ARCHIVE MIGRATION FOR EARTH OBSERVATION DATA HANDLED BY ESA

Vincenzo BERUTI, Eric DOYLE, Antonio FLATI, Bruno GRECO, Gian Maria PINNA
European Space Agency – ESRIN, Earth Observation Ground Segment Department
Via G. Galilei, 00044 Frascati (Rome), Italy
Phone: +39 (6) 941801
e-mail: vincenzo.beruti@esa.int

Abstract

Data preservation on physical media is bound to media technology, which evolves over time. Preservation of data on the same media where it was originally acquired must face the dual problem of the lifetime of the medium itself and of the long-term availability of the technology for reading it. In addition, the evolving market conditions make technology age and eventually turn anti-economic, or not competitive versus emerging ones, and that point alone in some cases may render the technology change very attractive. On the other hand, the migration of large amounts of data from an old medium to a new one requires time and poses a heavy burden on operation teams in terms of data copying and equipment maintenance, both required to support the multi-technology environment. This latter issue is even more complex if the migration has to take place in parallel to the handling of newly acquired data. Normally data archives are composed of two elements: data belonging to non operational missions, which are consequently stable in dimension, and data belonging to missions or programs still in operation, therefore adding on a routine basis new fresh data to the archive.
The key to successful operations is a constant attention to technological issues, careful benchmarking, financial analyses that take the future prospects into consideration and finally careful operation planning.
This paper provides an analysis of these archive migration related issues and describes the approach adopted, including the proposal of a set of guidelines and suggestions derived from the experience gained with the Earth Observation detached data archives handled by the European Space Agency, through the ESRIN center in Frascati (Italy).

Introduction

The European Space Agency currently holds large amounts of historical and operational data from various Earth Observation (EO) proprietary and third party missions. The volume and data rate of the various datasets span the whole spectrum of the EO missions. The ESRIN establishment located in Frascati, Italy is in charge of the acquisition, archiving and exploitation of the EO data received from both the ESA and third parties missions dating over a period of 25 years. The ESRIN EO Ground Segment is distributed among different acquisition and archiving centers. Archives are not concentrated in a unique location, but distributed in several European countries: this specific infrastructure is a source of particular requirements for the data archive management in general.

The table below shows the past and present EO missions operated by ESA and the datasets archived in our facilities:

Mission	Dataset archived	Time period
ENVISAT	ASAR,RA2,AATSR,GOMOS,MERIS,MWR, DORIS, SCHIAMACHY, MIPAS	2002 –
ERS-1/2	SAR, WindScatt, RA, ATSR, MWR, GOME	1991 –

	(only on ERS-2)	
TERRA/AQUA	MODIS	2000 –
Landsat program	ETM, TM, MSS, RBV	1975 –
SEASTAR	SEAWiFS	1999 –
IRS-P3	MOS	1999 –
Tiros/NOAA series	AVHRR	1978 –
Spot program	HRV	1987 –
JERS-1	SAR, VNIR, SWIR	1992 –1996
SeaSat	SAR	Jun-Oct 1978
MOS 1,1b	MESSR, MSR, VTIR	1987 – 1993
Nimbus	CZCS	1978 – 1980
AEM	HCMM	1978 – 1980

The total amount of data stored in the ESA EO archiving centers is presently about 1 Petabyte and the trend for the future envisages a large increase in this volume with the recently launched ENVISAT mission and other missions planned (Alos, Adeos II, Radarsat, Earth Watch and Earth Explorer missions, etc.).
The datasets are currently stored on various media ranging from old High Density Digital Tapes (HDDT), which in most cases now far exceed the nominal lifetime, up to Sony D1, DLT and IBM NTP Magstar tape cassette.

The difficulty and the cost to manage this large amount of data stored on various types of medium shall constantly increase, unless a different strategic approach is taken. Circumstances inducing such high costs are:

- the constant growth of the archive which includes new operational missions data in addition to the so-called historical data, belonging to missions not operated anymore,
- the volume of the archive, due to the low MBytes/cm^3 and corresponding high Weight/MByte factors, which have a direct impact over the storage and transport costs,
- the cost to perform operations with such a type of medium, including manpower and environment (to be controlled),
- the cost of the technology itself (recorder maintenance and media cost), only in part due to its obsolescence (in fact, most of the time the disappearance of a technology is driven by the market and the difficult procurement of the related media),
- the necessity to perform systematic data copying or data recovery tools at a certain frequency at high investment cost.

ESA archives history and set up

ESRIN recognized many years ago the need to preserve the value of the historical datasets as well as to prepare the future support for the archiving of data for future missions in a scenario which comprises all data lifetime, from acquisition to final archiving, to data exploitation, without ignoring the surrounding international environment and actions taken by the various actors in the EO world.
A preliminary prototype HDDT transcription system was installed at Fucino station in 1994 based on the CREO optical tape recorder. This system started transcribing the historical Landsat MSS data stored on HDDTs. This experience was stopped after a serious CREO recorder failure destroyed a reel of optical tape with more than 800 MSS orbits transcribed (data were still available on the original HDDTs).
Almost at the same time a study conducted by ESA led to identify the ideal supports for the two main classes of EO data to be handled by ESA, High Bit Rate data (more than 20 Mbit/s) and Low Bit Rate data (less than 20 Mbit/s). The identified recorders for the two types of data were at that time the Sony DIR1000 (D1 cassette) for HBR data and DLT for LBR data.

The first transcription system based on Sony DIR1000 recorders (4 in total, with also DLT output) in SCSI mode (ESA had to develop the driver) was installed at the end of 1994 and started transcribing the MSS data, substituting the CREO-based system. Similar systems for transcribing Landsat TM data started operations in January 1996 (4 in total).

A decision taken since the beginning was to transcribe the data in computer compatible format in order to be free in the future from having to perform the data synchronization every time. At the same time a Digital Browse was generated and sent to ESRIN for input into the catalogue and browsing system.

In May 1995 the first ERS-1/2 SAR transcription system with Sony DIR1000 output started operations and was followed by other 3 similar systems, including also output on DLT.
One of the goals of the transcription exercise was to start the recycling of the still usable HDDTs to perform acquisition at our ground stations. In fact, between ERS and third parties missions still using HDDTs for the acquisition, the cost alone of the HDDTs tapes was more than 3 Million dollars per year. Since 1995 only recycled HDDTs already transcribed have been used and no new HDDTs have been ordered. With the utilization of D1 and DLT cassettes the consumable figure dropped by a factor of 10 (about 6 passes average stored on one cassette against one pass on HDDT and lower cassette media).
From 1996 acquisition of satellites data was performed at all sites directly on DLT tapes, using direct ingestion systems, based on disc arrays.

For what concerns the Low Bit Rate data, the first experience was to store the AVHRR, NIMBUS etc. data on CCT, followed in 1986 by the optical disks. Data until 1992/94 were stored on Philips LMSI optical disks and on ATG afterwards. CZCS data were also archived onto ATG optical disks.
The ERS-1 LBR data were also at the beginning systematically transcribed from HDDT to ATG optical disks at Fucino station (were all HDDT's were received from all acquisition stations) and sent to the designated archive facilities. In 1993, due to the high cost of the OD solution, the transcription systems were converted to transcribe the ERS-1 LBR data on 8mm Exabyte cassettes.
With the launch of ERS-2 the scheme changed and a near real-time transcription of LBR data onto 8mm cassettes was performed directly at the acquisition stations.

Since 1998, the ERS-1 LBR datasets are transcribed from the original HDDT's onto DLT with a new transcription system. The ERS-2 LBR datasets are instead copied from the original 8mm cassettes onto DLT with the support of an automated archiving system. These two chains will generate a reference archive of all historical ERS-1/2 LBR data and will enable us to automate the global reprocessing required for some of the ERS instruments.

In total the number of HDDT's already transcribed on D1 or DLT is in the order of 55,000 for non ESA missions and about 100,000 for the ESA missions, with about 20% still to be converted.

At the beginning of 1997 a project was started in ESRIN to develop an archiving system for all missions handled by ESRIN, with a phased approach. The system is called Archive Management and is based on an AML/J library from ADIC with DLT slots for up to 1200 cassettes. The data server application based on the AMASS Storage Manager File System will manage the migration of the data from the on-line local storage space to the library and to the off-line archive.

The current strategy is based on the following approach:
- complete the installation at all sites of automated libraries and AMS,
- initiate conversion of archives on modern technologies (under evaluation: LTO, SDLT, StorageTek 9940A/B, AIT 3),
- record active mission fresh data on the new technology.

This approach will bring substantial savings in the operations of the ESA EO archive, especially when future migration to new media will be required. In this case the system itself can be set-up to perform a background unattended transcription while keeping the possibility to access the data transparently to the clients. Furthermore, the operational manpower needed to operate the archive and the data processing systems can be greatly reduced.

Lessons learnt and consideration for a common strategy

The experience gained in more than 25 years of operations, including lessons learnt and sometimes wrong decisions, has generated a number of issues and recommendations that are worth sharing with the archive community.
These recommendations are also in line with the strategy that ESA intends to implement for the future as an agreed approach to the archive problematic.
This approach is based on the following 5 basic guidelines:

- migration or partial migration of archives must be done as soon as it is justified by savings in operations,
- migration has to be supported by automation of operations,
- dual copy of archive, even if on different media,
- constant evaluation of new mature technologies,
- improvement of archive data access.

These issues are further elaborated in the following chapters.

Why archive migration is needed?

The medium that is chosen to archive data must have several characteristics, the most important being the ability to maintain the data stored for a long term period. Anyway, while the current media technology gives a theoretical possibility to store the data for more than 100 years (typically on optical supports), the problem remains in the rapid obsolescence of the device technology. A systematic data re-transcription (onto new type of media) must then be foreseen. To this respect, the fact of having the data recorded in computer compatible format brings a great saving in transcription operations. In fact, it permits the use of standard hardware (e.g. computers with standard SCSI interfaces). In addition, the usage of standard equipment ensures cost effectiveness when porting the data transcription software from one platform to another.

The cost of maintaining the data in a proper storage environment is not negligible. To this respect, the maintenance of the data stored on HDDT's was becoming prohibitive. The archive rooms for any type of magnetic medium have to be environmentally controlled, in order to maintain temperature and humidity, within the limits of tolerance. Clean air must also be ensured. The cost of maintaining a proper air conditioning system depends fundamentally on the MBytes/cm^3 factor of the media used to store. The media handling costs also depend on the same factor, such costs typically being produced by the high manpower to be allocated for the task and the specialized hardware to be used for the media maintenance (compare for example the cleaning and degaussing devices for HDDT's and the simple cleaning DLT cassettes). Typical MBytes/cm^3 factors are reported in the following table.

Medium	Storage Capacity [GB]	Volume [cm^3]	MBytes/cm^3 factor
HDDT	12	8800	1.4
D1	100	5900	17.4
NTP	20	360	56.9
DLT	40	360	113.8
LTO	100	360	284.4

When shipment of data between different locations (it is the case of ESA) is required, since the cost is proportional to the volume and weight of the media, further savings may be considered by shipping small lightweight cassettes or DVD types of media, containing more data.
The table below shows the relative cost for shipping several types of medium, in volume, and the cost per MByte shipped, with the HDDT made 100:

Medium	Shipment cost	Shipment cost per Mbyte
HDDT	100	100
D1	50	8
DLT	4	1.2
LTO	4	0.5
NTP	4	2.5
DVD	2	1.1

The possibility to have access to computer compatible formatted data, together with the storage onto smaller media, gives also the possibility to perform the automation of the data archiving, retrieval and re-transcription, by using automated media libraries and stackers.

Current automated libraries for the most popular type of media can greatly enhance the possibility to use the data at lower cost. When it is necessary to have large volumes of data retrievable automatically, the use of an automated library, based on low cost media such as magnetic cassettes or CD-ROM/DVD types, is advisable.

When coupled with specific software like a Hierarchical Storage Management (HSM) system, the usage of the automated library is further enhanced. An HSM can in fact greatly improve the performance in data storing and retrieving to and from an automated library by implementing a staging approach with automatic data migration between the different storage levels (off-line, near-line and on-line), promotion of most recently used data to the higher speed access level and demotion of less accessed data to the lower levels. In addition it provides services to perform automatic drives maintenance (e.g. with automatic cleaning when error rates exceed a predefined threshold) and media copy in the case of corrupted media.

It is clear that the possibility to have a big volume of data seen as being on-line can greatly reduce the media handling operations, reducing processing times and in the end highly decreases the operation costs. Also, in the future the cost of automated storage systems and media will continue to decrease, while the cost of managing data continues to increase. As already mentioned, one of the major costs in data archiving is given by the necessity to perform systematic data transcription from one medium to another (with frequency given by media lifetime or, more likely, by change in technology). The operational cost of copying the data from one medium to another is greatly reduced when the data are archive in an automated library.

Experience shows that archive migration is based on five major requirements:

- the reduction of operational costs: experience shows that the major cost in large data archiving projects is mainly due to the media, not the drives,
- the need to preserve the data integrity: the converging of purely archive oriented developments and data storage computer peripherals technologies is economically convenient, but sometimes with a degradation of the long term data integrity,
- today, in the computer environment, the technology changes rapidly and maintenance issues become very critical with hardware obsolescence,
- logistic and homogeneity of services: the increase of data to be archived poses problems in terms of space, media identification, data access, operational manpower support which sometimes limit the possibility of providing similar types of service in terms of speed, product

generation, etc. if compared with freshly acquired data from new missions, which benefit from more recent technologies and operation automation,
- data security: normally archives are unique and subject to a high risk of data loss for external causes (fires, floods, etc.). The availability of cheaper, smaller and more capable media justifies the archive migration and allows also the possibility of archives duplication. One way of ensuring the archive duplication at an acceptable cost is to maintain the previous technology data holding as a safe backup, in the case data need to be retrieved for whatever problem may occur.

- **How to migrate archives?**

The migration of archives on new technologies must ensure as a main objective the data integrity while data preservation is one of the key issues in today's EO world and for archives management in general. One of the immediate and logical questions may be: should the opportunity be taken to perform a selective data purging of the data set for data volume reduction based on an a-priori decision (excessive cloud coverage, redundancy, age) or is it better to preserve always all the data? How costly may such an operation be, as current algorithms for automatic feature detection are not safe and failure-proof enough? In some cases, the analysis of the net balance between data reduction and simple archive copy operations costs may lead to the decision to archive everything without any selection. Other questions may arise concerning data migration. Should one transfer the entire archive on a unique type of media? What other element of the archive is directly affected by the migration? How should data quality be ensured during the migration?

Preserving indefinitely the information contained in the data acquired by the EO satellites (the same problems apply to all types of information to be archived) means not only to archive the sensing data (containing the primary information), but also to store and be able to retrieve in conjunction with the data all the auxiliary information necessary for the data understanding and exploitation.

ESA has adopted some basic principle in archive migration:

- Transfer of data without data reduction is the baseline. The need of data stitching for overlap reduction and better data compacting, including the adoption of reversible compression methods may arise.
- Improvement of data quality and readability for future processing: data manipulation is accepted when oriented towards securing the data processing and products generation, improving the data quality and their readability.
- Accept non homogeneous archives in terms of media, derived by the need of embarking on new proved technologies life cycles, when the migration is justified by economical reason. The migration of a complete historical archive is slower than the modern technology cycle. So, even within the same mission, data may be archived on different type of media.
- Possibly standardize the archive access data format.
- The frequency of the migration depends from several parameters: availability of budget, proven reduction of operational cost, operational need, etc. A tentative figure might be in the order of 6-7 years, given the current trend of new technologies announcements.
- Transfer together with the data, on the same media, also all information needed for the data format reading and processing. With time, information sitting elsewhere in the system may become lost.
- Migration of archives very often includes the regeneration/conversion of the databases for the data inventory and media labeling.

- **Which parameters must be carefully observed in archive migration?**

Migration of archives is an activity that needs to be carefully analyzed and planned, when considering the impact in terms of investment, cost and operational efforts. The analysis involves several elements, all of them to be balanced and carefully considered:

- Impact on the running of operations that cannot be suspended or degraded.
- Overall cost to be budgeted for the operations to which several elements are contributing, in addition to the initial investment cost: manpower cost, media number, maintenance for old and new devices heavily stressed in the conversion phase, spares, time duration, mainly linked to the data input speed, assuming that the new technology data transfer is higher than the previous one.
- The related conversion of inventory databases may also mean selection of new commercial databases, cheaper to maintain and less costly for licenses.
- Automation of operations in input and output (adoption of robotic systems and related file management systems) may heavily reduce the manpower cost, allowing a non-stop operation, particularly during unmanned periods.
- Ensuring data integrity (no data loss or degradation) implies verification of the data after conversion has been performed with some basic control in screening the data in output and comparison with similar values of the data in input. Experience shows not to trust blindly the data copy functionality, especially when working with new technologies in high volume.

- **Why technology evolution must be constantly watched?**

Technology nowadays evolves with dimensions that only a few years ago could not be expected. This is mainly due to the increased variety of technologies (magnetic, optical, etc.) and the highly increased number of actors and companies in the specific field. A large contribution to the offering is given also by the merging of technologies oriented to the specific data archive requirements, of widely used consumer market devices and of technologies focused on computer data storage.

The technology awareness requires focused efforts and the need of in house experts, being a function which cannot be delegated to third parties: experience establishes some basic rules, which may appear obvious, but nevertheless always need to be kept in mind in the difficult competitive market, dominated by media announcements (magazines, articles, conventions, smart vendors, etc.)

Some recommendations are:

- follow directly the technology evolution increasing ones experience in the field, by participating to symposia, shows, demonstrations, etc, reading specialized literature, accessing the web for information and dialoguing with entities facing the same problems and may be looking for different solution,
- do not trust a technology without having done careful benchmarking,
- do not jump on very new technologies not yet proven and yet largely accepted by the market,
- do not blindly believe to various announcements and road maps (many technologies never maturate in reality)
- before adopting a technology be sure that the needed infrastructure is available (interface to main computers, commercial drivers from major companies, etc.)
- verify published performances with the characteristics of the archived data (record and block sizes, file lengths, data compression, etc.),
- verify availability of tools for operational diagnostic investigation/anomalies monitoring and cost of maintenance and spares (head cleaning, head duration, need of specific company support for normal operation, etc.),
- verify all the aspects involved in the mid-long term operations of certain technology: availability of maintenance support, supply of concerned media, road map evolution, diversified suppliers,
- avoid, whenever possible, to select equipment or media supplied by a unique manufacturer or not supported by large computer companies.

- **Which technology is the right one to select?**

Since for more than 20 years the only medium able to satisfy the requirements in terms of capacity and recording data rate needed for the EO missions data acquisition was the High Density Digital Tape Derived from the video market technology, this medium has also conquered in the past the leadership as a long term storage solution. In the 90's, a boost in the magnetic tape technology gave these types of recording media a new life in the EO world for data acquisition, recording and storage.
The following table gives an overview of the performance for typical magnetic media used or planned for EO data acquisition and archive.

Medium	Storage Capacity [GB]	Throughput [MB/sec]	Access Time (Sec)	BER	Media cost [$]	1GB cost ($)	ive cost [K$]	Road map (GB)
HDDR	12	~120 Mbps	120	10^{-9}	100	8.3	200	-
Sony DIR1000	100	30	15	10^{-13}	100	1	300	-
Quantum DLT	35	5	18	10^{-17}	60	1.8	5	-
AMPEX DCRSi	47	14	20	10^{-9}	100	2.2	150	-
IBM Magstar	20	9	10	10^{-17}	23	2.3	45	-
Sony AIT 3	100	12	15	10^{-17}	90	0.9	3	500
Sony DTF 2	200	24	10	10^{-17}	160	0.8	22	300
9940A Storagetek	60	10	2	10^{-17}	100	1.7	45	200
Quantum SDLT	160	15	15	10^{-17}	100	0.6	5	200
IBM LTO	100	15	13	10^{-17}	75	0.7	3.5	200

The enhancements of the magnetic tape technology, together with the improvements in all fields of technology necessary in the EO world, have completely changed the scenario. In addition, faster and cheaper computers with largely improved I/O capability are now able to perform via SW operations requiring dedicated HW only some years ago. Faster and bigger disks are now able to sustain high data rates for real-time recording.

Today it is not easy to select a new technology, due to the many choices offered by the market, making it difficult to evaluate particularly due to the wide price ranges.
Many factors contribute to the decision process and in many cases with also different bias, depending from the archive logistics, the available infrastructure, the volume and the compatibility with technologies already in place.

- Overall cost is ultimately the main issue: normally the total cost of the media is by far, in an acceptable technology lifetime period, superior to the investment in drives. Therefore, the trend of media cost and the competition on the market between several suppliers is a very important element of evaluation
- Preference should be given to technologies originally born for data archive, more focused on data integrity and normally providing larger infrastructure in hardware and application software availability
- Very advanced technologies, not yet completely matured, should be avoided. Do not pay too much attention to the data duration ensured by certain technologies (even hundreds years). Archives migration is required sooner for the reason stated above.

- It is advisable to select technologies supported by the consumer market (i.e. with video or sound application) or by the computers data storage industry (e.g. DLT, Exabytes, etc, were born as computer peripherals devices for data back up),
- Selection should be done on several parameters and ultimately a trade off must be made considering the total investment cost, the media utilization projected cost, the speed of data transfer, the access time, the availability of libraries and support software, the projected cost of operations, the merging of requirements of new missions to be supported and the historical missions data to be maintained. In general, operational costs for new missions are covered as a background job, as well as the major part of the archive conversion, once the same technology is selected.
- Before changing technologies, the archive service requirements must be re-analyzed and assessed. It might be the case that relatively small archive data sets can easily fit on magnetic disc arrays or CD-ROM/DVD libraries, exploiting the parallel evolution of those devices, not originally considered as archive devices. This solution may improve the archive access exploiting the direct access of those devices.

- **How to format the data?**

Experience shows that the technology conversion may also include a change in the data format and packaging. The data format conversion during the time of moving away from HDDT is a clear example.
The metadata associated with the information archived assume a fundamental importance for the long-term data preservation. The fact of having the data with the associated metadata available together permits an easier migration from one medium to another or, as it is more often the case, from a type of media to another, and at the same time guarantees a lower cost in the data retrieval and processing subsystems. It must be noted that the distinction between data and metadata is in some cases arbitrary.

Such types of problems, certainly not unique to the EO community but common in all space activities, are so important to promote the search for standards and recommendations. The Consultative Committee for Space data Systems (CCSDS) started in 1995 and under the auspices of the International Standards Organization endeavors to develop a reference model for long-term archiving. The Open Archival Information System (OAIS) aims to provide a framework for the understanding of archival concepts needed for permanent digital information preservation, expand consensus on the requirements for long-term digital information preservation and guide the production of OAIS related standards.

In general, the following guidelines should be adopted:

- Keep the data at the most primitive level possible in computer format, avoiding special data processing, which may transform irreversibly the data, not allowing future applications or new algorithms implementation,
- adopt whenever possible, one unique and common data format, at least for the historical missions. Sometimes it is not easy to change format for active mission during a mission lifetime,
- align format, whenever possible, to existing international standards if possible or to format commonly adopted in order to make easier any exchange of archived data with other organizations,
- package the data in a form that contains on the same media all information necessary for the products generation, possibly including auxiliary data, orbital parameters and eventually also all information needed to read and extract the data,
- be careful in adopting file management systems with their own proprietary format, since this may be a limitation for further archive migration.

- **What are the impacts of archive migration on archived data access?**

Archives are not stand-alone repository systems, but are part of an end-to-end structure for delivering products to users, which includes data identification, data retrieval, data transformation, data packaging, data quality check and data delivery.
The process of technology change and the related data conversion must satisfy the minimum requirement of ensuring at least similar services and data access at the level previously provided.
In order to reach this goal some basic issues must be kept in mind and taken in consideration:

- change of technology means reorganizing the media data content and therefore in order to allow data retrieval, also re-converting the data inventory data bases,
- the change of technology may impact each of the several missions processors input for products generation, if interfacing directly the archive devices,
- the change of format may impact also each of the several missions processors input for the products generation, if interfacing directly to the archive devices,
- it is advisable to provide for a unique interface upfront the processors, in order to minimize the above impact,
- processors also evolve with technology and software needs to be ported onto the new platform. It is advisable to consider at this point the advantages brought about by the new archive technology, which may allow new production strategies with new products and services (faster data access, faster throughput, etc.)
- new products and services require changes in the infrastructure for data retrieval, products ordering, packaging, etc.
- capability of faster data transfer may impact the overall architecture, with the adoption of SAN's etc. within a standardized way of retrieving and transferring data from archives to processors.

Conclusions

The ESA approach to the EO data archiving problematic has been based on a strategy aiming on one hand to ensure the preservation of the historical datasets under its responsibility and on the other to improve their exploitation, improving in parallel the data access. The approach is an integrated one and spans the complete data life cycle, from acquisition to storage to real and near real-time processing and dissemination to long-term archiving.
Some basic considerations can be derived for the future:

- archive maintenance includes data preservation and therefore also data migration,
- archive are part of an end to end system which includes data retrieval and data access,
- archive maintenance and exploitation is a costly activity and must be seen as an integral part of ground segment operation,
- archives cannot be based anymore on one homogeneous technology but face the technology evolution, as a constant migration process, when justified by economical reasons,
- selection of technology must be a trade off of several elements, of which the media cost might have the heavier weight,
- technology evolution must be followed very carefully and selection process cannot be delegated to third parties,
- archive migration must be fully automated adopting adequate libraries in input and output,
- a dual copy of the archive is advisable even if not based on the same technology.

The new available technologies and this global approach will bring great advantages to the EO user community, while keeping the costs of the data exploitation low enough to cope with the current constrained financial environment.