES

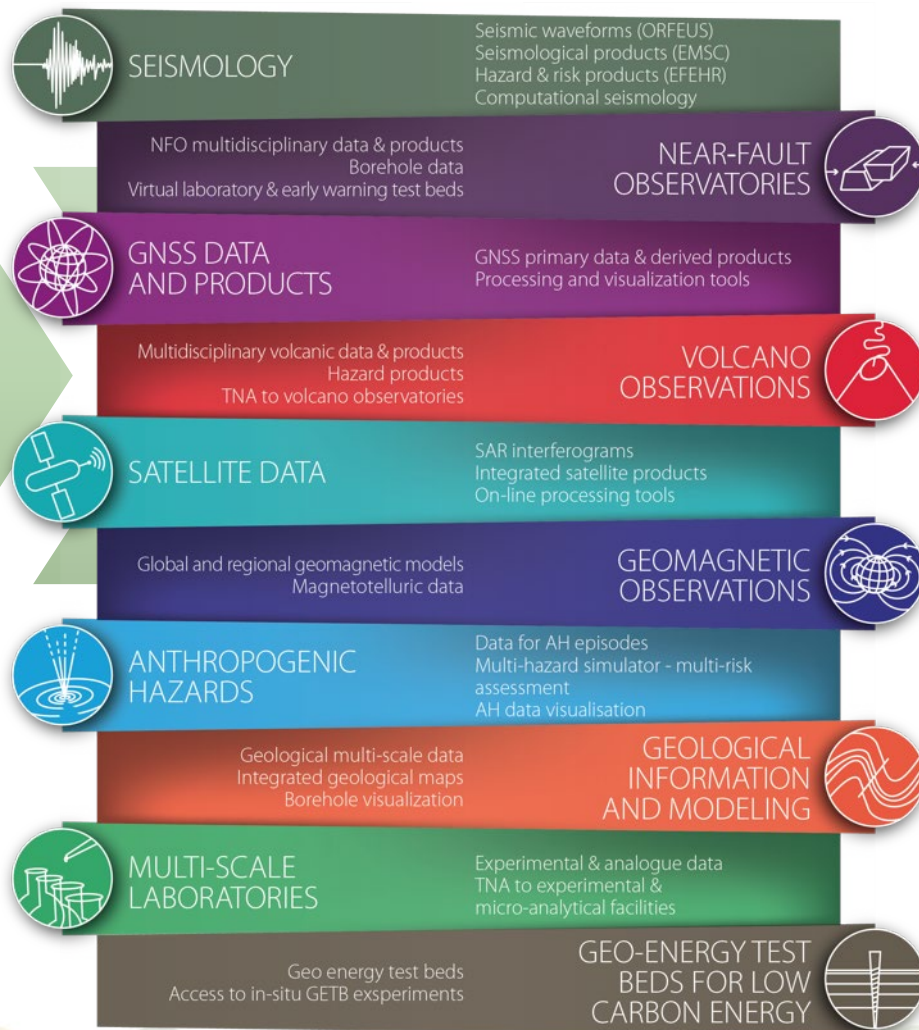
The EPOS experience & practice



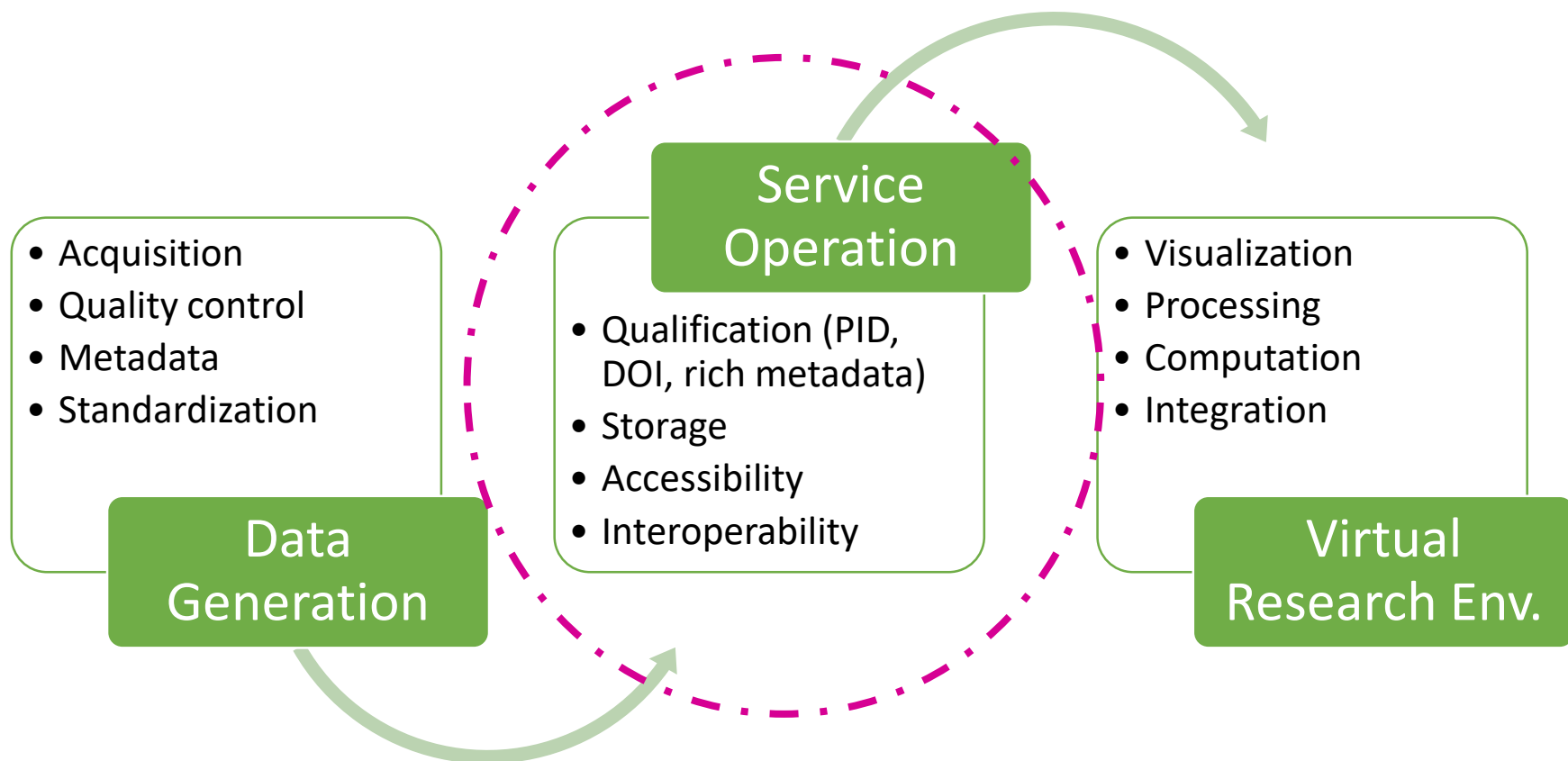
For Distributed pan-European RI

National , Regional
RIs

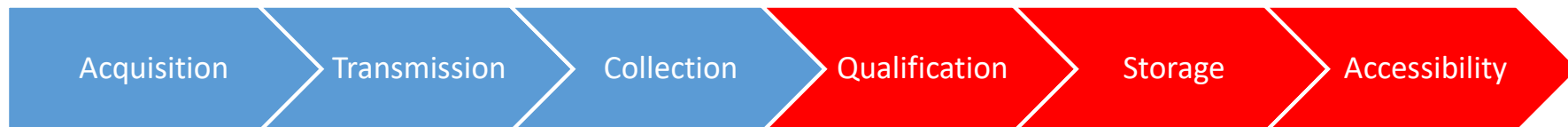
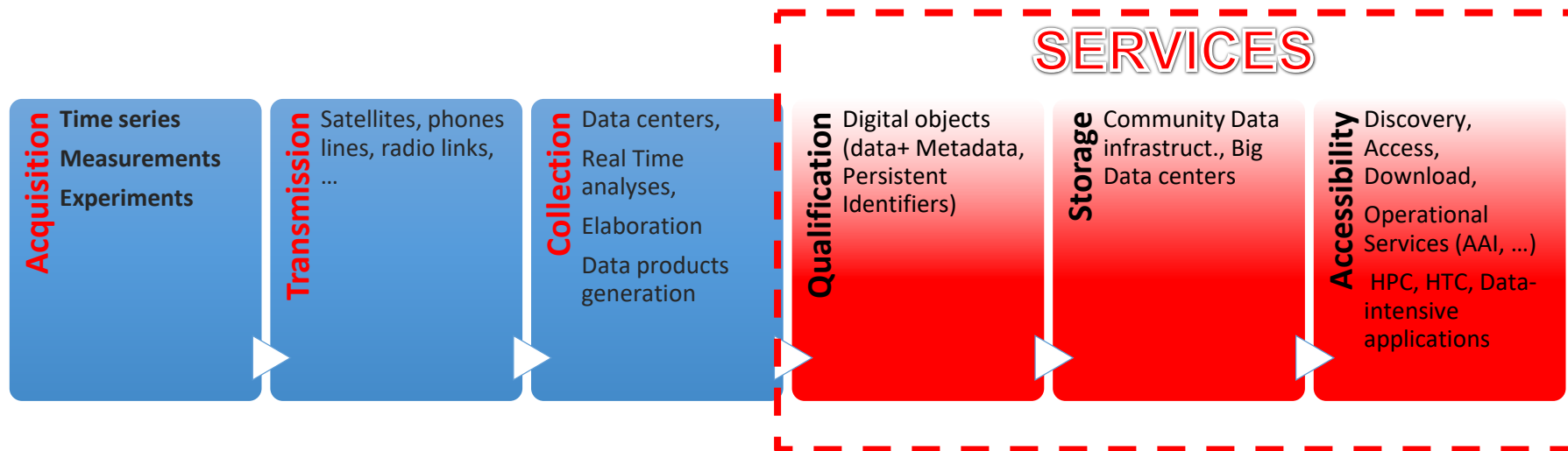
Engagement
of Data
Providers



The EPOS experience & practice



Data Timeline and Services



Data acquisition, validation & standardization

Data collection & preservation (PID, DOI)

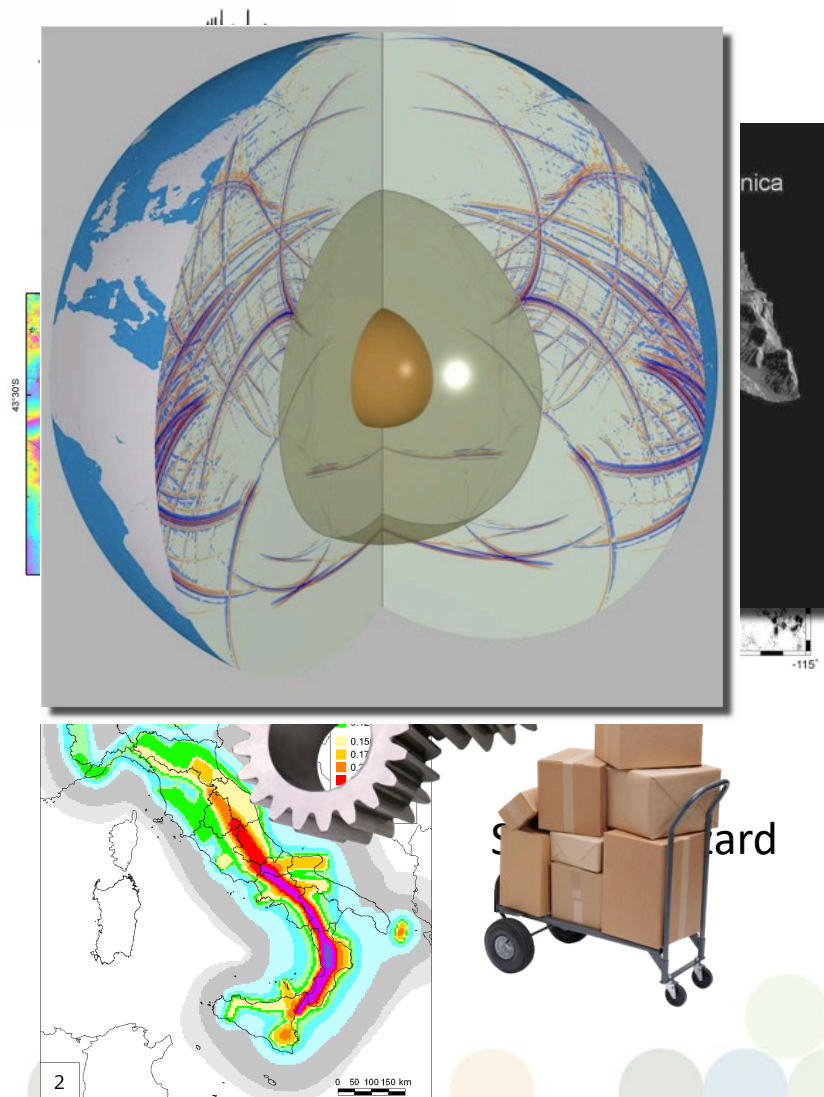
Accessibility, integration, computation

Data & service provision

- EPOS RIs produce a wealth of **standardized, quality-controlled** data for science and society.
 - Standardization and quality control implies **engagement of the scientific communities** to harmonize **data** and **metadata**.
 - **Data** heterogeneity: raw data, data **products**, software and **services**.
 - Monitoring networks and observing systems also serve for environmental **surveillance, alert**, and **geo-hazard** assessment
 - Data collected for **geo-resources** exploitation and anthropogenic hazard
- **Awareness of Data Value = Utility + Warranty**
 - Utility: data is useful, harmonized, standardized
 - Warranty: data is accessible and traceable

Access to Data Products (Taxonomy)

- **Level 0:** raw data, or basic data
- **Level 1:** data products coming from nearly automated procedures
- **Level 2:** data products resulting by scientists' investigations
- **Level 3:** integrated data products coming from complex analyses or community shared products
- **Level 4.** Software, IT tools



EPOS Data, Access, and IPR policy

Guiding principles:

- open access
- licensing
- no charges

Protect EPOS legally

Unrestricted use & access

Trace EPOS use & users

Balance: **Legal risk** : **Openness** : **Traceability**

Licensing
IPR
Terms & Conditions
Restrictions

Data & Service
Providers

Open Access
deposit terms

EPOS

Open Access
license

Data & Service
Users

Categorization

as needed for
legal aspects

Data & Data Products
Level 0,1, 2, 3

Tools & Software

Open

:

Restricted

:

Embargoed

mix and match as required

Anonymous

:

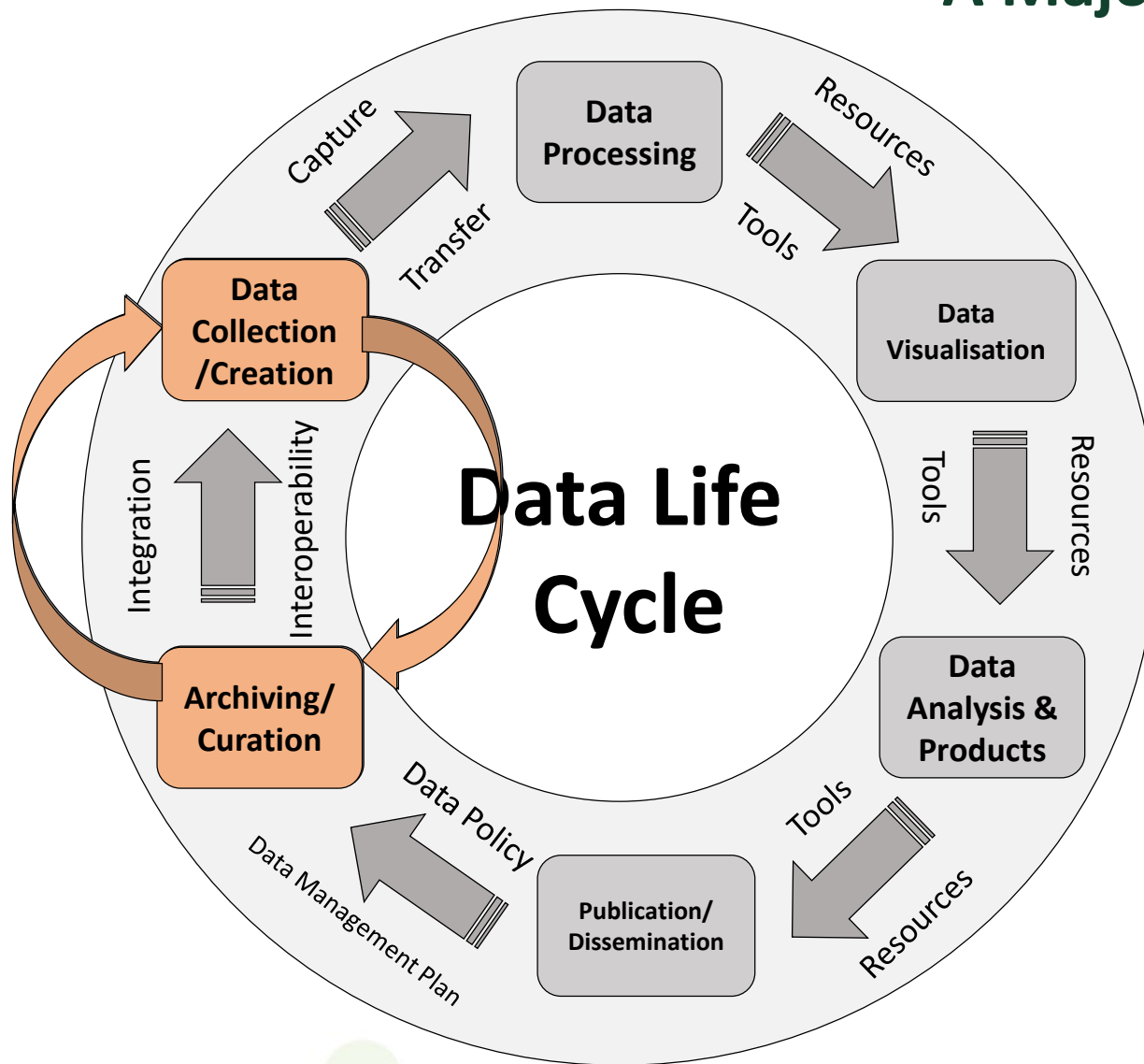
Registered

:

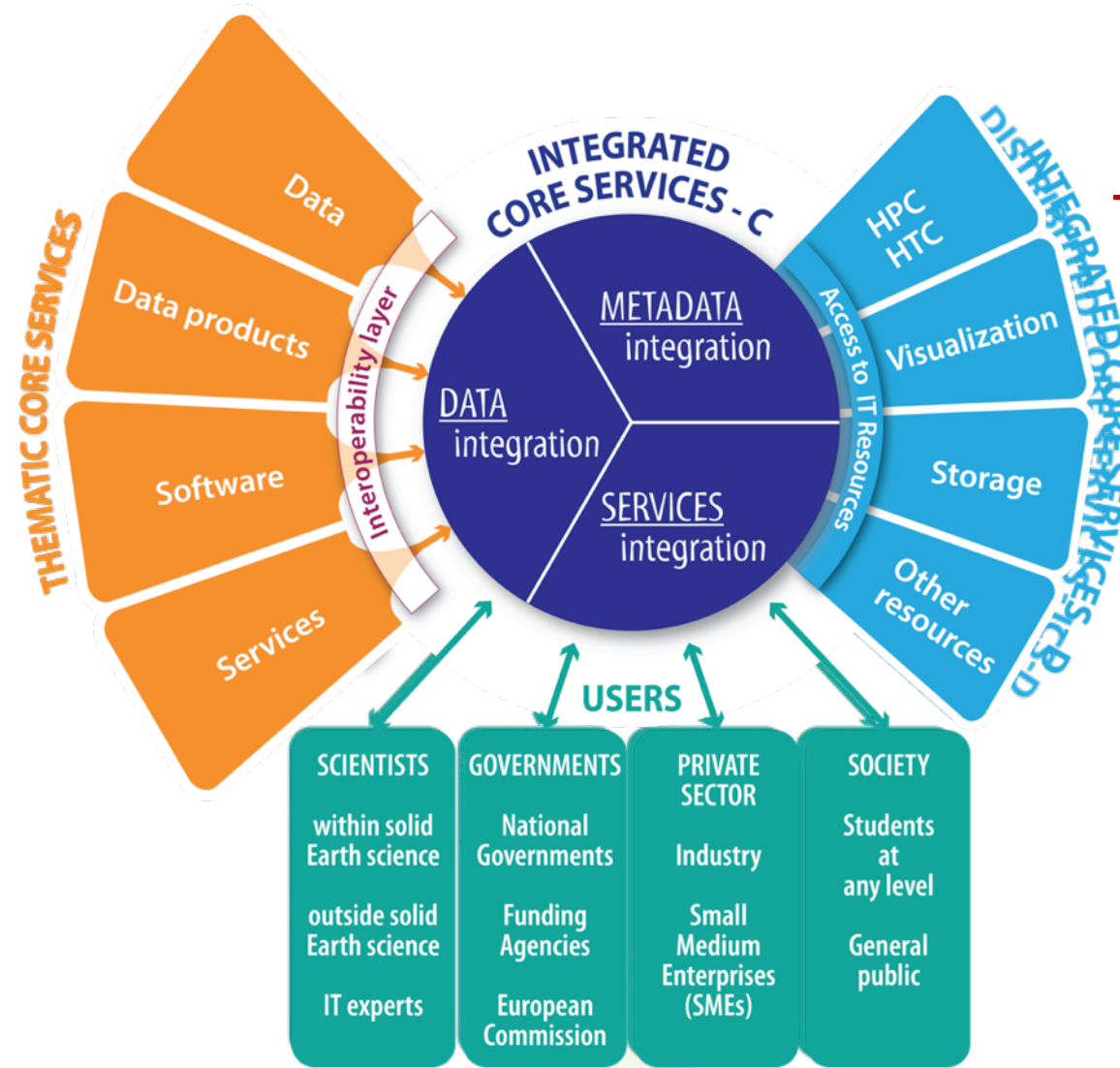
Authorized

Users





EPOS Architecture



The EPOS data & service provision to diverse category of stakeholders

SCIENTISTS

- **Data providers**
- Data users within the SE community
- Data users outside the SE community
- **IT experts**

GOVERNMENTS

- **National governments**
- **Funding agencies**
- **European Commission**

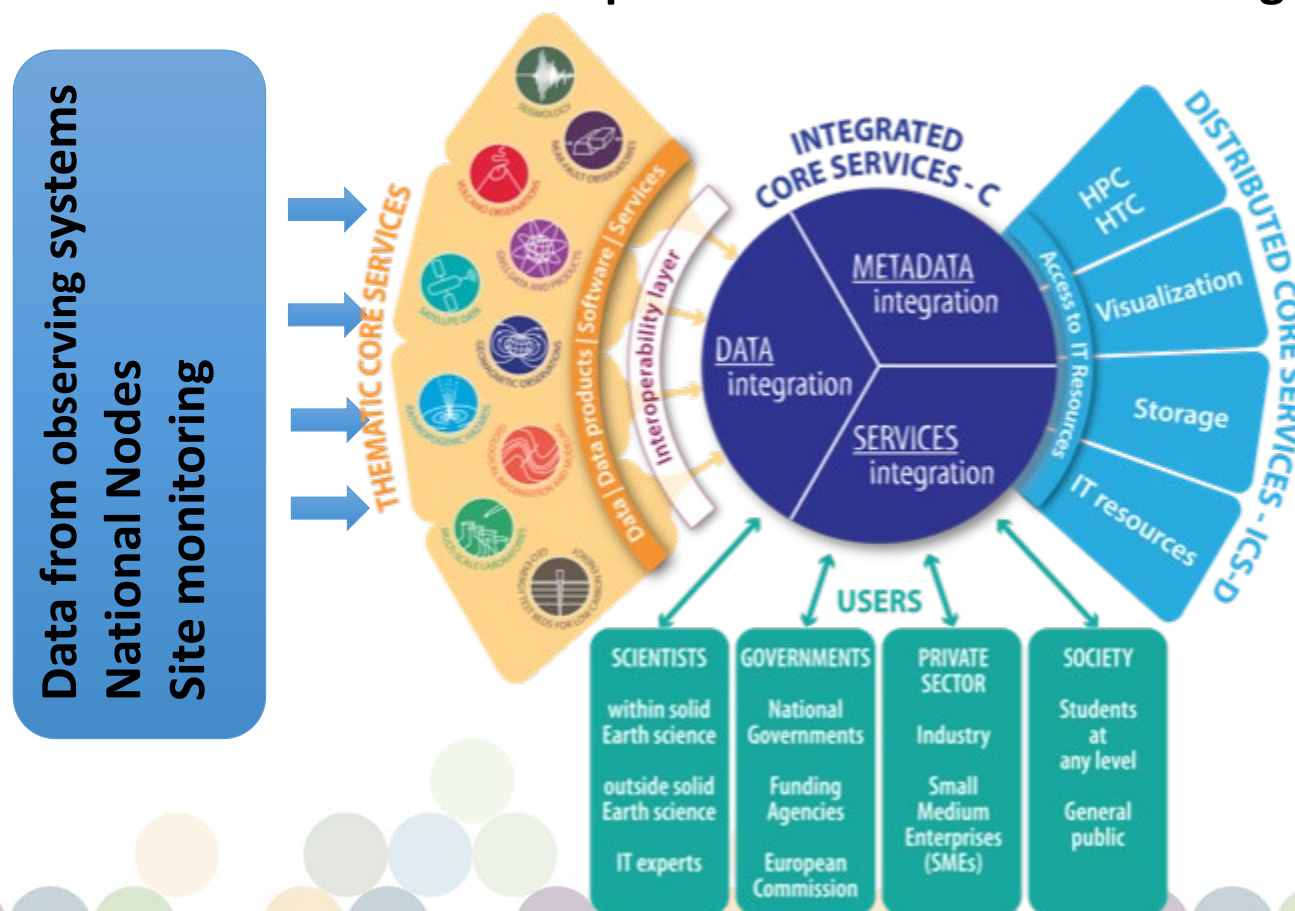
PRIVATE SECTOR

- **Industry**
- **Small Medium Enterprises**

SOCIETY

- **Students at any level**
- **General public**

According to its vision, the **EPOS mission** is to create a **single, distributed and sustainable, infrastructure** that **integrates** the diverse and advanced European National RIs for **solid Earth science** under a common framework to provide open access to data and products for diverse user categories



Real Time Services: an example from seismology

EPOS TCS:
community driven and governed services

Waveform Services



Waveform selection & access
Waveform metrics & Station Information
Strong Motion parameters
OBS data integration
Mobile Pool coordination & integration
Waveform modeling

Seismological Products



Earthquake Parameter Information
Macroseismic & Historical Event data
Seismological Products Platform
- rupture models / SiteCharTool / MT
- EventID / F-E-Region / ...

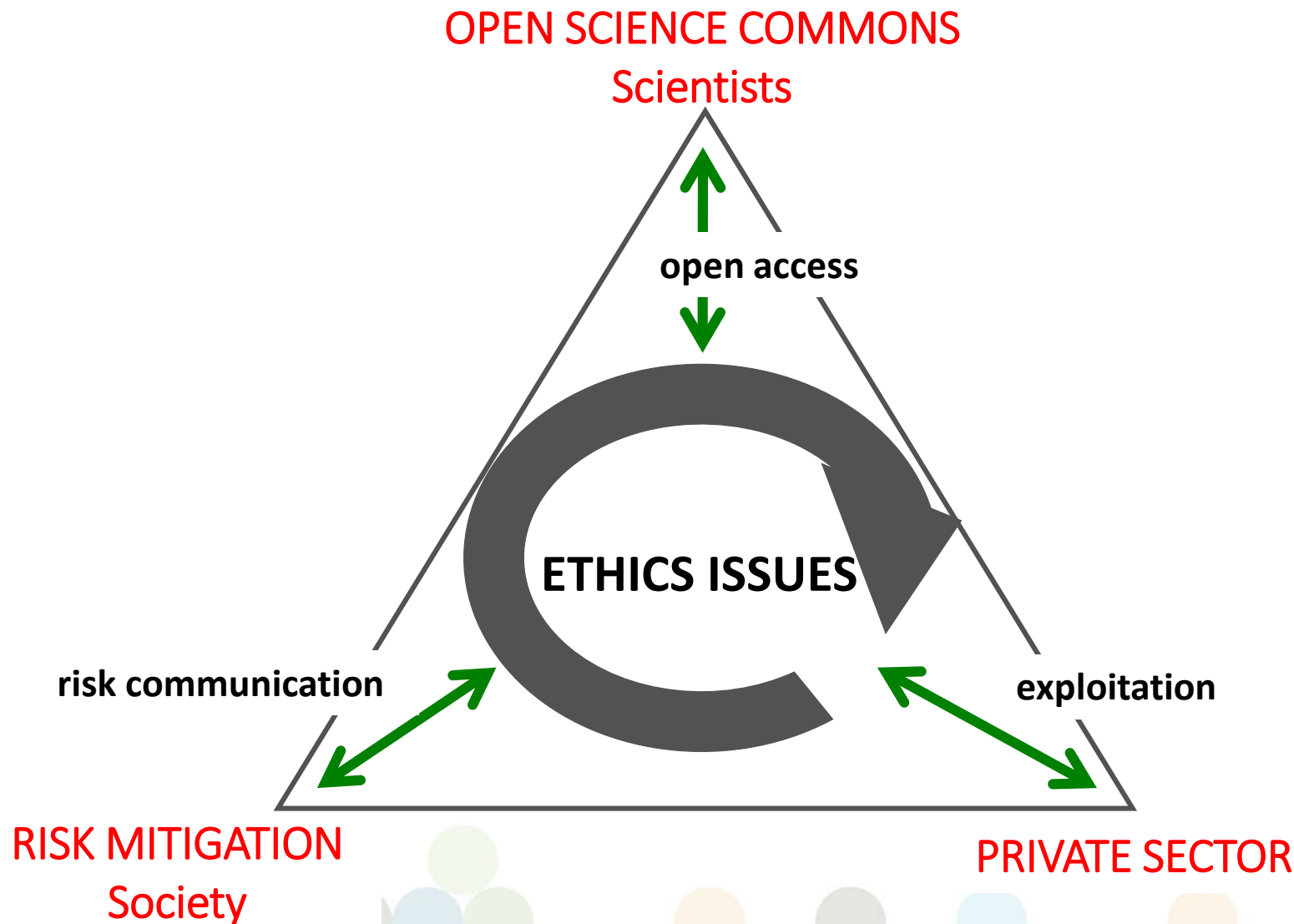
Hazard and Risk Services



Seismic Hazard Models
Seismogenic Faults
Ground Shaking Models
Geotechnical Engineering Information
Strong Motion records in buildings
Earthquake Engineering & Risk Services

From data to products

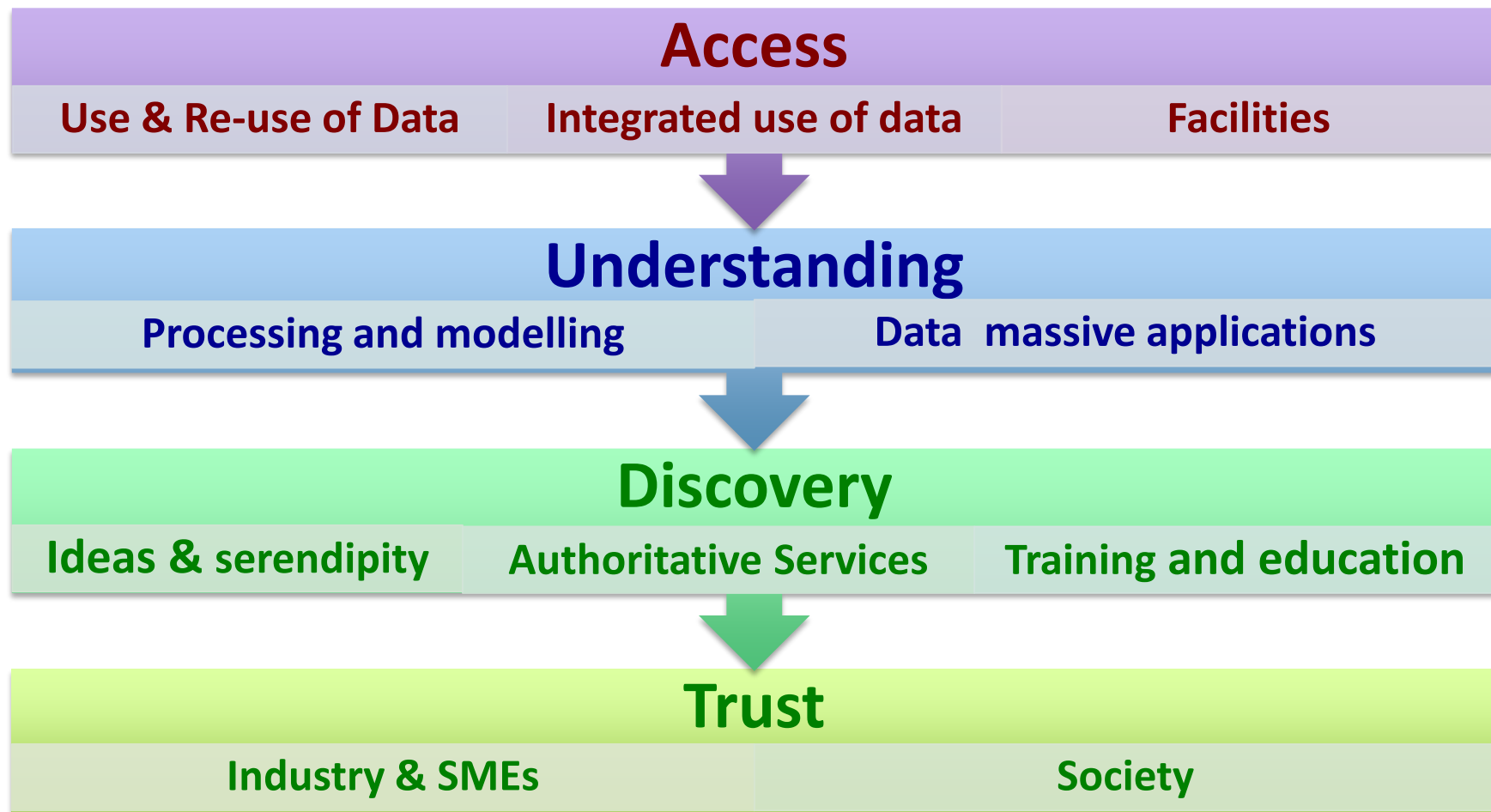




Summary – Points for Discussion

- **Solid Earth Science** is composed of different communities having a different tradition and awareness on **data management** and **open science** with peaks and gaps
- In solid Earth sciences (EPOS), **data sharing has enormous potential**. A **new typology of scientists** targeting **multidisciplinary problems** asks to have open access to data and scientific products
- The long-lasting RIs tradition in **data & metadata standardization** and **quality control** should be preserved as a necessary condition to make data FAIR+R.
- Challenges to tackle FAIRness in terms of available resources and solutions:
 - Traceability and **persistent identifiers (accountability of data providers)**
 - **Access procedures**, ideally by **automated means**, to data & services
 - Data **provenance** and traceability and link to publications
- To achieve the best results, **continuous orchestration** between scientific communities and ITs (e.g., AAI, PID) is needed
- **Interactions with industry** in Earth sciences require effective strategies and particular attention (**ethic issues**, use and re-use of scientific data)

The EPOS chain: high gain/high-*but manageable* risk



The Data Generation Univ

Observing Systems

Monitoring Networks

Experimental Facilities

Observatories

Researchers

Scientific Knowledge

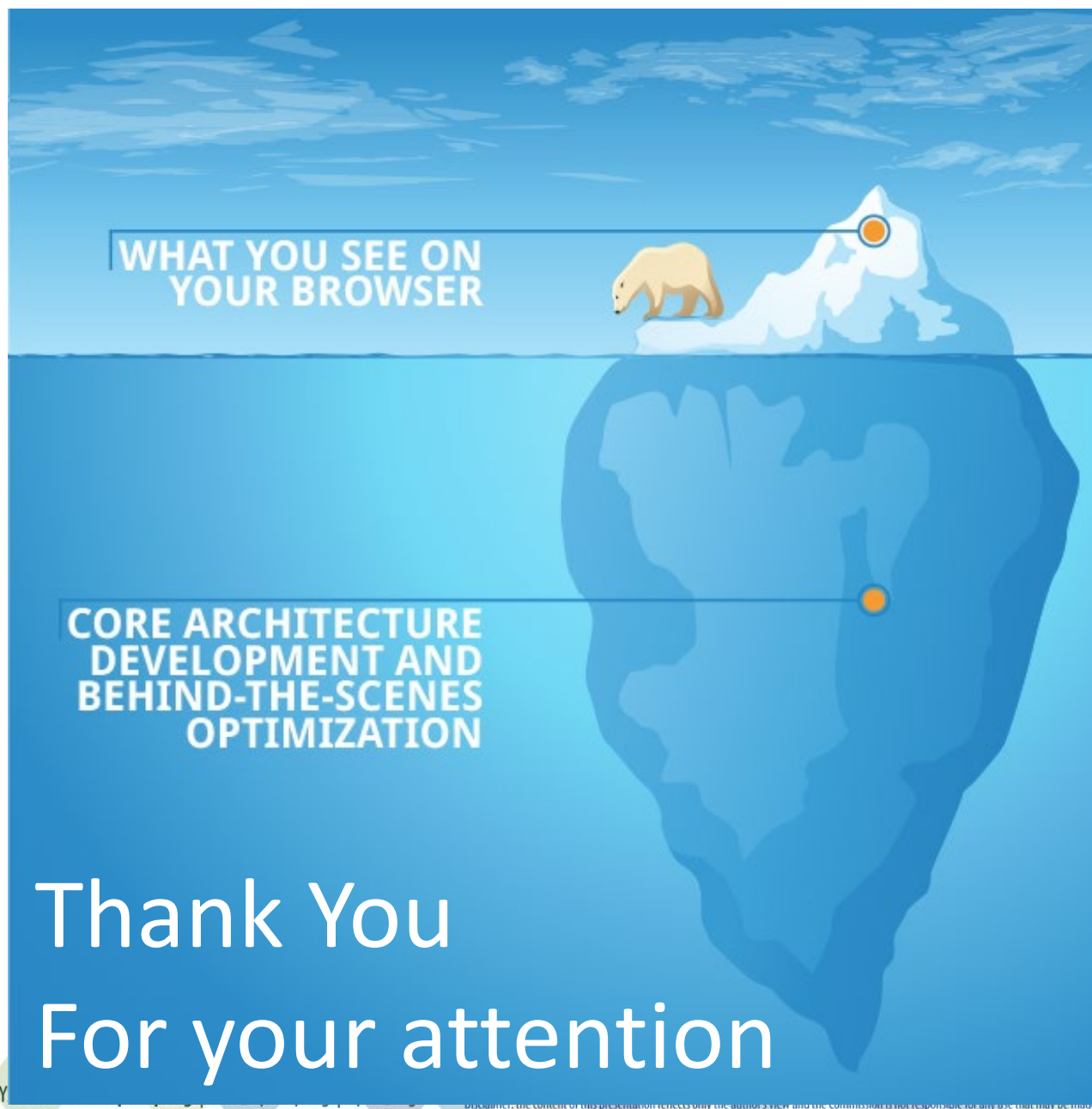
The Data Generation Universe



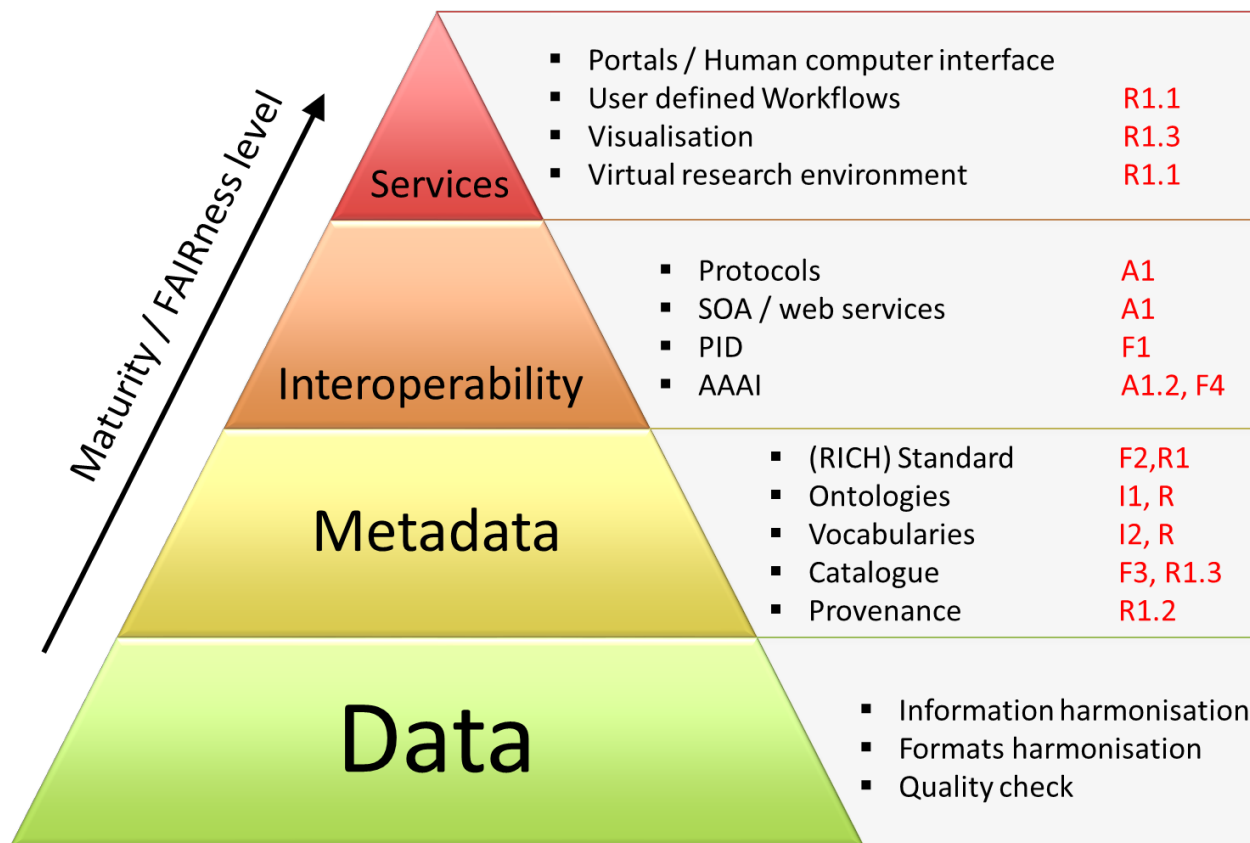
Research Infrastructures (RIs) sample a limited portion of the Data Generation Universe, while ensuring:



- Quality Controlled data
- Standardized Data & Metadata
- Metadata Curation & Integration
- Data Curation and Integration
- Services curation and integration
- Access to data and products
- Visualization of integrated data
- Access to multidisciplinary data
- Generation of new scientific products
- Data qualification
- Service qualification



Maturity / FAIRness “roadmap”



Findable 🔍
Accessible ➡
Interoperable ⚙️
Reusable ♻️
 +
Reproducible

Daniele Bailo EPOS