

DIGITAL MEETS HERITAGE

Implementation of Authenticity Evidence Record model for Supporting Preservation Scenarios



SCIDIP-ES
SCIENCE DATA INFRASTRUCTURE FOR PRESERVATION - EARTH SCIENCE

How SCIDIP-ES Infrastructure supports archive managers in LTDP

LUIGI.BRIGUGLIO@ENG.IT



Presentation Plan



- Premise
- SCIDIP-ES Infrastructure
- Dealing with Authenticity
- SCIDIP-ES Solution
- Some Scenarios
- Conclusion



SCIDIP-ES Infrastructure



SCIDIP-ES

EARTH SCIENCE

PRESERVATION

- **SCIDIP-ES** (SCience Data Infrastructure for Preservation – Earth Science) is a EU project (FP7-INFRASTRUCTURES-2011-2) Combination of CP & CSA started in September 2011.
- **SCIDIP-ES** is delivering long-term **preservation** services as part of the data **Infrastructure** for **e-Science** and specifically for **Earth Science (ES)**.
- **Infrastructure** is composed by Generic **Services** and by an evolving set of largely “discipline-focussed” **Toolkits** addressing: persistent storage, access and management



SCIDIP-ES Infrastructure

- **SERVICES**

RepInfo Registry	Persistent Identifiers
Gap Identification	Orchestration
Storage	

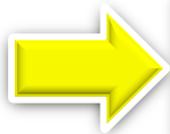
- **TOOLKITS**

RepInfo	Preservation Strategy
Authenticity	Packaging
Finding Aid	Certification
Process Virtualisation	Data Virtualisation



Dealing with Authenticity

- **OAIS RM** – Magenta Book July 2012:
 - *“The degree to which a person (or system) may regard an object as what it is purported to be. The degree of authenticity is judged on the basis of evidence”*
- **InterPARES** clarified how the evidence has to be collected during the whole lifecycle of DR (*both before and after preservation begins*)



*Key issue is collecting the appropriate **evidence** for all the **events** that may affect authenticity*

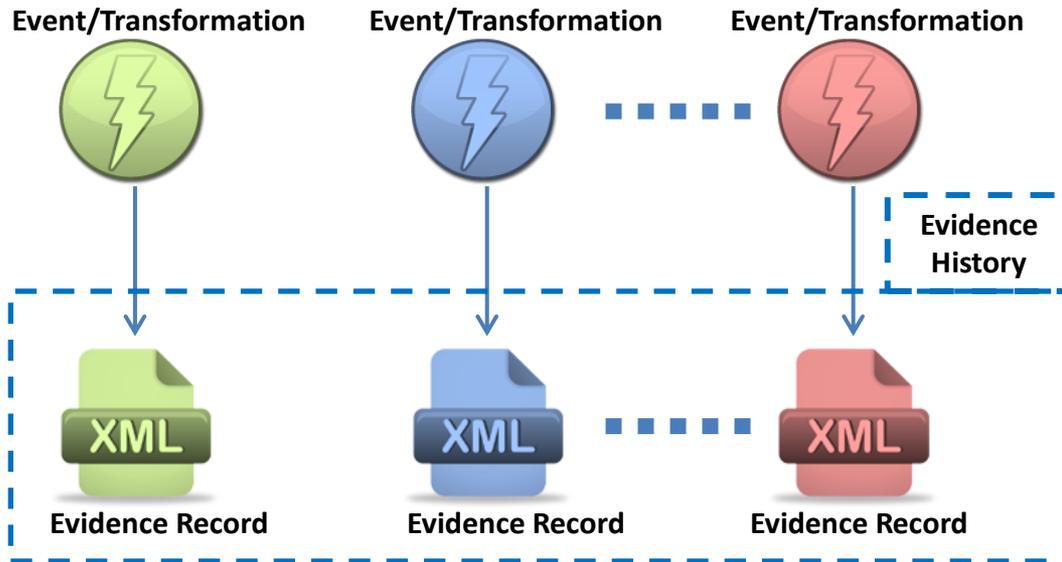
Dealing with Authenticity



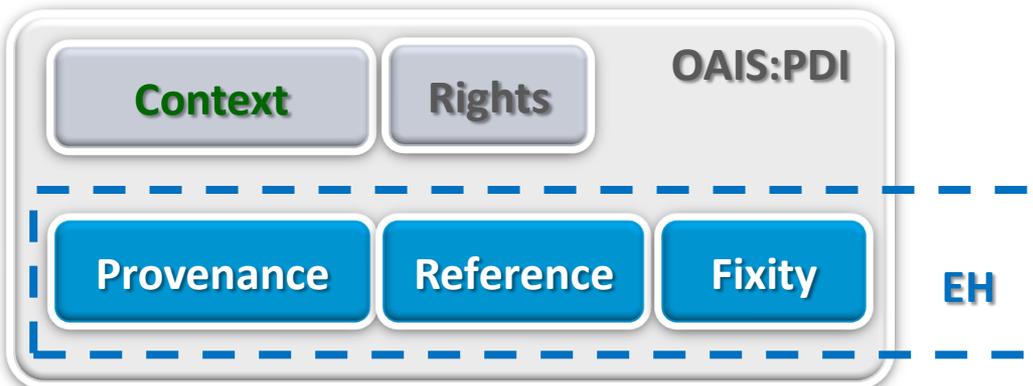
EU funded NoE aimed at bringing coherence, cohesion and continuity to research into barriers to the long-term accessibility and usability of digital information and data

- **APARSEN** proposes a methodology for the management of the authenticity of Digital Resources (DR):
 - **Formal authenticity model**: to represent the DR lifecycle and the management of authenticity evidence
 - **Operational guidelines**: to guide the process of instantiating the model in a specific environment
 - **Case studies**: carried out to tune the methodology and test its effectiveness in a set of heterogeneous environments

SCIDIP-ES Solution



- **SCIDIP-ES** manages **evidences** of events and/or transformations on digital objects (called digital representations in **PREMIS**)
- **Evidences** are provided by structured documents called **Evidence Records (ERs)**
- Each transformation is documented by its own ER
- Whole set of ERs provides the **Evidence History (EH)** of all transformations carried out on digital representation, that is **provenance**
- **SCIDIP-ES** supports the archive manager to capture and manage key information of the **OAIS** Preservation Description Information (**PDI**), that is the **EH**



SCIDIP-ES Solution



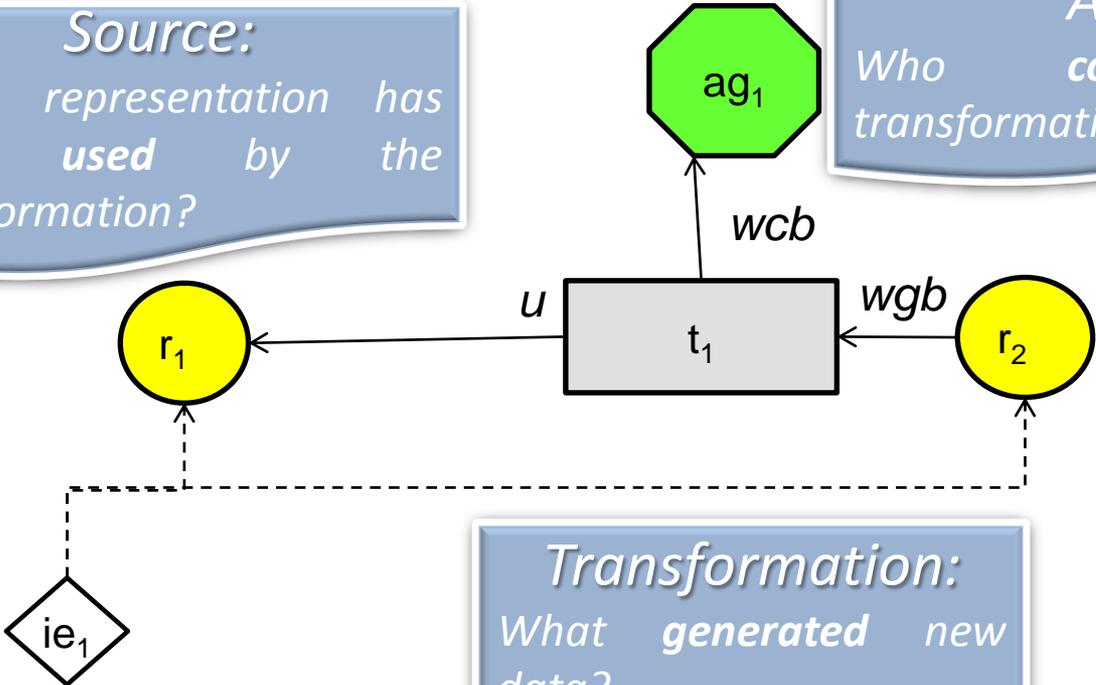
Source:
Which representation has been used by the transformation?

Agent:
Who controlled the transformation?

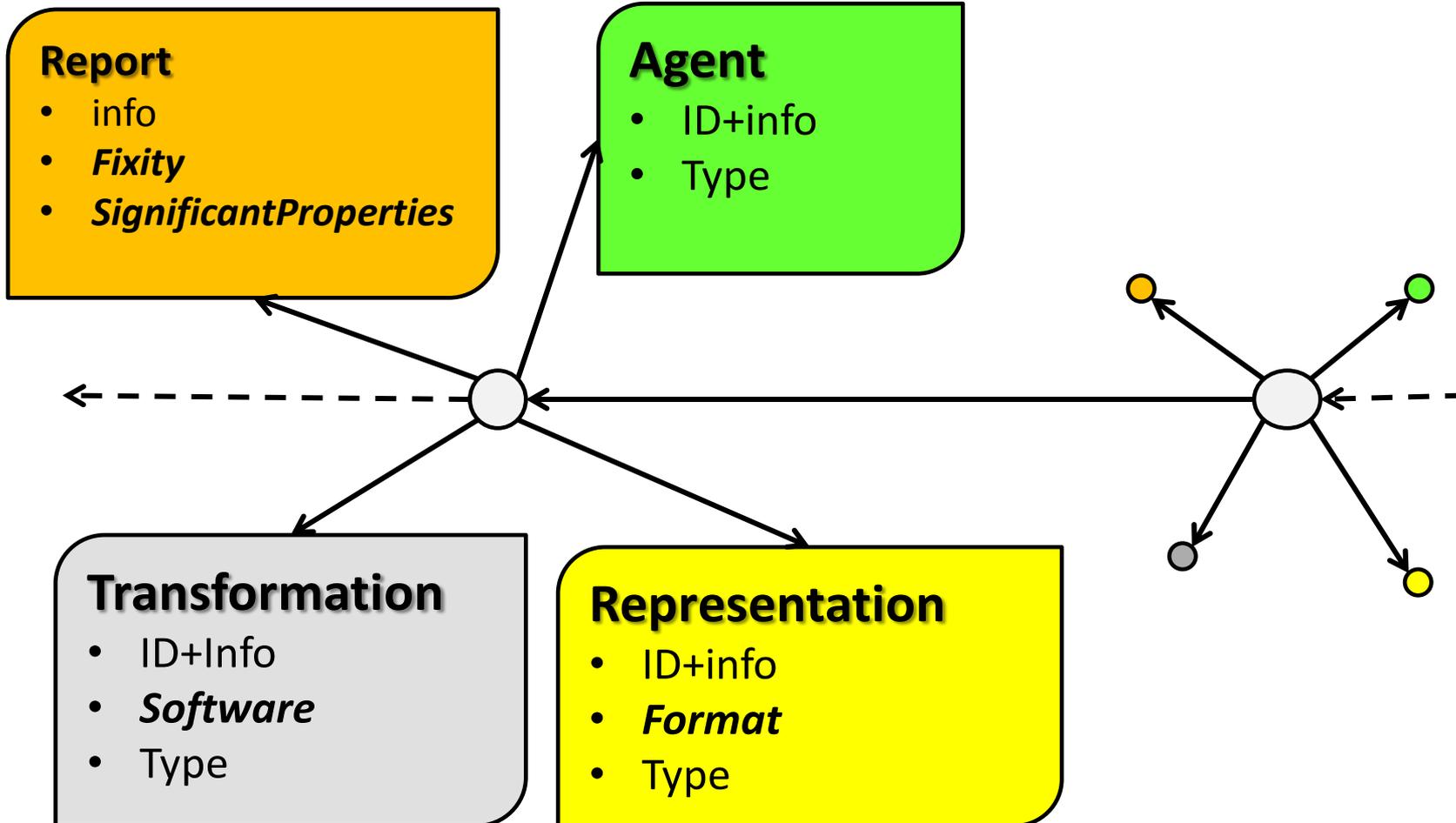
Representation:
digital representation of an Intellectual Entity and result of a transformation

Transformation:
What generated new data?

Intellectual Entity:
abstraction of digital representation



SCIDIP-ES Solution



SCIDIP-ES Solution



Open Provenance Model formalism has been adopted for modelling lifecycle of digital object as a ***provenance graph***

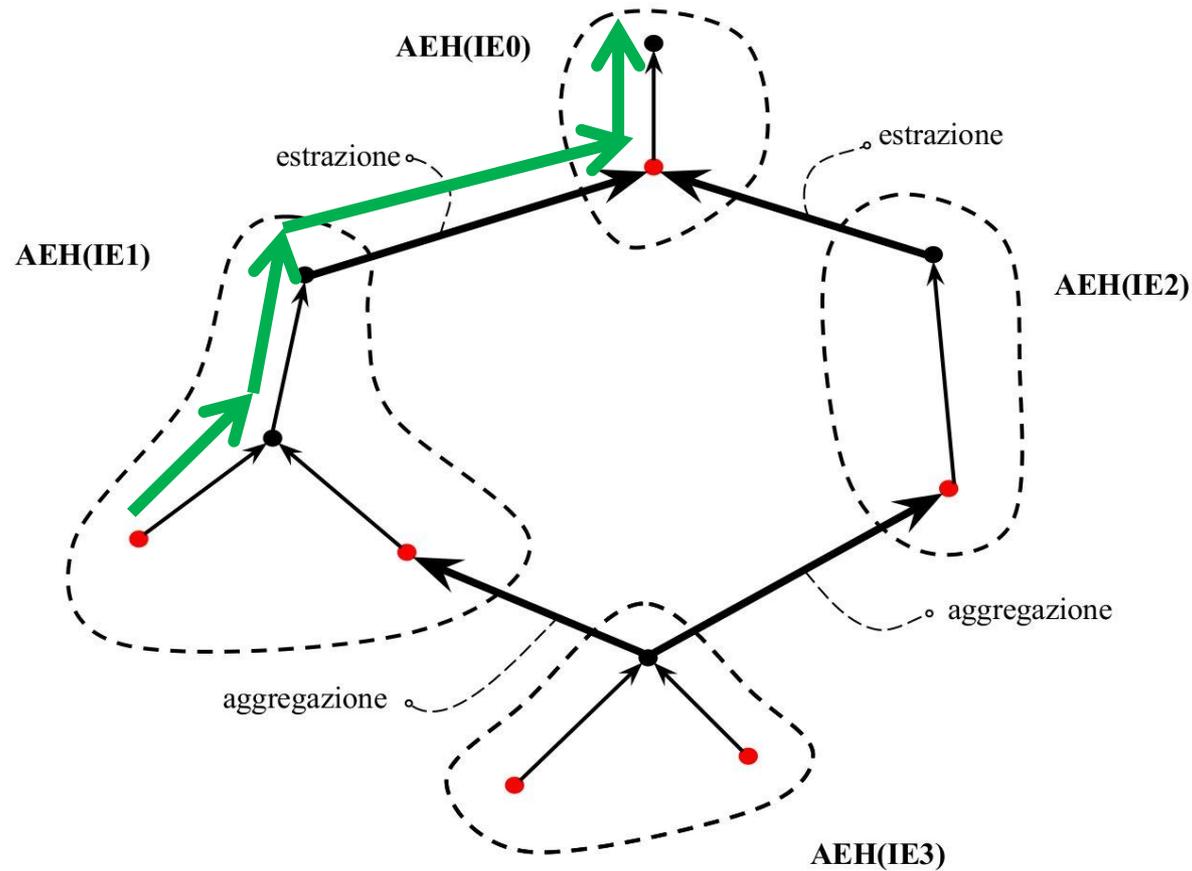


Standardized XML-based structures to represent both the provenance graph and the authenticity evidence

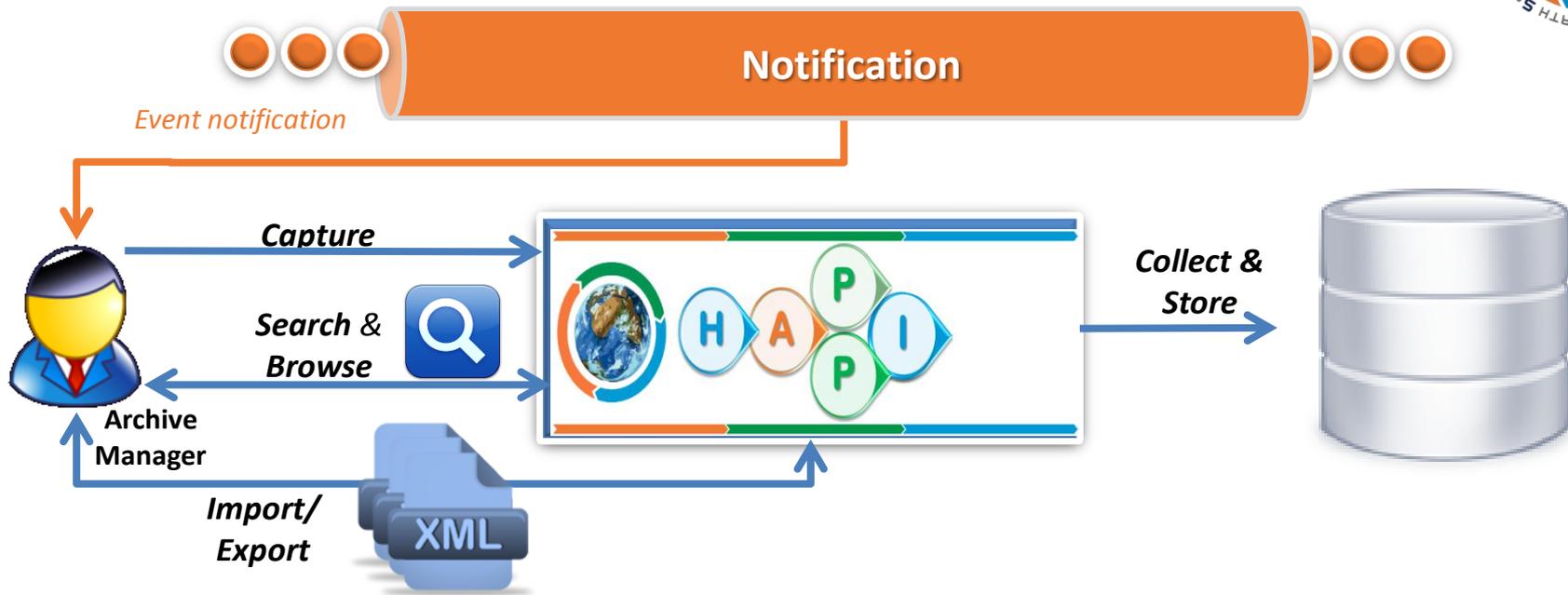
Common Dictionary for ensuring **Interoperability** among different repositories in managing the authenticity evidence

PREMIS

Example of Evidence History



SCIDIP-ES Solution

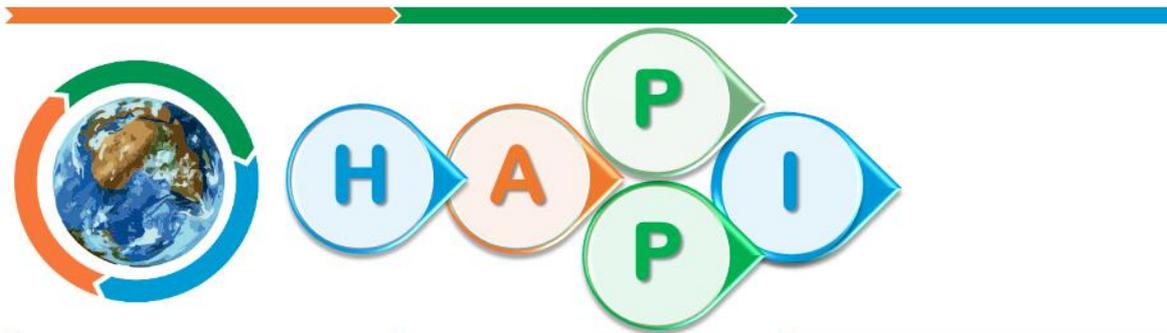


- **SCIDIP-ES HAPPI (Handling Authenticity Provenance and Persistent Identifiers)**
 - **Manage** Intellectual Entity
 - **Capture** Evidence Record Documentation (OPM1.1 and PREMIS2.2)
 - **Collect** Evidence History in a scalable database
 - **Search/Browse** Evidence History
 - **Import/Export** Evidence History

SCIDIP-ES Solution



Archive
Manager



Authenticity

Provenance

Integrity



Evidence
Manager



Intellectual
Entity
Manager

Pers. Identifier

Persistence



Some Scenarios

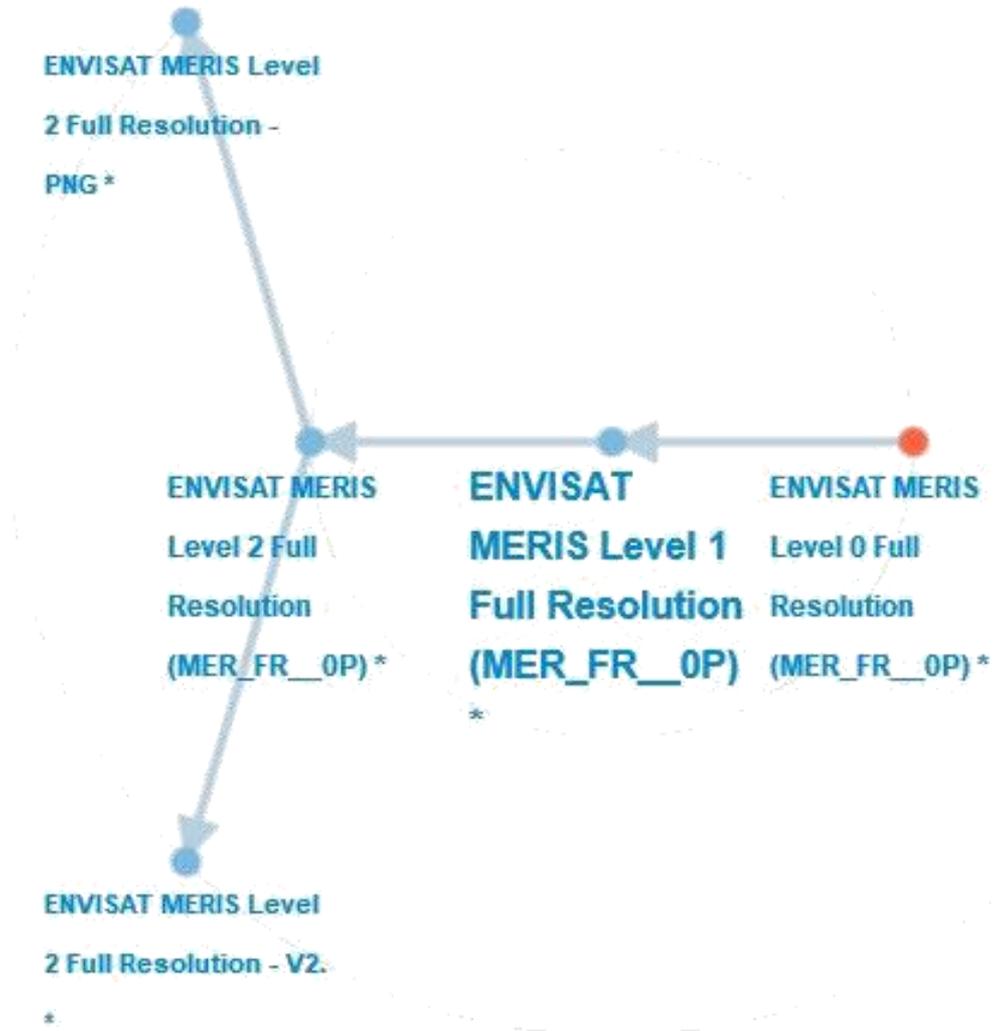


- Different missions/campaigns for Earth Phenomena, Earthquakes, Geophysics and Geomarine Observations
- They need to trace Provenance of data:
 - **Who** processed and controlled data
 - **When**
 - used data **Source**,
 - **How** (algorithm and software)



Istituto Nazionale di
Geofisica e Vulcanologia

Some Scenarios



Some Scenarios



00:00
2013/09/04



ENVISAT MERIS Level 2 Full Resolution (MER_FR__0P)

Was controlled by Luigi Briguglio

Was generated by Transformation from L1 to L2

Annotation: "Captured L2 dataset from ENVRI portal"

For further details see [evidence report](#)

[link to dataset](#) from ESA

00:00
2013/09/02

ENVISAT MERIS Level 1 Full Resolution (MER_FR__0P)

00:00
2013/09/02

ENVISAT MERIS Level 0 Full Resolution (MER_FR__0P)



Conclusion

- **SCIDIP-ES** is delivering a data preservation **Infrastructure**, specifically addressing needs from Earth Science Community
- SCIDIP-ES Infrastructure provides a set of Generic Services and Toolkits which can support archive managers, of other communities too, in LTDP processes
- **Authenticity, Integrity, Reference and Provenance** are addressed by the toolkit SCIDIP-ES HAPPI and adopts widely recognised standards such as **OAIS, OPM and PREMIS**
- ES community is **validating** the SCIDIP-ES Infrastructure in their premises
- Generic Services and Toolkits of SCIDIP-ES Infrastructure may be evaluated in different communities (e.g. **Cultural Heritage**)

DIGITAL MEETS HERITAGE



Thanks for your kind attention



www.scidip-es.eu

