



Alliance Permanent Access to the  
Records of Science in Europe Network

---

# Exemplar Introduction to APARSEN Meta-Data Packing Tools and Services

Matthias Hemmje – FTK

APARSEN-EGI-Community-Forum Training on Data Preservation, May , 22.05.2014

# Successful Reference Innovation: DP Infrastructure for Earth Observation Data



# Proof-of Concept Outreach and successful Pilot Take-Up by SCIDIP-ES Innovation Partners



**ISPRA**  
Istituto Superiore per la Protezione  
e la Ricerca Ambientale



UNIVERSITA' degli STUDI di ROMA  
TOR VERGATA  
UFFICIO PARCO SCIENTIFICO



# Pilot Scientific Data Domain: SCIP-ES Stakeholders, i.e., Prospective Users

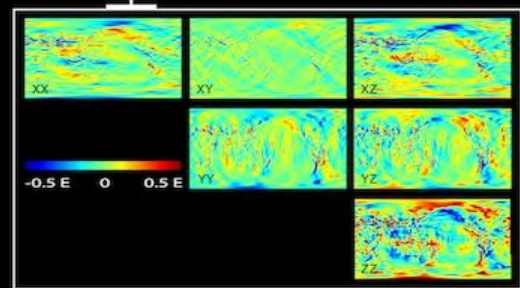
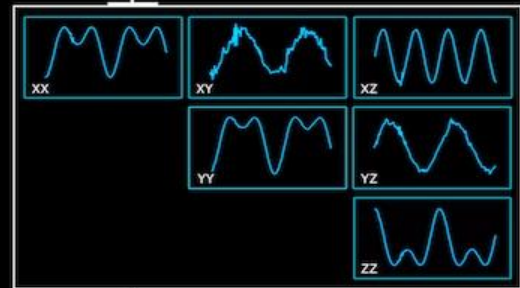
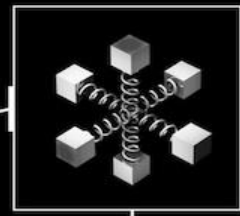
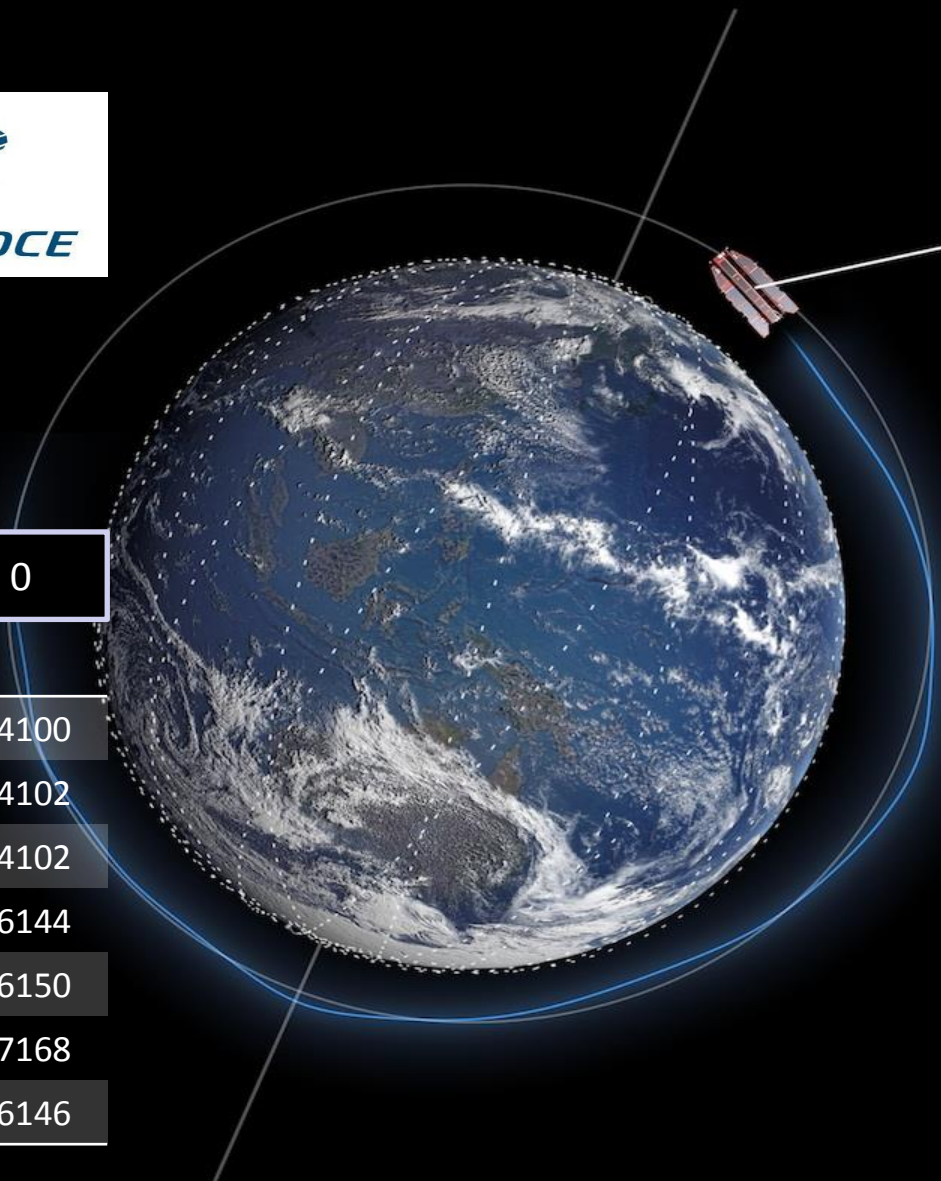


# Pilot Scientific Data Domain: Satellite Data Lifecycle - Acquisition of Data



Level 0

001	4100
002	4102
003	4102
004	6144
005	6150
006	7168
007	6146

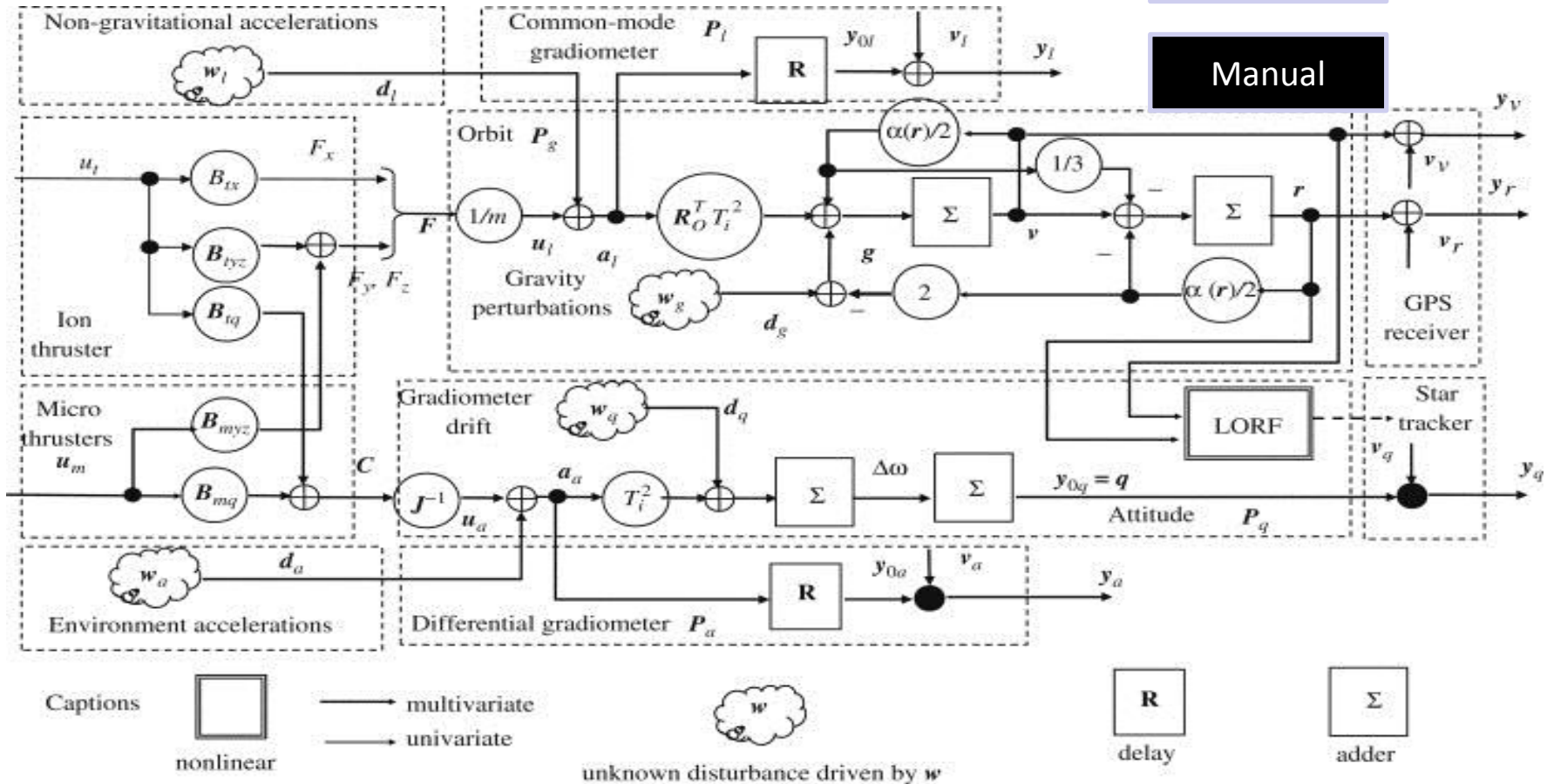




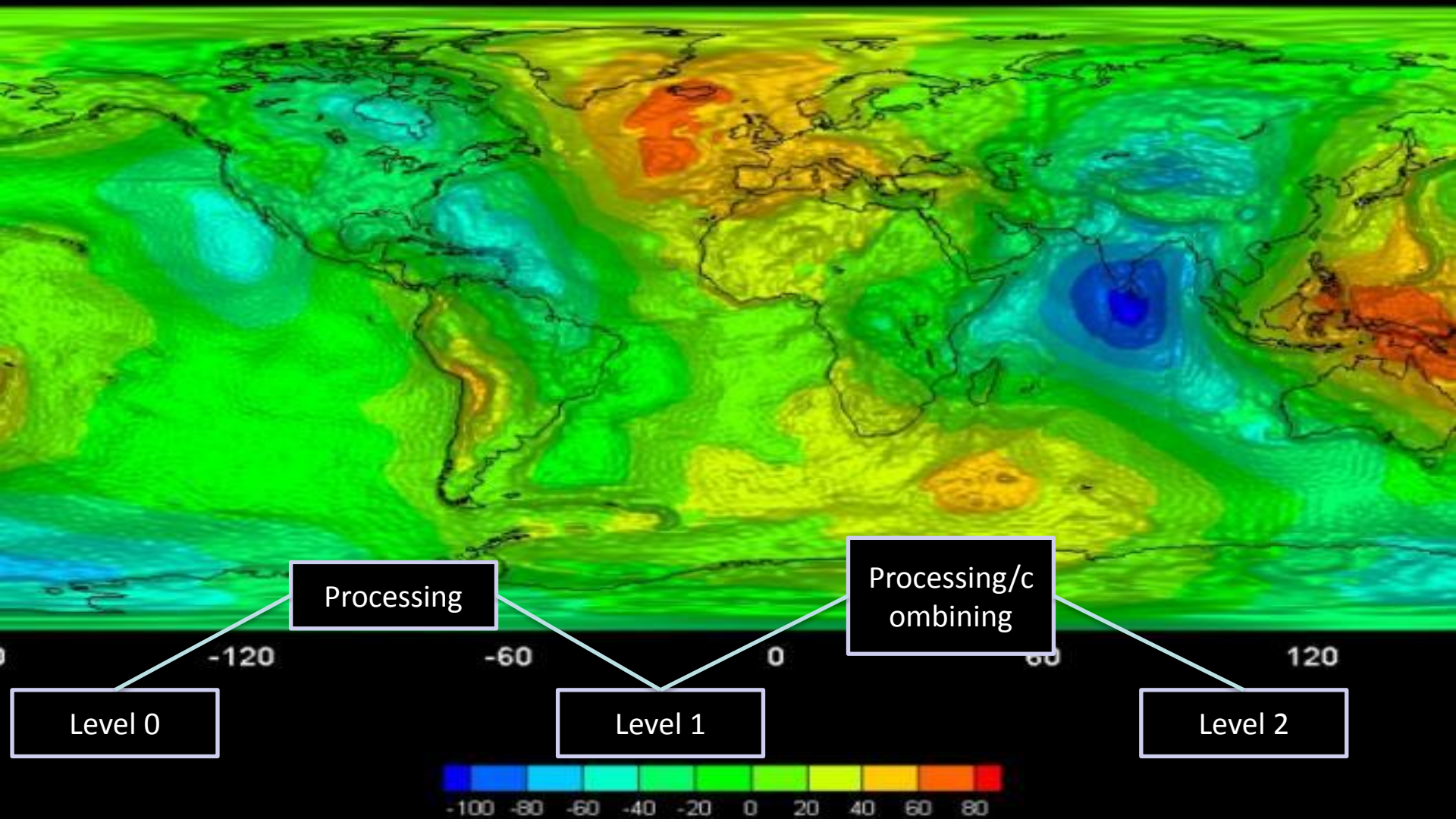
# Pilot Scientific Data Domain: Data Lifecycle - Processing of Data

Algorithm

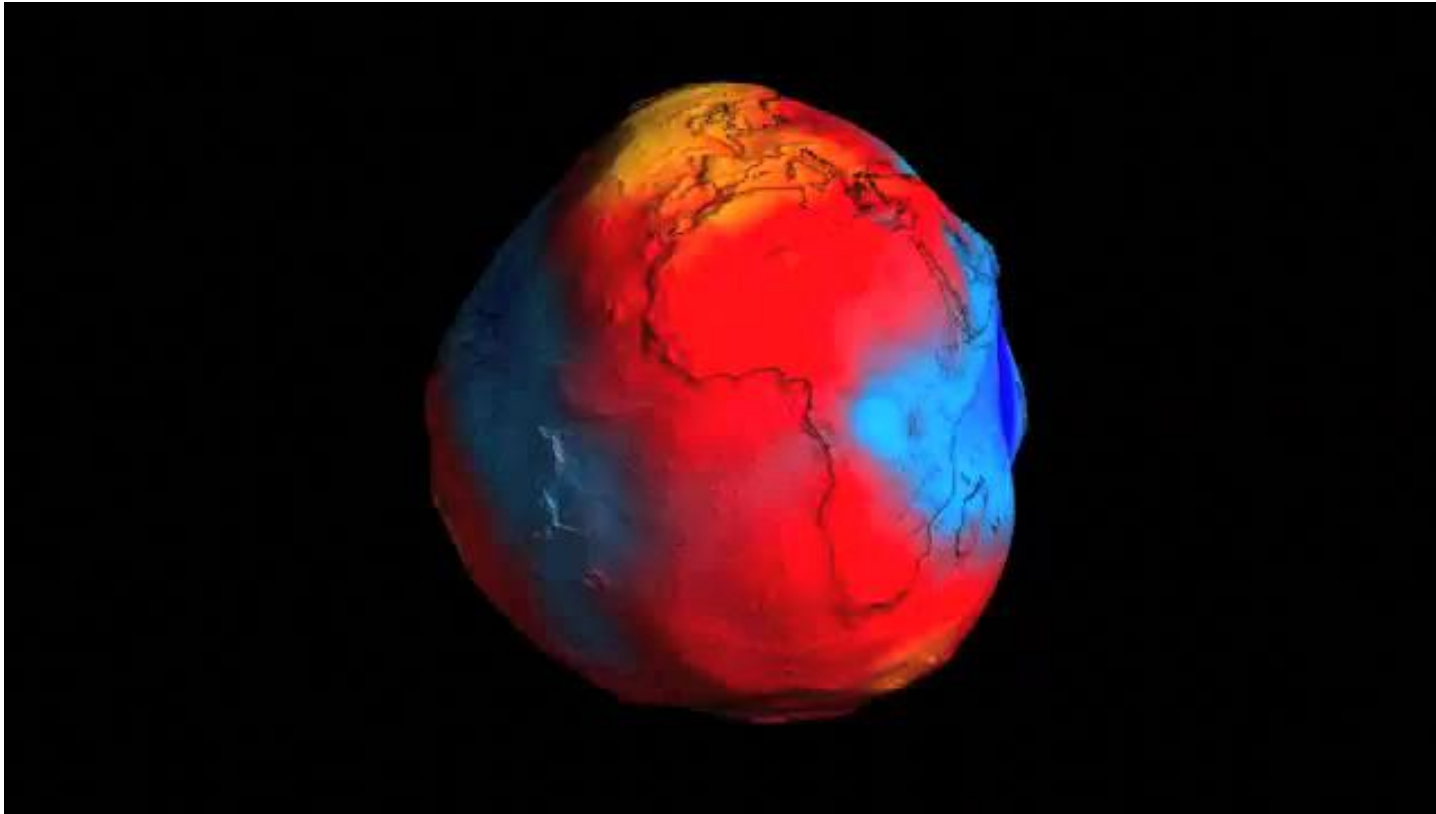
Manual



# Pilot Scientific Data Domain: Data Lifecycle - Combination of Data



# Pilot Scientific Data Domain: Data Lifecycle - Reuse of Data



## Publications

**3 AFRICA EXPLORATION**

Many geoscientists are using satellite gradients for North-Central Africa. Control for height... [View full article](#)

**Fig. 3 (below):** Different habits in Chad/Niger/Chad Basin. Observe which structures are not highlighted. See also the cross-section area of GOCE.

**Fig. 2 (below):** The profile of Bouguer anomaly. Bouguer  $\rho = 2670 \text{ kg/m}^3$ . The arrow is the Chad basin. Left: north of Chad basin and at 8000m depth.

**CONCLUSIONS**

- GOCE data in Exploration geophysics is important to improve the spatial resolution corresponding to the 100 km resolution of the Bouguer anomaly.
- Application of higher frequencies in terms of error or volume is a measure of data processing at level 10.
- Application control of 2D and 3D topographic geophysics allows quality control of newly acquired data globally.

**REFERENCES**

- W. K. Miller, J. J. G. 2011
- W. K. Miller, J. J. G. 2011
- W. K. Miller, J. J. G. 2011

**2 AFRICA EXPLORATION**

Many geoscientists are using satellite gradients for North-Central Africa. Control for height... [View full article](#)

**Fig. 2 (below):** The profile of Bouguer anomaly. Bouguer  $\rho = 2670 \text{ kg/m}^3$ . The arrow is the Chad basin. Left: north of Chad basin and at 8000m depth.

**CONCLUSIONS**

- GOCE data in Exploration geophysics is important to improve the spatial resolution corresponding to the 100 km resolution of the Bouguer anomaly.
- Application of higher frequencies in terms of error or volume is a measure of data processing at level 10.
- Application control of 2D and 3D topographic geophysics allows quality control of newly acquired data globally.

**REFERENCES**

- W. K. Miller, J. J. G. 2011
- W. K. Miller, J. J. G. 2011
- W. K. Miller, J. J. G. 2011



# Piloting Specific Use Case Demand: Knowledge Preservation for Satellite Data

Level 0

Level 1

File specs.

Desc. Info

Processor

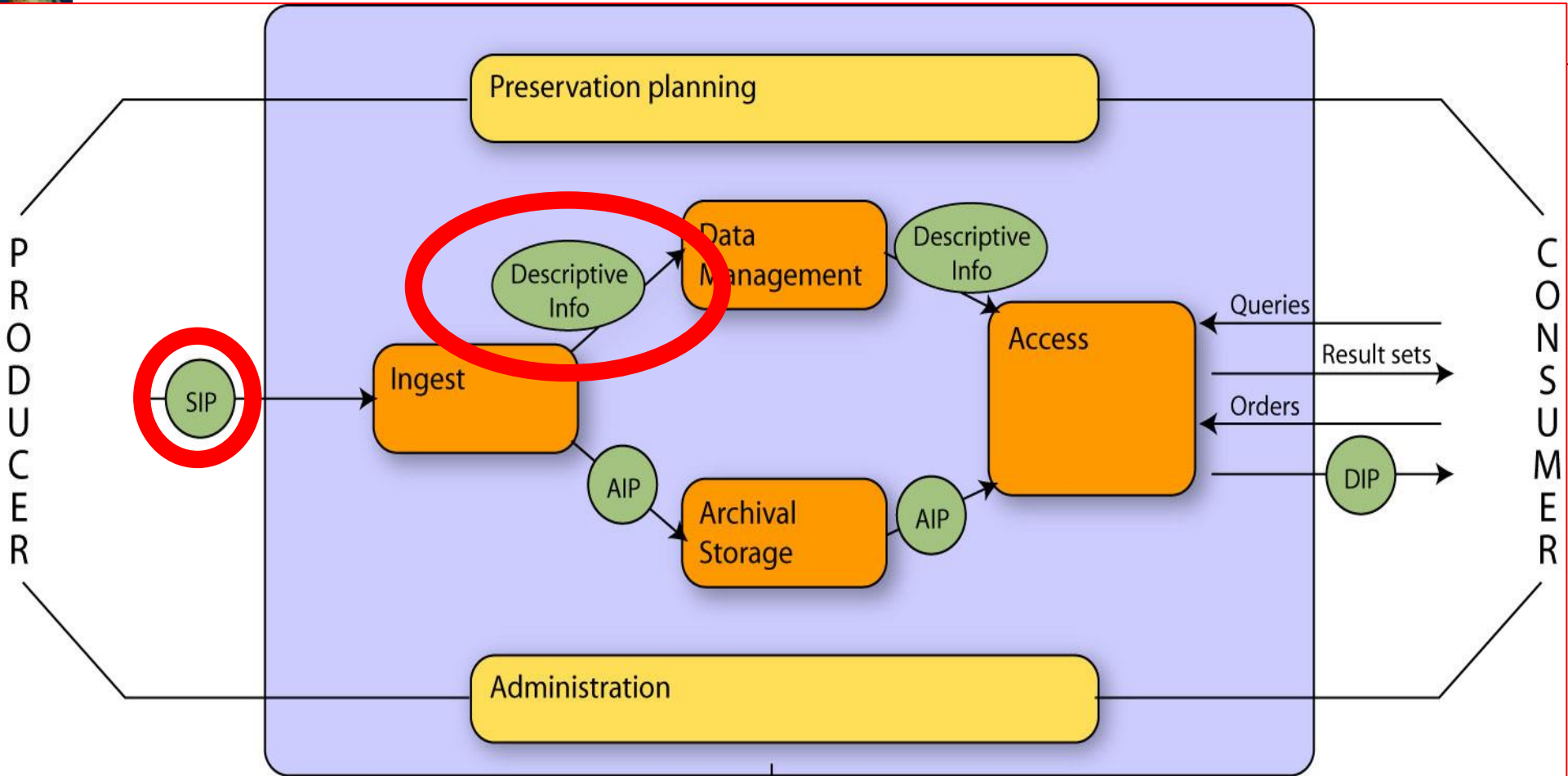
Algorithm

User Manual

Publications

- ***The preservation of data (the “bytes”) is useless without the preservation of the knowledge associated with the data (e.g. the “quality”, the process to generate them)***
- We must:
  - Ensure and secure the preservation of archived data and associated knowledge for an unlimited time span.
  - Ensure, enhance and facilitate archived data accessibility.

# APARSEN Strength: Starting Point OAIS

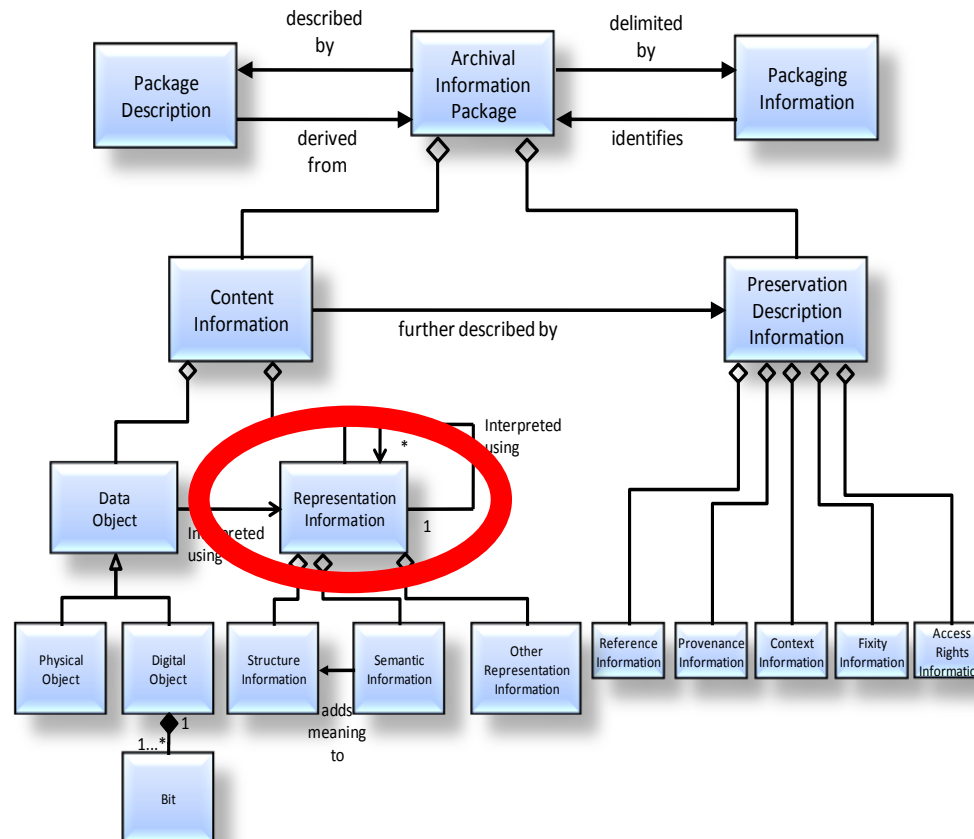


SIP: Submission Information Package  
 AIP: Archival Information Package  
 DIP: Dissemination Information Package

MANAGEMENT

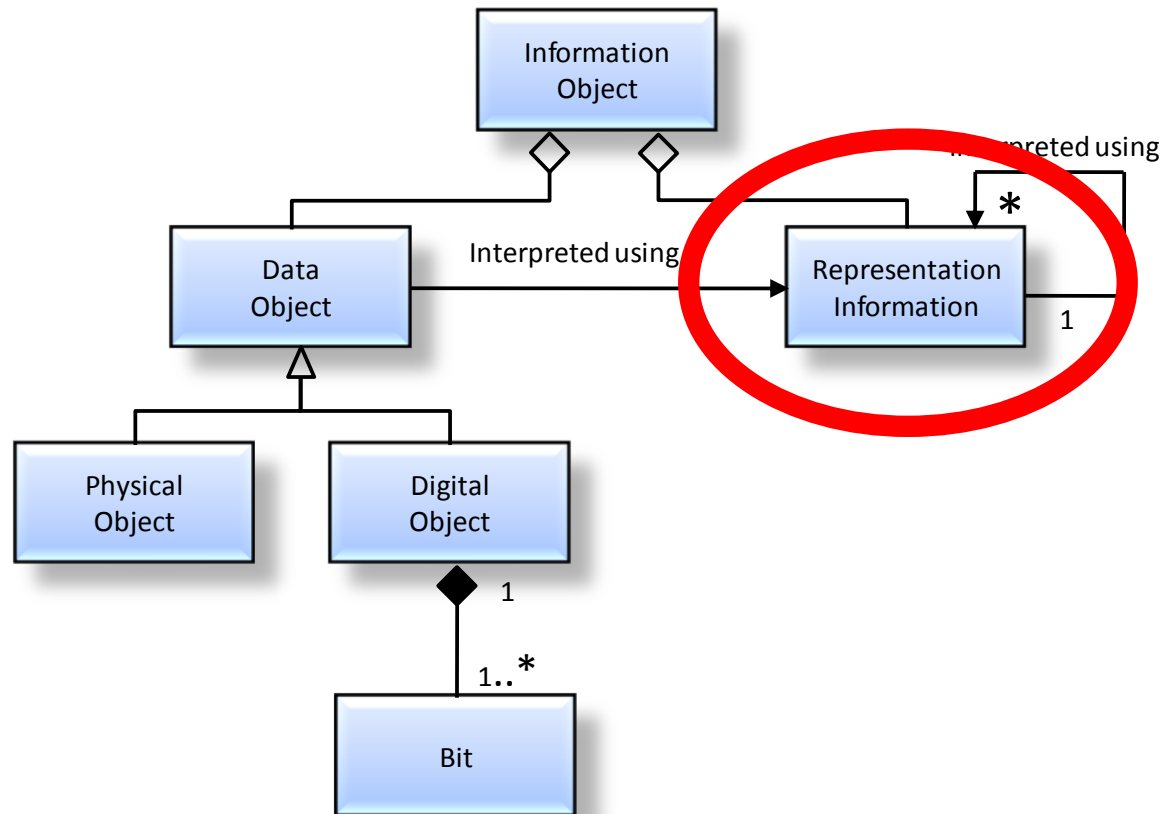
# Unifying Open Standard – Managing DP Meta Data

## Challenge: Meta-Information for Preservation (I)



# Unifying Open Standard – Managing DP Meta Data

## Challenge: Meta-Information for Preservation (II)

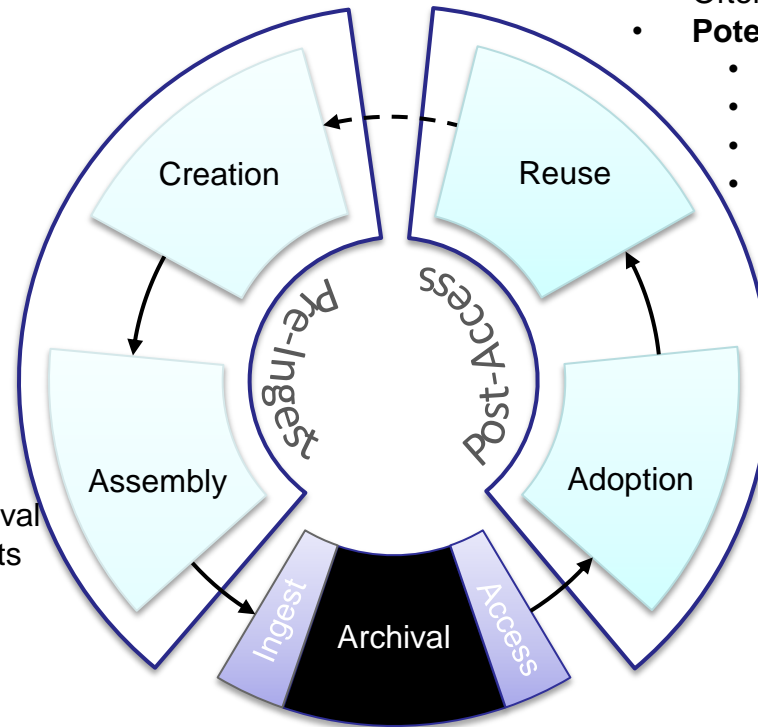




# DP Meta-Data in the Archive-centric Information Life-Cycle Distribution, Processes and Integration of Archive Systems

- **Production** of new digital objects
- **Use** according to original purpose
- Archival often right after creation, in parallel to use

- **Exploitation** of digital object by consumer
- Often **re-purposing** of digital objects
- **Potential outcome**
  - Creation of new digital objects
  - Revision of digital objects
  - Extension or update of metadata
  - Annotation

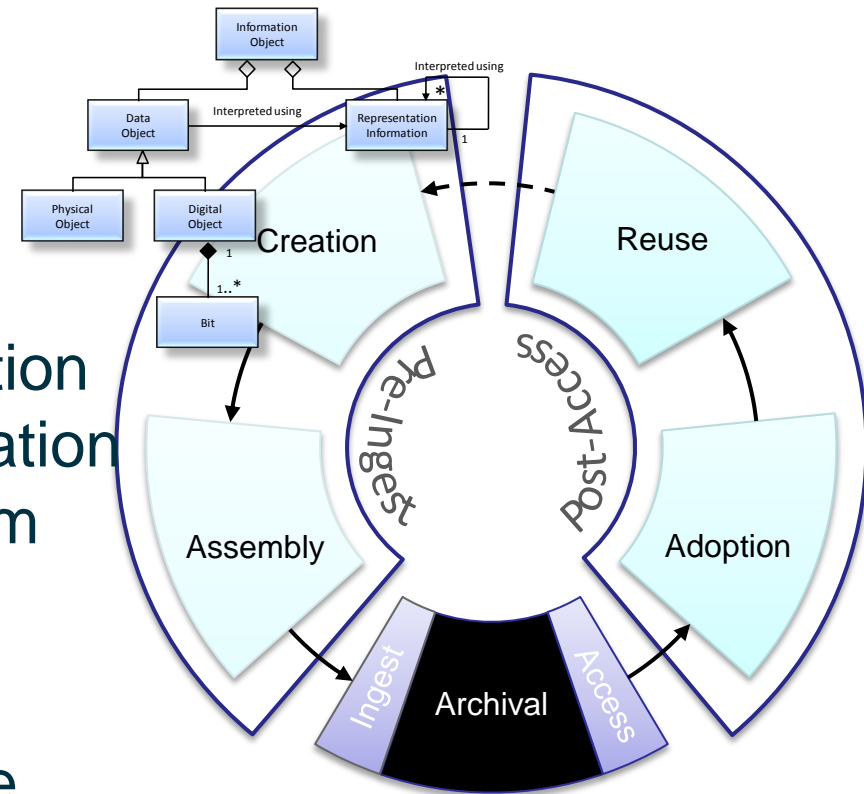


- **Appraisal** of objects relevant for archival
- **Compilation** and enrichment of objects to preserve
- Creation of Submission Information Package (*SIP*)

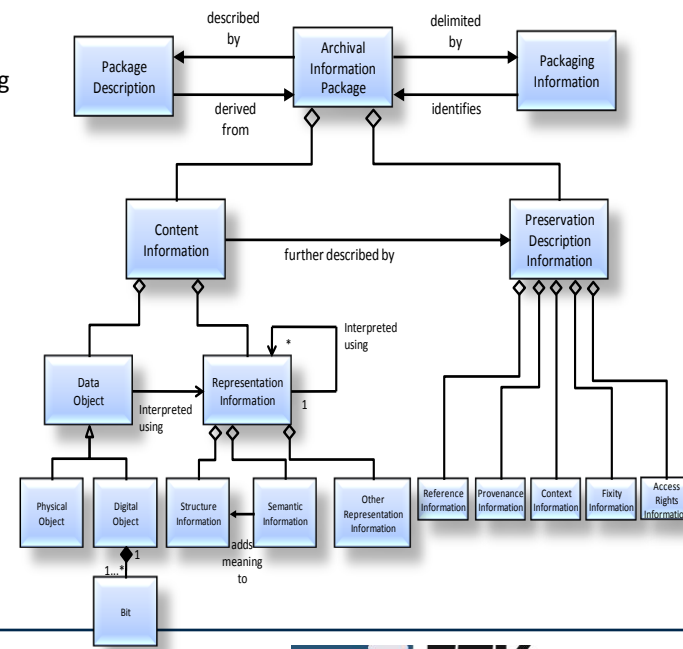
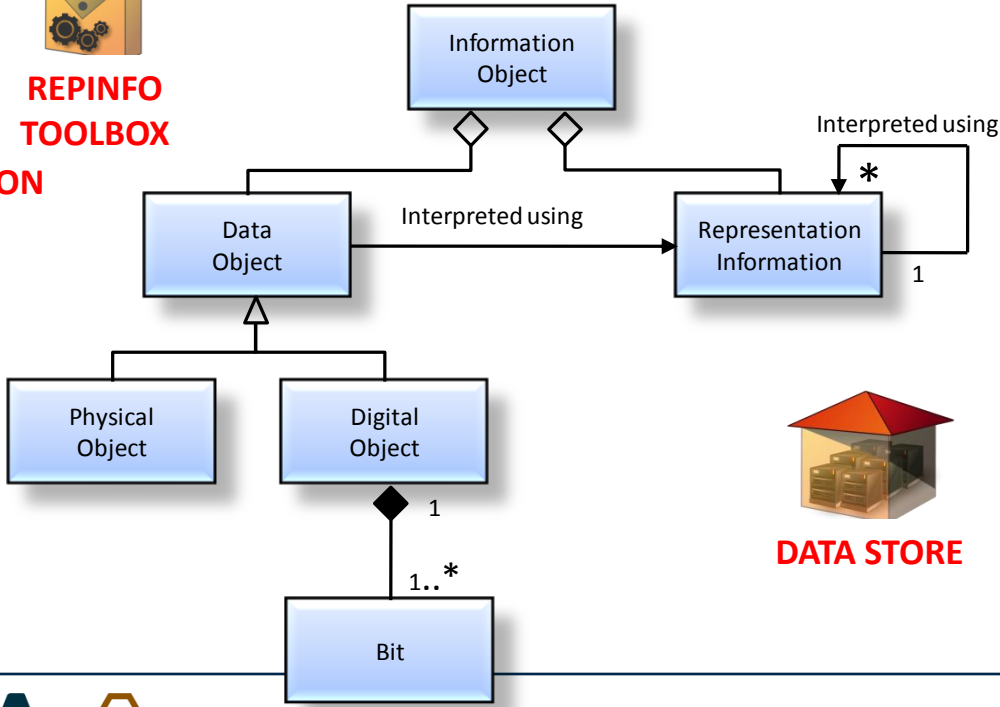
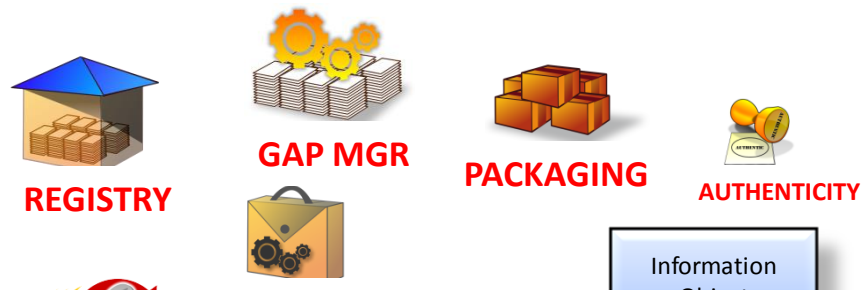
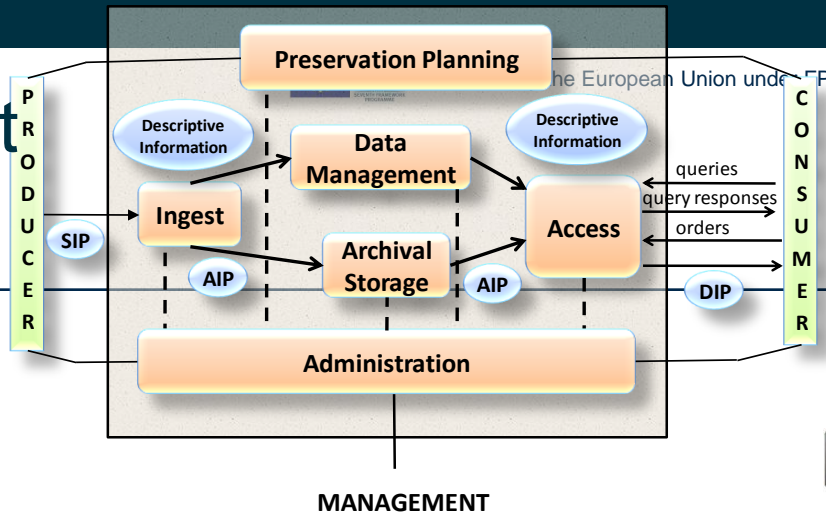
- **Receiving and Examination** of Dissemination Information Packages (*DIP*)
- **Adaption and integration** of digital objects into working environment
- **Recontextualize** digital objects and accompanying information for prospective reuse

# APARSEN USP: Archive-Centric Knowledge Preservation Lifecycle

- Move through Repositories along Life Cycle
- For Meta-information Preservation and re-use: harvest Representation Information or Extract/Transform
- Alternatively move to another/next repository/lifecycle



# Exemplar Innovation Pilot AIP-Toolkit (SCIDIP-ES)



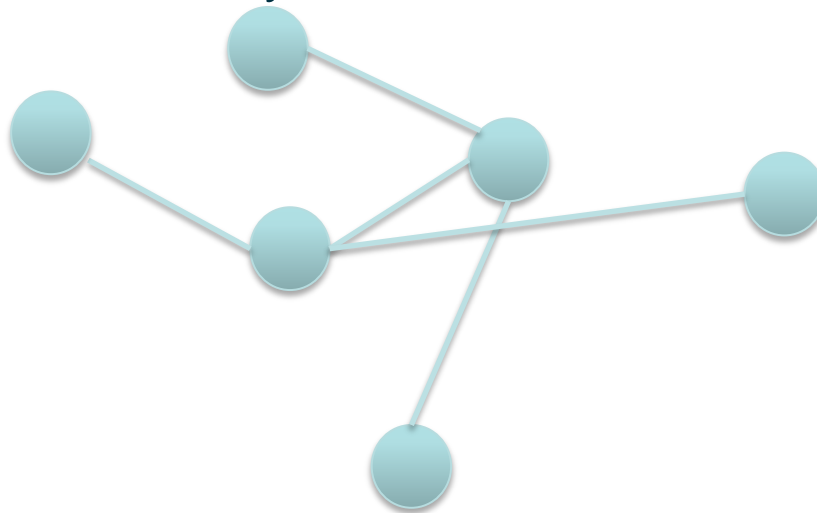
# APARSEN Tools and Services: Registry of Services and Tools for DP

- Various Registries exist, among others: *COPTR*, *DCH-RP tool-registry*, *APARSEN*
- Differ in: methodology, provided data
- APARSEN Evidence-Based DP Tools Registry/Repository.
  - Provision of evidence, where the software has been found effective in preservation.
  - Besides links and description to S&T
    - ▶ Description and links to applied testbeds
    - ▶ Description of scenarios
    - ▶ User ratings
- In terms of OAIS, those evidence information is **Representation Information** (the information that maps a Data Object Into more meaningful concepts).



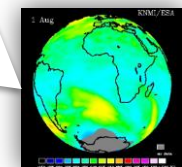
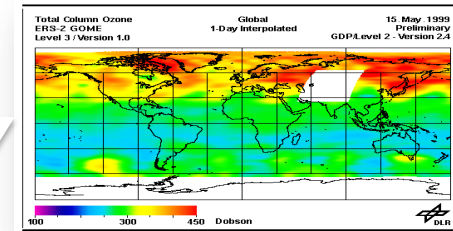
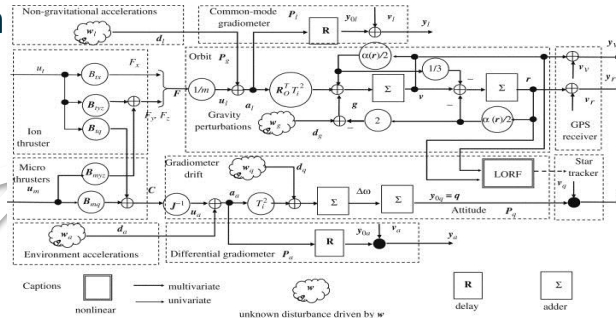
# APARSEN Tools and Services: DP in SCIDIP-ES

- In terms of OAIS, this evidence information is **Representation Information** (*the information that maps a Data Object Into more meaningful concepts [CCSDS]*).
- **Representation Information Network:** The set of Representation Information, that fully describes the meaning of a Data Object.



# APARSEN Tools and Services: DP in SCIDIP-ES

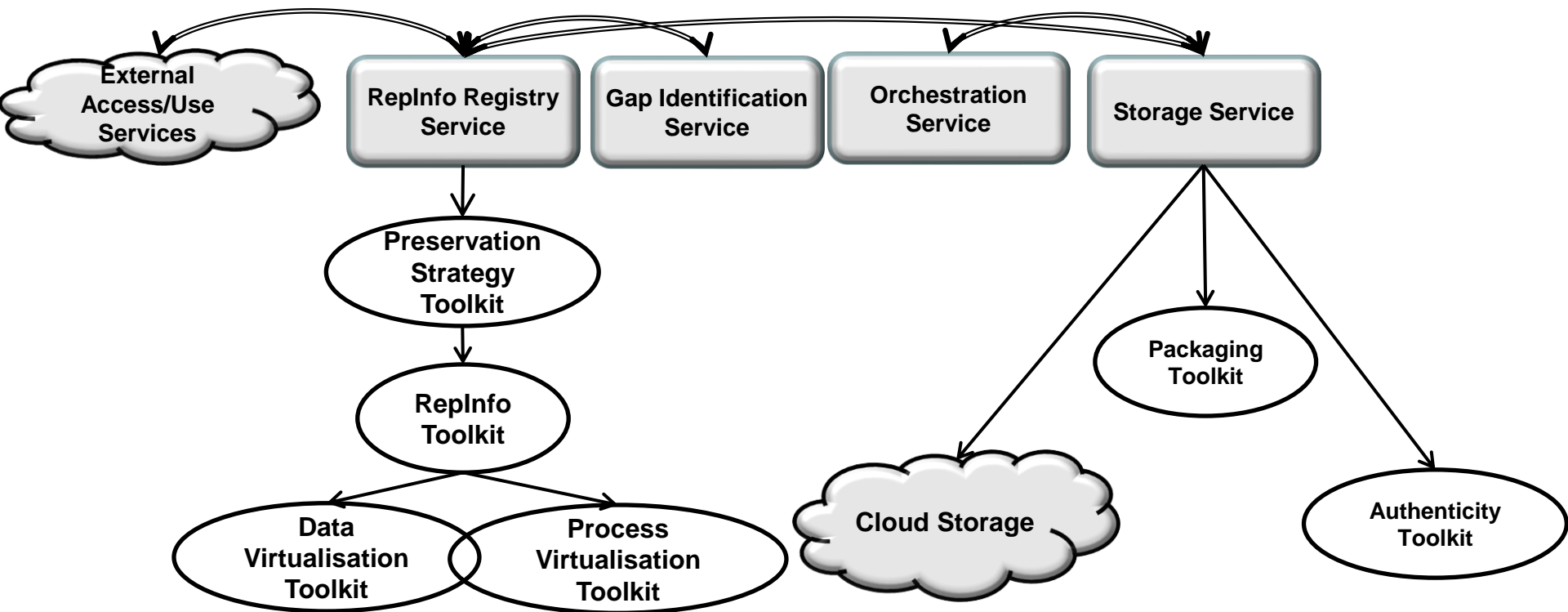
- SCIDIP-ES project: building an e-infrastructure for Earth Observation Data. Several joint activities between APARSEN and SCIDIP-ES.
- ESA Preserved Data Set Content
  - Data Records (from raw data to high level products)
  - Processing software (e.g. processors for visualization)
  - Mission Documentation



latitude	longitude	Ozone	Time
132	50	34.9	12/03/1999
178	50	45	12/03/1999
190	50	78	12/03/1999

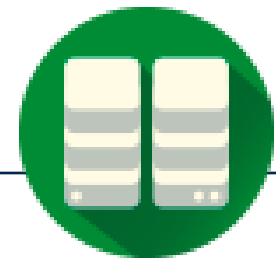
# APARSEN Tools and Services: From Registration to Packaging&Storage

- The preservation of EO data (the “bytes”) is useless without the preservation of the knowledge associated with the data



# APARSEN Tools and Services: RepInfo Registry Service

- **Repinfo Registry Service: A Container of Knowledge**
  - Structure: headers, footers, instrument measures, annotations, fixed parts, variable parts
  - Meaning:
    - Basic information (The so-called “metadata”: acquisition time, lat/lon, etc.)
    - Auxiliary Information, such as Sensor Description, physical parameters measured by the sensor, etc.
    - Usage Information: how this digital object can be exploited by different user communities
  - Rendering Information: describing what additional software can be used to display/process/edit the digital object.





# APARSEN Tools and Services: Gap Identification Service

- Know that a change has happened
  - Hardware, software, environment, tacit knowledge
- Understand the implications – is there a “gap in understandability”
- Decide on the “best” course of action for preservation
- If Transformed – how to maintain authenticity
- Alternatively hand it on to another repository
  - As “Archival Information Packages”

Orchestration  
Service

Gap Identification  
Service

Preservation  
Strategy Toolkit

Authenticity  
Toolkit

Packaging  
Toolkit

# APARSEN Tools and Services: Archive Information Packaging Tool (I)

- Select S&T through Dashboard:



ScienceData Infrastructure for Preservation - Earth Science (SCIDIP-ES)

**Packaging Toolkit**  
 Packaging supports the creation and manipulation of Archival Information Packages.

**Description**

Packaging is about: Foster reuse of preserved material is a key motivation for the preservation of digital objects. Provision of sufficient content for later access and comprehension of preserved material is a critical task. The packaging toolkit supports the creation and manipulation of Archival Information Packages, Archival Information Units. Furthermore the toolkit enables the integration of representation information from various SCIDIP-ES registries.

- [create or modify AIP](#)
- [create or modify AIC](#)
- [create or modify AIU](#)
- [browse information packages](#)

- Select RTD by Use Cases:

## USE CASES



### ARCHIVE SETUP

Follow this link to discover how to configure the SCIDIP-ES preservation services. ...



### DATA ACCESS

See how SCIDIP-ES services provide added value to users, while preserving data in the long term ...



### ARCHIVE EVOLUTION

See how services and toolkits prevent data loss when things change in the archive ...

# Archive Information Packaging Tool (II)



## SCIDIP-ES

SCIENCE DATA INFRASTRUCTURE FOR PRESERVATION-EARTH SCIENCE

### SCIDIP-ES Packaging

[create or modify AIP](#)

[create or modify AIC](#)

[browse Information Packages](#)

## SCIDIP-ES

SCIENCE DATA INFRASTRUCTURE FOR PRESERVATION-EARTH SCIENCE

### ES Packaging

▼ create new AIP

1: provide an AIP title

2: upload digital object

✕ Clear All

ois book.pdf  
Done

Clear

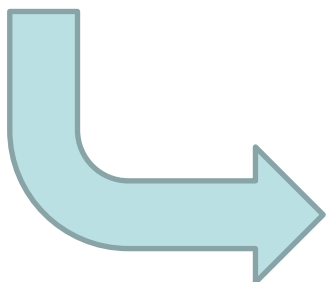
3: add manifest and create AIP

add

⤴ create new virtual AIP

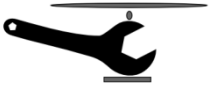
⤴ modify existing AIP

[back to menu](#)



## Create new AIP





# APARSEN Tools and Services: Archive Information Packaging Tool (III)

### AIP constituents

- Archival Information Package
- Content Information
  - Digital Object
- Preservation Description Information
- Representation Information

### values of digital\_object

826b33be-42cc-4885-bb3d-1630045712e3

oais book.pdf

### operations on AIP constituents

enter new reference

select reference

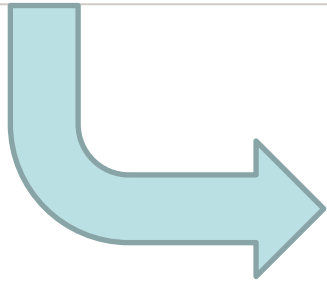
select reference

AIP title: Test

AIP ID: bf22bea9-7e95-4cb6-b672-130347501a6a

input format:

output format: OAI-ORE





# APARSEN Tools and Services: Archive Information Packaging Tool (III)



## SCIDIP-ES Packaging

### list of information packages

packagename	packagetype	action
AIP_OAI-ORE_abc_33c6a4bc-4c2b-4505-abb5-0a374aca347c.zip	OAI-ORE	
AIP_OAI-ORE_abc_41a9b6c7-e5d4-4e36-ab21-b83826506f8e.zip	OAI-ORE	
AIP_OAI-ORE_Test__bf22bea9-7e95-4cb6-b672-130347501a6a.zip	OAI-ORE	

[back to menu](#)

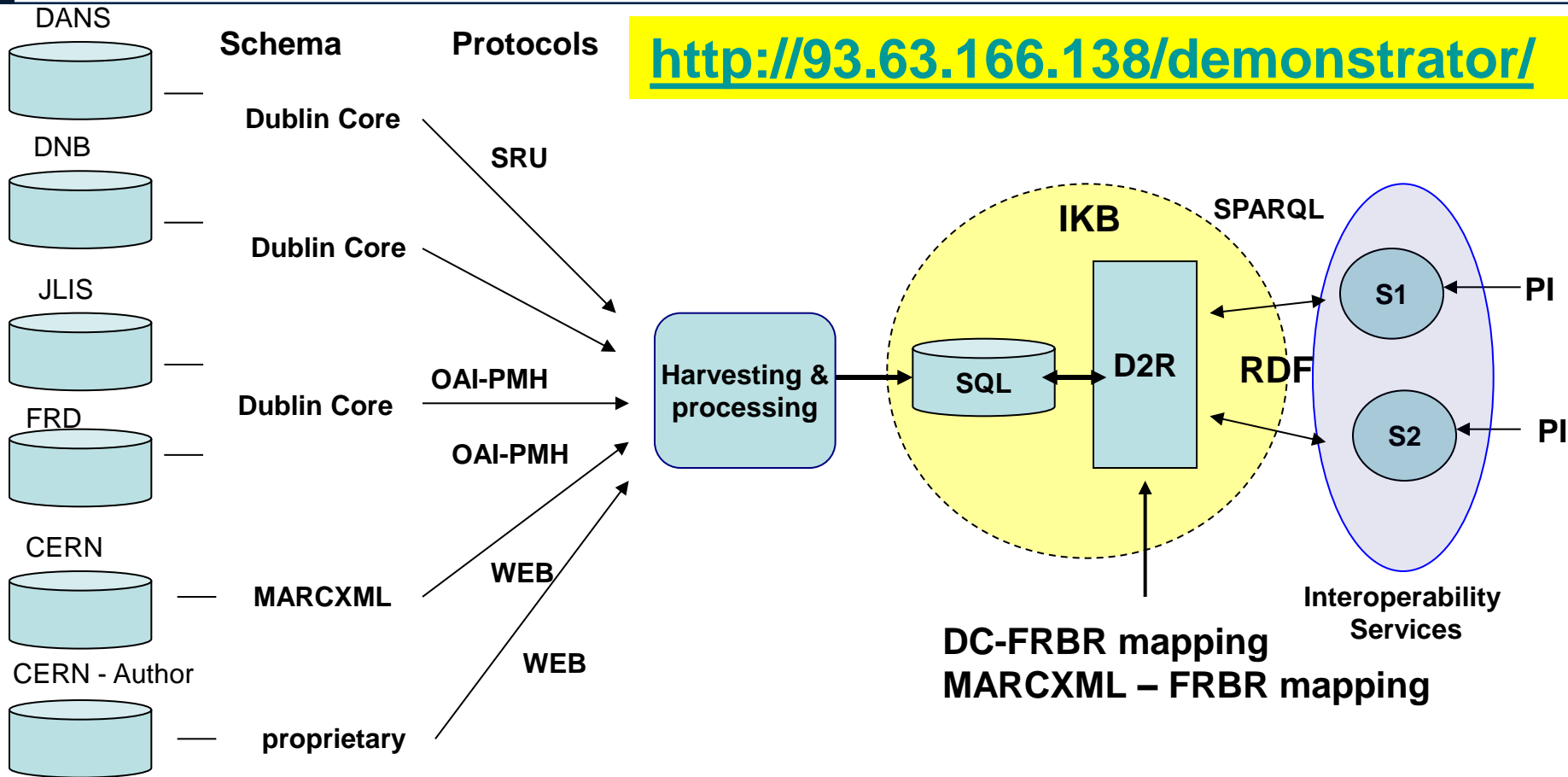
# APARSEN Tools and Services: Registry of Services and Tools for DP

- APARSEN vision
  - create and populate a knowledge base with preservation related software.
  - consolidation of distributed located preservation related software in one access point.
  
- Need for **Registry Alignment** to share community specific knowledge about existing DP Services and Tools
  - to share efforts and knowledge.

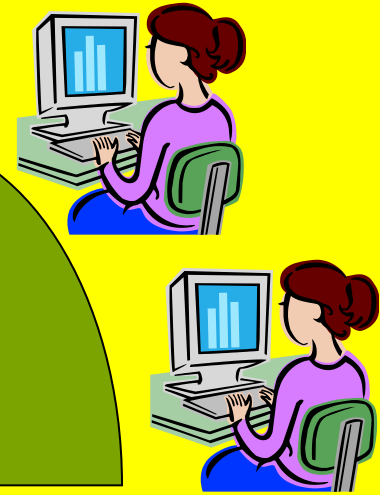
# Further Exemplar Offerings beyond Meta-Data Packaging

# Further exemplar offerings of the APARSEN VCOE: Persistent Identifier Interoperability Framework

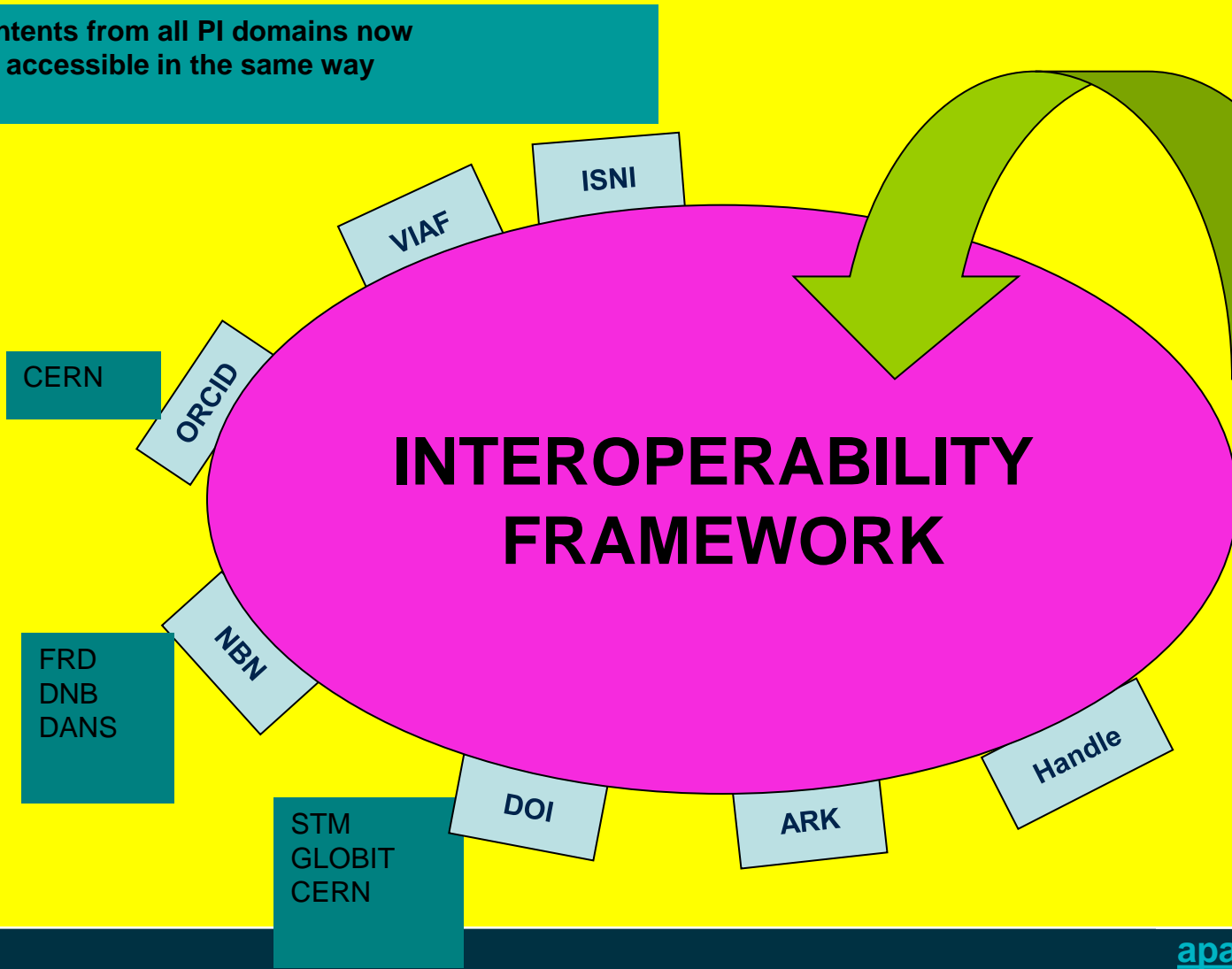
<http://93.63.166.138/demonstrator/>



# Further exemplar offerings of the APARSEN VCOE: Persistent Identifier Interoperability Framework



Contents from all PI domains now are accessible in the same way

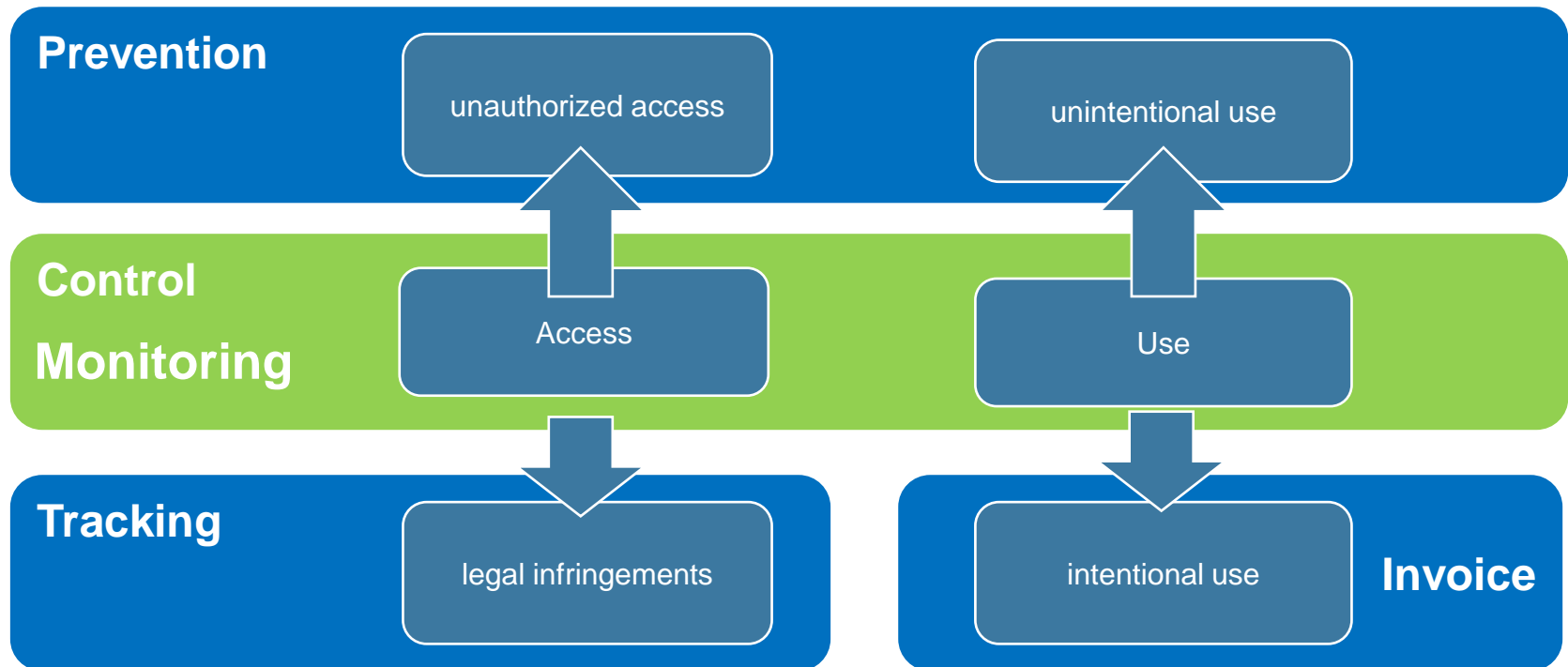


New services cross-domains for users requirements



# Further exemplar offerings of the APARSEN VCOE: Digital Rights Management (DRM)

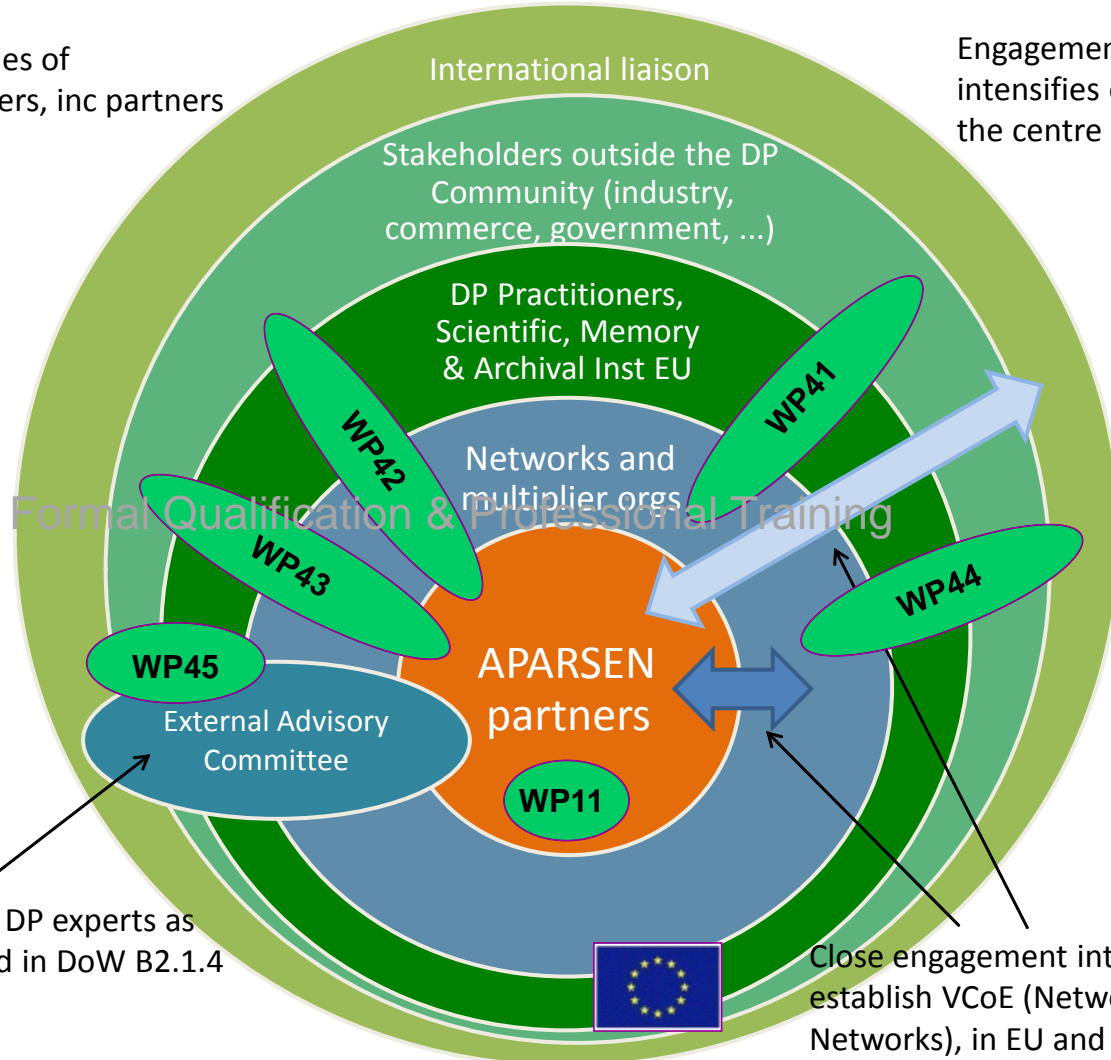
- Layers and Service Categories in DRM



# Further exemplar offerings of the APARSEN VCOE: Web-based Training (WBT) and Continuous Professional Education (CPE)

6 Categories of Stakeholders, inc partners

Engagement intensifies closer to the centre



Individual DP experts as mentioned in DoW B2.1.4

Close engagement intended to establish VCoE (Network of Networks), in EU and beyond

# Further exemplar offerings of the APARSEN VCOE: Webinars and E-learning, Online Training Portal for CPE

APARSEN webinar on Virtual Centres of Excellence

**RSEN**

Alliance Permanent Access to the  
Records of Science in Europe Network

0:00:27 / 2:00:16

YouTube

# Further exemplar offerings of the APARSEN VCOE: Interactive Map Resource Registration and Mediation

## Interactive Map of actors and stakeholders in Digital Preservation

Show Organisations Help

Select all | Deselect all

- Funding/policy
- Research
- Preservation
- Publishing

Please select the **categories** to show on the map. Click on the markers on the map, or hover your mouse to get a tooltip with a brief description.

Select All Europe | Deselect all

- Montenegro
- Albania
- Macedonia
- Turkey
- Belarus

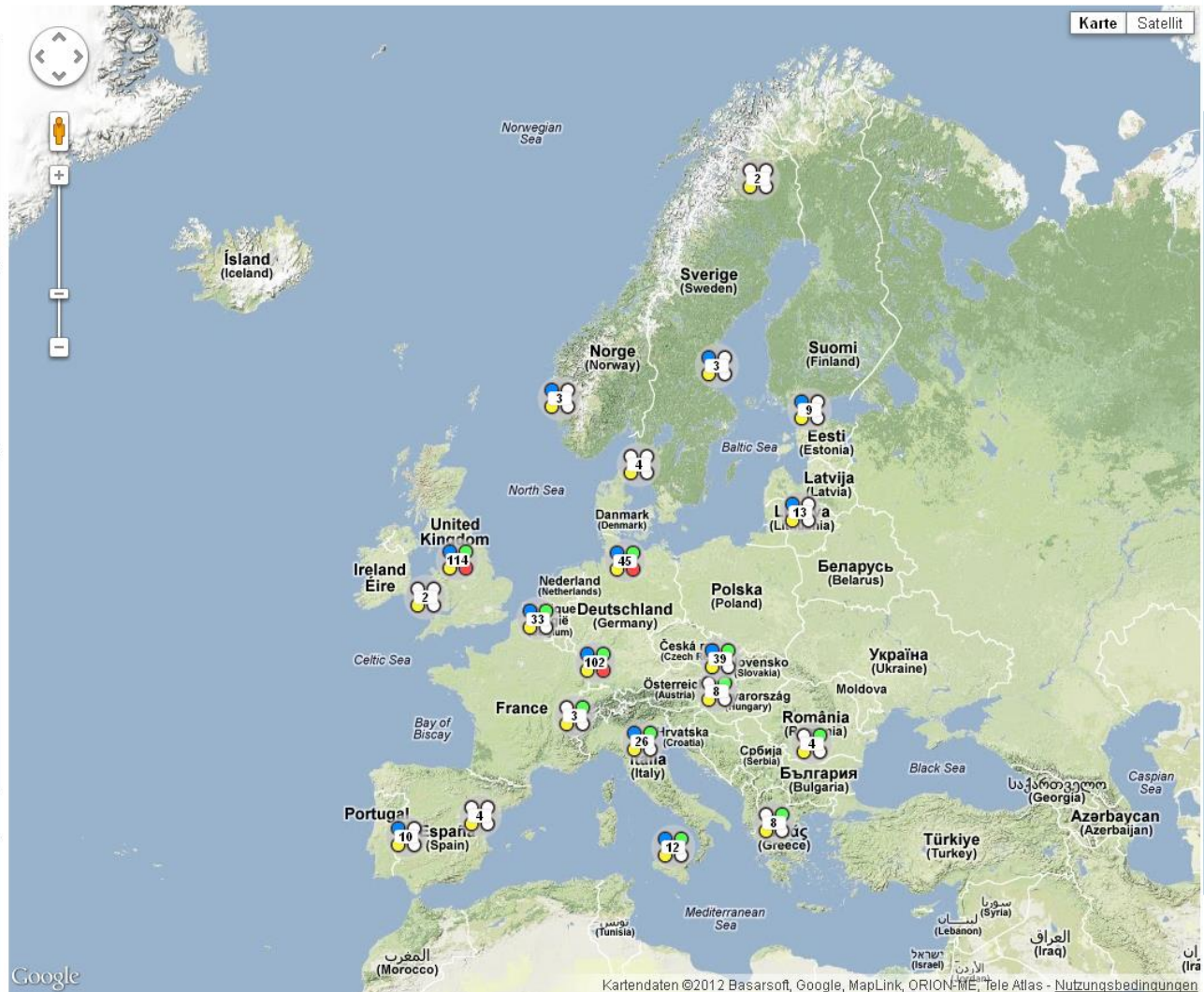
Only organisations from selected **countries** will be shown on the map.

Select all | Deselect all

- Agriculture and nutrition
- Behavioral sciences
- Humanities
- Life sciences
- Medicine
- Physical sciences
- Socio-cultural studies
- Social sciences
- Technology
- Psycholinguistic
- <no Community>

Please select the **communities** to show on the map.

**Credits**  
The interactive map includes information given by Digital Preservation Europe (DPE). We would like to thank the organisation of DPE for supplying their data set.





# Working towards the Virtual Centre of Excellence

Keeping digital resources accessible, understandable and easy to find

[View more](#) →



Search



[Login](#) →

[Register](#) →

## Excellence in Digital Preservation

### About Digital Preservation

Digital encoding has become the dominant way in which we create, shape and exchange information, but this dependence carries many risks.

- As digital technology advances, hardware and software may become outmoded and information may become inaccessible.
- Other users may be unable to understand or use data.
- Access and use restrictions may make it difficult for others to re-use data.
- The ability to identify the location of data may be lost.
- The current custodian of the data may cease to exist.
- The ones we trust to look after the digital holdings may let us down.



Today's research community must assume responsibility for building a robust data and information infrastructure for the future. Digital preservation is too big an issue for individual institutions or even sectors to address on their own. The answer to these challenges calls for coordinated approaches on both national and international level.

### Downloads



**APARSEN brochure on Trust**  
*(hi-res) (see here for more from APARSEN)*

[Download](#) ↓

NEW



**Summary of the studies, thematic publications and recommendations**  
*(see here for more from ODE)*

[Download](#) ↓

[View full list of downloads here](#)

Community



Training



Knowledge Base



Consultancy





# Thanks for your attention

## Any Questions?

Prof. Dr.-Ing. Matthias L. Hemmje  
APARSEN Stream Leader Sustainability  
Member of the Executive Board  
[mhemmje@ftk.de](mailto:mhemmje@ftk.de)

FTK Research Institute for  
Telecommunication and Cooperation e.V.  
Martin-Schmeißer-Weg 4  
D-44227 Dortmund  
Germany