Project Acronym: Presto4U
Grant Agreement no: 600845
Project Title: European Technology for Digital
Audiovisual Media Preservation

D6.1: TRANSVERSAL CoP IMPACT ANALYSIS

Project funded by the European Community in the 7th Framework Programme

Table of Contents

Scope

The Presto4U project addresses the issues related to long-term digital preservation in the audiovisual domain. It works on the understanding of underlying problems related to different *Communities of Practice* (*CoP*s) and to develop solutions and guidance adapted to their situation.

This document makes an initial Transversal Impact Analysis among the *Communities of Practice* working in the domain. It allows to identify the detailed situation at the beginning of the project and to understand the main obstacles that prevent audiovisual collection-holders from advancing in the preservation of their assets. It is based on a global analysis of problems common to all communities and of local identified problems in different aspects of the community's activity. The Deliverable will be continued during the project, in order to analyse the impact of the project's results in the uptake of actions within the communities and to measure its evolution and the incidences it will have within the PrestoCentre.

Executive summary and Wrapper

The Presto4U project aims to develop long-term digital preservation of assets for all the communities concerned with audiovisual production and use. It works in the understanding of the totality of problems associated to digital preservation. In order to address the diversity of the domain, it has been split in nine *Communities of Practice (CoPs)*, each grouping particular practice use-cases, similar technical problems, and common activities. To complement the community's analysis, the project studies the existent technology (as well as its evolution) and potential tools that could be used within the *CoPs*. A study of standards associated to the audiovisual domain is equally carried on, in order to determine the most adapted ones.

The audiovisual world represents one of the greatest and long-lasting impacts of twentieth century's technology. The impact of audiovisual has permitted the creation of cinema, radio and television, and permitted the invention of totally new creative domains. Furthermore, with the arrival of digital technology, audiovisual has entered the personal domain, and today content creation and sharing is one of the greatest drivers of Internet and social relations.

Millions of hours of audiovisual material exist in analogue and digital forms, and a real challenge exists regarding their preservation of, so that humankind will be able in the future to access the works of the past as we access today the writings, paintings and works of art of ancient times. With audiovisual preservation, issues seem to have become more complex than it used to be with physical objects; analogue contents fade out after some decades and digital contents become inaccessible if not properly organised in a preservation process. Thus the risk of loss is higher than it ever was that we lose a great part of our audiovisual world in next decade.

Solutions exist; they are not miraculous as could be expected, but they consist in a series of actions, which can guarantee long-term in an efficient and reliable way. Many content-holders have already launched important digitisation and preservation programs, and more will come in the future; however if there is not quick action, many communities are at great risk of loss because they lack the knowledge and the means to effectively ensure preservation.

The Presto4U project works on all the issues related to digital preservation on the long-term, and provides a reference knowledge point for preservation and for strong technical recommendations regarding actions, tools and guidelines for effective action. The objective of this transversal Deliverable is to give a cross vision through the 9 Communities of Practice, on their initial situation and needs as they have been brought up by the work of *CoP* leaders and experts. This is probably not an exhaustive analysis, however most of the existing problems are brought up and analysed. What looks like a local problem related to a community is often in fact a larger problem shared by several communities.

This Deliverable analyses then the general context of the audiovisual domain and the different variables that prevent or slow actions towards long-term preservation. These obstacles, or "drawbacks" as sometimes called, are not only technical; even if many technical problems are found: the initial analysis shows how variables like the mission of an archive, the legal situation, the economic environment an even essential components as staff can have a crucial impact on decision making and in advancing towards an effective action. Every archive, every content-holder is a different world in itself; while

sharing global concerns and generic problems such as digitisation or format choice, each member has a particular profile and situation which makes them different, while similar, from the others.

The work is divided in two main areas, independently of the general context and the intervening variables: the common issues and problems which concern all or most communities, with issues like digitisation, content management, restoration format and standards or the legal situation; then the singular issues and problems for each community, sometimes shared by a second community.

Two major profiles have been observed within the study; those called here "well established communities", as broadcast or film archives, which have a long tradition in content conservation and preservation. These well established communities have partly undergone through digitisation actions, and are directly concerned with preserving their assets. The second profile corresponds to communities which develop a strong activity related to audiovisual production and use, however they don't have any archival or preservation mission, and only slowly discover the nature and size of the preservation problem as well as the difficulties encountered in content description and use within a legal framework.

These two profiles are a strong issue for the project, since the first ones may be highly experienced with archiving and preservation; while the second ones are concerned with reuse and distribution through Internet and community sharing of assets.

The fact of creating a particular community around individuals and their contribution to the archival heritage, as well as the identification of their strong preservation issues, not at all covered up to now, brings to the project a new dimension, since the analysis leaves the traditional archival domain and brings the problems to an individual level where all these issues have a different dimension, and where the concept of preservation is only starting to exist. In all cases content-owners have mostly lived through critical loss situations, where parts of their assets were lost due to the lack of a preservation policy and tools. This situation should happen less often in the future: on one side tools and information is there and in many cases circulating effectively; on the other side consciousness of approaching preservation as a whole integrated process is growing, and digital preservation experts emerge all around as well as competence centers as the PrestoCentre, who contribute to the global knowledge and dissemination of results.

Finally the economic environment is presented, with the practical issues related to financing (independently of the major issue of finding the money to do the work) and how they interfere and should be considered by archives and content-holders. It also presents new perspectives, as preservation as a service or public – private partnerships, which can contribute to the improvement of preservation and the undertaking of actions impossible otherwise.

During the following year, the project will analyse the contribution of the different activities of the project to the uptake of results and to the acceleration of digital preservation actions. The first year has shown the great expectancy of many content-holders who have found in the Presto4U project an environment where to share their concerns and to start building around the community, working on common issues and making decisions with a potentially strong impact on the future of preservation.

1 General Overview of the Transversal Impact Analysis

1.1 Introduction

Audiovisual is made of Sound and Moving Images fixed on analogue carriers or defined as digital data. This simple definition covers a huge variety of activities and domains in which the same definition has multiple meaning in terms of nature of content, quality, age and value. All domains share common access and preservation problems of technological, economic or legal nature. There are also distinctive issues, possibly affecting a domain imposing specific issues for preservation or sometimes unsolved problems.

In order to narrow down the variety of approaches, the Presto4U project has defined *Communities of Practice*; bringing together institutions, individuals or companies, who not only may share a same vision on their activity but work with similar tools, methodologies, metadata and legal issues. This approach permits to focus issues on common grounds and to compare problems, which may seem very different, but may be part of a common definition or description applied in a different context.

The objective of this Deliverable is to make an initial analysis of the situation of the different communities in order to extract important information for the continuity of the project. It gives an initial state of the situation among the different communities after the first feedback rounds made by *CoP* Leaders, in contact with their identified Experts, or through the knowledge and comprehension leaders have of their community.

The Deliverable identifies:

1) Common issues concerning several or all communities: many of the identified problems are common, or cover several or all communities

2) Specific problems related to a community: some problems are only found in one or two communities and cover local issues, possibly totally absent in other communities

3) The economical capacity to undertake actions in the different communities: problems are not only technical but cover a wide range of aspects, influencing the take-up of digital preservation

1.2 The overall vision

The Presto4U project concerns the preservation of the audiovisual domain in all its diversity and facets. The strategy of structuring the project through nine different communities of practice has permitted to address particular aspects of this vast domain and will permit to analyse specific aspects and difficulties as well as general problems preventing and slowing down digital preservation. When analysing from a global perspective the main drawbacks of today's situation, two main aspects stand out, probably constituting the main problems of the domain or even of any content holder be it audiovisual or other:

1) Getting into the digital world; thus digitising analogue contents with all the increasing difficulties this action implies, due to the fragility of original carriers and the obsolescence of playing machines

2) Understanding digital preservation and its implications as an archiving process on the long-term; even if there is a clear consciousness about the fact that it <u>has</u> to be done, there is not a clear perception on <u>how</u> it has to be done.

Other important issues are associated to those initial two obstacles; the economic situation or the legal environment may block any action, however clear the process may be for the content holders. Sometimes however, fortunately not too often, there exists a total ignorance of the incurred risks of content loss if appropriate actions are not undertaken; this situation was much worse five years ago and important European projects have contributed to increase the level of awareness[1]; however a global pedagogy still has to be continued on the risk of loss, and on the evolution of technical obsolescence.

1.2.1   To be missioned or not

The first months of the Presto4U project have revealed a global dichotomy in the profile of members of the Communities of Practice. Most of the members are content holders, thus implying that they have audiovisual assets representing some form of value to them, be it historical or economic or both. However the reason of their existence can be very different and will have strong implications on the way they understand, analyse and apply digital preservation. The two identified profiles are:

1) <u>Institutions with a preservation mission</u>: These institutions have some kind of reason why to keep audiovisual contents and are supposed to keep them for a certain amount of time.
   Their profile can be very diverse: archives (national, regional, private), libraries, museums, associations, foundations, all committed to the conservation of their assets and have different levels of responsibility for performing it. This means that if something is lost some higher instance may make them accountable for the loss. Missions are not always clear; they may be clearly established national missions (as in national archives or libraries) or be a *de facto* mission complementary to other responsibilities (a national broadcaster who has contents not belonging to them, but for which they have to ensure preservation)

2) <u>Institutions or companies,</u> who produce, inherit or receive audiovisual contents, with no specific preservation mission. This is the result of some kind of activity, private or public, generating contents with or without economical value. If there is no economic value associated, these contents may be donated or even destroyed; if there is an associate economic value, it may be vital for an enterprise to keep those contents available (as for musical production or video production companies). This means that often these institutions or companies will have to assure some form of archiving in order to assure the accessibility of the contents.

1.2.2 The variables that condition and constraint Preservation action

To these two profiles are associated a number of variables which may constraint independently or simultaneously any action regarding Digital Preservation:

1) The first and more important variable is whether or not, when an institution has an archival and preservation mission, thus a responsibility regarding the contents they keep, they are supposed to receive the means to fulfil this mission. This implies specific staff and tools to perform the action (which may not always happen as expected, creating sometimes dramatic preservation situations); and in some cases having a permanent and long-established financing, allowing to undertake long-term projects and preservation plans. This is not only for national or regional archives, it also concerns many private companies who establish preservation actions and budgets in order to keep and reuse their internally produced assets.

2) The second variable is cost associated to the action. There may be consciousness and will to undertake many actions, but financing may not be there. Digitisation and preservation often need a continuous financial input through a long period of time (years), and this may be difficult or impossible to obtain.

3) The third variable is the condition of associated rights regarding the collections. In many cases archives develop a commercial activity based on the rights they own on their assets, potentially representing a good complement to a regular funding and allowing undertaking new actions. However in many other situations institutions have to keep important collections (as *Cinemathèques*) with no associated right attached to the content, so have no commercial outcome opportunity for their contents.

4) A fourth variable concerns the staff and equipment needed to develop an archival and preservation strategy. The technological shift in our society from analogue to digital has created a natural obsolescence of millions of private or public contents, often donated to institutions, libraries, museums or archives, and these reception institutions may not have the knowledge regarding handling and storing audiovisual contents, nor the staff to monitor or the machines to read them. In this context many collections may be doomed to disappear.

5) A fifth variable concerns the difficulty of decision-making regarding the technical solution or the preservation format to be adopted. In a continuously changing environment it is difficult for anybody who is not a technical expert to identify the most adapted solution and the best quality/cost ratio. Between rumours and technology announcements by major constructors, a great uncertainty exists, tending to prevent decision-making in the expectation of a promised future solution, which may never come.

Finally there are geographical and geopolitical variables. Archiving is not generally considered a priority action amongst the problems a country or society may have; and even if a legal environment may exist for heritage preservation, there are little or no funds to undertake such an action, continuously postponing it to better days! The situation in Europe is highly unbalanced, between countries investing in digital preservation, and countries maintaining their assets the best they can, waiting for the opportunity to do it

someday. To this may be added other local issues like climate conditions (you have more chances for longevity if your archive is in the Arctic Circle regions than under the tropics).

When these variables are added to the two previously explained obstacles, it defines a quite critical situation, in which getting into the digital world and staying in it may be a major challenge for a content holder. Independently of all these reasons, this situation creates a very high uncertainty; where content holders and deciders don't clearly know or understand where to go and which decisions to take.

Fortunately digital preservation is a reality for many contents holders and the community is slowly but surely advancing to digital management of audiovisual assets on the long term, even when funding may be low. The cost associated to technology is regularly decreasing which tends to simplify some of the issues; and information regarding digitisation and preservation is becoming easier to access. Decision-making regarding formats and technology also tends to simplify as the world is moving towards uncompressed video contents and some solutions adapted to small institutions or structures start to appear as well as online services. However, strong impediments such as absence of rights or mission, may totally prevent audiovisual contents from entering in the digital world.

1.2.3 The implications of Digital Preservation

It may be clear what digitisation and its implications are, and many deliverables of the Presto4U project as well as many European projects from the past, analyse these issues and propose guidelines, technology and solutions. It is however far more complex to explain Digital Preservation and its impact on content managing and archiving. The archiving science, developed through centuries of successful conservation, is based on the preserving of physical objects.

Physical preservation is a rather passive action in which, if you keep your objects in a secure location under suitable light, temperature and humidity conditions, your contents may last for a very long time provided the material is robust enough to endure time degradation (paper is not as robust as stone). In the twentieth century, passive preservation was highly improved through precise means for controlling temperature and humidity and adapted to the nature of the artefacts. This kind of preservation has a cost; however a constant cost with few fluctuations and few interventions on the objects themselves.

In the digital world, we have the impression of dealing with objects, since most contents contained in objects can be represented in a digital form and if not (a monument for example), digital representations can be obtained that create the impression of being in presence of the physical object. However a digital object, or file, is not physical and so will not behave as physical objects are supposed to; it is just a series of numbers organised through a specific code (the format) that will allow the appropriate code transcriber (the software or player) to transform it into an image sequence, text, sound or picture. The illusion is quite realistic since it gives us the impression – and an impression of facility regarding the actual manipulation of physical objects, that we are handling objects. Digital objects mimic physical objects while being totally different in nature and behaviour.

At the same time the robustness of Digital is based on the simplicity of series of numbers, which need to be transformed into perceptible objects. This data, however effective for

object representation, is rather fragile and can be altered. Thus any change in the structure or the order or the interpretation code of the numbers may result in a deformation or inaccessibility of the digital object. These series of numbers have then to be regularly checked, verified and controlled, as well as the interpretation code and the representation code in order to maintain integrity of the content. These series of numbers are defined as files; depending on a coding format, which is accessible through specific software.

A second situation associated to the description of objects through numbers, is the fact that these numbers have to be stored somewhere, meaning that some kind of carrier has to keep these numbers in a structured and accessible way. Magnetic hard-disks, digital tape, optic devices, flash memories, all are specific carriers for digital information: however they are also subject to accidents, obsolescence, and loss of information or other kinds of damage.

Preservation actions are made necessary by these two issues: on one side the integrity and legibility of information, on the other the permanence and accessibility of the carrier. Files may be damaged, formats become obsolescent or are simply "improved", software changes at a high speed and carriers present problems, obsolescence and inaccessibility. Digital Preservation methodologies deal with these issues, analyse them and establish rules and behaviours to be adopted in order to guarantee access under any lapse of time following the production of the content.

This implies checking regularly all the points listed above, and undertaking a series of actions if any kind of problem appears. There are also associated strategies that reduce the incurred risks, such as making several copies and keeping them in different locations, but this doesn't change the nature and regularity of the actions to be undertaken. Some years ago all these actions were slow and sometimes hand-controlled; today efficient tools exist that reduce the risk of losing contents or accessibility to contents due to environment changes. This is called Digital Preservation and is considered as being an Active Archiving methodology in opposition to Passive Archiving. Active Archiving means continuous checking and acting on digital contents in secured environments with controlled access. There is a permanent cost associated to this, as well as to regular actions needed when contents have to be trans-coded or copied from one format to another or from one carrier to a new one, in a process called migration.

There has been a constant dream to go back to a Passive Archiving situation, where a digital format would be accessible forever and specifically designed carriers would have a thousand year life perspective. However no clear and efficient solution has arisen yet, also due to the fact that, however efficient, a technical solution has to be adopted by a large community of users to have a chance to be maintained over decades. Often adopted technical solutions are not the better ones, just the most popular ones.

Another trend in Digital Preservation is based on the expectation that in the future, new technology may permit to address old contents. This implies to keep digital contents in their initial format (while maintaining integrity) and keeping a description of the interpretation code and software so that the access environment can be emulated in the future. This may work on some contents provided they are relatively simple; however it is also a component which needs to be analysed and can provide interesting solutions in the future.

The Digital world is, on the other hand, an excellent technological environment for human productions: contents can be copied easily with practically no loss, transmitted at high

speed from one location to another, stored in small and easy to access devices and transformed, modified and reorganised easily and efficiently. The main concern in Digital Preservation is how to keep the Digital world and its productions safe and accessible within a continuously moving technical environment; Digital Preservation is the solution for that. It is based on the concept that, if technology moves, contents have to move and adapt in order to stay compatible with the technology changes.

Digital Preservation is then the only solution for long-term access to contents. It may seem complex to apply for small structures or institutions, however large institutions have already adopted this approach, which works well once it is installed and guarantees a high level of security for the future. Other solutions are slowly appearing concerning small and medium collections and it is imaginable that in a near future other solutions may appear for small collections or individuals. It seems very feasible that Services will develop providing sufficient guarantee and risk management to make any content owner confident they have secured the long-term life of their contents.

1.2.4   Other aspects of Digital Preservation, metadata and access

Securing the preservation of Digital assets is the main and essential objective of Digital Preservation. Two other issues associated to it are content description and online access.

Audiovisual contents are time-based, this brings complexity for their description and retrieval. It is important within an audiovisual document to know what it contains, its structure, components or featuring people in order to find a specific sequence, image, situation or person. Other related details need to be known, such as who made the content or contributed to it and has some right, the technical state of the content, and where is it located on a shelf or within an information system. The full set of these vital information pieces was formerly called Documentation and is more commonly called today Metadata. Each content has or should have a set of associated metadata, to allow retrieval and usage of a specific content.

Metadata is a well-established domain of practice and research (linking different metadata schemes has become a challenge in many domains of activity), and there are different ways of presenting and structuring metadata. Metadata may describe the content itself (structure, scenes, locations) as well as the legal right holders, the production companies, the information created during the digitisation process (technical metadata) as well as essential identification information (who, when, where). A non-exhaustive list gives us:

- Descriptive metadata (necessary elements for search and discovery; Title, Date, Language…)

- Rights metadata (necessary for reuse of material; Author, copyright, Intellectual property rights, associated rights)

- Description of physical items (for physical objects; form, location, state…)

- Technical metadata (description of digital items; format, bit-rate…)

- Quality metadata (state of the content, identified defects, restoration needs…)

- Analogue preservation metadata (information about preservation actions performed on analogue items)

- Digital preservation (information about preservation actions performed on digital items)

- Identifiers (information allowing to have unique numbers for items; barcodes, ISAN, EIDR…)

- Extracted metadata (information regarding scenes, speech to text, OCR, face-recognition, or any type of information automatically extracted from a content)

A large historical and well-organised audiovisual archive has all these information, however this doesn't mean it is needed in all cases and situations. Metadata's aim is not to give the largest and deepest possible description of a content; it is based on use; meaning that metadata structure and deepness depends on the nature of the archive and the kind of usage. In a broadcast archive there may not be image description (what kind of objects are there, or the location description; countryside, sea…) however there will be a very precise description of right-holders, names of people appearing in images, identification of music played…). An archive selling stock-images based on the nature of the images (objects present in the image, location, type of movements…) probably keeps not much information about rights, those having previously been bought in order to freely commercialise the images.

Producing the metadata is yet another issue! Historically audiovisual contents were documented by hand, based on the retrieval needs associated for their use. Hand documentation is very effective and archives based on this method have strong and coherent descriptions of their contents. However documenting by hand is a time consuming process and it has become impossible to describe accurately the huge amount of contents produced daily.

The trend in last years has been to apply technology to extract the necessary information needed to identify and retrieve contents. No clear solutions exist yet and international research has worked strongly in content recognition and retrieval projects as well as information extraction actions in order to improve descriptions or to simply have an idea of what is within an audiovisual document. The produced information, by hand or automatically extracted, has to be structured and kept through organisational schemes in efficient and fast accessible environments; this is why content repositories have associated data-bases related to each content, where different levels of information can be stored in an organised way, in order to associate to each program or file a text based description. Metadata databases as well as content files are associated together in systems permitting to manage contents in a structured and organised way, called DMAMs for Digital Media Asset Management systems.

DMAM's have a managing function allowing content holders to organise their collections and easily find whatever program they are looking for. Advanced systems have strong retrieval functions for discovering new material based on content analysis and information extraction. A DMAM may also be capable of managing digital preservation, i.e. checking the integrity and consistency of files and doing specific actions when problems are encountered.

Metadata does not describe totally a content, it is an attempt to capture the main features

important for further reuse; this is why content discovery technology is more and more applied in order to enrich descriptions through content search, in order to enrich already described collections and improve their usability.

Having a Digital information system implies that there are uses associated to the digital contents. The main use lies in access to contents in open or secured online spaces for general access or commercial activity. In the last years the effort put on digitisation and digital archives was fostered by a need for access, sometimes more important than the preservation itself. Archives and content holders have changed their perspective through the access possibility, not only giving a new visibility to otherwise totally inaccessible contents, but also creating a new marketplace for content exchange, and user interaction as well as providing an image of openness to any institution making their contents accessible.

These new issues have also generated new problems regarding rights protection and the perspective of multiple copies of the same contents. This has somehow slowed down many initiatives; however, in the last 18 months, major content holders have started uploading contents in dedicated environments within UGC websites (User Generated Content sites, e.g. YouTube) in order to diffuse restricted amounts of contents and also to generate small revenue associated to publicity and number of hits per user on their content.

Even if for many content holders it is practically impossible to put contents online due to the fact that they do not own any rights, there is a slow tendency to permit access at least to parts of the collections in order to give institutional visibility and in many cases (as in broadcast) give back to the citizens what their taxes have helped to produce.

1.3 Conclusions

Communities of Practice are in unequal situations regarding these general considerations. The legal environment in which they exist as well as the existence of an archival mission, have strong consequences in the undertaken actions. The first analysis in this document will show the common issues among them and the ways in which communities may act in order to deal with preservation issues.

In order to simplify the presentation, the nine communities are grouped in smaller themes, each sharing common interests and problems.

- <u>Sound, music and art</u>: **Sound & Music archives** and **Art & museum objects, artists and their representatives**
- <u>Professional actors</u>: **Broadcast**, **Video production and postproduction**, **Film collections and filmmakers**, **Footage sales libraries**
- <u>Research and Education</u>: **Research and scientific collections** and **Learning & teaching repositories**
- **Personal Contributions**

The Sound, Music and Art theme, shares common concerns with diversity and complexity

of the objects (physical analogue or digital) they conserve - artistic works for the majority - and are strongly concerned with the re-performance or presentation of the works they collect.

The Professional actors group brings together all the world of video production for commercial uses, it includes film collections, with or without a commercial perspective.

The Research and Education group shares the use of contents for research and teaching contexts, where specific issues apply concerning reuse of contents.

A single member theme is represented by Personal Contributions, an emerging community concerning content holders, individuals as well as services.

The second part will deal with specific problems identified so far within the communities. Sometimes, only one community initially presents a problem, but after further analysis it may appear that the proposed solutions will have an incidence on the activity of other communities.

Finally the economical capacity to undertake actions will be presented in order to present the strong differences between *Communities of Practice.*

2 Methodology and general context

The work leading to this Deliverable consisted in an analysis of the initial situation as described in D2.1 *Knowledge Schema and Community requirements* and a deeper understanding of the issues related to each Community after eight months and expressed in D2.2 *Communities of Practice establishment*. Further analysis was obtained during the Dec 4 2013 Workshop in Paris and discussions and interviews with the *CoP* leaders and experts.

This Deliverable will constitute the starting point for the longitudinal impact analysis to be done until the end of the project, thus allowing the analysis of the impact of the project on the established communities and the research uptakes it may lead to. It will study the initial needs of the *CoPs* as well as the impact and sustainability the project will have on the designed communities, and permit a global coordination regarding needs, methodologies and actions.

2.1 Contextual problems

One of the main objectives of the Presto4U project was to establish the 9 Communities of Practice. Through these communities, the audiovisual domain is well covered and represented; there are however important differences between historically established communities and the more recent ones, for which the project has engaged in helping building and following.

Broadcast, film or sound communities existed long before the project. They are generally structured in associations or federations and have started a long time ago to bring their problems in common and to work on issues concerning the whole community[2].

Other communities, while sharing concerns, as the Art and Music communities, or the Education and Research communities, are only discovering the need for Preservation and only recently have started analysing the situation and undertaking actions in this direction.

This brings some unbalance between the experiences and perceptions of problems; while established communities have in their majority faced the general preservation issues and undertaken digitisation, new communities are structuring their knowledge in order to address from a general and common point of view the issues concerning their activities. This difference can be seen in the feedback from the communities and how they undertake discussions and analysis of their problems.

Established communities have a good perception of their preservation problems, have already started acting towards digitisation and in most cases are accustomed to deal with preservation issues. Growing communities have worked in information sharing and gathering, analysis of the situation among members and identification of the main problems that hamper digitisation and preservation.

When analysing the existing information produced so far within the *CoPs*, it is easy to identify recurrent subjects that are regularly presented as an obstacle for their activities. All communities are concerned with the essential fact of digitising and preserving audiovisual contents, so whenever there is audio, video or film concerned; there will be many common

issues regarding the technical aspects of long-term conservation. More general issues concerning rights or content management will also appear, with clear differences and potential perspectives, or unsolvable issues, depending on the type of Community and the level of awareness.

2.1.1 Getting Digital

Before even getting into Digital Preservation, a general concern is just getting into the digital world: digitising collections. Independently of financial issues, there are still many uncertainties in this domain; organisational issues, cost evaluation and planning, format and carrier choices, or finding suitable Service Providers to perform parts or all of the work, are still a major issue for all communities, even before thinking of having a DMAM system to manage and give access to contents! This is probably still a strong problem for many contents holders and with time, the risk of losing analogue contents is growing rapidly. It can be said that the necessary information on how to act and decide is already widely disseminated, through numerous European and national projects who have worked on these issues and brought insight and comprehension on guidelines, methods and services for getting into the digital realm. However there are still many obstacles and decisions difficult to make, independently of finding financial resources, which have never been very high in the domain.

An efficient Centre of Competences as the *PrestoCentre* is not only a result of these initiatives, but also a reference point for anyone seeking the necessary information and advice. And it covers not only analogue digitisation, but also early born-digital contents whose formats or readers have become obsolete. Its aim is to provide orientation, information and services to accelerate the digitisation process and to do it in the perspective of long-term digital preservation.

2.1.2 Being Digital

Once the digital breach has been overrun, generally content owners discover there is another problem concerning the longevity of their digital assets. The problem is two-fold:

- How do I keep my new digital files accessible through time?
- How do I ensure that digitally born contents are managed and kept safely?

Two similar issues, with however some differences mainly in the way they are approached: when analogue contents have been digitised, there is a consciousness developed through the process, of the complexity and fragility of audiovisual contents. Also, important decisions have been taken concerning digital formats and the storage systems in which they are kept. When contents are born-digital, content holders often don't have previous experiences in content management and tend to think that there is no problem associated to digital contents, that they will last forever and no special attention has to be put on them.

Indeed a strong pedagogy on the risks of digital corruption has still to be undertaken in many communities, which have built up collections, used them and after some time discover or painfully learn the risks of not addressing long-term conservation policies and actions.

The second strong issue with digital contents is what to do once you have all that digital stuff: how to manage it, structure it, organise, access it and keep it in a management and preservation environment? Here is where DMAMs enter as well as all the related issues: metadata, transcoding, reformatting, restoration and, of course, long-term preservation.

It is not simply by buying an information system capable of managing content-holder's assets that the solution can be found; it is by clearly analysing the context of production, management and expected uses, that efficient systems can be implemented in order to guarantee the best possible and cost-effective solution. In other words, understanding the problems, analysing the context and evaluating the means are the main issues, which lead to Digital management of Preservation.

2.2 Technological problems

From a global point of view there is the general shared issue about contents, programs, files, their different formats and how to be capable of addressing them and following technology changes. All communities have some kind of problem related to this issue and are faced with some kind of problem for which they need a solution. The second strong common issue is what to do to with the contents? Where to keep them, which information system, metadata scheme and all the issues regarding content management? The third strong issue is how to interact with the surrounding world and communities: should I be compatible with other collections, do we need to use same standards, formats, publishing schemes and tools? As we will see there are different ways to answer these questions since the way they are focused may strongly differ.

One issue lies in the problems; a totally different issue is how content holders are prepared to face them. There is indeed a technological rift, defining two strong trends or profiles among the nine communities. On one side well established archives with a tradition of managing audiovisual contents of different types with policies, missions, and sometimes rights on the contents they hold. On the other side there are activity sectors, generating through their activities audiovisual contents, and once the accumulated material starts reaching important volumes or when some unexpected loss appears, they need to develop a preservation an migration strategy in order to keep and ensure the longevity of their material. The main issue with these activity sectors is that they often don't have the mission, the needs and the know-how to address digital preservation, and are thus faced with problems for which they are not necessarily prepared.

Established audiovisual archives working in digital preservation, have organised their work through a workflow. This workflow, which has inspired the OAIS[3] archive model, is based on the Reception – Delivery trend: on one side you receive contents, on the other side you have to deliver them for use. This model already implies the already expressed fact that audiovisual archives are based on reuse of material, and most of their strategies depend on this perspective. Actually an audiovisual archive is less an archive than, for example, national archives are, in the sense that the later have mainly a conservation responsibility, while audiovisual collections, while assuring conservation, are directly concerned with access. Between Reception and Delivery there are a series of actions, allowing this process to take place: description, data management, storage, administration and preservation planning. To these issues quality improvement issues may be added, in the delivery process, like transcoding, restoration and more recently quality upgrading.

OAIS Functional Entities
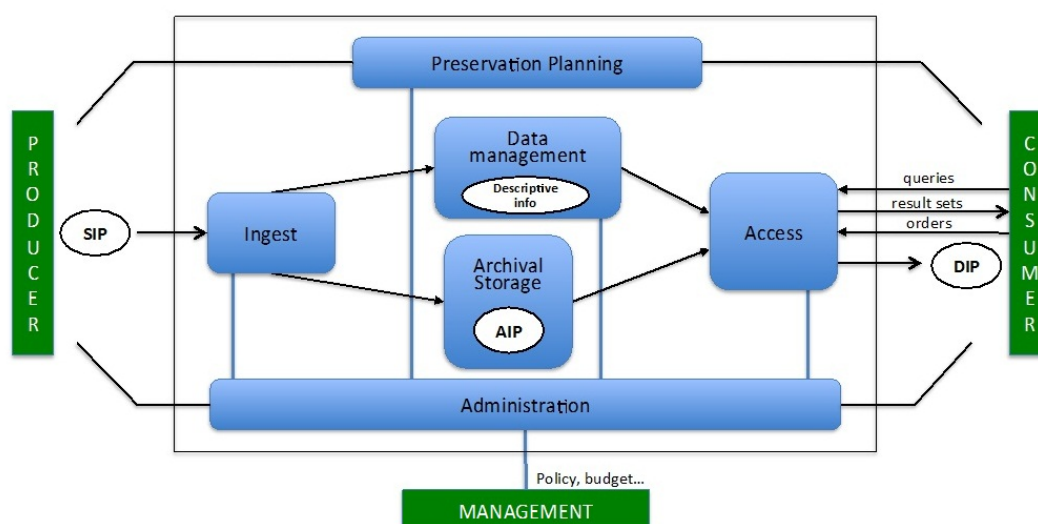
SIP = Submission Information Package
AIP = Archival Information Package
DIP = Dissemination Information Package

Established archives have then a clear workflow in which there may be local problems or issues, however the process as a whole is well understood and managed.

For activity sectors, accumulating audiovisual material and needing to develop a preservation strategy, the situation is less clear; they understand there is an issue related to long-term conservation, however they have not enough experience and practice in order to address it. They are facing the problems other communities faced some decades ago, when production contents could be kept for reuse and after a certain amount of time collections started growing and becoming unmanageable unless a strict process and tools were put to work in order to structure, describe and organise contents. In earlier days, long-term preservation was not a subject of concern; analogue contents were supposed to last and the archive organisation was based on effective documentation (cataloguing as it was called) for retrieval of contents based on a reuse model. Audiovisual archives developed in that period (in the same way libraries or archives developed this methodology centuries ago) methodologies and processes to guarantee effective management and fast access to all their contents ("fast" is a relative word here, since in analogue days this meant sending somebody to the good shelf to find the good box containing the program).

Today activity sectors of some *CoPs* face both problems: content management of more or less large collections, and preservation planning and management. These two constraints are increased by the community factor, where there is a need to interact and exchange with other content holders members sharing similar responsibilities and activities, and the fact that their global aim is to make contents accessible on Internet.

2.3 Legal problems

Legal issues are the second domain of concern for content holders; the situation among them can be totally different, often related to the technological situation since many of the undertaken actions may depend on which rights the content-holders have and what benefits can be obtained. The two extremes of the situation are being in possession of rights regarding the held collection, or having to ensure a conservation and preservation responsibility without any rights attached to the contents.

Audiovisual productions imply a complex legal environment due to the numerous sectors of activity concerned with production. Production generates large numbers of right-holders, who have contributed through their work to the final audiovisual product or program. It is not only the number of right-holders that makes right management complex in the audiovisual domain; it is the nature of production and reuse of material that generates a quite unique situation. In the audiovisual domain there is the concept of work and the concept of excerpt. A program can be seen as a unique object used and sold in its totality as a book or a painting would be (it is what happens in the feature cinema domain). However in many sectors of activity (broadcast, documentary) the concept of excerpt is very strong, as parts of audiovisual material can be used in new contexts. A good example of this is TV news, where elements from different sources, contemporary or archive-based, are integrated in the same program. This already implies a complex legal description of a news program; however parts of that program may be reused in the future in other news programs or new productions, thus creating a cascade of chained rights resulting in the need of an archaeological-like work in order to identify the original right-holders, potentially consuming a considerable amount of efforts.

To make the situation even more complex, when a content-holder authorises through right clearing and payment the use of an excerpt in a new program, he is not selling the rights but allowing the use of the contents on a defined geographical perimeter and for a fixed time duration. It is a kind of rental service, potentially difficult to follow since, if a producer sells his product to a third party, the third party should go back to the original right-holder to negotiate an extended use of the content. This complexity may increase with international exchange and through the different legal environments prevailing in different countries as well as the different definitions given to similar words.

The context is then complex and even if important projects as PrestoPRIME and now Presto4U are working on the simplification of international exchange, mainly through multilingual rights glossaries and through rights ontologies to simplify the understanding of legal context. From a practical point of view, managing rights implies highly trained staff and rich databases where all the information related to productions is kept. When a content-holder sells the rights to use a program, he then has to distribute the income between the different right-holders in function of their importance and participation to the production.

There is a large circulation of contents in many domains of audiovisual activity, and for many content-holders, commercial activity is essential to their survival since it bring the necessary income to maintain and develop their business. From a Digital Preservation perspective, content holders may have the rights but not the financial capacity to develop archiving strategies. On the other hand many contents holders have an archival mission (otherwise stated not limited in time), but do not hold any rights related to the contents they keep, thus preventing any commercial activity related to the rights or publication of the

contents online. National archival institutions and *Cinémathèques* clearly are in this situation[4]. Many institutions can only put contents online when they are in the Public Domain.

2.4 Economic problems

Technological and legal issues are directly related to the economic situation of the collection-holder. Either content-holders are in the Public or in the Private sector.

When institutions are in the Public sector they tend to have a regular financing through State funding. State funding may provide the main bulk of money needing to ensure the preservation mission these institutions have; however a strong technology shift as analogue to digital migration can be difficult, or in some cases unattainable due to the extra financing required.

For Private content-holders, the situation can be very critical from a Preservation point of view. Many of them have simply collected contents produced by their activity, and try to keep their contents within the commercial circuit; however this may not generate sufficient revenue to ensure preservation planning and actions. There are some new solutions appearing in the market, as the fact of subcontracting preservation to third parties (as will be seen in Chapter 5); however, independently of cost, the technical and procedural difficulties observed in audiovisual preservation may be much worse here, since many content producers are not at all concerned with archiving and lack the knowledge and experience.

Probably for both public and private partners the main obstacle comes from the complexity and difficulty of planning and making cost-calculation for digital preservation. Long-term planning and budget distribution through time imply a regular functioning environment, this is extremely difficult to provide due to the variability of public and private funding, mainly in difficult economic periods. Of all the problems, funding is probably the most crucial one since the cost of long-term preservation implies a constant and regular cost, often unattainable for many content-holders.

2.5 Two distinctive attitudes

The analysis and reactions of *CoPs* leaders and experts, done at the beginning of the project, can be grouped into two major tendencies depending on if content-holders have an archival mission or if they are simply holders concerned by preservation:

1) Established archives and memory institutions, who are conscious of the implications and actions related to long-term audiovisual preservation. They have a good understanding of technical, legal and economic issues and have identified specific needs preventing today part of their activity to be correctly undertaken.

2) Content-holders with no archival experience (sometime keeping very large collections of audiovisual contents) or mission who have to build up a preservation strategy and don't have the knowledge and/or staff to undertake it. These communities see audiovisual preservation as something they have to do, however they don't want this activity to interfere in their regular activity. This is why they tend

to look for turn-key systems, capable of managing their contents and managing long-term preservation quasi automatically, as a back activity requiring little control and effective through time.

Both tendencies share legal complexities and economic needs, the main difficulty will be for them to get into the digital world, this will have a first and very strong economic impact in the functioning of any institution.

2.6 Conclusions

In the next chapters a detailed analysis of technological needs was presented. The two strong trends presented up to now are a fairly good knowledge of preservation actions as can be found in established audiovisual archival institutions; and a need for effective, easy to use technology to take care almost automatically of preservation of contents for institutions or activities dealing with audiovisual content however not being an archive with preservation missions.

Some local problems exist which will be analysed and dealt with during the Presto4U project. The wide diversity of encountered problems, where some of them like financing are really difficult to deal with, give a good vision of the complexity and state of advance of the audiovisual community as a whole. Circumscribing problems to *Communities of Practice* is a good strategy, since it creates a feeling of shared issues among the members and group disperse problems within usage conditions.

3 Common issues concerning several or all communities

The intention in this chapter is to analyse problems whose generality may allow them to be addressed globally. It is the mainstream of audiovisual preservation, where all communities gather and share knowledge and issues common to most of them.

3.1 Technological issues

As explained in chapter 2, there is an issue, common to all communities, and this is how to keep contents. The main implication here is related to the migration process or the fact of performing the indispensable actions that will permit a content to be re-accessed.

3.1.1 Migration, the problem all communities share

Migration, considered as the general process, leading to the transfer of some content to another carrier or format is indeed the highest shared concern. Migration covers many different actions, and may be globally defined as: "performing some action on a content that will not modify fundamentally the original content while modifying the access environment".
The different situations in which the word is used are:

- Analogue to analogue migration: this process transfers analogue content from one format to another. No more used except in the Cinema domain where film copying for distribution and preservation is still being done, however since cinema is shifting to digital this kind of migration is diminishing strongly in volume, potentially causing a problem for the future for film preservation.

- Analogue to digital migration: also called digitisation, it defines the process that takes content from an analogue format and carrier to a digital format. In the preservation area it implies going from an analogue master to a digital master. This is the main action that many archives and content holders have been performing in the last 15 years in order to preserve their contents from chemical degradation. This activity is dependent on the availability of analogue playback devices, less and less available, and on the possibility of still accessing the original analogue content. It is important to stress that digitisation is a part of digital preservation, when digital contents are kept as files and integrated in an information system (or DMAM) capable of managing, checking and transferring files. Volumes of analogue content are impressive (millions of hours) and quickly disappearing.

- Early born-digital to contemporary digital migration: It is a different kind of digitisation, which implies bringing the first digital formats to a current accessible digital format. It is a highly complex task since it implies finding or emulating the initial playback systems and environments. Many former digital carriers systems are definitively inaccessible. The amount of contents is not so high, however a number of original productions have been lost before due to un-stabilised formats and carriers.

- <u>Migration from a file format to the same format on a different carrier</u>: One of the basic actions of the process called migration; it follows a cycle of carrier and CMS obsolescence with a cycle of 5 to 7 years between each migration, in which a new carrier is used to store the same content while keeping the same format. It can also be done to "refresh" content; meaning that some problem has been identified on a storage device or on the carrier, and a copy is done in order to ensure the integrity of the file.

- <u>Migration from a file format to a new file format</u>: It is a major shift in the migration process and also a major decision for any archive in its lifecycle, it is necessary when a format becomes obsolete. It has to be severely tested mainly when using lossy-compression formats, in order to be sure that the process will not degrade the initial quality of the content or introduce visual artefacts in the contents. It may help reducing storage volumes.

- <u>Migration as an industrial process for longevity</u>: It implies carrier migration and format migration in a controlled and robust information system, with processes for quality control, error and integrity checks as well a strategy for disaster recovery using external storage systems or copies. It is the strongest and more effective method found up to now in order to ensure long-term access to audiovisual contents.

As seen, the same word describes quite different applications and tends to generate some confusion among users and communities. While established digital archives are well acquainted with the process and implications and have adopted effective preservation strategies; many content-holders ignore the methods and actions to undertake, and even more when it comes to individual storing of home-produced contents, all at high risk of loss.

Migration is a common issue for each community, with however local differences, which may make digitisation and preservation a particularly difficult task. Any improvement in this domain is useful to everybody, since whatever activity is done there is always a migration issue associated.

3.1.2   Metadata organisation and extraction

Metadata is the most crucial component of an archive. Contents with no metadata are almost un-usable, and alongside with increase in volumes, content description has become the central activity of archival institutions. The situation is far from being homogenous; well-established communities have rich descriptions, helping retrieval and access; while in many other domains metadata can be extremely simple or inexistent, making content retrieval and usage a difficult task.

Poor or defective descriptions and metadata are a clear problem shared by all communities. Many content holders don't know what they are keeping and a strong challenge for them is to understand and evaluate what they are keeping, this has a strong incidence on the preservation policy, financing and strategy. The time-based nature of audiovisual content presents a further difficulty, since contents often have to be described

on a sequence basis. As suggested earlier, there is no perfect description of an audiovisual content, but different approaches depending on the nature of the collection, associated rights, production and reuse context, and the nature of interactions with the collection. The most appropriate or needed set of metadata will depend on the profile of the archive or collection , however for all contents and communities, identification, quality and preservation metadata are indispensable from a long-term conservation approach.

Communities here have either well-advanced description information, often manually generated (including if possible production-associated metadata, the information gathered during the production process which often is not included in the archival process), or need tools for enriching the existing information or providing information essential for content retrieval and reuse.

Technology could provide solutions to improve the documentation or to use poor metadata for search purposes. Two main issues are automatic segmentation of contents (identifying the different segments of an audiovisual content) and clustering into coherent segments based on common visual, sound or environment properties. Once segmentation is done, precise time-based descriptions can be introduced highlighting the most useful information for reuse, depending on the content production context. All communities share concern for segmentation associated to time-based descriptions, since this structuring helps retrieval of content.

Many technologies contribute to automatic metadata extraction, with two possible kinds of uses:

1) Automatic extraction as a complement for manual annotation; this is a very promising domain since it reduces the amount of time needed for manual description.

2) Automatic extraction as a retrieval process; where extracted information is gathered and cross-analysed in order to find contents through semantic clustering and structuring.

Different technologies contribute to these two approaches based on some kind of feature recognition and clustering; among the leading technologies issued from research can be mentioned:

- Speech to Text technologies: deliver a textual transcription of speech. It has different levels of accuracy depending on the used algorithms and the quality and dimension of the word reference databases. Speech to text is used to give rough descriptions of sound-based contents, or as a navigation and retrieval tool for large collections. It is often associated to semantisation, the process of clustering words into semantic structures, thus simplifying the understanding of topics within a document. The process of transcription is in the whole very effective; however issues as language models or intonation deviations, as well as background noise, may interfere strongly with the accuracy of results.

- <u>OCR: Optical Character Recognition</u>, allows identifying written characters on images and recognising words and names. Very effective for person identification since often the name of a person is associated to the presented image.

- <u>Face identification and recognition</u>: locates where there is a face on an image and, based on a reference face database, identify the person.

- <u>Object recognition</u>: allows image clustering and search based on object identification within images, it may be applied to recognise locations, objects, logos…

- <u>Content comparison</u>: allows searching of identical or similar contents among huge image databases; very useful for clustering images, searching for copies of a same image.

These promising technologies are already at work within information systems and on the net; as precision and calculation time improves, they are increasingly used to generate metadata for identification and navigation through contents. It is more difficult to apply them to huge databases of content like in audiovisual archives, where before introducing them it has to be checked that they do not generate a high level of retrieval noise, delivering irrelevant results.

3.1.3   Audio and Video Restoration

Among the common problems to all audiovisual contents is the issue of quality; the questions asked are:

- How did the image originally look like (generally unknown)?
- I can see defects; can I remove them?
- Can I go back to the original quality?
- Can I improve the quality to make contents look better?

Independently of the Archival process, quality issues are very important for reuse and commercial activity. Poor looking or sounding content will be considered unusable even if the content in itself is very rich or has a historical value. Research and technology providers have made great advances in this domain, however there are still pending problems, making digital restoration a continuous challenge for research and development.

In last years the effort was mainly put on eliminating perceptible defects or defaults within audio and video contents. A strong trend today appears with quality upgrading because of the need for high quality images for HD production and viewing. Even if a SD content will never become an HD image, there are potential levels of improvement.

3.1.4   Digital Media Asset Management systems

The key of Audiovisual and Digital Preservation! A DMAM first describes a hardware and software environment to manage contents, with a series of established tasks and

decisions related to ingestion, annotation, cataloguing, storage, retrieval and distribution of contents. It is the neuralgic tool controlling contents, their descriptions and actions. A DMAM may be improved with archival functions, thus taking in charge the control of content integrity and performing all the migration tasks necessary for longevity (there is still an external action to be undertaken by archival managers regarding storage systems and carriers that may change through time, as well as the management software itself).

There is a long tradition for DMAMs (even if initially were not called this way) since whenever contents are managed digitally there is a need for an environment capable of managing the contents themselves (the essence), different versions of contents, their descriptions, rights and any information needed to use or publish contents. Many institutions and archives started building their own management systems at the end of the nineties, bringing together different functional blocks and linking them in order to have a unified vision on collections and to centralise all actions and knowledge.

Institutions even had (and many still have) several management systems adapted to the management of specific tasks and environments; for example a management system for controlling location of contents on shelves and to follow the movements of physical media within a digitisation plan. Or systems oriented towards rights management, where specific databases contain all the necessary information and documents for right clearance and retribution to right holders. Through time there has been a strong and understandable tendency to bring all systems together in a unique system capable of managing all the actions needed in an archive, thus limiting redundancy of programs, documentation or efforts. Progressively building a DMAM can be a problem since some functions may not have been initially anticipated, and may result in later system dysfunctions or in process slow-downs.

Industrials have been proposing integrated DMAMs since the beginning of 2000, and a variety of offers exist, with different performances and profiles. Installing a DMAM is a major decision for a content holder, since before choosing a tool it is necessary to formalise the needs in terms of storage, management and usage, in order to find the best adapted tool and the proper dimensioning.

Independently of homemade or industrial DMAMs, there are also "ready to integrate" environments. The PrestoPRIME project produced an effective OAIS compliant environment called P4[5] (PrestoPRIME Preservation Platform) with tools and services for preservation management.

Media management systems are generally oriented towards large collections of contents and less well adapted to small or medium collections (they actually don't exist for personal collections). A first need has been identified here by audiovisual production communities, not yet engaged in preservation, for the need of simple tools adaptable to small and medium activities. This request for adapted systems is accompanied by a wish formulated by the Research and Teaching communities, that management systems should not only be easy but should take care almost automatically of preservation actions in order to ensure longevity and access.

This depends on a well-established understanding of the need for preservation tools and management; however many actors of these domains feel that the effort should not prevent them from doing their regular activities while ensuring that, in the background, preservation is being done. This constraint is very strong and implies that new ways of

thinking and applying digital preservation are imagined in order to cover these identified needs. A totally different solution would be to outsource preservation to third parties; this is promising for the future, however it may be very expensive for some communities and not necessarily adapted to their needs.

3.1.5 Formats and Standards, understanding and use

Many uncertain elements surround the selection of formats and standards. Their diversity constitutes a challenge for any inexperienced content-holder (and even for experienced ones), looking like a maze of possibilities, with no clear instruction on how to decide or even the need for them.

Formats: they constitute the first determining decision a content-holder has to confront when addressing digitisation and migration. Where am I standing and towards where am I moving? It could seem that the choice of a format is a question of common-sense, adopting the best one in terms of quality and amount of storage; however the tendency is to adopt the most popular formats, used by a majority of actors in a community. A format is a compromise between quality and storage, the ideal situation would be to use raw images with no compression; however until recently this was totally unthinkable because of the huge amount of storage it would imply. Today, as storage capacity is less of an issue, uncompressing starts to be an option for simplifying formats issues in the future.

When considering audiovisual archives, they tend to be at the end of the production chain, and inherit the format choices of the production environment. Historically they would accumulate production contents, catalogue and document them in order to make them accessible for future reuse. With migration due to carrier and format obsolescence, archives were confronted with the need of choosing the most adapted format in order to maintain the quality of their assets and somehow stabilise future migration needs (by choosing a format that may prevail for a longer time). The choice is in itself difficult since a large variety of production formats exists, with different initial qualities, and used for different purposes. In the past there was also a strong difference between *professional* formats (e.g. Betacam) versus *amateur* formats (e.g. VHS), and the difference in terms of quality was considerable. Today the quality difference is practically inexistent and amateurs produce high quality images in the same formats used by professional production. This separation is also important in terms of archiving; professional actors only produce in professional formats and keep professional carriers; while art and music collections as well as research and teaching collections hold a large variety and mix of professional and amateur formats thus preventing the adoption of common formats at the community level.

In past days the format was closely linked to the carrier (VHS is both a format and a specific cassette) while in the digital world formats and carriers are different elements of the existence of files. Choosing a format is different from choosing the carrier on which it will be stored. However the tendency today is to be as little as possible linked to proprietary formats and to choose open formats and well-established carriers.

Standards: they establish proved models for different kind of activities and bring homogeneity among actors of a domain. Standardisation is an active and increasing activity, introducing references and quality objectives. D4.2 brings a clear description of audiovisual standards and how the interact with uses and activity. Standards are related to software and very useful when making choices. Standards are stable and reliable since they are the result of a long process of validation and verification and are organised in order to provide interoperability between tools. For software development, standards are indispensable since they provide reference objectives and guarantee compliance to other tools and existing formats.

The situation among *CoPs* is variable regarding standards. Well established archives who have been developing their activity for a long time are less compliant to standards; they have progressively build their environments, procedures and quality control systems which have proved to be reliable for their activity. They may use *de facto* standards (more an established tradition than a standard) or develop a compliance approach towards standards trying to adapt their activities to them. However as long as institutions have internal workflows and don't need to exchange with other institutions or bodies there is no strong need for standard compliance.

The situation is much more complex for institutions or companies starting an archival activity or structuring their existing activity in order to be compliant to standards. The increasing number of standards makes any choice complex and, as in all domains, the initial difficulty is to define the needs with respect to the activity and to find the most adapted solution. This is why the Presto4U project develops a strong activity regarding standards, in order to analyse existing standards, explain their use and integration, and also to work actively with standards body when specific needs are identified, e.g. the work done by several partners with the MPEG standardisation bodies[6].

Codecs, transcoding… Other issues are related to former encoding formats and involve mainly Codecs (tools to encode and decode audiovisual content into a standard compliant format), or transcoding (decoding content in one format to encode it in another different format).

Formats are defined by standards, however software is needed to encode audiovisual contents and make them compliant to the format. Often a format is defined before the tools exist and once the format is defined, encoder development starts, led by industry. The encoders may have different encoding qualities or be more reliable and it has been seen that later versions of an encoder can be faster and more efficient in terms of quality.

Transcoding is a different issue; it can be transcoding for delivery, where an original high quality content is transcoded to another format following the delivery requirements (e.g. audio Wav files delivered in MP3), however the original master, or "mezzanine" formats as they are often called, is not modified. In this case it is normally a downgrading in quality, without significant content modifications. Another more crucial issue regards migrating from a mezzanine format to a new one, due to format obsolescence or extended storage capacity, allowing keeping the master files with a higher data rate. Mezzanine migration

has drawbacks, mainly when moving from a lossy compressed format to a new lossy compression format (MPEG2 to MPEG4) where a deep analysis of the degradation and artefacts due to first formatting has to be performed, in order to analyse how these defects may be amplified within the new format.

However, the tendency, at least in production and broadcast, is to move towards lossless compression and even in the future to move towards uncompressed signals. The format, as seen earlier, is a compromise between quality and cost (mainly storage cost); for archives having been for a long time in the digital world, sometimes dramatic choices had to be done some 15 years ago, important parts of their collections may be stored in highly compressed rates. As with DMAMs and tools, the tendency here is to move towards non-proprietary formats in order to have a full understanding of what the format proposes, based on the standard it relies on.

3.1.6   Open-source and free tools

The final difficult domain concerning technology is the choice of tools. Different possible choices are possible for archives and content-holders, depending on their technical infrastructure, dimension and complexity of the tasks to be done. The different models are:

- Using industrial systems, with the advantage of proposing a certain number of tools already integrated within a management system, checked and verified and with a maintenance service associated to it.

- Acquiring separated tools and integrating them internally into an existing system, or building a system from scratch based on existing commercial tools.

- Sub-contracting to integration companies who will make the selection process and integrate the tools and services based on the customer's requirements

- Working with open-source or free tools and adapting them through internal or external development to their needs

There should be a last possibility, developing your own tools; however this is practically not done any more, except for large institutions with powerful IT departments and an established experience in audiovisual software development.

As shown, the range of possibilities is very large; the choice will depend on the complexity of the tasks to be undertaken, and the finances attached to the activities. Hybrid solutions can also be imagined, bringing together different tools and making them interoperable through different development patterns.

It is clear that Open-source and free tools are a potential source of economy for the development of systems for content management. Generally free tools tend to be quite simple and not adapted to complex tasks; however interesting tools may be available.

Open-source is a very attractive software environment for internal system development; the principle is that a community of users contribute and participate to the evolution and improvements of software: this is a strong feature regarding robustness and adaptability to changing environments. However it also implies that an open-sourced tool is adopted by a community, which is not always the case. Many good examples exist but many tools have been left dormant and require strong development efforts to make them usable in a professional environment.

This is also one ambitions of the Presto4U Consortium, to identify and check usability of a number of existing tools in order to provide guidance and choice possibilities for content-holders looking for local or general tools for their activities[7].

3.2 Legal issues

Legal problems have been addressed in chapter 2.3, through a comparison of institutions in possession of exploitation rights against those institutions who only have a mission with no rights associated to this.

When analysing the community's issues regarding legal situation, there is a third position, ignorance of rights associated to the content. This is often the situation in research and teaching communities, where contents are used, published and distributed within no clearly defined legal framework. There are legislations allowing usage of published material, like the US *Fair use*[8] concept, strongly applied for research and education objectives. Most countries tolerate a certain use within these contexts; however it is far from being a rule.

An increasing number of content producers, mainly in the research and academic domain use Creative Commons[9] as a legal environment, so that information circulation is not restrained by strong copyright or legal protections. It is less the case with audiovisual archives, however there is a tendency to move towards such solutions, making content usage and circulation easier.

3.3 Conclusions

From this initial cross analysis of communities appear a series of unsolved or difficult to address issues, common to most content-holders. These represent the global issues of the audiovisual domain and are present each time preservation is addressed. In the following chapter a presentation of specific needs of each community will be done, in order to address sometimes small but important problems preventing the advancement of Digital Preservation.

4 Specific issues among the nine CoPs

After analysing global issues concerning all communities a close-up will be done here, community by community, to identify problems preventing or slowing down digitisation and long-term preservation uptakes.

4.1 Introduction

The specific community's problems can be the result of quite different causes; technology is not the only issue presenting strong problems, other issues as the legal environment, the economical context and the mission can be as strong as the technology one. To make the situation more complex, it is often not a unique problem that slows or stops action: it is a series of small problems of different nature that, when accumulated, create a strong impediment to progress. Digitisation and Digital Preservation are relatively simple areas of activity; the processes are known, many tools are there, and archival science has existed for long time. However, unawareness and lack of knowledge on how to organise digitisation and long-term preservation may make them look as totally unaffordable for institutions and content-holders outside of the archival domain.

4.2 Broadcast CoP

Broadcast archives have done huge efforts to move digital; their high reuse of material implies a fast and effective archival system permitting to find any content in a reduced amount of time. In the broadcast domain a strong accelerator for digital file management has been the migration of production environments towards digital production, mainly for the News. This generates digital files that can be easily stored near the production environment, and reused quite easily. It also puts a strong pressure on digitisation, since the analogue archive is considered too slow to answer to production needs.

Independently of financing, some problems affect this Community; among which we may find:

- The real benefits of digital production can be achieved only when ALL production has become digital. A mixed environment of analogue and digital production brings increased complexity for production staff. This situation is still very present among broadcast archives, which have not yet totally migrated to digital file management.

- Difficulty of linking broadcast archives between them. Broadcasters are one of the oldest communities, and archiving has started a long time ago based on local rules and practices. Metadata schemes can be quite different and often not standard-compliant. Strong efforts will be needed in the future to establish effective exchange protocols.

- Broadcast archives may not possess rights on many contents; however the archive has a preservation responsibility on them. This creates an uneven balance regarding the commercial perspective for the archive, who, without being a national archive, has to act as a de facto national heritage centre.

Problems are quite general and show the strong efforts done within this community in past years. Format choice and storage systems are quite stabilised and tending to be more homogeneous than in other communities.

4.3 Sound and Music archives

Within the Sound and Music community, the main localised problems come from the Music community. Sound archives are quite advanced in digitisation and preservation and share most of the concerns of broadcast archives, which also hold very large collections of radio material.

For Music archives, two issues are encountered, the second one being a crucial one in terms of preservation and reuse:

1) Music Museums possess more and more audiovisual material and they lack the knowledge and the management tools for describing, updating and migrating this material. Museum oriented archival standards are not really adapted to audiovisual contents, and integration efforts have to be done (or parallel databases kept) in order to deal with these contents.

2) Complex production environments: Music production works on proprietary sequencing and editing tools used for musical mixes. This concerns all domains of musical production: popular, classic and electroacoustic. Each production may contain up to thousands of audio-files, linked together, with different associated controls as dynamics, sound processing, repetitions… These environments have a strong rate of reuse, mainly in the popular music domain, however no preservation tool or approach exists for these productions, and production companies lack the necessary knowledge to understand the structure of these productions, as well as to guarantee even a medium-term longevity.

This second issue is also found within the Art domain and within all media art and scenic arts. It is the consequence of bringing together many elements contributing to the final work. In some cases there are even physical objects that have to be preserved in parallel. The major difficulty is linked to the possibility in this domain to replay works in the future, keeping the elements as well as the knowledge and formalising all the software devices so that an artistic event or a new version of a work can be shown again in the future.

4.4 Video production and Post-Production

This is a highly diversified domain, with different content-holder profiles, ranging from small auto-producers, to large commercial and even national institutions concerned with production and post-production. These may include:

- Private companies (freelancers, professionals with different backgrounds, B2B and

bottom lines are the main drivers)

- Public institutions (academics, policy makers, public service and publishing are main drivers)

This community is highly concerned with commercial activities based on the assets holders produce and modify. However in many cases the archival situation is extremely fragile due to the cost associated with digitisation and archiving, and often the lack of knowledge associated to the process.

Among the initial problems that can be spotted, we find:

- The need for a credible archive storage and retrieval model that would enable producers to exploit rights more easily. The same model might also allow producers to see benefits currently lost to them in reusing their own 'stock footage'.

- The safe, cheap, simple storage and retrieval of file-based video assets is the single greatest need of any producer heading into the tapeless future. Strong demand for tools or methods that would make it easy, simple and flexible to create versions for different destinations.

- Content tracking. Keeping traceability on contents in order to follow their use and reuse. This implies also more visibility of content as they move through the workflow. The use of fingerprints and watermarks is an important issue here.

- Moving away from videotape (and film) and gaining benefits from file-based production. There is abundant technology for media asset management, file formats and wrappers, file transfer methods, digital archiving, yet comparatively little standardisation.

- Using the Cloud as an Archival environment. Many questions exist here and this looks like a potentially good solution for this community for effective archiving while minimising internal costs.

In Video Production and Post-Production, there is the de facto archive issue, companies keep their own or somebody else's productions and so become progressively an archive. In some domains, there is a link between production and archiving, in others archiving may or may not come. A Museum is concerned with conservation, a production company is not supposed to be. There is consciousness that production has an economical value (sometimes very high) however this does not necessarily imply that there will be a preservation policy associated to it.

4.4.1 Film collections and filmmakers

Film collections are the oldest existing archives concerning audiovisual material. Film has already undergone a strong technology shift, quite as important as the shift from analogue

to digital is, and this was at the end of the 1920s when silent film was abandoned with the arrival of sounding film. This created a huge black-hole in which it is considered that about 80% of existing production was destroyed or abandoned because considered as having no more interest for entertainment.

Until recently the motion pictures industry considered that film was the best carrier providing the best quality and a proved longevity. However in recent years the production of motion pictures has moved towards digital film, which proposes high-resolution quality and simplifies distribution and exploitation. All of a sudden film as a carrier was exposed to the risk of disappearing as a major production and preservation media. A great risk of loss of film may happen in the next years due to the fact that film conservation and reproduction is an expensive process (10 times more expensive than audio). This opens the perspective of a new black hole situation, as a hundred years ago since a large part of film collections may not be preserved in time.

Among the identified issues:

- A variety of responsibilities for content keeping: small filmmakers generally struggle to keep their own content, because they don't have proper storage conditions for real film. They also struggle with digital technology. There is a partial overlap here with the *CoP* for individual collections, because a really small filmmaker is an individual with few digital technology resources beyond those available to any individual. Film collections exist to keep content made by others and do it very well, but face a revolution in moving from reels on shelves to digital cinema packages (DCPs) in some kind of mass storage.

- The digital cinema packages (DCPs) for distribution are lossy. EDCine promoted using a high-data-rate lossy JPEG2000 encoding ('hardly lossy') as a mezzanine format, so that versions of lower data-rate could be computed efficiently. Filmmakers now have identical technical problems with video production: the output is files, which need mass storage and digital preservation.

- Film collections typically have no intellectual property rights (IPR) on content in their collections, with the important exception of films that are made under a nation subsidy programme. For those films, the national collection can negotiate or even dictate specific rights regarding access. Filmmakers have the same moral rights (where that concept is recognised) as any other creator of new works.

4.4.2   Footage sales libraries

In the footage sales *CoP*, a distinction should be made between archives making preservation and developing on top of it a footage sales activities, and companies making a business out of footage sales, with no clear preservation strategy attached to their activity. The first ones count on commercial activity as complementary revenue for their general activity, while the later depend on their activity and have little resources to afford

digital preservation.

The general profile defines a community with a large variety of media and formats, lack of standards and workflows for preservation, lack of systematic use of descriptive metadata, lack of commonly adopted DRM systems, and lack of knowledge concerning digitisation and preservation. Many actors in this domain think that making a digital copy is enough to assure preservation (a point of view shared by other *CoPs*).

Clearly this Community overlaps activities of other Communities as the Video production and the film *CoPs*, however the particular complexity of content sales, in terms or rights, pricing, tracking and reuse issues, justify a specific community to deal with these issues.

4.4.3   Research and scientific collections

Research is directly concerned with audiovisual material, a unique source of study, and for capturing information related to scientific activities. It is present in all domains from science to humanities, and has grown strongly in recent years thanks to the possibility of content sharing and publishing.

The main problems research has regarding producing and using audiovisual collections are probably metadata and rights. Collections need to be documented in order to be reusable. The problem of documentation would break down into: who does the work and what standards are followed? The general area is too diffuse to make it easy to talk about documentation standards. There is no precise documentation environment and the issue is that nobody wants to work on the documentation, because creative people prefer to move on than to provide detailed information about what they have already done.

One of the main concerns of this Community is datasets for study and experimentation, free of rights and usable by different members of the domain. In scientific studies, when an article has been published, the results have to be sharable and experimentation redone with the same set of content. This can be a strong barrier in many sectors of activity because of the lack of sharing possibilities.

This Community started structuring itself with the Presto4U project; as within the Teaching community, an important amount of awareness actions have to be developed in order to explain the risks of improper use of contents and the consequences of loss of content for scientific activities due to lack of preservation actions. Only from there can the community start organising issues for preservation, independently and as a group of users.

As for the learning and teaching repositories *CoP*, Creative Commons presents a good usage perspective for content production and exchange. It provides the proper reuse environment and allows scientists and educators to exchange content and related studies. It is a good future perspective; however a huge amount of material is already protected and will be difficult to liberate unless strategic actions at European level are undertaken.

4.4.4   Learning & teaching repositories

This *CoP* shares many issues with the *Research and Scientific CoP* and an important part of their action consists in explaining risks and benefits of proper right management and the advantages and difficulties of long-term preservation of digital contents. They also share the large variety of formats, analogue and digital, professional and amateur, present in both Communities.

Both communities also would like preservation and migration to become a simple to manage process, where once the process is launched, automatic tools and services apply to collections and don't become the central activity of the Community. This implies the use of adapted DMAMs, conceived for medium size collections and permitting the content management on one side and preservation management as a back activity.

There is a need for clear standards and forecast on the evolution of Internet and its uses, mainly because most of their production is Internet published and shared. This is particularly strong in the domain of metadata, where descriptions are not homogeneous and no specific standard has been adapted yet by the community.

4.4.5 Art & museum objects, artists and their representatives

This *CoP* shares many issues with music production archives and collections, with specific issues due to the fact that contents may or may not be produced by audiovisual professionals and many different technologies may contribute to the existence of a work of art. Among the salient issues of this Community:

- Collecting institutions have no control over the technology used to create a work. Diversity is thus a problem, and with diversity comes obsolescence. When literally anything can come into a collection, the requirement to preserve technology becomes enormous.
- Video is a main issue, and for older works another main issue is presentation on CRT displays.

- Classification is another challenge (what is the work, how is it described?), as well as documentation, maintenance of the technology needed to exhibit the work or preservation of the content of the work (which itself is a problem of definition: is it the signal on a videotape, or is it more than that?).

- Moving the signal (there is a signal somewhere in 'the work') from an obsolete carrier into a file on modern digital storage, and thereby losing the original artefact; or somehow keeping the original artefact in playable conditions; or trying to make a perfect copy of the original, a 'new master'.

- There are two basic problems with taking an existing work and 'going digital':
    o 1) Loss of the artefact;
    o 2) Loss of the original method of exhibition. It is too simple for engineers to say that all obsolete forms of technology could be simulated by appropriate use of new technology.

Independently of the complexity and diversity of formats, a main issue regarding an art collection is its intrinsic value depending on the physical object itself. If some element is modified the work may lose a part or even all of its value. Another associated issue is the quality of contents; an image or a sound can be of very poor quality, however if it was intended that way by the artist, it should not be modified, or restored and even less improved.

4.4.6   Personal contributions

This is the youngest domain from an archival point of view. It concerns contributions made by individuals to institutions as well as the problems people face with their own collections at home. The problems are similar to those of all other communities, however there is a strong lack of guidelines and established practices to help institutions to receive contents and to manage legal, descriptive and commercial issues.

Personal contributions tend just to be contents; with no metadata, sometimes no clear ownership and no understanding from the individual's point of view of the value or interest they may have for an archive (many individual content-holders tend to overestimate the value of their donation). There is then from a practical point of view a need for:

- Guidelines for agreements with personal content-holders.

- Metadata schemes that will bring the minimum amount of needed information produced by the donator.

- Commercial partnership guidelines.

- Preservation issues (your content will be digitised and sent back to you with a digital copy of it). The receiver will preserve a digital copy.

For institutions there is the problem of integrating personal contributions within a larger collection frame. Lots of cataloguing and structuring work has to be done in order to organise collections in sub-collections or themes.

From a user point of view, where a content-holder is confronted with the preservation of his digital assets, the main problems are:

- Lack of dedicated tools or services for archiving hundreds or thousands of objects (an archival tool that would manage the essence linked to a metadata database, manage the life-cycle, make essential checks of integrity).

- Lack of archiving guidelines or know-how (how should I organise my collection, name the items, describe them?). Lack of knowledge about quality and formats (which is the best format for long-term, which quality definition?).

- Lack of interest in keeping or structuring their holdings, in this case they may look for institutions or associations willing to receive their contents and digitise them while making them accessible.

- Preventing the "drawer" effect, where individuals keep their digital contents on a DVD in a hidden place instead of making them accessible to others through online publishing.

Probably the main problem for personal archiving is the lack of tools and services for an efficient archiving. Since it is a domain in expansion, these tools may appear in the future. Social media also is playing a lot in this domain, online user-generated sites permit contents to be accessible and create the illusion that contents are being somehow preserved.

4.5 Conclusions

As mentioned earlier, problems are diverse and often linked in such a way as preventing actions to be undertaken. Awareness is however progressing and Presto4U will strongly contribute to this through understanding and fostering problems and solutions. The last chapter of this deliverable will analyse some economical implications that may constitute obstacles or opportunities for the future.

5 The economy of Digital Preservation

It was expressed earlier in this document the difference between "passive" and "active" preservation; both have costs however they don't behave in the same way through time. Cost analysis is a highly important issue in preservation however it may be approached from different angles, diverging from the historical "in-house" activity of archives and repositories. This chapter presents several issues related to cost and revenue of preservation as well as commercial and partnership perspectives.

5.1 The cost through time

Active or passive, preservation has always a cost. The difficulty lies in how cost is distributed through time.

For passive preservation, once the functioning and staff costs are established there is a slow regular increase in cost through time depending on the increase rate of collections and sometimes some one-off expense related to infrastructure evolution (as new buildings or installation of an air conditioning system).

For active preservation, distribution of costs through time looks more like a chain of mountains, starting with a very steep and long mountain called digitisation, this is the passage from the analogue to the digital world. This process may take years for large collections and can imply doubling the functioning budget for years and hiring temporarily or definitively new staff with different qualifications. Once the big mountain is over, the archive goes somehow back to a regular situation however with huge internal changes:

- Staff has changed
- Technological environment is different
- Regular check and control actions have to be planned
- Databases have to be built and interconnected

After a period of time that can last from 5 to 7 years, another mountain of smaller dimensions will appear called migration, meaning the transfer of contents from the first information system (which is becoming obsolete) to a new one: this often implies extra work and staffing as well as buying the new environment and the carriers needed to make if function. These small mountains will reappear every 5 or 7 years with maybe a higher peak if some transcoding has to be done at some point.

The good news is that with time the process starts to run smoothly, risks are minimised and migration becomes a kind of routine except from a financial point of view.

This implies that archives need to have irregular budgets through time in order to follow the preservation "landscape" and prepare for major events at precise moments. This is possible in many cases, however it can be crucial for many archives mainly due to difficulties in anticipating the purchase of information systems or increase of staff. These will be needed at a precise moment in the archives preservation plans.

5.2 The difficulty of managing punctual important budgets

A situation currently appearing in many countries and organisations is the arrival of an extra financing budget on one year, not carried on the following years. These punctual increases are of course welcome and useful, however progresses may be lower than desired, since they will only cover a small part of the needed investment. A difficulty added to this situation is the fact that in some administrations budgets cannot be transmitted from a year to the following one, meaning that the structure will have to spend the increase in a very short amount of time.

This situation can be very negative when it comes to large budget increases, since it implies that there will be a need for external services, requiring call for tenders, further lengthening the process. In order to prevent these uncomfortable issues; archives have to prepare beforehand their preservation and migration plans in order to be ready to apply them as soon as some extra financing appears.

5.3 The difficulty of cost calculation

Calculating costs is a complex exercise mainly for analogue to digital migration. Costs are highly variable depending on the conservation state of contents and the difficulty and time needed to do the job[10]. Many institutions have mainly budgets but no dedicated cost analysis procedures to evaluate the real cost of their internal actions. Cost calculation is a major activity undertaken in order to anticipate future needs regarding preservation and digital preservation. Many variables have to be considered, however it can be efficiently structured in order to separate the fixed expenses (staff, electricity, heating…) from variable costs: technology acquisition, carriers, installation of new equipment, infrastructures.

An effective anticipation of costs through time allows building annual budgets in advance in order to prepare institutionally for the expenses. Even if the exercise is difficult there are some positive aspects: the cost of technology goes down year after year so today's anticipation may be lower a few years later.

5.4 Budgets and commercial activity

For commercial institutions working in the B2B domain, it has been seen that preservation can be a challenge due to the cost it implies on the long-term and the complexity in managing it. However these commercial actors manage to make business out of their contents. It is then a strong temptation for archival institutions to commercialise their contents when they have the appropriate rights. The question here is how much revenue can an archival institution expect to make out of their contents?

For B2B activity this can be an important source of revenue with some variability depending on the nature of the archive, the potential value of its contents, and the facility with which the right clearance can be undertaken (content-holders may have rights but not a structured database to manage them). A rough estimation is that a well-organised, digital archive, with a good Internet selling platform could expect to make from 25 to 30% of their

revenue through selling contents to professionals. For B2C this percentage drops down to a very low level ranging from 2 to 5% of the annual budget.

B2B activities can be a good complement for the activity of an archive, however it is far from covering the costs of digitisation and preservation of assets. For B2C benefits are marginal and this activity is more done for improved image than from a financial standpoint.

5.5 The cost of outsourcing

Outsourcing is a potential solution for Preservation. Services are only starting to exist proposing a complete set of actions dedicated to the preservation of audiovisual assets. For the moment, services are oriented towards archiving in the strict sense of the word, keeping valuable information in a secure location, with sufficient guarantees of long-term accessibility (up to 50 years). This is not yet a preservation service, keeping and checking contents within a usage perspective as well as providing access. However actors in the domain are moving towards such types of solutions.

Outsourcing Preservation can be an interesting approach for content-holders wishing to provide a long-term perspective to their assets, while not possessing the capacities and knowledge to do it internally. This is a major evolution in archival perspective; archives always considered that their responsibility was to be close to the contents they keep and to constantly keep an eye on them. This is impossible in a digital world, and even if initially archives wanted to keep an eye on their hard disks or data-tapes, this attitude has strongly smoothed down through time and it is thinkable today to externalise this responsibility. For *CoPs* not directly engaged in Digital Preservation, this can be an effective approach, where preservation will look like an effortless action.

This has a cost, but more easily manageable since migration and checking are then the Service Providers' responsibility and he will charge for an annual cost per Gigabyte associated to a number of files to preserve and thus provide a fixed cost evaluation for an archive through time.

5.6 Living in the Cloud

The Cloud is another way of outsourcing Preservation. The Cloud as a concept has many definitions; it can be an internally designed environment where contents are distributed among a number of identified units providing distributed and redundant storage. This situation already exists in many companies and may look totally transparent for users.

Dedicated Clouds as services is an important future development for archiving, since, as well as for archiving companies, the responsibility of the service provider is identified and established through a service level agreement[11].

The big Cloud is another issue: the idea that contents are chopped in pieces and distributed in computers all around the world generates a sense of lack of control to content-holders. However there are already many services on the Cloud, many of them dedicated to storage. A storage environment is not a preservation environment: there are

high risks of loss since the provider has no responsibility in securing the contents through time.

Cloud services have the advantage of being potentially more economic than dedicated preservation services, they constitute a promising issue for the future, mainly for small and medium collections and for all those who want to externalise this task.

5.7 PPPs

Public Private Partnerships have generated intense discussions in past years, mainly through the agreements Google made with national and private libraries regarding book digitisation, delivering to content-holders a digital copy of their books while Google kept and used another copy[12].

However PPPs can be a good economical choice for content-holders in order to advance towards the digital world. In exchange of digitising assets the digitising company obtains some kind of commercial rights allowing the monetisation of their efforts. This is highly dependent on the rights the content-holders own on their contents. If they own exploitation rights then it can be extremely positive to externalise the commercial exploitation of contents through specialised Service Providers or platforms dedicated to this kind of activity.

Cost calculation is a strong issue here: what is the potential cost of the efforts? In how much time can the private partner recuperate investment? PPPs are long-term agreement, and require patience and good negotiation skills.

5.8 Conclusions

Economical issues are presented as an obstacle for all collections. Archiving and audiovisual content exploitation generate revenue but at low speed and in a long-tail perspective. Probably the best attitude is to have a clear understanding of the implicit costs of digital preservation and to analyse the costs and benefits of any action.

Digitising and Preserving are not quickly done actions; they take time to organise, to secure the appropriate funding, to train and prepare staff and to decide the functioning and use of the archive or collection once they are digital. The advantage of the digital world is to provide a new environment for publication, distribution and exploitation and thus to bring a totally new perspective for value.

Glossary

Term	Definition
CMS	Content Management System
CoP	Community of Practice
DCP	Digital Cinema Package
Digitisation	Transferring contents from an analogue format to a digital format
Digital Preservation	Having an established strategy and policy for long-term preservation
DMAM	Digital Media Asset Management system
HD	Video High Definition
Migration	Transferring content from one format and/or carrier to another one
OAIS	Open Archive Information System
SD	Video Standard Definition

Document information

Delivery Type	Report
Deliverable Number	6.1
Deliverable Title	Transversal CoP Impact Analysis
Due Date	Month 12
Submission Date	2/1/2014
Work Package	6
Partners	INA (leader), B&G, DFI, TATE, TV2, LUCE, KCL, CNR, IT Innov
Author(s)	Daniel Teruggi, Luca Bagnoli
Reviewer(s)	Jean-Hugues Chenot
Keywords	
Document Identifier	Deliverable_D6.1_Presto4U_02_2013_V1.0
Dissemination level	PU=Public
Document Status	Release

Project Acronym	Presto4U
Project Full Title	European Technology for Digital Audiovisual Media Preservation
Grant Agreement	600845
Project Coordinator	Beeld en Geluid
Contact Details	Sumatralaan 45, 1217GP Hilversum, The Netherlands. msnyders@beeldengeluid.nl

Document Status Sheet

Version	Delivery Date	Comment	Author
0.9	28/12/2013	Final Draft	Daniel Teruggi
1.0	02/01/2014	Release	Daniel Teruggi

References

Presto4U Consortium D2.1; CoP Knowledge Schema: A report describing the methodology for the unified collection and sharing of knowledge of the CoPs.

Presto4U Consortium D2.2; CoP Establishment: Collated report describing the objectives, the strategic context and the success criteria of each CoP, and membership of the expert working groups, their roles and responsibilities, approach and working methods, a work plan and deliverables, and a high level communication plan.

Presto4U Consortium D2.4; CoP Progress Report year 1: Collated report bringing together the individual CoP reports of year 1, describing their long term digital preservation technological needs, barriers and suppliers.

Presto4U Consortium D3.2; Research 0utputs Assessments v1: This report will describe the research outputs identified with the   potential to address CoP needs, together with results of the first year's analysis work on potentially useful   technologies.

---

[1] The FP6 PrestoSpace project www.prestospace.eu and the FP7 PrestoPRIME project www.prestoprime.eu have strongly contributed with tools and methods as well as global awareness through the technical and methodological analysis of audiovisual preservation.

[2] FIAT/IFTA for broadcast, IASA for sound and video, FIAF for film, AMIA for moving images are the leading ones in the domain. They are well established since the sixties and meet annually to discuss their issues. There is some overlapping among them, however they have strong identities and followers.

[3] OAIS, Open Archive Information System http://en.wikipedia.org/wiki/Open_Archival_Information_System

[4] A particular legal situation concerns Ina – *Institut National de l'Audiovisuel,* in France, a National Establishment however capable of developing Commercial and Industrial activities. A substantial part of the annual budget is provided by the French state, the rest is obtained through commercial and industrial activities. This kind of establishments is called EPIC (Industrial and Commercial Public Establishment). www.ina.fr

[5] https://www.prestocentre.org/library/tools/p4

[6] Presto4U D4.2, Interim Report on Audiovisual and Preservation Standards: Interim report describing relevant audiovisual   and preservation standards, providing evidence of use, guidance on their technological applicability and   suitability, and feedback to standards bodies.

[7] Presto4U Consortium D4.5; Interim Report on Software Tools Catalogue: Interim report, describing the user and design requirements   for a catalogue of software tools, processes, techniques and methods used at different stages of the research   lifecycle in long-term preservation of digital audiovisual media.

[8] http://en.wikipedia.org/wiki/Fair_use

[9] https://creativecommons.org/

[10] Some useful tools can be found at the PrestoCentre, like the Storage Planning Tool https://www.prestocentre.org/library/tools/storage-planning-tool-0

[11] An interesting document analyses the potential impact of the Cloud in the broadcast domain: https://www.prestocentre.org/library/resources/coming-storm-report-impact-cloud-broadcast

[12] http://books.google.com/