

The ESA Earth Observation Long Term Data Preservation (LTDP) Programme

Vincenzo Beruti, Mirko Albani ⁽¹⁾,

⁽¹⁾ *ESA-ESRIN*

Via G. Galilei, CP 64,00044 Frascati, Italy

E-Mail: Vincenzo.Beruti@esa.int, Mirko.Albani@esa.int

ABSTRACT

The main objective of the ESA's proposed Long Term Data Preservation (LTDP) initiative is to guarantee the preservation of the data from all EO ESA and Third Parties ESA managed missions on the long term, also ensuring their accessibility and usability, as part of a joint and cooperative approach in Europe aimed at preserving the EO European data from member states' missions. The need to ensure the preservation of the Earth Observation data has been expressed by practically all environmental monitoring programmes and recently again through the Climate Change Initiative.

Following one year consultation with space Agencies and workshops with the owners and holders of other Earth Observation data archives, ESA member states, as part of ESA's mandatory activities, approved a three year initial programme with the aim to establish a full long term data preservation concept and programme beyond 2011. Long term data preservation includes the continuous consolidation and technical evolution of archives, archive management systems and data access systems to guarantee the basic data preservation and proper data accessibility. Beyond and even more importantly, archived data can be used only if also the processing chains, the algorithms and the data access technology are maintained and evolve such that users can actually receive and process the data products always with up to date technology. Archive management includes as well interoperability, standardization issues, archive data security and archive certification processes.

At the beginning of 2008 ESA started the set-up of a cooperation framework with other European space agencies and EO satellite operators to address LTDP issues from a technical point of view and to pursue a stronger coordination at European level. Over the last year a set of European LTDP common guidelines has been established. These initial guidelines will be consolidated and promoted within CEOS and GEO, and constitute the basis for the ESA's EO data preservation approach and for the further cooperation with other European EO data archive holders.

Up to now all long term data preservation activities at ESA, including investments and operational support, have been considered part of the missions' budget and have been executed as background task. This approach is no longer adequate for the future due to the expected increase of magnitude of the archives dimension in the next years and to the enormous demand for archived data access and exploitation.

This paper will present the ESA LTDP programme and the main activities planned for the next years focusing on the set-up of the European LTDP Framework.

INTRODUCTION

The need for accessing historical Earth Observation (EO) data series strongly increased in the last ten years, due to an increased awareness from the users and the need to satisfy requirements mainly for long term science and environmental monitoring applications. This trend will increase even more in the future in particular for the growing interest on global change monitoring which is driving users to request time-series of data spanning 20 years and more, and due also to the need to support the United Nations Framework Convention on Climate Change (UNFCCC). Content of EO data archives is extending from a few years to decades and therefore their value as a scientific time-series is continuously increasing. Hence there is a strong need to preserve the EO space data without time

constraints and to keep them accessible and exploitable, facing and solving the problem of technology obsolescence. The preservation of EO space data can be also considered as a responsibility of the Space Agencies or data owners as they constitute a humankind asset. In the next decade, the wealth of information currently locked inside the global data archives must be fully exploited and re-analyzed on a global scale (Figure 1 and Figure 2 show examples of information extraction from long-term data series). This challenge relies on full accessible and exploitable archives.

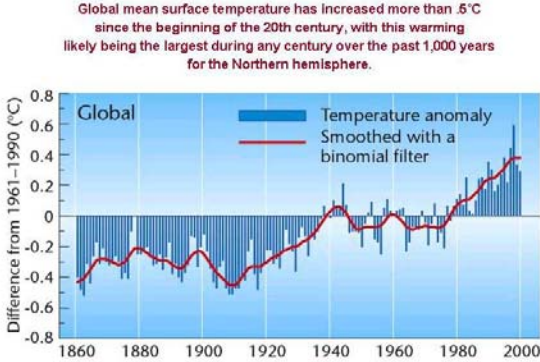


Figure 1: Surface temperature increase

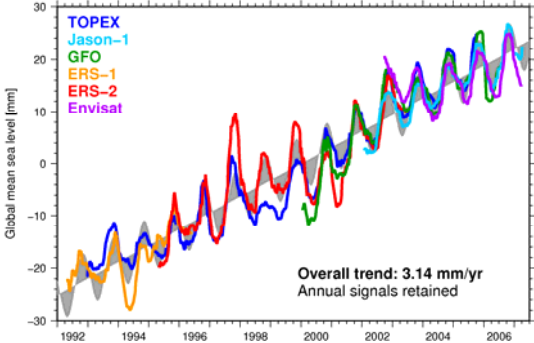


Figure 2: Global sea level raise (courtesy of Remko Scharroo)

The application areas heavily benefiting from the EO long term data archiving exploitation are wide and span from European and Global Environment (e.g. Forest Monitoring, Soil Monitoring, Urban Development, Air Quality Monitoring, Ecosystems Monitoring and management for protection of terrestrial, coastal and marine resources) to Management of energy resources (e.g. solar, etc., Development and Humanitarian Aid Health including the understanding of environmental factors affecting human health and well-being, Food security including sustainable agriculture and combating desertification, Water resource management through better understanding of the water cycle, Civil Protection and disasters monitoring (e.g. Flood Prediction and Mitigation, Landslides, Subsidence, Volcanoes Monitoring), including as well Global Climate Change (e.g. Systematic Climate Observations, Drought Monitoring, Monitoring of the Atmosphere, etc.), Global Security and Sustainable Development.

European* EO archives challenges

The large amount of new Earth Observation missions upcoming in the next years (Figure 3) and the enormously enhanced data sensing and transmission to ground capabilities will moreover lead to a major increase of EO space data volumes. This fact, together with the increased demands from the user community, marks a challenge for Earth Observation satellite operators, Space Agencies and EO space data providers regarding coherent data preservation and optimum availability and accessibility of the different data products.

* In the scope of this document Europe is intended including Canada



Figure 3: Current and future European and Canadian EO missions (excerpt)

Error! Reference source not found. shows in broad values the current dimension of the ESA archive per mission and the expected volume for the next years. The data holding growth boosted since the launch of ERS and in particular of the Envisat mission and will be increasing even further with the contribution of additional Third Party Missions (TPM) and with the GMES program. The plans of new ESA missions indicate 5-10 times more data to be archived in next 10-15 years. Similar trend is also monitored at all National archives.

Traditionally in Europe, there has been poor cooperation in this field with no common approach for long term preservation and access to EO space data even if cooperation and sharing are key aspects to be pursued for the benefit of the user community. Preserving today's science records is fundamental but single organizations have difficulties to afford data preservation in the long term.

ESA has recognized that a cooperative and harmonized collective approach on Long Term Data Preservation (LTDP) in Europe is absolutely needed in order to coordinate and optimize European efforts in the LTDP field and to ultimately result in the preservation of the complete European EO space data set for the benefit of all European countries and users and with a reduction of overall costs.

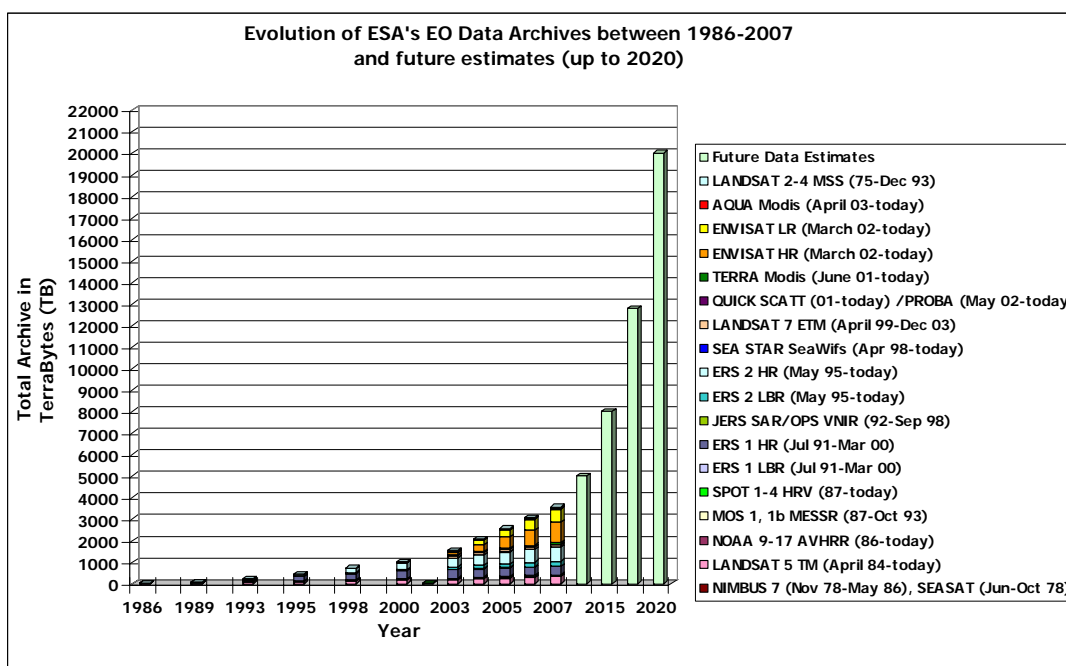


Figure 4: ESA expected volume of archives

ESA Executive has presented to Member States in 2006 an ESA EO long term data preservation strategy, aiming at two major achievements:

- Pursue a European long term data preservation Framework involving possibly all EO data operators/ owners in Europe aiming at a common and coordinated approach.
- Ensure a long perspective to the ESA EO data preservation, including an enhancement of the user access, of the archives exploitation and of the related services.

EUROPEAN EO LTDP FRAMEWORK KEY PRINCIPLES

As part of the strategy implementation, ESA initiated a coordination action to share among all the European (and Canadian) stakeholders a common approach for targeting the preservation of all European (including Canada) EO space data for an unlimited time-span ensuring and facilitating their accessibility and usability through the implementation of a cooperative and harmonized collective approach among the EO space data owners.

The main goals of the European EO Long Term Data Preservation Framework are to:

1. Preserve the European, and Canadian, EO space data sets for an unlimited time-span.
2. Ensure and facilitate the accessibility and usability of the preserved data sets respecting the individual entities applicable data policies.
3. Through the adoption of a cooperative and harmonized collective approach among the data owners (LTDP Framework) based on the application of European LTDP Common Guidelines and sustained through cooperative (multi-source) long term funding schemes.
4. Ensure, to the maximum extent, the coherency with the preservation of other non-space based environmental data and international policies.

The European LTDP Framework is open to all possible members and is to be intended as a collaborative framework consisting of distributed and heterogeneous components and entities cooperating in several areas to reach a harmonized preservation of the European EO space data set. The framework is based on the contribution of European EO space data owners through their ideas and possibly their infrastructure in accordance to commonly agreed LTDP Guidelines and should follow a progressive implementation based on a stepwise approach (short, mid, long-term activities) for a progressive development and implementation of technology, methodology, standardization, operational solutions and data exploitation methodologies as key aspects for the set-up of the framework.

The adoption of standards (e.g. for data access interfaces and formats, procedures, etc..) and common technical solutions can also allow to significantly reduce preservation costs.

The European LTDP Framework should be sustained through a cooperative programmatic and long term funding framework based on multilateral cooperation with multiple funding sources from at least the European EO space data owners.

A European framework could also trigger the availability in the long term of additional permanent funding sources (e.g. European Commission) and can increase the possibility for any European (including Canada) EO space data owner to preserve missions data beyond their funding schemes into the cooperative and distributed framework.

THE LTDP COMMON GUIDELINES

A Long Term Data Preservation Working Group with representatives from ASI, CNES, CSA, DLR and ESA was formed at the end of 2007 within the Ground Segment Coordination Body (GSCB), with the goal to define and promote, with the involvement of all the European EO space data and archive owners, a set of European LTDP Common Guidelines and also to increase awareness on LTDP. The insurance and facilitation of access, respecting the individual entities applicable data policies, and exploitation of

the archived data are also part of the guidelines whilst data access policies are on the other hand not addressed.

During the 1st Earth Observation Long Term Data Preservation workshop in May 2008, the draft guidelines and the framework were presented and debated by all European and Canadian EO data owners, data providers and archive holders. A consolidated LTDP Common Guidelines document has been produced addressing the following nine main themes and defining for each the “Guiding Principle” and the “Key Guidelines”:

- Preserved data set composition
- Archives maintenance and data integrity
- Archives operations
- Data security
- Data ingestion
- Data access and interoperability
- Data exploitation and re-processing
- Standardization
- Data Purging/Appraisal

An extensive public review process of the guidelines document is now open with the goal to collect additional comments and feed-backs from the EO space data owners and archive holders and to issue a commonly agreed version of the document by the end of 2009. The LTDP guidelines constitute a basic reference for the long term preservation of EO space data. The guidelines should be adopted for old missions with a step-wise approach and straightforward for new missions and projects. The application of the identified guidelines is strongly recommended also following a step-wise approach starting eventually with a partial adherence.

COOPERATION ACTIVITIES

The initial areas of cooperation related to LTDP framework have been identified in:

1. **Policies** for the consolidation and issue of the European LTDP Common Guidelines, and adherence to them, and for the definition and application of a purge alert / appraisal procedure to EO space data.
2. **Technology, methodology and developments.** The aim is to jointly evolve archive and data access technology through studies/pilots sharing the acquired know-how and infrastructure and to share knowledge/experience exchanging information to favour technical cooperation (cross participation into reviews, share of solutions, products, developments, etc).
3. **Standardisation** activities in close link with international bodies (e.g. CCSDS, CEOS, OGC, INSPIRE, EU initiatives, GEO).
4. **Operational solutions** setting-up the principles for a common European distributed archiving concept aiming at the creation of an interoperable network of archive centres possibly reusing infrastructure of the different entities (as a single archive) in the long term.
5. **Data exploitation** through the definition and implementation of joint EO historical data exploitation programmes.

ROADMAP VISION

The roadmap for the set-up of the European LTDP Framework can be articulated in three main phases to follow the best practices defined in the LTDP Common Guidelines and to progressively cooperate in the concerned areas.

1) The objective of the first phase is to reinforce the LTDP approach at each entity, to start the cooperation among agencies and EO space data owners (e.g. in methodology, standardization, sharing of

information, etc...) and to define the future European organisation of LTDP in a very long term perspective.

2) In the mid term perspective (second phase) cooperation should be strengthened through the implementation of common activities among European EO space data owners with the goal to achieve an interoperable network of archives (e.g. share of solutions and systems, coordination of common technology developments, adoption of standards...), but also improving operational services according to user needs (i.e. Climate changes monitoring operational systems, etc.). During this phase, additional entities in Europe will be attracted and become part of the European LTDP Framework with common technical views and solutions and standardized services.

3) In the long term perspective cooperation should be further extended through the sharing of infrastructure (e.g. common and shared access points, interoperable and transparent data access and infrastructure), allowing to have a unique network of data, shared resources for data reprocessing and products generation and a common and harmonised security levels and layers..

ESA EO LTDP PROGRAMME

It is worth to remind that the prime objective of the ESA's proposed Long Term Data Preservation strategy is to guarantee the preservation of the data from all EO ESA and Third Parties ESA managed missions on the long term, also ensuring their accessibility and usability.

Up to now all long term data preservation activities at ESA, including investments and operational support, have been considered part of the missions' budget and have been executed as background task. This approach is no longer adequate for the future due to the expected increase of magnitude of the archives dimension in the next years and to the enormous demand for archived data access and exploitation.

For this purpose, ESA member states, as part of ESA's mandatory activities, approved in November 2008 a three year initial LTDP programme with the aim to establish a full long term data preservation concept and programme beyond 2011. ESA will implement the high priority activities in the next three years focusing on data preservation and enhancement of data access. In the 2009-2011 period and with the limited funds, four main tasks will be addressed: Analysis and studies, Implementation (ESA archives), European LTDP Framework Coordination, European LTDP Programme Preparation. The LTDP work-plan for the initial period 2009-2011 is shown in Figure 5.

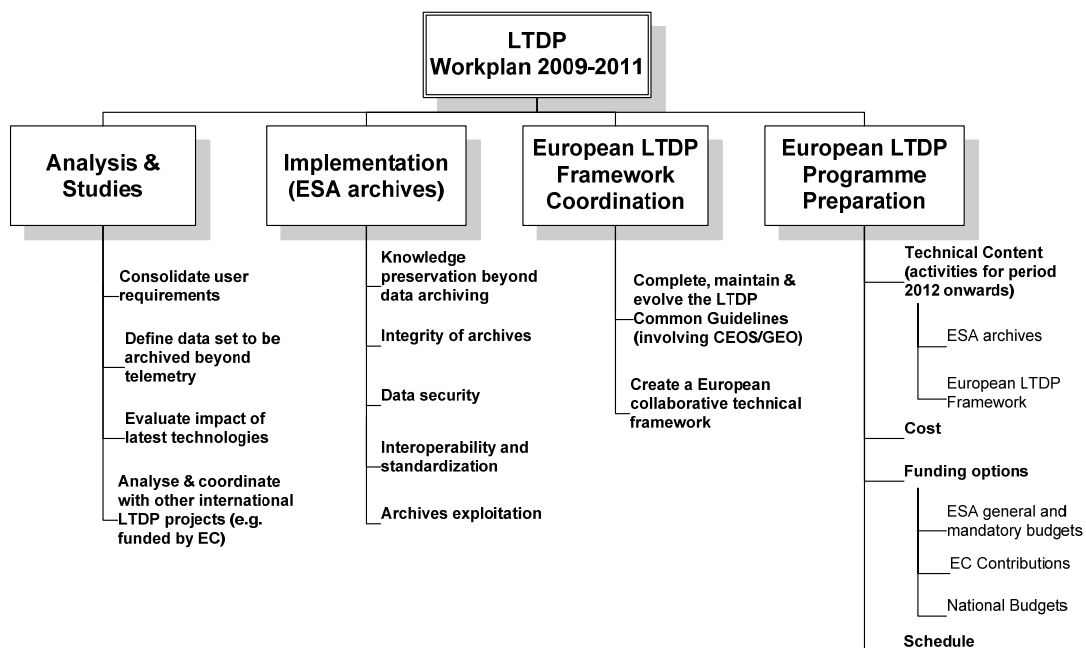


Figure 5: LTDP Work-plan 2009-2011

Analysis and Studies

The implementation plan is structured in four main activities. The first one is aimed at consolidating user requirements, including those from the Climate Change Initiative and GMES, and at reviewing the LTDP approach and concept for EO data.

Currently only raw data and/or level 0 data of all operated satellites are archived at ESA facilities, plus the necessary auxiliary data for processing. Additional data should be archived (e.g. calibration data, algorithms, documentation, etc.) in order to guarantee knowledge preservation and the possibility for later re-processing and re-calibration. A second activity aims at defining, through a study in cooperation with other European EO data archive owners and with the involvement of the scientific user community, the composition of the “data set” to be archived in order to guarantee knowledge preservation.

Recently and still in the prototyping phase, several options have been evaluated for distributed archiving, product online archiving and distribution, temporary net-based archives closer to the main data demand and exploiting a large number of distributed and shared processing capacities. As a third element, and together with the other European EO data archive owners, these options will be further consolidated and tested and the technologies expected for the next years will be structured into a joint technical architecture plan.

The fourth activity aims at the analysis, in cooperation with other ongoing international initiatives (e.g. projects funded by the EC), to provide technological developments & methodologies/procedures in the LTDP area or that could have impacts on EO LTDP. In this way, technological efforts in data preservation are re-used and applied to the maximum extent, also to EO data.

Ongoing activities

An LTDP Programme Management support contract is closed to be formalized with Industry aiming at several purposes, including the support for the preparation of additional studies and industrial implementation in the field of Archive security, archive certification, networking, etc., including support to the preparation of the programmatic approach for the period above

An open tender has been recently issued for a technological study (LAST) covering, as main activities, the collection and analysis of the requirements from ESA and EO partners, the definition of the LAST fields of study, an archiving Technology Survey: extensive analysis, some recommendations of technologies in short, mid and long term and bench marking / prototyping activities on the recommended technologies

Implementation (ESA archives)

In parallel to the analysis and study activities, and exploiting the out coming results, a set of basic and high priority LTDP activities will be implemented in the ESA archives in order to ensure and secure the preservation and integrity of the ESA archived data and to enhance their accessibility and usability with recent technology. Planned activities are based on the application the LTDP Common Guidelines.

Ongoing activities

In the context of this programme, ESA is in the process of upgrade the Tape libraries at most of the EO archive centers , replacing the STK Powerhorn Tape Libraries based on 9940 B technology with the new SUN SL3000 based on T10000 B based technology.

Other initiatives are in progress or will be soon initiated in order to recover data from historical missions, to enhance the on line archive data access, to implement faster data reprocessing, to explore adoption of archive virtualization for data distribution, to adopt new metadata search methodologies including data mining techniques.

European LTDP Framework Coordination

Ongoing activities of the LTDP Working Group will continue, with the involvement of all European data owners and archive holders, to coordinate and favor the set-up of the European LTDP Framework:

- LTDP guidelines will be consolidated and promoted within CEOS or GEO in order to be agreed and recognized at international level.
- Technical cooperation activities are planned (e.g.. definition of LTDP users requirements and composition of the data set to be archived to guarantee knowledge preservation, standardization issues on format for archive, data exchange, products archives, etc.).
- Liaison with international bodies and initiatives will be established (CEOS, CCSDS, GEO, EU projects). The international context of the European LTDP Framework is shown in Figure 6.

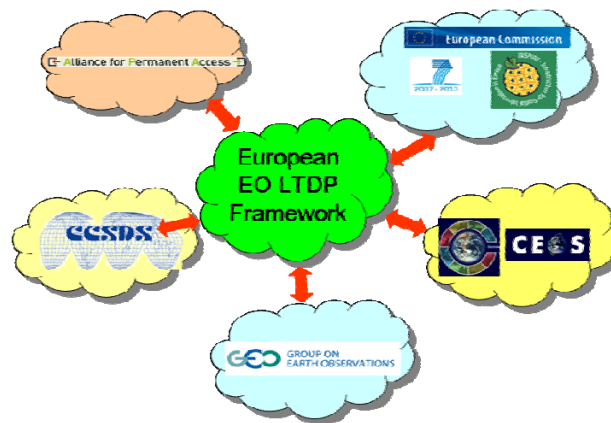


Figure 6: International context

European LTDP Programme Preparation

An LTDP programme proposal for the period beyond 2011 will be prepared and presented. To this end the medium and long term LTDP activities will be defined and cost and programmatic aspects including different funding options will be evaluated.

CONCLUSIONS

The ESA LTDP Programme has been initiated and is progressing, targeting very clear objectives:

- To provide an adequate response to the user needs for a long term EO data preservation and access, ensuring data integrity and access with a very long perspective.
- To define and implement a European framework that involves all concerned European actors in order to face together the future challenges, including funding issues.
- To enhance the ESA infrastructure in order to adequately respond on a solid ground to all current and future archive technical requirements constantly exploiting the technological evolution.
- To promote the initiative at international level for an increased awareness about EO long term data preservation aiming at a coordinated world wide approach.

The ESA LTDP programme for the period 2009-2011 will allow the Agency to continue the coordination of European efforts and initiatives in the LTDP field and to start the application of the LTDP Common Guidelines to its own missions implementing the high priority activities focussing on data preservation and enhancement of data access.