

Digital Preservation for Long-Term Environmental Monitoring

ESA-ESRIN, the European Space Agency Centre for Earth Observation (EO), is the largest European EO data provider and operates as the reference European centre for EO payload data exploitation. EO data provide global coverage of the Earth across both space and time showing the world through a wide-enough frame so that large-scale phenomena can be observed locally with great accuracy and gathering data from sites not easily accessible for ground-based data acquisition facilities.

EO data acquired from space have become powerful scientific tools to enable better understanding and management of the Earth and its resources. More specifically, large international initiatives such as GMES and GEO, supported by the European Commission, all European National Institutions and by many other international organisations, focus on coordinating international efforts to environmental monitoring, i.e. to provide political and technical solutions to global issues, such as climate change, global environment monitoring, management of natural resources (e.g. air quality, marine environment, forest ecosystem) and humanitarian response.

A long-term vision

Requirements for accessing historical archives have been strongly increased over the last years and the trend is likely to increase for the long term science and the long term environmental monitoring. Since the time-span of EO data archives extends from a few years to decades, their value as scientific time-series increases considerably regarding the topic of global change and likely in the near future it will be necessary to re-analyse on a global scale the information currently locked inside large thematic data archives.

A concrete example of benefit of applying a long-term data preservation policy concerns the monitoring of the current state of the oceans, for which global data and model synthesis are needed. Improvements in scientific assessments may come in areas like sea level change, sea surface temperature, variations in global oceanic chlorophyll distribution and related primary production as key variables in carbon cycle studies or radar altimeter significant wave height climatic trend.

Concerning the users of these data, at present several thousands Earth scientists, researchers, environmentalists and climatologists have online access to EO missions' metadata (millions of references), data (in the range of 3 to 5 PB) and derived information. Their main objectives are to study the global climate change (Scientific Community and Principal Investigators users, e.g. via the ESA's dedicated initiative) and to check the status of the instrument and the quality of data (ESA internal scientists/technologists).

In one sentence, it is clear why so many users are calling for the need to preserve the EO data and keep them accessible without time constrains. On the other hand, major constraints and obstacles remain, as follows:

- Current EO data archiving and preservation are mostly limited to the satellite operation lifetime;
- The data volumes are increasing dramatically;
- More and more EO missions data can be called "historical" and more and more operators are faced with the decision if and how to preserve these data;
- EO data preservation and access policies, if existing at all, are different for each EO mission and each operator or Agency.

Moving towards data preservation, approach and strategy

To respond to the urgent need for a coordinated and coherent approach for the long term preservation of the existing European EO space data, ESA in cooperation with the European (including Canada) EO space data owners has established a European-wide dedicated programme and actions for permanent access to the scientific information (including data and publications) by the definition of:

- an overall strategy for the long term preservation of EO data;
- a common policy for the harmonization of European archives for ESA and Member States national EO data holders (including Canada).

It is worth note to recall on the main goals of such an initiative, which are:

- to preserve the European (including Canada) EO data set for an unlimited time span;
- to ensure and facilitate the accessibility and usability of the preserved European EO data set;
- to possibly ensure the coherency with the preservation of other non-space based environmental data and international policies.

ESA-ESRIN is also participating to a number of international projects partially funded by the European Commission and concerned with technology development and integration in the areas of long term data preservation and distributed data processing and archiving. More specifically, relevant contributions to the long term preservation of EO data come from the projects:

- CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval), that will research, implement and disseminate innovative solutions for digital preservation based on the OAIS reference model and
 - PARSE.Insight (Insight into issues of Permanent Access to the Record of Science in Europe), that aims to develop roadmap and recommendations for developing an e-infrastructure to maintain the long-term accessibility and usability of scientific digital information in Europe.
- Both these projects build on an insight and analysis of the current status of EO datasets for what concerns their current representation, maintenance and archiving. Finally (and coherently with its effort in PARSE.Insight) ESA is also one of the founding members of the Alliance for Permanent Access to the Records of Science whose aim is “to develop a shared vision and framework for a sustainable organisational infrastructure for permanent access to scientific information”.

Further information and resources:

<http://www.esa.int/>

<http://www.gmes.info/>

<http://www.earthobservations.org/>

<http://copi.esa.int>

<http://www.casparpreserves.eu/>

<http://www.parse-insight.eu/>

<http://www.alliancepermanentaccess.eu/>