Data Policies with ESA: Long Term Data Preservation and Stewardship supporting Open Science and Open Access at ESA.

Mirko Albani (ESA EOP) : mirko.albani@esa.int
Rosemarie Leone (ESA HIF) : rosemarie.leone@esa.int

THE EUROPEAN SPACE AGENCY MANDATE

- The European Space Agency (ESA), has the mandate to assure the long term preservation, sharing and exploitation of space data and its associated knowledge.
- ESA Convention: particularly, article III (sharing) and annex I article III (inviolability)

OPEN ACCESS AT THE EUROPEAN SPACE AGENCY

- Since early 2017, an Open Access policy for ESA's information and data facilitates broadest use and reuse of the material for the general public, media, the educational sector, partners and anybody else seeking to utilise and build upon it.
- ESA has since been releasing a continuously growing number of contents under the Creative Commons IGO licensing scheme, with the Open Access compliant Creative Commons Attribution-Share Alike 3.0 IGO or, in short, CC BY-SA 3.0 IGO licence as the standard.
- Other free and open content such as data from ESA's Planetary Science Archive (PSA) (http://open.esa.int/esa-planetary-science-archive) or ESA Earth Observation missions such as e.g. Copernicus Sentinel http://open.esa.int/copernicus-sentinel-satellite-data/, Envisat, ERS or Earth Explorers (http://open.esa.int/esa-earth-observation-data) can be found at the relevant links

http://open.esa.int/

EUROPEAN COMMISSION
Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs

Space Policy, Copernicus and Defence
Space Data for Societal Challenges and Growth

Legal notice on the use of Copernicus Sentinel Data and Service Information

The access and use of Copernicus Sentinel Data and Service Information is regulated under EU law.[1] In particular, the law provides that users shall have a free, full and open access to Copernicus Sentinel Data[2] and Service Information without any express or implied warranty, including as regards quality and suitability for any purpose.[3]

ESA UNCLASSIFIED - For Official Use

ESA Data Policy
for ERS, Envisat and Earth Explorer missions

(Simplified version)

- **OBJECTIVES OF THE ESA EARTH OBSERVATION DATA POLICY**

The purpose of this Data Policy is to:

- adapt the existing ESA Earth Observation Data Policies to the "Joint Principles for a

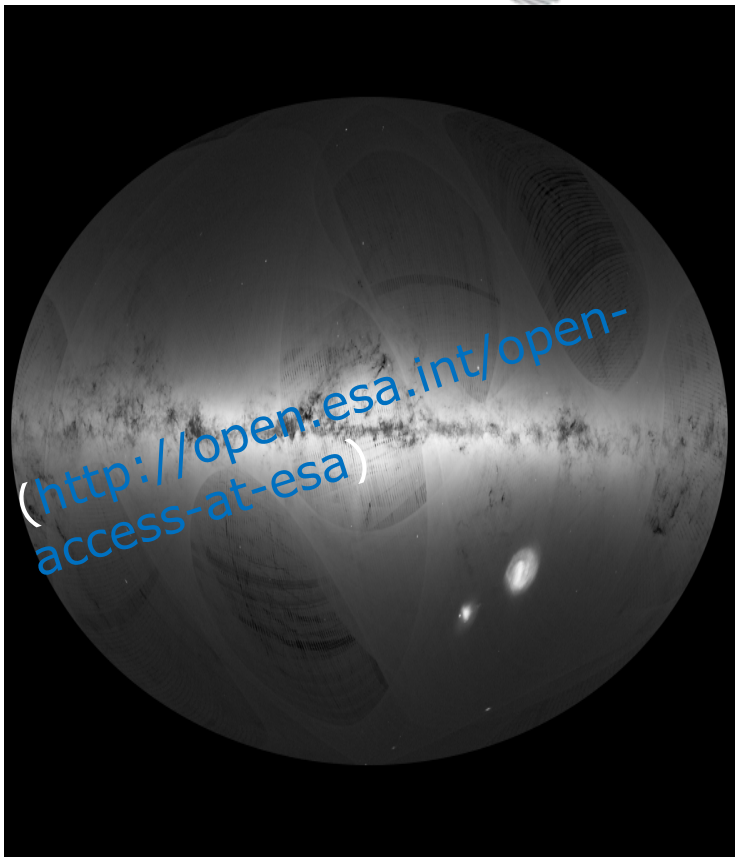
OPEN ACCESS AT ESA: FAQ

Why are not all of ESA's images and other content released in an open and free fashion?

Many of ESA's images, videos and other contents are produced with partners, for example, in science and industry. Priority is given to material that is either fully owned by ESA or for which third-party rights have already been cleared.

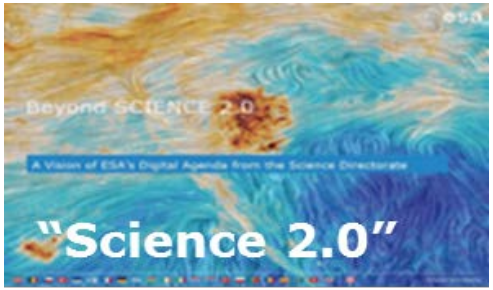
What is Open Access?

Generally speaking, Open Access stands for free and unrestricted online access to research results and findings. Usage rights are often granted via Creative Commons Licences. There is not one, but various statements and definitions of Open Access, such as the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, the Budapest Open Access Initiative or the Bethesda Statement on Open Access Publishing.



OPEN SCIENCE AT ESA

- Open Science is being driven by advances in ICT and ultra-rapid digital technologies, combined with a growing demand to do science for society and in society.
- Initiative for digital innovation and revolution are boosting space mission scientific research, pre-commercial and commercial big data application services and innovation all around the world. Citizen scientists are already being invited to contribute to scientific tasks such as the validation of data, the gathering of in-situ observations or the classification of remote sensing images.

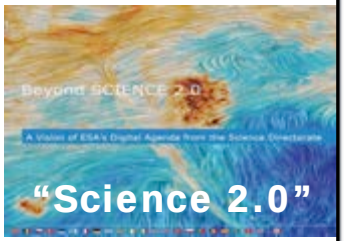


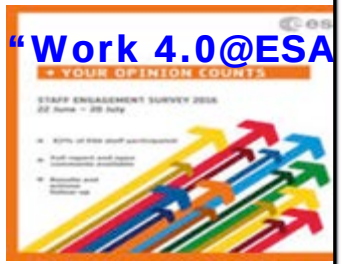
OPEN SCIENCE ENVIRONMENTS EVOLUTION

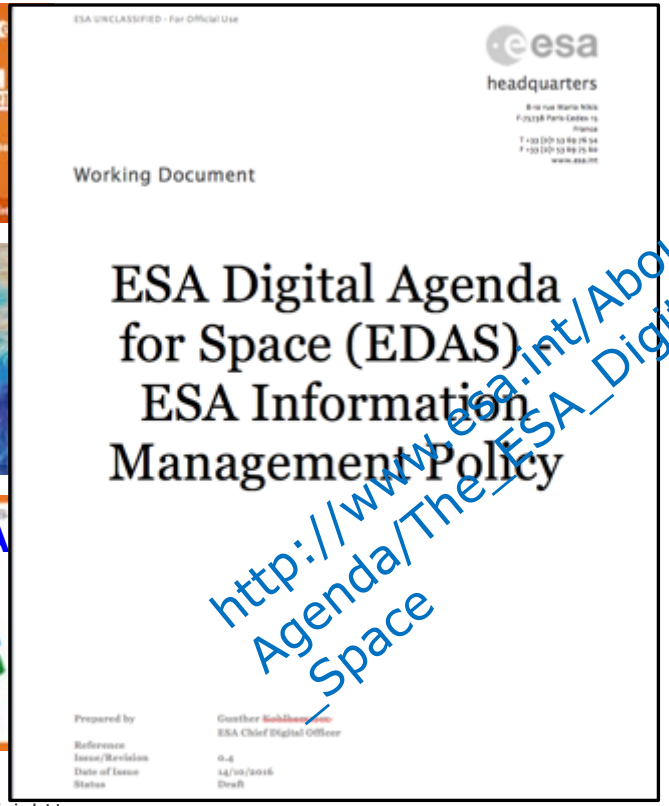
Open data access, Open Publications, Virtual Living Labs, Thematic Collaborative Environments, Data Intensive science, Cloud-based data analytics, Crowdsourcing & Citizens science activities, App camps, Hackhathons, Open Source Toolboxes,

Hackhathons, Open Source Toolboxes, workflows for processing, Advanced training of new class of data scientist, E-learning, Scientific outreach on social media, Data Visualisation

THE VALUE OF DATA AND INFORMATION IS FROM DERIVING INSIGHTS FROM IT







http://www.esa.int/About_Us/Digital_Agenda/The_ESA_Digital_Agenda_for_Space

- Open Science is a key component of ESA’s Digital Agenda for Space.
- ESA’s evolving information management policy increases these opportunities.
- Space 4.0 in Europe will strongly depend on the appropriate management of the huge value in space data & associated knowledge
- ESA shall “Hold ALL of its information and data digitally and online accessible”

EDAS INFORMATION MANAGEMENT POLICY

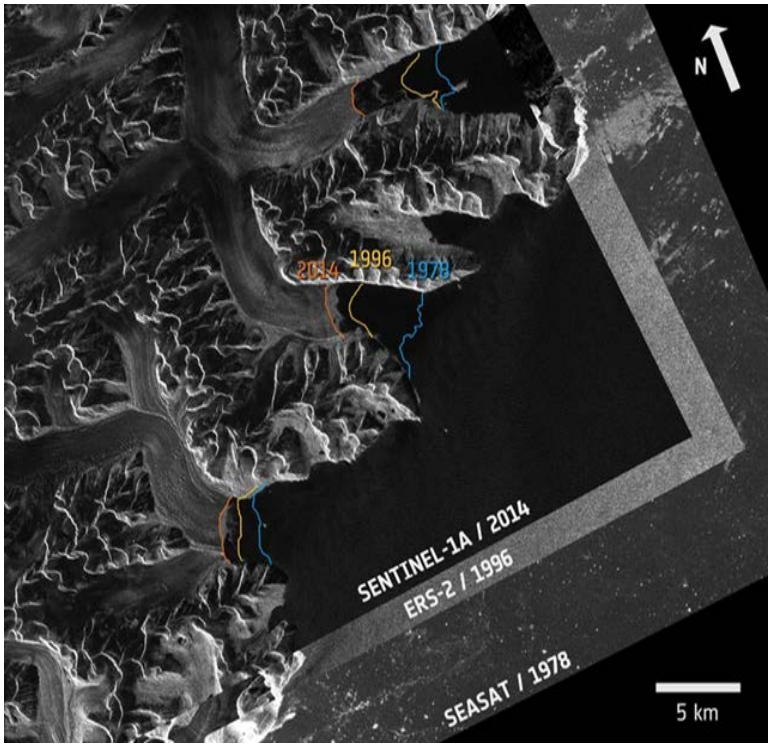
High Level Classification of ESA's Information:

- "Space Data+" as far as processed and managed by ESA
  - Instrument Data from ESA Spacecraft & Third Party Satellites
  - Telemetry Data and auxiliary data about satellite and instruments
  - Campaign Data, Calibration & Validation Data
  - Higher level data derived from the above data types
- Software and IT tools in use for ESA
- ESA's Technical Information
- Technical and Managerial Information and Reports from Industry
- Technical and Scientific Information and Reports from Scientists
- ESA Management Information
- Personal Data

Addressed by LTDP+ for the EO Data and Associated Knowledge, being extended to others (Joint Activities)

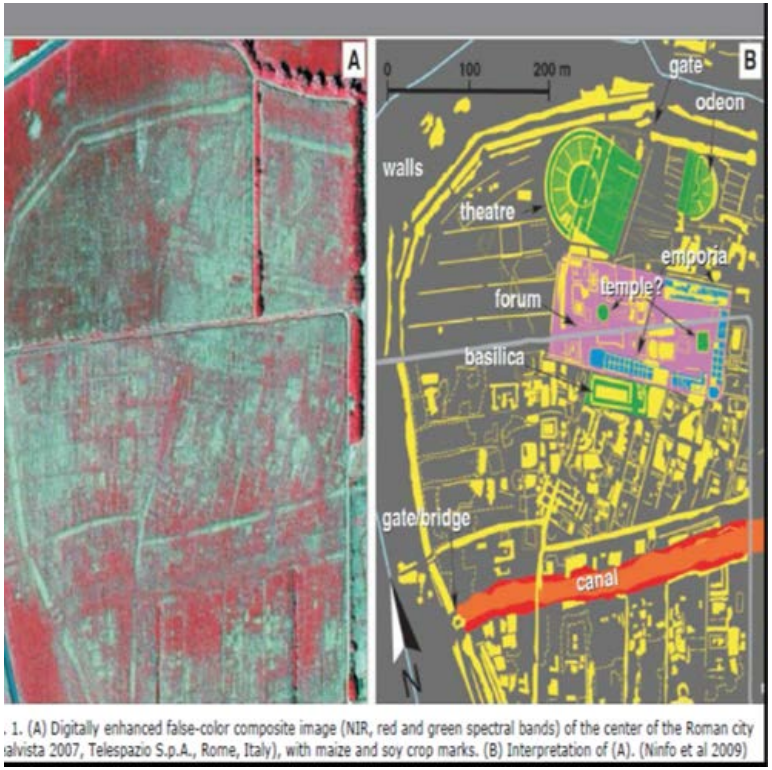
Space 4.0 in Europe will strongly depend on the ***appropriate management of the huge value in space data and information*** held and shared by ESA and its partners across member states with industry and with the science communities.

THE VALUE OF INFORMATION FROM SPACE
EARTH OBSERVATION SPACE DATA HUMANKIND ASSETS



- Monitoring of global change processes is leading to the increasing demand for long-term time series of Earth Observation (EO) data spanning 40 years, or more.
- Earth Observation data are necessary to support international activities such as the United Nations Framework Convention on Climate Change (UNFCCC).

EO SPACE DATA CULTURAL HERITAGE APPLICATION: ARCHEOLOGICAL PROSPECTION FROM SPACE



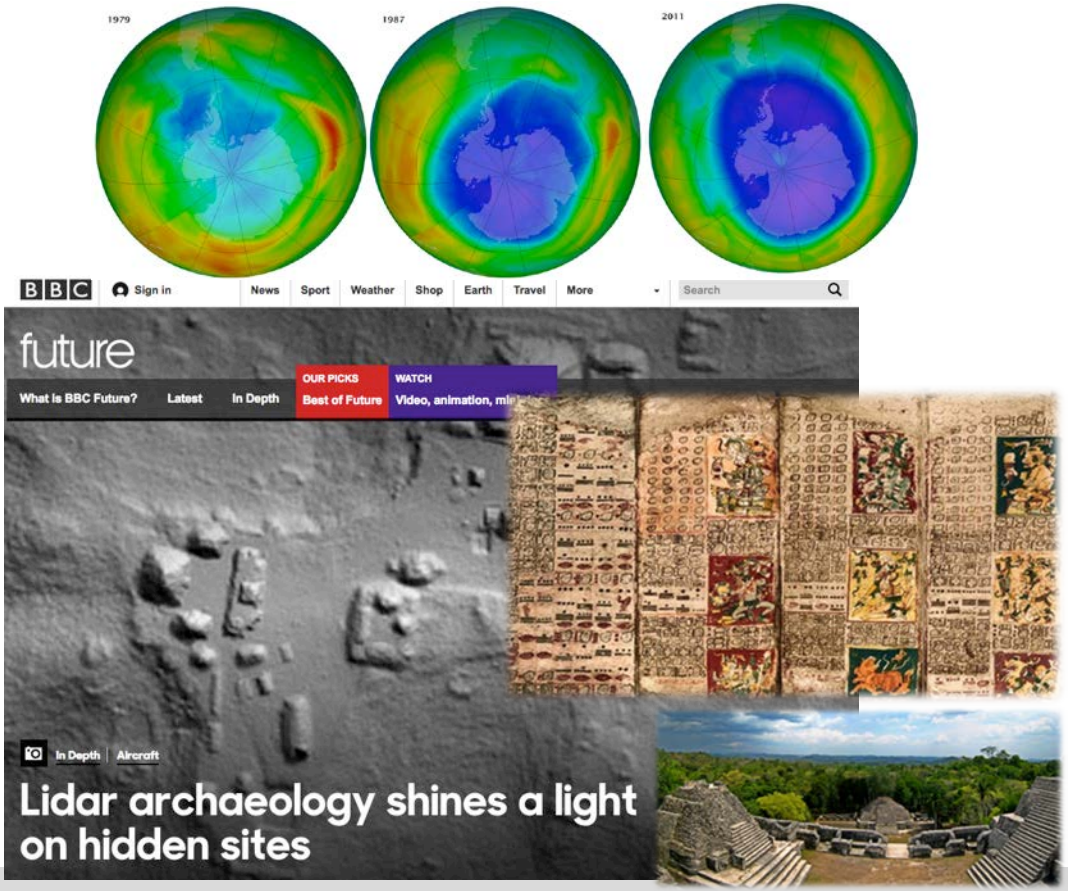
. 1. (A) Digitally enhanced false-color composite image (NIR, red and green spectral bands) of the center of the Roman city alvista 2007, Telespazio S.p.A., Rome, Italy), with maize and soy crop marks. (B) Interpretation of (A). (Ninfo et al 2009)

ESA UNCLASSIFIED – For Official Use

CROSS FERTILIZATION OF HERITAGE DATA

- Heritage (past) data are crucial to monitor evolution in time through comparison with current ‘live’ data.

- New solutions and technologies pave the way for unlimited potential for preserving, discovering, sharing, and exploiting heritage data assets many of which accessed so far by restricted groups of specialized scientists.

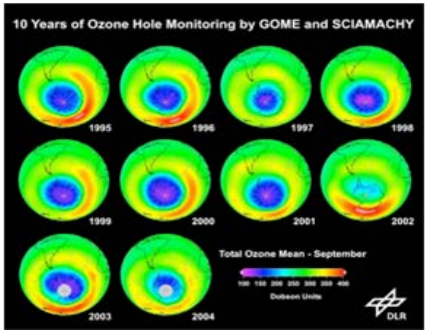


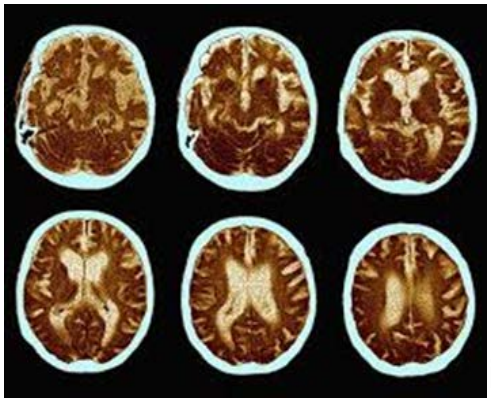
CULTURAL HERITAGE:KNOWLEDGE IN THE DIGITAL ERA

- Humanistic culture and scientific communities are facing a new era where discovery has no geographical and disciplines boundaries.
- Digital culture introduces innovative scenarios for “usage and reusage forever” of knowledge.
- Open data access and linked data create new opportunities across disciplines.

HUMANKIND KNOWLEDGE ASSETS COMMON NEEDS FOR EARTH AND LIFE SCIENCE: PRESERVE, DISCOVER, ACCESS, FEDERATE, EXPLOIT

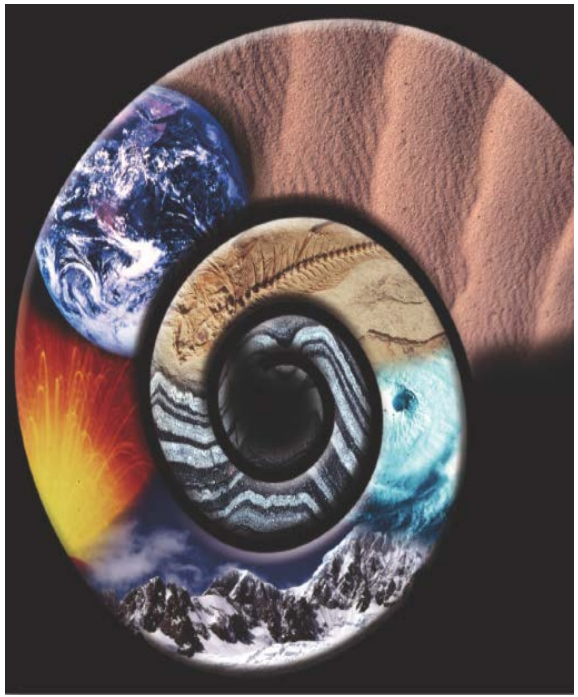
Ozone Depletion (Envisat)





SPACE AND EARTH SCIENCE KEY CHALLENGES

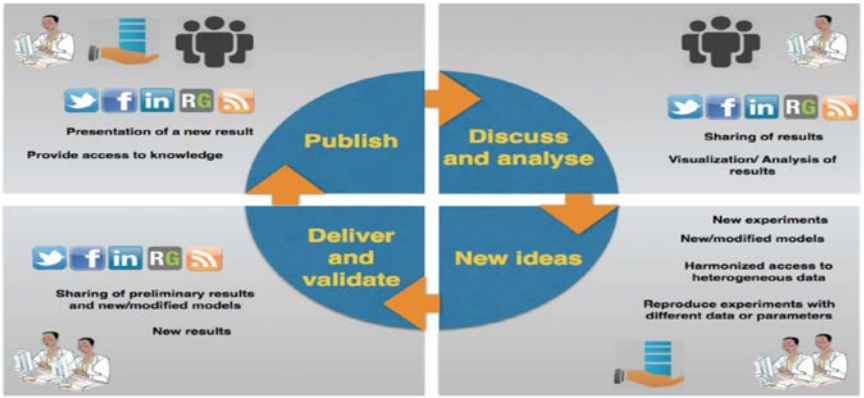
- Interdisciplinary Research
- transparency, reproducibility and accountability of research results
- Global challenges management and environmental policy decision-making and near real time response
- Citizen science to raise public environmental awareness

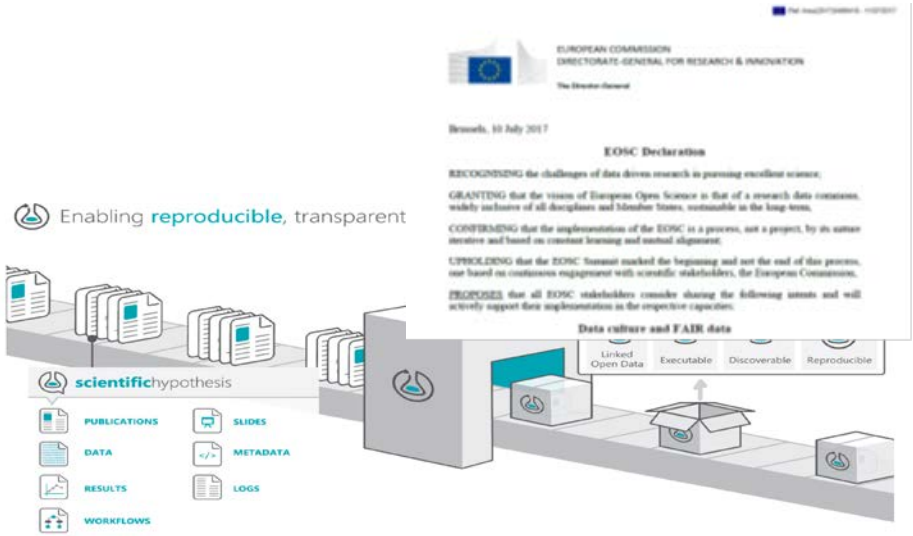


FAIR PRINCIPLE AND RESEARCH LIFE CYCLE

Ensure that scientific research and operational applications results in Earth Science (and wider) are preserved, re-used and shared effectively within and among different communities.

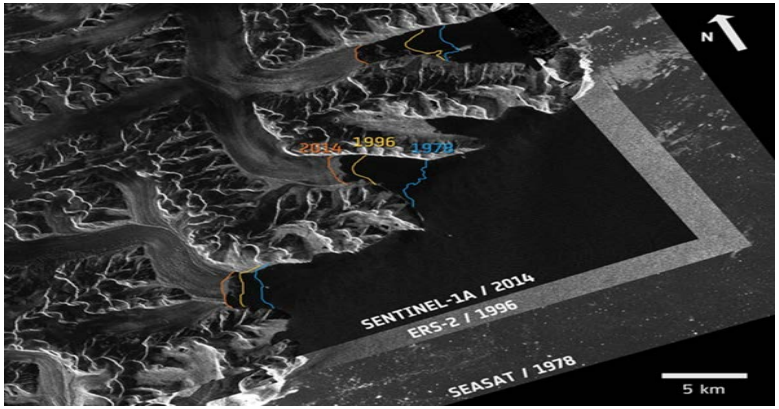
Through Research Lifecycle Management based on the innovative use of Research Object approach and technology

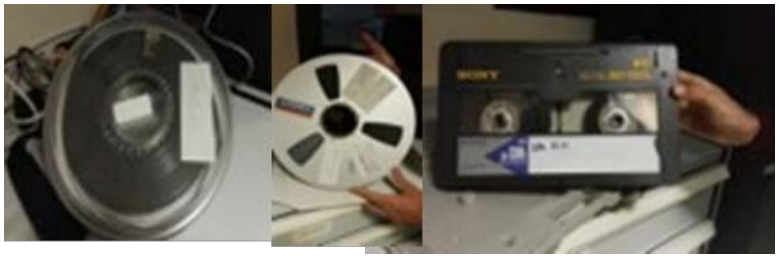




THE ROLE OF THE LTDP+ PROGRAMME

- On the opposite side of disruptive technologies, the ultra-rapid obsolescence of digital technology, fragmentation of resources, proliferation of standards are major threats for data, information and knowledge assets.
- The long term data preservation and stewardship programme has the mandate to assure the preservation, discoverability and exploitation of space data and information assets for future generation as invaluable, unique (non-repeatable) resources from Space, ensuring data accuracy, quality, consistency and security in the long term.
- IT infrastructures solutions and services harmonization, federation and interoperability is playing a crucial role for the valorization of long term data records and knowledge in accordance to FAIR principles.





LEVERAGING ON INTERNATIONAL COOPERATION, R&D AND STATE OF ART TECHNOLOGIES

PV 2015 CONFERENCE

SciDataCon 2016

CODATA

iPRES 2016

ESA UNCLASSIFIED - For Official Use

CCSDS

ECSS

ISO

Projects

scidip-es

everest

Heritage Data Programme (LTDP/LTDP+)

esa

GROUP ON EARTH OBSERVATIONS

GEO GROUP ON EARTH OBSERVATIONS

GEO-XII – 11-12 November 2015 Document 10

Data Management Principles Implementation Guidelines

This Document is submitted to the GEO-XII Plenary for information. Feedback from the GEO community, and in particular, from those who have responsibility for data management is being sought. Principles will be further elaborated and updated next year.

Canada

DLR

gscb Ground Segment Coordination Body

esa

CSA ASC

LTDP Long Term Data Preservation

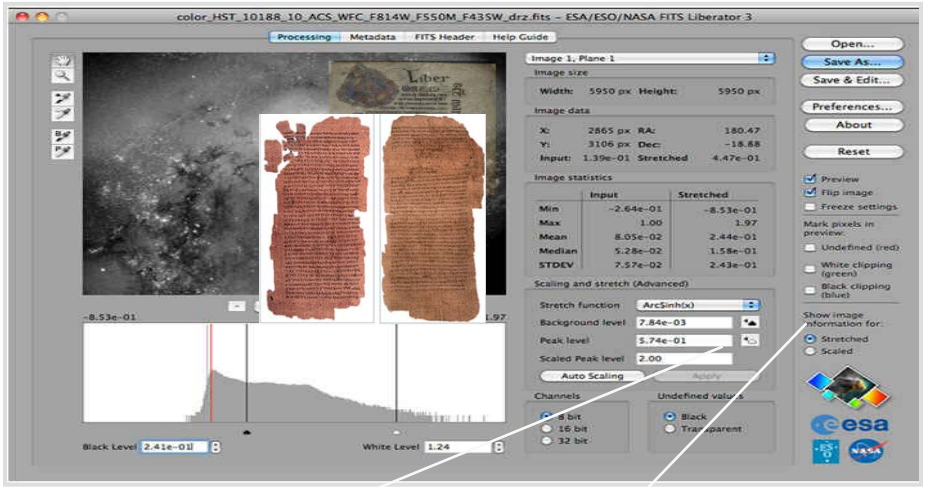
cnes

EUMETSAT

UK SPACE AGENCY

CEOS

FROM ANCIENT MANUSCRIPTS TO SPACE IMAGES

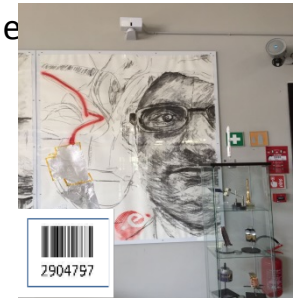


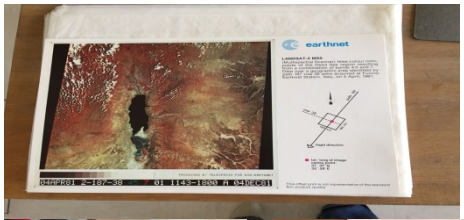
	TIFF	FITS
Architecture	32 bit	64 bit
Max File size	4 GB	Unlimited
3D Visualisation	None	Yes
Format	Proprietary	Open Source
Updated	1998	6 monthly

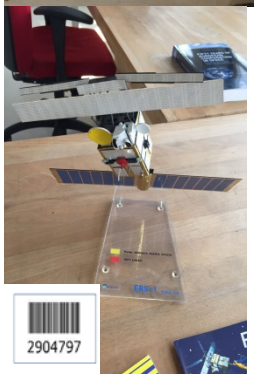
SPACE SCIENCE AND SPACE HISTORY INTERDISCIPLINARY CROSS -FERTILISATION

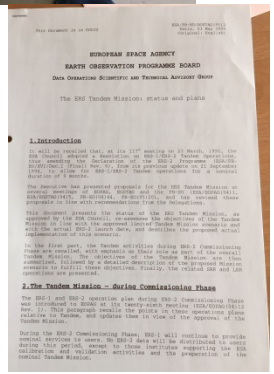
Digitalising and making accessible :

- text and images on paper and digital documents (TIFF, PDF → FITS and PDF A Searchable),
- Photos (JPEG → FITS),
- Video (mp4),
- Audio (mp3),
- 3D CAD (major CAD software proprietary
- data bases (major proprietary formats,
- Software (package)
- physical objects (3D Scan, 3D/2D Photos) e
  - art, pins, events gadgets
  - satellite mock-ups, 1:1, 1:10
  - comet sample









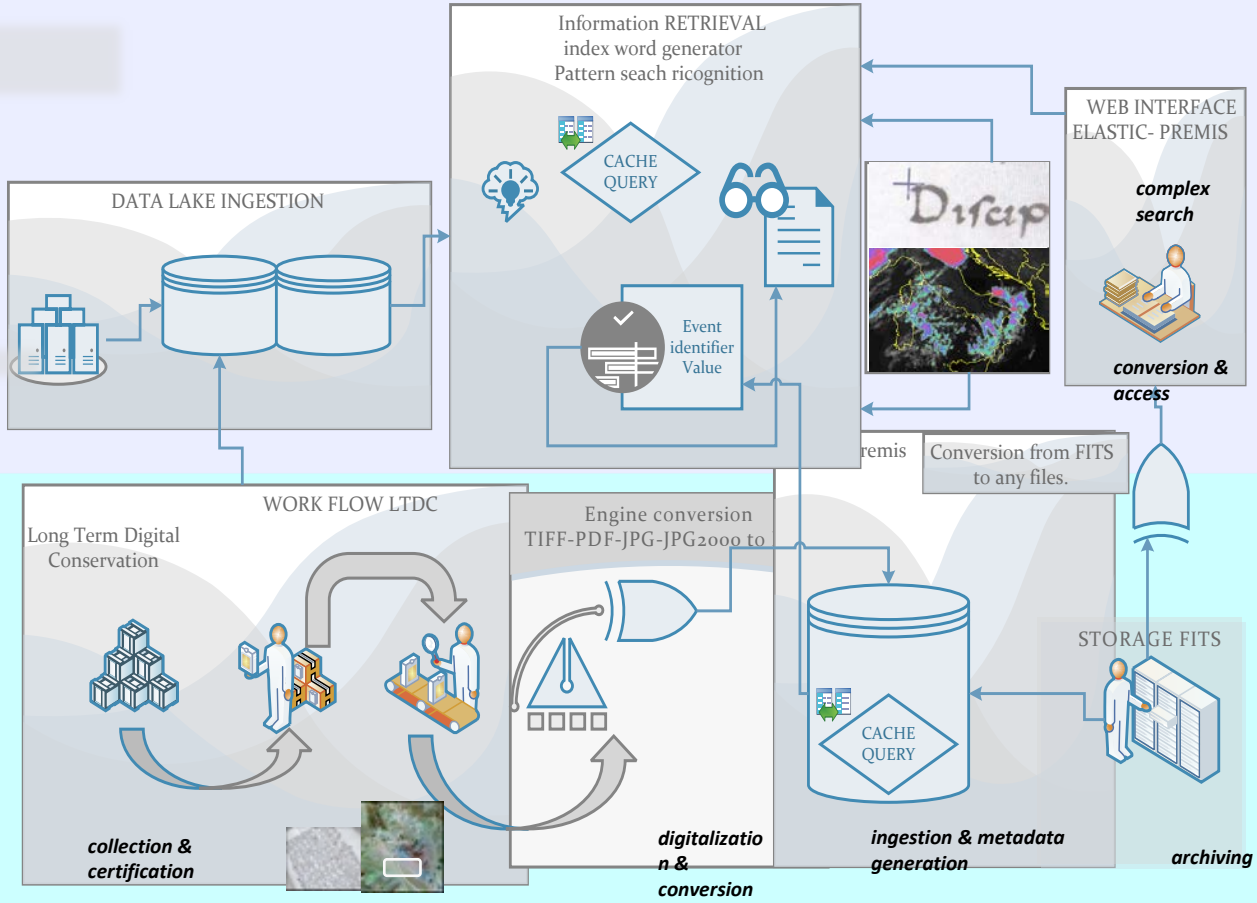
DATA CURATION, DATA STEWARDSHIP AND DATA SCIENCE - NEW PROFILES FOR OPEN SCIENCE

Federation and Interoperability

- Existing catalogues, inventories, data repositories (e.g. RSS, TEPs), EDAS Systems
- Distributed Catalogues Federation
- Exploitation Platform
- Virtual Research Environment

Discoverability, Accessibility, and Exploitation

- OCR and Full Indexing
- Elastic Search
- Discovery Metadata
- Pattern Recognition
- Machine Learning
- Open Data and Linked Data
- Semantic Web
- Big Data Analytics



Open Science

- Knowledge Management
- Information Life Cycle Management
- Record Management
- Certification and Trust
- Security
- Ethics
- Reward and Incentives

Digital Preservation

- Multi-format digitalization technology and scanning systems
- Archiving
- Preservation management
- File Formats (FITS, PDF/A)
- Metadata Management

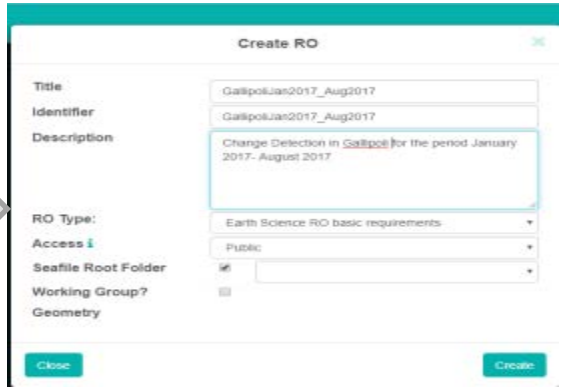
Evaluate how human activities can cause *Posidonia* meadows regression

Level 1: Land Monitoring runs the WPS in the VRE in the Apulian region and creates a RO with the results

WPS running

Results

RO creation



Level 2. Sea Monitoring runs a workflow developed to detect *Posidonia* regression using the Virtual Machine

Posidonia distribution in 1986

Posidonia distribution in 2006

Diff analyses result

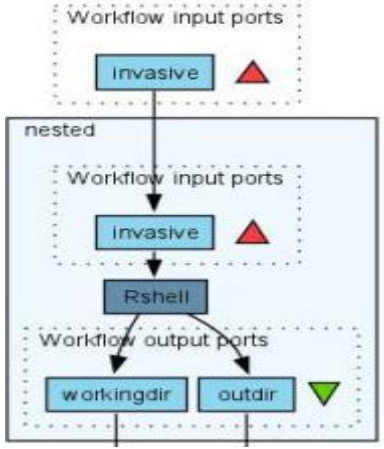
Level 3. Visual comparison between the results from LM and SM analysis.

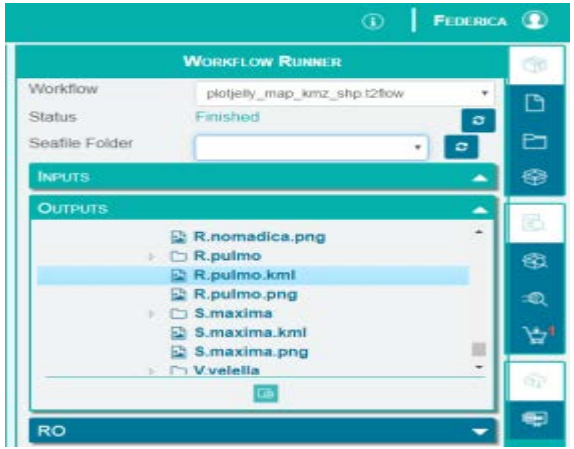
Conclusions: from this first analysis there appears to be a correlation between the human activities detected by LM and the *Posidonia* regression offshore Gallipoli detected by SM

EARTH SCIENCE VRE SERVICES CATALOGUE

On line Data processing – optimization of IT resources and collaboration using a virtual lab

Taverna Workflow EVER-EST Workflow runner EVER-EST Virtual Machine





Implementing FAIR data principles through the adoption of Research Object able to encapsulate, share and reproduce the entire research cycle

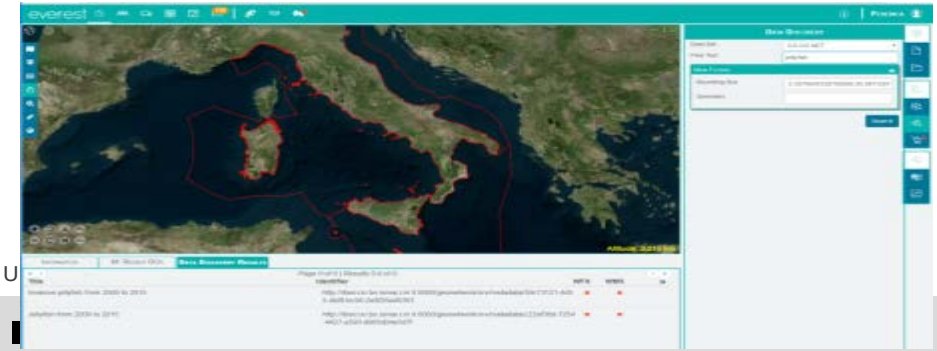
LOOKING FORWARD FOR THE OPEN SCIENTIST IN 2030

Data sharing and Harmonization- reduction of data and knowledge fragmentation.

EVER-EST ROHUB and Collaboration sphere

2030 global marine repository accessible through a Virtual Lab

Easy data Discovery Re-Use and Re- Purposing of open data



2030 data discovery on multiple database

SHAPING THE FUTURE OF CITIZEN SCIENCE – A DAY OF LIFE IN 2030