

France

Christophe Dessaux

Sonia Zillhardt

Ministère de la culture et de la communication

A1. DIGITISATION AND ONLINE ACCESSIBILITY

A1. Progress on the systematic gathering of information about current and planned digitisation of books, journals, newspapers, photographs, museum objects, archival documents, audiovisual material and the availability of overviews of such digitisation in order to prevent duplication of efforts and promote collaboration and synergies at European level.

Since 2001, there exists a catalogue of national digitised collections of cultural heritage that covers, at a national level, all the digitised collections as well as the involved institutions, whatever their sources of financing. Integrated since 2004 in the European MICHAEL project, this catalogue covers 1363 collections and 513 institutions since February 2008. It is accessible to the public via the *Patrimoine numérique* (Digital Heritage) <<http://www.numerique.culture.fr>> website that is integrated in the Culture.fr gateway ("Collections" portal).

The catalogue aims to provide a comprehensive inventory of the projects carried out in digital heritage. It is enriched through a network that is built around the "digitisation" steering committee of the Ministry that brings together the network of cultural institutions. The obligation to feed the catalogue is set down in the subsidy contracts for projects of the national digitisation plan.

A version intended for the institutions of University education (NUMES) is in development by the Ministry of Higher Education and Research (MESR), in coherence with *Patrimoine numérique* catalogue. It aims to create an online national inventory of digitisation projects in establishments depending on the MESR: universities (including university libraries) large research institutions (including the CNRS); and it will favour the structuring of the national digitisation documentation map of higher education and research (see question A2).

Furthermore, a national report on digitisation progress has been laid down every year until 2006 in the framework of the Minerva programme. At last, in the context of the NUMERIC study launched

by the European Commission in 2007, a work on the statistics on digital heritage matters is under way within the "digitisation" steering committee of the Ministry of Culture and Communication.

A2. Progress on the development of quantitative targets for the digitisation of analogue material in archives, libraries and museums, indicating the expected increase in digitised material which could form part of the European Digital Library and the budgets allocated by public authorities.

1) National cultural heritage digitisation plan

An annual call for projects on digitisation

In the framework of the national digitisation plan coordinated by the Ministry of Culture and Communication (the "digitisation" steering committee set up in 2006 with representatives from archives, libraries, museums, heritage), a call for projects is launched annually (last call in August 2007 to finance projects to be carried out in 2008). This call is endowed with 3 million euros a year. It is organised around six thematic programmes aimed at structuring the digitised content according to national priorities: territories; people; the French and the languages of France; art and archaeology; architecture; contemporary creation (see details in question A5).

The objective of the 2008 call is set within the objective of providing access to content promoted by the European Digital Library:

- develop a structured global offer of digitised content through thematic regional programmes
- significantly strengthen the offer of digitised cultural resources made up of coherent collections of critical size
- promote free and open consultation of cultural heritage by all web users
- favour individual uses by making emblematic heritage collections available.

2) Mass digitisation – books (BnF): a new contract in September 2007

The rapid technological evolutions and declining costs have allowed for a change of scale in the treatment of collections: as of 2006 the BnF (French National Library) began a mass digitisation of its heritage collections. Two projects, that will end in 2008, have already allowed for the digitisation of 30 000 printed documents and the text conversion of 60 000 documents already digitised in image mode available on Gallica. In September 2007 a new contract for “mass” digitisation was awarded by the BnF to the Safig company. It covers the text mode conversion of 300 000 documents over three years. The first works treated by Safig have been available since April 2008, the growth will then be of some 8 000 documents per month. In late 2010 the BnF will make available, via its online library (Gallica 2) over 400 000 print documents, in both image and text mode, representing over 45 million pages. In addition to these 400 000 documents there will be the press clippings from the five year press digitisation plan (3,5 million pages) as well as the documents stemming from the digital conservation of the BnF: specialised collections, printed or audiovisual.

3) Audiovisual collections (INA)

By increasing the government funding of the INA (National Audiovisual Institute) by 22.6% over the 2005-2009 period the state has given the INA the means to carry out a clear policy of diffusion of its collections on the internet while guaranteeing its preservation plan.

The professional collections of public TV channel archives, whose conservation depends on the INA, goes back to the 1940's for radio and the 50's for television. They are therefore on old formats that have altered over the years. The films and the soundtracks, that were based on cellulose acetate are the victims of a chemical reaction called “the vinegar syndrome.” The old magnetic tapes have undergone various forms of damage linked to heat and humidity. Furthermore, the playback equipment, namely the heads of the videotape players, are hardly manufactured anymore.

On a total of professional collections of some 1 500 000 hours, 830 000 hours are endangered

in various ways (damage to the media, obsolescence of the playback equipment...)

To meet the challenge, in 1999, INA launched the Preservation and digitisation plan (PSN) to guarantee the conservation of “old” collections, to facilitate exploitation of the collections, provide better access and to better ensure its mission of valorisation of heritage.

In late 2007, 470 000 hours of programmes had been transferred by INA from analogue to digital formats: 95 000 hours of television stem from recent collections; 375 00 hours are covered by the systematic preservation/digitisation of endangered collections (PSN), including 140 000 hours of radio and 235 000 hours of television.

45% of the total of endangered collections are now digitally saved, representing 22% of the Film branch, 100% of video and 28% of radio.

Since 2004 the whole of these digitised programmes are accessible online worldwide on the professional site “inamediapro.fr,” open to 6400 accredited clients, with a third outside of France.

Since 2006 the site “Archives pour tous” (“Archives for All”) on ina.fr has opened its radio and television archives to the public and provides 18 000 hours of content online (14 000 hours of television and 4000 hours of radio representing over 100 000 documents), with over 80% accessible for free.

4) Heritage collections of higher education and research establishments – the digitisation guideline for digitisation of higher education and research.

The Ministry of Higher Education and Research (MESR) finances digitisation projects in the framework of their contracts with universities and large research institutions that depend on it. These projects, to date, essentially concern heritage documents that are copyright free, like, for example the “Liber Floridus” project on the illuminated medieval manuscripts of the higher education and research libraries that is the object of a partnership with the Ministry of Culture and Communication – the “department for books and reading” and the BnF.

If the MESR has not, to date, set down quantitative objectives for these projects, it has carried out an expertise: pertinence, programming in the context of the four year contracts and results at contract term. Follow up indicators were set down with concerned

institutions allowing for adapted evaluation of projects. It also supports, in the context of the contracts, the acquisition of adequate material in institutions that present a sustained activity of digitisation and valorisation of their heritage collection (inter-university libraries, common documentation services).

Digitisation master plan

In order to avoid redundancy, to federate expertise and equipment and to valorise results at a national and international level, the MESR launched a digitisation master plan in 2008 with the objective of organising a digital documentary map of higher education and to articulate it with the other major national and international projects. The digital documentary map is based on the organisation along thematic axis of specialised subjects. Each axis is steered by a referral institution in charge of the scientific coordination of the digitisation carried out by the higher education institutions in the given subject or theme. Its role is to propose a shared digital plan for setting up thematic portals that provide shared and open access.

A3. Progress on partnerships between cultural institutions and the private sector in order to create new ways of funding digitisation of cultural material.

The digitisation of cultural material is publicly financed. This is namely the case for the large public institutions depending on the state (BnF, INA, Louvre, etc...) whose funding is part of the institution's allocated budgets. This is also the case for digital collections conserved by higher education and research establishments. In this framework, the support for the digitisation projects that participate in the construction of the European Digital Library should be included in the support programmes of the European Commission, as has been initiated in the current call for projects of the eContent+ programme. The national digitisation plan has a present budget of 3 million euros per year that finances projects on the basis of a call for projects. The majority of these projects are carried out in partnership with local government bodies that provide 50% public financing. In 2007 specific financing (10 million euros) for the European digital library was allocated in the BnF budget. The funds come from the "Centre National du Livre" (CNL - National Book Centre) that obtains

its funds from two taxes: one on the revenues of the major publishing houses, the second from on duplication devices on the French territory ("redvance sur l'emploi de la reprographie").

1) An agreement between the BnF and SNE to experiment a common access to works under copyright and works in the public domain.

The new dimension of cultural works written in the context of online libraries entails the search for a new partnership between editors. This must allow for access to documents covered by copyright, alongside the public domain documents. The BnF, with its strong commitment to the development of digital libraries, has been working closely with the "Syndicat national de l'Edition" (SNE - National Publishing Syndicate) since 2006 to define the conditions of online publishing of protected works, within strict respect of national, EU and international legislation on intellectual property.

A BnF/SNE workgroup comprising representatives from various French publishing houses was set up to this end. It commissioned a report on the definition of an economic model from a consultant specialised in digital libraries (Numilog). Numilog considered that in matters of access to works there has to be a paying model to ensure the balance of the publishing sector. He proposed a model of a "digital bookstore" based on individual licences for access to books that reproduces the sales model of printed books. The digital libraries, be they Gallica 2, the European digital library or the Francophone Digital Library, could thus offer full text search functions both for public domain or copyrighted works. Web users could thus freely consult the whole works from public domain as well as the notices, or even short excerpts, of protected works, with the permission of copyright holders. Virtual reading of protected works would be done through the intermediary of registered distributors, chosen by the publishers. Subscription forms to book collections or temporary reading rights could be proposed. The sale of paper copies would thus also be possible through bookstores. The recommendations of the study were approved by the BnF and the SNE board. Following the adoption of the economic model, a technical charter was established. It outlines the standardisation choices needed for the inclusion of data on copyright protected works.

Following this step, the BnF and the publishers carried out an experimentation of the economic model on the <http://gallica2.bnf.fr> site created by the BnF which has been presented to the public at the 2008 Paris Book Fair (Salon du Livre) from March 14 to 18 2008. Both a portal and digital library, it contains some 50 000 public domain books and at least 10 000 copyrighted works made available by publishers that wished to participate in the experiment.

This project is an experimentation of the legal digital offer: it is open to all publishers and e-distributors along the following lines:

The model proposed access to copyrighted works on e-distribution platforms that are authorised by the work's publishers. The publishers can thus play the role of e-distributor and ensure themselves the online communication of their works.

The experimentation will last one year from the last Paris Book Fair in March 2008. At this time thousands of contemporary publishing documents will be referenced on Gallica 2. Following the experimental phase, that will allow to better observe usages, an evaluation will be carried out by the different partners. Gallica 2 contains a federated search engine that allows users to make a unique search on the heritage collections and contemporary works. For the documents belonging to digital collections of the BnF, the search will be on metadata and plain text, if it exists. The visualisation and download will also be done inside Gallica 2. Access is free.

For the contemporary works referenced on Gallica 2, the retained model proposes search on metadata and plain text. The title of the work appears in the list of results, accompanied by a short excerpt that gives the terms of the user's search within the context.

The user can choose to learn more by clicking on the title. The enriched reference of the title then appears (table of contents, cover illustration, back cover text, detailed description of the work). The user can click on a link to the platform of the e-distributor where, according to the case, different free or paying services are proposed: excerpts, access to the whole document, download, online subscription rental or varying offers according to the e-distributor.

The graphic presenting the access platform to the digital collections of Gallica 2 is available at: http://www.bnf.fr/pages/catalog/pdf/gallica2_schema.pdf. The e-distributors that wish

to participate in the experimentation first have to sign a contract with the BnF. They must namely provide the agreements concluded with publisher's whose works they distribute. They must be able to set up an OAI repository where the BnF can harvest the document metadata using the OAI-PMH standard protocol: descriptive reference, table of contents, large format cover illustration, back cover text, full description of the work, indicator of copyright status... A URL link leading to the digital location of the document must also be provided. At the launch of the experimentation a given document can only be available on one e-distributor. Lastly, the e-distributor must provide content in simple XML format. The exchange of encrypted data will be done with the FTP protocol. This content will only be used to index the document and help the web user find it.

A technical charter has been set down by the BnF. It outlines the standardisation choices necessary for the inclusion of data on copyrighted books on the Gallica 2 access portal (document content transmission formatting, metadata format, data entry protocols, etc...).

In order to encourage publishers to participate in this experiment, the CNL (National Book Centre) proposes, in the form of a subsidy, a support for the digital distribution of copyrighted documents in the context of the test phase of the French prototype for the European Digital Library. The "Digital Policy" commission of the CNL regularly examines the demands. A key selection criteria of the CNL is the compatibility of the e-distribution and publisher offer with the documentary policy of Gallica 2.

2) A BnF/France Telecom technological partnership

In January 2007 the BnF and France Telecom signed a draft document of a future agreement intended to promote French and European digital cultural heritage to facilitate broad public access, namely through the project of the European Digital Library. A key aspect of this partnership deals with a contribution of technological competence by France Telecom R&D in matters of document organisation, of information gathering and access to results as well as treatment of natural languages applied to text. The "document network" theme was given priority treatment by the BnF through the elaboration

of a synthetic report on the overall metadata available. Within this theme the work will focus on the geo-localisation (recognition of geographical named entities, ex: toponymic entities for the association of documents and places).

The goal for the treatment of data provided to France Telecom by the BnF will mainly be to create links between different documents. These links can be direct from one document to another (for example, automatic recognition of a citation of a document by another), links between documents passing through a reference framework (entries of geographical authorities for geo-localisation, for example), such documents could also be supplied by the BnF.

The objective for 2008 is to test the treatment results and to demonstrate the potential for innovation in terms of services, to create demonstration interfaces, in the framework of the partnership, at targeted points deemed particularly interesting, such as, for example: the user path history in the document network; geo-localisation or map based access to a digital collection. The next phase of the BnF/FT work will focus on possible formats for data exchange, storage of links generated by the treatment and building interfaces, as well as the question of intellectual property and the possibility of integrating the tools developed in the context of this partnership within the Digital Library.

3) An agreement signed with INA by the TF1 private television channel for the world-wide exclusive commercialisation of its news archives for professionals.

TF1 decided to confide the INA with the world-wide exclusive exploitation of its news archives for the 1989-2006 period, destined for professionals. INA will digitise and progressively market the material worldwide over the next 12 years. This marketing mandate concerns all newscasts and news reports produced by TF1 over this period. The agreement complements the mandate that was signed in 2006 with INA on the 1982-89 news archive collection. This TF1 collection enriches the INA archives and will provide over 20 000 more hours of TF1 news programmes. It is thus over 30 000 hours that are available with the digitisation of the archives. This collection will also be available to the general public since it will be published online, as the digitisation progresses, on the public websites of both the TF1 group and the INA.

This agreement confirms the policy of marketing of "third party" collections (other than those made up of public television and radio archives) that INA has been engaged in: over the last three years numerous audiovisual producers (such as CAPA, AFP video...) have given INA a mandate for the sales and marketing of their archives.

Inamediapro.com is the B2B service of the INA destined for professionals (journalists, TV producers using stock shots or streaming, DVD publishers...). It allows them to watch the archive content, to segment excerpts and to make orders. This service, by late 2007, contains more than 450 000 hours of television and radio programmes, making the world leader in the field.

4) A high-level work group on the diffusion and digital reuse of public digital data set up in 2008 by the Ministry of Culture and Communication.

At the crossroads of politics, economics and law, the questions pertaining to the digital distribution of public data have long been the object of extensive study, regulatory and scientific work both on national and European levels. In 2007 the creation of the "Agence pour le patrimoine immatériel de l'Etat" (APIE - Agency of Immaterial Heritage of the State) marked the beginning of an inter-ministry reflection, essentially on the question of marketing of public data. Given the specific nature of its cultural heritage, the Ministry of Culture and Communication (MCC) wishes to carry out its own reflection on the digital diffusion and reuse of public cultural data, taking into account its mission as a cultural public service. This study will be cross sector contribution to this inter-ministry task.

It has therefore been decided to set up a high level work group in order to define the doctrine and operational modes of the ministry in matters of digital diffusion and reuse of public digital data. This group is composed of ministry representatives, companies and institutions of the sector and qualified personalities. It will start from the recommendations of the Ory-Lavollée report of 2002 and will shed light on the technological, legal and economic evolutions of the past years, drawing on international comparisons and the existing findings and results.

This workgroup, in 2008, will propose a "charter" for the diffusion and reuse that will seek to coordinate the actions of the Ministry of Culture and Communication

an the actors in the field. Once validated by the minister, the guiding principles of this charter will be declined in as many operational tools for the strategy of the ministry.

A4. Progress on the set-up and sustaining of large scale digitisation facilities.

The digitisation operations are essentially technical or are sub-contracted to cultural institutions in the framework of public tenders to private suppliers.

1) Mass book digitisation (BNF)

An operation of mass digitisation of its heritage collections (books) is being carried out by the BnF in the context of a contract launched with the SAFIG company in September 2007 (see question A2). Two projects, that will end in 2008, have already allowed for the digitisation of 30 000 printed documents and the text conversion of 60 000 documents already digitised in image mode available on Gallica. In September 2007 a new contract for "mass" digitisation was awarded by the BnF to the Safig company. It covers the text mode conversion of 300 000 documents over three years. The first works treated by Safig will be available as of Paril 2008, the growth will then be of some 8000 documents per month. In late 2010 the BnF will make available, via its online library (Gallica 2) over 400 000 print documents, in both image and text mode, representing over 45 million pages. In addition to these 400 000 documents there will be the press clippings from the five year press digitisation plan (3,5 million pages) as well as the documents stemming from the digital conservation of the BnF: specialised collections, printed or audiovisual. Furthermore in the context of its policy of preventive preservation, the BnF is progressively abandoning film based media for digital: the internal production pipelines are being strengthened and the production, if the rights permit, will feed Gallica.

2) Audiovisual collections (INA)

For the digitisation of its audiovisual collections, the INA is also involved in mass digitisation (Preservation and digitisation plan, see question A2). The INA is a unique research lab that develops new technologies such as the "Signature" technical process for legal management and the protection of internet content. A leading European pole for education and training in the field of sound and images, INA is today a major European actor in professional audiovisual education. Since October 2007, Ina'Sup, the higher school for audiovisual and digital content ("l'école supérieure de l'audiovisuel et du numérique") proposes two masters: Digital audiovisual production and distribution and Management and preservation of audio-visual and digital heritage.

A5. Progress of cultural institutions, as well as publishers and other rightholders to make their digitised material searchable through the European Digital Library. Describe in particular the progress on the application of common digitisation criteria by cultural institutions and, where relevant, private companies, in order to achieve interoperability with the European Digital Library and to facilitate cross-language searchability.

1) The common digitisation objectives defined in terms of content at the national level

The priorities in terms of the content presented below are organised to coherently digitise the overall media formats (printed documents, sound, photography, maps, manuscripts...).

- the national digitisation plan of the Ministry of Culture and Communication covers the overall domains (archives, libraries, museums, monuments, archaeology...) and media. It is structured around the six thematic programmes destined to organise the digital content according to national priorities: territory; people; French and the languages of France; art and archaeology; contemporary creation.

Domain	Collections
	Programme 1: The territories
Geography	Land register maps Maps and plans River, lake and sea

History of cities and towns	Deliberations registers of over 100 years
Local history	Regional and local press Archival search instruments Documentary files and dossiers Sound archives
National History	Emblematic documents of the history de France, Marie-Antoinette World War I and II Memory of slavery Literature manuscripts Medieval manuscripts
Programme 2: The people	
Genealogy	Civil status archives (excluding registers) Parish registers Name lists (genealogy)
Society	Social group memory Memory of immigration Archives of the work and jobs.
Intangible Heritage	Know how and crafts Song, traditional dances, tales and oral traditions
Programme 3: The French and the languages of France	
Linguistics	Oral corpus
Programme 4: Art and Archaeology	
Fine Arts	Reproductions of videos, paintings, sculptures, posters, furniture Graphic arts cabinets Decorative arts (wall paper ...) Photography Naïve art Portraits (writers, personalities, caricatures...) Musical instruments Games and toys
Archaeology	Underwater archaeology Stone art et and cave art
Science and technology	History of science et technology
Performing arts	Music manuscripts

Programme 5 : L'architecture	
Historic monuments	Architectural heritage (world heritage sites, European heritage label, castles, cathedrals, classical monuments...)
Architecture	20 th century heritage
Gardens	Documents, photographic collections....
Programme 6: Contemporary creation	
Music	Contemporary music Music of the world Orchestral music
Contemporary art	Works by creators
Design	Works by creators
Fashion and textile	Heritage and contemporary collections
Art books	Contemporary collections
Experimental film	Works by creators
Theatre	Photographs, videos, archives
Circus, fair and street arts,	Heritage and contemporary collections
– the book digitisation by the BnF in the context of Gallica is carried following a documentary charter http://www.bnf.fr/pages/infopro/numerisation/pdf/Charte_Gallica.pdf that defines the priority digitisation criteria to ensure geographic,	chronological, linguistic and content domain coverage and guarantees the balance with other print digitisation operations at the national level (networked digitisation). In the context of the contract for 100 000 yearly printed documents, the following themes are privileged:
Programmes on national heritage	History of France: regional history, history by period, historic biographies The institution of sociology as a field of knowledge History of education Emergence of social question and the foundation of social security History of French Protestantism and its influence in Europe. Reviews in social, artistic and literary sciences Natural science in the XIXth century: French contribution to the development and spread of knowledge in this domain. French medicine in the XIXth century: disease and its treatment French poetry and theatre Enfantina (children's literature) Art fairs Reference works on art Publisher and bookstore catalogues

Genealogy

Languages:

- History and diversity of languages
- Learning languages, both orally and in writing
- Translation of masterpieces of several national literatures
- Translation of the classics of antiquity

Daily press

Legal documentation

Genealogy

A French library for a Europe open on the world

(French works and documents dedicated to Europe)

Reception of European philosophy in France

Universal exhibitions and fairs in the XIXth century, showcases of progress and instruments of the spread of knowledge.

Sinica (documents written or translated into French on China)

Colonial questions

– in terms of audiovisual material the INA has embarked on an exhaustive digitisation operation of its collections, presented in structured way to the public via the “Archives pour tous” (“Archives for All”) site on Ina.fr: themes (architecture, fine arts, performing arts, film, economy, ecology, geography, history, justice, literature, media, music, domestic and international politics, sciences, human sciences, society, sport, television), personalities, eras and radio and television programmes.

2) Access to digital heritage: 3 complementary tools ready for integration in Europeana

Resources of the archives, libraries, museums, historic and archaeological monuments: the Culture.fr/Collections portal of the Ministry of Culture and Communication

The “Collections” portal, access service to the resources of archives, libraries, museums and heritage departments via the Culture.fr portal, was launched online in August 2007. This version integrates 29 new resources, bringing the total number of accessible heritage records to 3 million with 2 million illustrated with images. It includes the advanced search functions based on semantic functionalities and multilingualism (English, French, Spanish). The Ministry of Culture and Communication, in association with public institutions (Louvre, George Pompidou Centre, Quai Branly Museum, Réunion des Musées Nationaux...) and regional administrations, is pursuing the opening of its digitised resources to the general public via the “Collections” search engine included in the <<http://www.culture.fr>> portal. This “unique access point” covers the most diverse domains: architecture, archaeology, furniture, painting, sculpture, engravings, drawings, illuminations, live

performance, ethnology, maps and plans, archives, genealogy, photography, events... From a unique search request “Collections” searches in 29 different documentary sources and displays the overall results by order of pertinence.

Printed collections: the Gallica digital library of the BnF
The Gallica site (<<http://gallica.bnf.fr>>) was produced in 2000 and today provides access to over 250 000 digitised works, including press collections, and 100 000 images. An overhaul became necessary to remain within the evolutions of the Web. The beta version with the Gallica 2 interface (<<http://gallica2.bnf.fr>>) has been online since October 2007 and will co-exist with the previous version of Gallica until 2008 when all functionalities are operational and all the documents of the old version have been transferred to the new. The documents stemming from present digitisation programs will also progressively enrich the available corpus (100 000 print documents in text form every year). The site will also integrate the data on recent works available in digital formats via e-distributor sites (free or conditional access, see question A3-1) as well as documents from BnF partner libraries. In late 2008 Gallica users will also benefit from the new functionalities of personalisation and sharing in the Web 2.0 spirit. In late 2010 some 400 000 copyright free documents will be available in text and image mode. To these 400 000 thousand printed documents will be added press clippings stemming from the five year press digitisation plan (3.5 million pages) as well as documents from the digital preservation pipelines of the BnF: specialised

collection, printed and audiovisual. The conception of Gallica 2 took into account the user remarks collected during the BnF study (see: Report on the usages and the expectations of Europeana users, *Noémie Lesquins, June 2007* available at: http://bibnum.bnf.fr/usages/BnF_Europeana_EtudeUsages2007.pdf).

The present version of Gallica 2 opened in October 2007 provides access to books that are already present in Gallica: it has a full-text based search engine and offers readers the possibility to create a personal space where they can save their documents and insert page markers.

Audiovisual collections: the “Archives for all” portal of the INA

Since 2006 INA has opened its archives to the general public on the ina.fr portal, “Archives pour tous”. Ina.fr is today a leader of Video on Demand in terms of volume and quality of content: over 100 000 documents can be consulted online, corresponding to 18 000 hours of programming (14 000 hours of television and 4000 hours of radio). This offer is enriched by over 50 000 hours a year and has been enlarged to the online education field with “Learn” (“Apprendre”), an audiovisual learning tool. Over one million visitors a month surf the site.

3) Standards and recommendations to guarantee interoperability

The digitisation operations carried out at the national level fall under the inter-ministry interoperability recommendations (RGI - référentiel général d’interopérabilité) steered by the “Direction générale de la modernisation de l’Etat” (DGME – Directorate general for state modernisation).

The national digitisation plan of the Ministry of Culture and Communication aims to support initiatives that facilitate access to heritage collections by a large public. Beyond the obligation to publish digitised collections online, project heads are invited to make collections available through the common national, regional or thematic tools. The access through the national aggregate portal “collections” (see above) of the Ministry of Culture and Communication is privileged, in view of the portal’s integration in the European Digital Library.

Furthermore, coordination and diffusion actions of best practices are also being carried out to favour

the inclusion of interoperability standards. In this context, a national meeting on the OAI protocol was organised on October 8th, 2007. The BnF, within its network of associated partners, provides specific training on OAI, putting its expertise in the field at the service of its partners and validating the conformity of OAI repositories with the recommendations and standards on the matter. The Ministry of Culture and Communication (direction of the Archives of France, with the participation of digitisation steering committee) has published online a guide that outlines the specifications for digitisation. The goal is namely to take into account the recent experiences and the latest technical evolutions in order to favour the quality of digitised resources and their integration in access portals.

Concerning books and publishing, a technical charter has been set down by the BnF in the context of Gallica 2. It provides the normalisation choices necessary for the inclusion of data on copyrighted books in the Gallica 2 portal (content and document transmission formats, metadata and protocols for data harvesting, etc).

In the projects of the institutions that it finances through its four-year contracts, as with the contracts that it leads, the Ministry for Higher education and Research recommends the use of open technologies: the adoption, at the outset, of open technical and documentary standards that are recognised at the international level (XML and data models) and, further down, the international exchange data exchange protocol OAI-PMH.

A6a. Progress on mechanisms to facilitate the use of orphan works.

A6b. Progress on mechanisms to facilitate the use of works that are out of print or out of distribution.

A6c. Progress on the availability of lists of known orphan works and works in the public domain.

A6d. Progress on the identification of barriers in your legislation to the online accessibility and subsequent use of cultural material that is in the public domain – and the steps taken to remove them.

The Superior council of literary property (“Le Conseil supérieur de la propriété littéraire”) consulting organ for the Minsiter of culture, will soon examine a draft project with the goal to find procedures that improve the situation of orphan works.

The report will first define orphan works, quantify the phenomenon sector by sector and recommend possible actions according to the varying contexts between the work's categories.

It will also be preventive through mechanisms for the identification of the works.

Already, the project has revealed that orphan works are marginal in film, music and to a lesser degree, in audiovisual, and that they are more numerous in writing and still images, which could lead to diversified propositions according to the sectors.

The law on intellectual property contains legal dispositions that could resolve a number of situations. The article L 122-9 states: "In the case of a clear abuse of the use or non-use of copyright by the representative of the deceased author cited in article L. 121-2, the high court can order any appropriate action.

This also in case of conflict between the stated representatives, if there is no known rights-holder or in case of absence or decease.

The court can be seized by the Minister of Culture" (often the case for post card collections).

The article L. 211-2 of the same law states a similar mechanism for neighbouring rights: "Beyond any person that can justify a reason to act, the Minister of culture can seize the justice authority, namely if there is no known rights holder or in case of absence or decease.

It is, of course, essential to not confuse an anonymous work with an orphan work. In this regard French law provides rules that can facilitate the management of such works.

According to the article L. 123-3 of the code of intellectual property the anonymous works fall into the public domain after seventy years following their publication and can be exploited without the difficulties posed by the orphan works.

If they were to be in the same category as orphan works they would entirely lose their status of anonymous works.

The article L. 113-6 of the same code also foresees that the authors of anonymous works are represented in the exercise of their rights by "the editor or original publisher, as long as they have not made known their civil identity or quality".

The rights attached to these works are thus defended by a representative of the author or the rights holders.

B. DIGITAL PRESERVATION

B1. Progress on national strategies for the long-term preservation of and access to digital material. Describe the organisational approach, indicating the roles and responsibilities of the parties involved as well as the allocated resources. Describe the specific action plans outlining the objectives and a time-table for the specific targets to be met.

1) A national reference guide in the building stage

In the framework of the national "digitisation" steering committee, a specific effort has been undertaken since early 2007 to co-ordinate the actions of the ministry concerning the long term preservation of heritage. It aims to establish a referential that transversally brings together the existing recommendations of the different sectors and actors. This national reference guide will take the form of a reference web site, easily updated, and accessible on the "digitisation" site of the Ministry of Culture and Communication (online launch in July 2008).

2) The "Distributed preservation and archiving system" (SPAR) of the BnF

The recent acceleration and growth of digital collections (over 100 teraocets per year today) and the diversity of their formats, confronts the BnF with the challenge of their conservation: it is essential to securely and lastingly store the digital objects on a solid and effective foundation.

After one year of study the BnF launched the SPAR project, a full-fledged digital warehouse. Its conception is based on international norms that are recognised in the world of preservation of digital information (OAIS model).

SPAR is more than a simple storage space for secure data. It makes multiple copies of digital objects that allow one to anticipate further copies before definitive loss. But it also allows, through the precise and complete recognition of input data formats, to guarantee the continuity of access by making necessary transformations in case of technological obsolescence of computer restitution tools. Providing this guarantee implies a constant process of tech-watch on the formats, prototyping and tool testing. The overall arsenal is inherently foreseen in the conception of SPAR. Furthermore, SPAR allows to step back at any

given time to restore objects to their original formats. SPAR is a system put at the service of the community. It must guarantee that restored documents have not been altered. To do this, SPAR tags each stored object with a digital marker. In order to guarantee the rights of "communicability" of the used digital objects, SPAR relies on expert system of rights management that allows it to calculate the user licences of the digital objects and to apply necessary restrictions according to the user profile (researcher at the BnF, web user...) SPAR does not have the vocation to only by an internal tool at the BnF. With a will to mutually share expertise and costs, the BnF will open its system to other partners and institutions, thus providing the service of "third party archivist" of digital heritage. SPAR is a broad-scoped project that will be span a number of years. After having acquired the infrastructure in 2005, the BnF, in June, launched a call for projects for the production of the software components and adopted an orientation in favour of open source software to ensure maximum independence. SPAR will be operational in early 2009.

3) The development of electronic archiving

The pilot project of the PILAE platform steered by the directorate of French Archives (DAF)

The DAF, with support of the General directorate of state modernisation (DGME) is presently producing a pilot platform that will function during a transitory period from 2008-20012 for the National archives in order to receive, preserve and communicate the born digital archives produced by the central services of the state (material from data bases, bureaucratic documents or produced from electronic documents management systems, electronic message services, immaterial data-flow such as public tenders...). The pilot implements the standard for archival data exchange; allows one to receive the transferred archives in this format and control a part automatically according to the conditions set down by the agreements established between the different actors of the archive process, partly in manual form by the archivists of the national archives; to automatically convert input data to lasting formats, documents that are not in the archival target format; to securely conserve the data by replication on two distant sites and, finally, to transmit them via download. In the first stage the pilot will be operational for the producers

of the archived documents in the ministries, for the archivists in the different ministries and the archivists at the national archives. The pilot could also be reused as a model for other administrations that wish to implement an archiving platform in their own environment.

In the framework of the program for the development of electronic administration and the establishment of a general interoperability reference guide on the national level, the launch and follow up of the initiative "IF-08 digital archiving" (DAF and DGME): elaboration of the reference guide (secure archiving policy, exchange standards for archival data).

The DAF is also involved (as a project manager, The Archives of the state of Victoria in Australia being the head of the project) in the eArchiving project that aims to define an exchange standard between producers and archive services that was accepted by the UN/CCEFACT in October 2006. The standard had previously received the support of the ICA (International Council on Archives). A questionnaire for archive services around the world was distributed in December 2006 by the UN/CEFACT, ISO and the ICA. On this basis, Andrew Waugh, publisher of the project, wrote a first version of the standard in September 2007. An international call for comments is presently being launched.

The actions of the Ministry of Higher Education and Research (MESR).

The "Centre Informatique National de l'Enseignement Supérieur" (CINES - National Computer Centre of Higher Education) has been designated by the MESR for the lasting archival of data, namely the data resulting from digitisation operations. The archival platform of the CINES (PAC) was designated in late 2007 as the tool for ensuring a the lasting archival of digital documents from the research sector.

B2. Progress on exchange of information with other Member States on your strategies and action plans.

The exchanges with member states of the EU take place both in framework of procedures set down by the European Commission (participation by the MCC and the BnF in the member states' experts group,

participation of the Directorate of French archives in European archivists group EBNA), in the context of professional networks active at the European level (CENL for the libraries, EURBICA for the archives...) and through the participation in European projects supported by the Commission.

Thus, French participants (from the Ministry of Culture and Communication, the French National Library – BnF – and the National audiovisual Institute - INA) since September 2007 are participating in thematic EDLnet network (network for interoperability of museums, archives and audiovisual) in charge of developing the prototype, called Europeana following the proposition by France.

Two French members (Martine de Boisdeffre for the European branch of the International council of Archives and Christophe Dessaux for the MICHAEL association) are members of the "EDL" foundation, destined to manage the future European digital library and that brings together the European associations and federations of heritage institutions (libraries, museums, archives, audiovisual). The INA and BnF also joined the foundation in late 2007.

The French teams also participated in ongoing projects of contribution to the European Digital Library.

- MICHAEL, access portal to the inventory of European digitised collections. Its present broadening will allow for the extension of the existing service to provide access to the collections of 18 countries in all as of May 2008 (French participation: Ministry of Culture and Communication).
- The BnF teams are actively participating in the TELplus project that, since October 2007 is dedicated to the improvement of accessibility, search and collaborative functionalities of the future European Digital Library. The BnF is co-ordinator of WP3 (improving accessibility) and has a significant participation in the WP1 (OCR of digital content) with a contribution of some 7 million pages with a high quality OCR.
- In the audiovisual domain the INA coordinates or participates in numerous European and international projects destined to innovate in the domain of audiovisual heritage preservation and valorisation. Among these projects, some like K-Space, Mesh or Caspar bring together dozens of international partners (institutes, universities, private establishments, etc...). The INA coordinates

the PrestoSpace project (34 partners from 9 countries) for the development of concrete solutions for the large scale preservation and digitisation of audiovisual archives according to criteria of quality, speed and cost effectiveness.

At last, French partners are involved in the following proposals:

- In the framework of the 7th PCRD and projects relative to centres of excellence, the BnF is participating in the IMPACT project, proposed by the KB (Netherlands) for improving OCR procedures.
- Project for creating a European archive portal: France contributed to the project of the European portal (APENET), submitted to the eContent call closed on October 4th. The need for such a portal was set down in the recommendation of the Council of November 14th 2005. It was reinforced by the launch of the European Digital Library. The European archive portal should allow, if the project is accepted, to make accessible the search instruments and digitised documents conserved by the different European countries. The idea is to give a greater legibility to a complex European history, namely marked by numerous conflicts and population movements. The resources thus made available will also be open for consultation on the European digital library portal. The specific archive portal could represent an intermediary link.
- Finally, still within the October 4th eContent+ call, and in the context of the European digital library, France has participated in the ARROW proposal (BnF participation on the integration of copyright protected content) and ATHENA (participation of the Ministry of Culture and Communication, namely for the development of museum sector integration) presently in negotiation. There is also the INA participation in the VideoActive project for a European access portal to radio-television archives

The set up of a legal deposit for websites by the BnF is carried out in close collaboration with other libraries and national European archives brought together in the International Internet Preservation Consortium (IIPC). A large number of national libraries in the EU have joined the consortium participating in the software development and the effort for standardisation that favours interoperability and the coherence of the strategies adapted by the member states

associated in IIPC. The BnF, that ensures technical steering of the consortium particularly oversees the coherence and the interoperability of the solutions retained in the sector, both at a European and international level. It also participates in the promotion and the valorisation of web archiving within the international and European professional instances.

B3. Progress on legal provisions for multiple copying and migration of digital cultural material by public institutions for preservation purposes.

The French law "Copyright and neighbouring rights in the Information Society" (« Droit d'Auteur et Droits Voisins dans la Société de l'Information » loi n° 2006-961 du 1er août 2006 dite DADVSI) transposes into French law the European directive 2001/29/CE on "the harmonisation of certain aspects of copyright law and related rights" and amends the code of intellectual property (CPI).

In matters of long term preservation of digital heritage, the law now permits: "the reproduction of a work, executed to the ends of preservation destined to preserve the conditions of its onsite consultation by libraries accessible to the public, by museums or archive services, on reserve that these seek no economic or commercial advantage."

The law also makes provision for the following exceptions:

- The representation or reproduction of excerpts of a work (excepting certain categories) in the context of education and research, in exchange for a fixed charge (as of January 1st 2009).
- The temporary reproduction (storage) of works in the context of technical procedures allowing the use or transmission of the work via a network: this concerns the buffering done by certain internet access providers to optimise access to resources on the network,
- The reproduction and the representation (by persons or institutions open to the public such as libraries, museums, archives...) strictly reserved for personal consultation by the handicapped persons,
- The reproduction or representation, in whole or in part, of a graphic, sculptural or architectural art work by the written press, audiovisual or inline, with the sole goal of information and in direct relation with the work, providing the author be clearly cited

The same DADVSI law introduced in the Code of heritage an exception for reproduction rights favouring organisations in charge of legal deposit. In virtue of this text the author can not forbid the depositing organisations from:

1° Consulting the work on premise by duly accredited researchers on individual consultation strictly reserved for research ends,

It is the same for neighbouring rights holders:

«The artist, singer songwriter, producer of phonograms or videograms the audiovisual communication company can not forbid the reproduction and the communication to the public of documents cited in article L.131-2 under the conditions foreseen in article L 132-4.»

The text also states that the producer of a data base cannot forbid the extraction or reuse through making available the whole or part of the data under the same conditions.

B4. Progress on policies and procedures for the deposit of born-digital material. Please note how you are taking into account developments in other Member States in order to prevent a wide divergence in depositing arrangements.

The dispositions of the heritage code concerning the legal deposit (article L131-2) include digital documents: "printed, graphic, photographic, sound, audiovisual, multimedia documents, irrespective of the production technique, publishing and distribution, will be subject to a compulsory legal deposit, from the moment they are put at the disposal of a public via material media, irrespective of the nature of the media.

At the BnF it is the Audiovisual department that is in charge of receiving the legal deposit of all multimedia documents or data-bases distributed on a media. INA is in charge of legal deposit for radio and television, the CNC (Centre National de la Cinématographie) for film.

The decree 2006-696 of July 13th, 2006 makes provision for the BnF to be able to "demand deposit of a digital file, in substitution of the deposit of a printed, graphic or photographic document". At last, the title IV of the DADVSI law cited above (law n° 2006-961 of August 1st 2006) opened the legal framework for archival of internet sites of the French domain "Are also submitted to legal deposit the signs, signals,

writing, sounds or messages of any nature that are the object of a public communication in electronic form.”

This disposition is also cited in the heritage code (article L.131-2). The law foresees the “collection through automatic means” (Internet site harvesting), but also covers the possibility to determine other means in agreement with website publishers. The organisations in charge of this legal deposit are the BnF and the INA. The law foresees that a code or access restriction cannot hinder the collection by the organisation in charge of the legal deposit. The BnF intends to carry out this legal deposit through the addition of the three following procedures: mass automated harvesting of the French “.fr” domain; targeted, though automated, collection of sites along criteria that guarantee the continuity and enrichment of the legal deposit. These automated collections will be carried out along a calendar laid down by the BnF and which will vary for targeted collection according to the update frequency of the publications; in case of failure of the collection procedures the deposit of a copy or extract of the electronic media by the publisher or producer can be envisioned according to criteria to be determined between the concerned parties.

The realisation of the “Web archiving” project is also set in an international cooperation movement. The BnF ensures the technical steering of the IIPC consortium (International Internet Preservation Consortium), created by 11 national libraries in 2003 following its initiative. It brings together the most innovative realisations in the world. In February 2008, the IIPC had some 38 members including representatives from: France, United Kingdom, Spain, Italy, Germany, Netherlands, Denmark, Norway, Sweden, Finland, Austria, Poland, Czech Republic, Latvia, Croatia, Slovenia. This configuration ensures a strong convergence of interests and policy of EU countries in terms of web archival, rejoining those of further countries in North America, Asia and Australasia. The diversity of IIPC members favours collaboration between different public institutions (libraries, archives, universities...) as well as partnerships with private companies or foundations, such as the Internet Archive, the British Hanzo Ltd company or the European Archive foundation. The cooperation is first focused on document exchange with each

institution (library or national archive) concentrating on the collection of its national websites.

This is followed by common sharing of financial means and expertise for the development of standards and tools destined for the entire international community. These applications mainly concern the automated collection of websites, indexation software, consultation interfaces and preservation strategies.

Thus, all software developed by the members of the consortium, such as the Heritrix harvesting robot are open source software.

Furthermore, the new web archival format WARC was essentially conceived by the members of IIPC. It is undergoing ISO standardisation.

Lastly, the consortium constitutes a forum for exchange of expertise and best practices that is extremely useful on technical, organisational, documentary and legal levels.

Both in its role as technical and administrative coordinator as well as its own technical choices, the BnF is clearly anchored in this movement of international cooperation.

In the EU, it also developed a privileged partnership with the British Library (co-financing of the “smart crawler” project) and the Royal Library of Denmark (the BnF projects to use the open source archive management software Netrake Suite).

B5. Progress on legal provisions for the preservation of web-content by mandated institutions.

The law on “Copyright and neighbouring rights in the information society” (DADVSI title IV) of August 1st 2006 also applies the legal deposit to “the signs, signals, writing, sounds or messages of any nature that are the object of a public communication in electronic form.” The BnF must therefore integrate these new “born digital” editorial forms in its collections, in the spirit of legal deposit over the last five centuries: constitute a heritage that represents the national editorial dynamic, without judgement of the quality or form of the content.

Large scale automated (instantaneous, yearly) harvesting of the Internet has been carried out since 2004 in partnership with American Internet Archive using the Heritrix robot. The BnF has also laid down, since 2006, the computer infrastructure allowing it to manage, in autonomous fashion, mid sized thematic collection. The pilot projects of election site

harvesting “Internet on campaign” (carried out from October 2006 to July 2007 in collaboration with the Science Po University, the Forum for Internet Rights (le Forum des droits sur Internet) as well as eight municipal libraries of publisher legal deposits), served as a laboratory for the establishment of a complete chain for site treatment (collection, indexation, preservation).

Web archival has thus become a daily affair at the BnF. The collections gathered to date represent over 110 Teraoctets of data, some going back to 1996, or 10 billion documents.

In light of this new mission the priorities of the public establishment for the next three years will concern:

- opening public access: an upcoming decree will outline the conditions for communicating internet archives to the public. As of 2008, in a new reading

room, researches should be able to start exploring this new scientific material and experiment new research services (search by URL using the WayBack Machine, full text indexation...)

- the growth of the BnF's production capacities, allowing the institution to carry out in-house large scale “instantaneous” harvesting with the capacity to sustainably preserve the results as of 2009
- new harvesting campaigns
- the extension to other domains of digital collection that will allow for the archival of files that the automated harvesting, for technical or legal reasons, can not collect
- carry out research, improve and create collection, indexation and archive access tools adapted to the needs and demands of tomorrow's researchers.

*Version officielle du rapport d'étape français
sur la politique de numérisation du patrimoine*

A. DIGITISATION AND ONLINE ACCESSIBILITY

A1. Progress on the systematic gathering of information about current and planned digitisation of books, journals, newspapers, photographs, museum objects, archival documents, audiovisual material and the availability of overviews of such digitisation in order to prevent duplication of efforts and promote collaboration and synergies at European level.

Depuis 2001, il existe un catalogue national des collections numérisées pour le patrimoine culturel, qui recense au niveau national l'ensemble des collections faisant l'objet d'une numérisation ainsi que des institutions impliquées, quelle que soit la source de financement. Intégré depuis 2004 dans le projet européen MICHAEL, ce catalogue recense 1.363 collections et 513 institutions à fin février 2008.

Il est accessible au public via le site *Patrimoine numérique* <<http://www.numerique.culture.fr>>, intégré au portail Culture.fr via le moteur "Collections".

Le catalogue vise à recenser de manière exhaustive les projets menés dans le domaine du patrimoine.

Il est alimenté par un réseau s'appuyant sur le comité de pilotage "numérisation" du ministère, qui comprend les directions sectorielles du ministère et fédère le réseau des institutions culturelles. L'obligation d'alimenter le catalogue est inscrite dans les conventions de subvention des projets du plan national de numérisation. Une version destinée aux collections des institutions dépendant de l'enseignement supérieur (NUMES) est en cours de développement par le ministère de l'enseignement supérieur et de la recherche (MESR), en cohérence avec le catalogue Patrimoine numérique. Il vise à mettre en place un inventaire national en ligne des projets de numérisation dans les établissements relevant de la tutelle du MESR : universités (y compris les bibliothèques universitaires) et grands établissements (y compris le CNRS) ; et il favorisera la structuration de la carte documentaire numérique nationale de l'enseignement supérieur et de la recherche (voir question A2).

Par ailleurs, un rapport national sur les avancées de la numérisation a été établi chaque année jusqu'en 2006 dans le cadre du programme Minerva.

Enfin, dans le cadre de l'étude NUMERIC lancée

par la Commission européenne en 2007, un travail sur les statistiques en matière de numérisation du patrimoine est en cours au sein du comité de pilotage "numérisation" du ministère de la culture et de la communication.

A2. Progress on the development of quantitative targets for the digitisation of analogue material in archives, libraries and museums, indicating the expected increase in digitised material which could form part of the European Digital Library and the budgets allocated by public authorities.

1) Plan national de numérisation du patrimoine culturel: un appel à projets de numérisation annuel

Dans le cadre du plan national de numérisation coordonné par le ministère de la culture et de la communication (comité de pilotage "numérisation, archives, bibliothèques, musées, patrimoine" mis en place en 2006), un appel à projets est lancé annuellement (dernier appel en août 2007 pour financer des projets à réaliser au titre de 2008). Cet appel est doté de 3 millions d'euros par an. Il est organisé autour de six programmes thématiques visant à structurer les contenus numérisés selon des priorités nationales : les territoires ; les personnes ; le français et les langues de France ; l'art et l'archéologie ; l'architecture ; la création contemporaine (voir détail dans la question A5).

Les objectifs de l'appel 2008 s'inscrivent dans les objectifs de mise à disposition des contenus promus par la bibliothèque numérique européenne :

- développer une offre globale et structurée de contenus numérisés par des programmes thématiques ou régionaux ;
- augmenter significativement l'offre de ressources culturelles numériques constituée d'ensembles cohérents et de taille critique ;
- promouvoir une consultation libre et ouverte du patrimoine culturel numérique par tous les internautes ;
- favoriser les usages individuels en rendant accessibles des fonds patrimoniaux emblématiques.

2) Numérisation de masse - livre (BnF): un nouveau marché en septembre 2007

Les rapides évolutions technologiques et la baisse des coûts permettent un changement d'échelle dans le traitement des collections : dès 2006, la BnF s'est lancée dans une numérisation de masse de ses fonds patrimoniaux.

Deux marchés, qui s'achèveront en 2008, permettent déjà la numérisation de 30 000 documents imprimés et la conversion en mode texte de 60 000 documents déjà numérisés en mode image et consultables sur Gallica. En septembre 2007, un nouveau marché de numérisation « de masse » a été notifié par la BnF à la Société Safig. Il porte sur la numérisation et la conversion en mode texte de 300 000 documents sur trois ans. Les premiers ouvrages traités par la société Safig seront consultables dès avril 2008, l'accroissement se fera ensuite au rythme de 8 000 documents par mois. Fin 2010, la BnF mettra à disposition du public via sa bibliothèque en ligne (Gallica 2) près de 400 000 documents imprimés, simultanément disponibles en mode image et texte, ce qui représente plus de 45 millions de pages. A ces 400 000 documents imprimés, s'ajouteront les fascicules de presse issus du plan quinquennal de numérisation de la presse (3,5 millions de pages), ainsi que les documents issus des chaînes de sauvegarde numérique de la BnF : collections spécialisées, imprimées et audiovisuelles.

3) Collections audiovisuelles (INA)

En augmentant la dotation publique de l'Ina de 22,6% sur la période 2005-2009, l'Etat a donné à l'Institut National de l'Audiovisuel (INA) les moyens de mener une politique volontariste de diffusion de ses fonds sur Internet tout en garantissant son plan de sauvegarde.

Les fonds professionnels des archives des chaînes publiques, dont la conservation incombe à l'Ina, remontent aux années 1940 pour la radio et aux années 1950 pour la télévision. Ceux des Actualités Françaises couvrent les années 1940. Ils se trouvent donc sur des supports anciens qui ont été altérés au fil du temps. Les films et les bandes sons, dont le matériau de construction était l'acétate de cellulose, sont victimes d'une réaction chimique appelée « syndrome du vinaigre ». Les bandes magnétiques anciennes sont soumises à toutes sortes de dégradations liées à la chaleur et à l'humidité. En outre, les matériels de lecture, notamment les têtes de lecture des magnétoscopes ne sont pratiquement plus fabriquées.

Sur un total de fonds patrimoniaux professionnels de l'ordre de 1 500 000 heures, 830 000 heures sont menacées à des titres divers (dégradation du support, obsolescence des équipements de lecture...)

Pour relever ce défi, l'Ina a lancé en 1999 un Plan de sauvegarde et de numérisation (PSN), pour garantir la sauvegarde du fonds « ancien », pour faciliter l'exploitation du fonds et en améliorer l'accès et pour

répondre à sa mission de valorisation du patrimoine

A fin 2007, 470 000 heures de programmes ont été transférées par l'Ina de l'analogique au numérique : 95 000 heures de télévision sont issues de fonds récents ; 375 000 heures relèvent de la sauvegarde / numérisation systématique des fonds menacés de dégradation (PSN), dont 140 000 heures de radio et 235 000 heures de télévision.

45% du total des fonds menacés sont désormais sauvegardés en numérique, soit 22% de la filière Film, 100% de la filière Vidéo et 28% de la Radio

Depuis 2004, l'ensemble de ces programmes numérisés est accessible en ligne dans le monde entier sur le site professionnel inamediapro.fr, ouvert à 6 400 clients accrédités, dont un tiers de clients étrangers.

Depuis 2006, le site « Archives pour tous » sur ina.fr ouvre au grand public un accès aux archives de la radio-télévision et offre 18 000 heures de contenus en ligne (14 000 heures de télévision et 4 000 heures de radio, soit plus de 100 000 documents), dont plus de 80% est accessible gratuitement.

4) Collections patrimoniales des établissements d'enseignement supérieur et de recherche - schéma directeur de la numérisation de l'enseignement supérieur et de la recherche

Le ministère de l'enseignement supérieur et de la recherche (MESR) finance des projets de numérisation dans le cadre de son dispositif de contractualisation avec les universités et grands établissements relevant de sa tutelle. Ces projets concernent à ce jour des documents patrimoniaux libres de droit, comme par exemple le projet "Liber Floridus" consacré aux manuscrits médiévaux enluminés des bibliothèques de l'enseignement supérieur et de la recherche et qui fait l'objet d'un partenariat avec le ministère de la culture et de la communication - direction du livre et de la lecture et la BnF. Si le MESR ne fixe pas, à ce jour, d'objectifs quantitatifs à ces projets, il en réalise une expertise : pertinence, programmation dans le cadre du contrat quadriennal et réalisation aux termes du contrat. Des indicateurs de suivi sont mis en place avec les établissements concernés pour permettre une évaluation adaptée des projets. Il soutient également, dans le cadre des contrats, l'acquisition de matériels adaptés dans des établissements qui présentent une activité soutenue de numérisation et de valorisation de leurs collections patrimoniales (bibliothèques interuniversitaires, services communs de documentation).

Schéma directeur de la numérisation

Afin d'éviter les doublons, de mutualiser les compétences et les équipements et de valoriser les réalisations au niveau national et international, le MESR lance en 2008 son schéma directeur de la numérisation, avec pour objectif d'organiser la carte documentaire numérique de l'enseignement supérieur et de l'articuler avec les autres grands projets nationaux et internationaux. La carte documentaire numérique s'appuie sur une organisation en axes thématiques ou disciplinaires spécialisés. Chaque axe est placé sous le pilotage d'un établissement référent, chargé d'assurer une coordination scientifique de la numérisation conduite par les établissements d'enseignement supérieur pour cette discipline ou thématique. Son rôle consiste à proposer un plan de numérisation partagé à moyen terme. La diffusion des données numérisées est facilitée par la mise en place de portails thématiques permettant un accès fédéré et ouvert.

A3. Progress on partnerships between cultural institutions and the private sector in order to create new ways of funding digitisation of cultural material.

La numérisation du patrimoine est financée sur fonds publics. C'est le cas notamment des grands établissements publics dépendant de l'Etat (BnF, INA, Louvre, etc), où ces crédits sont inscrits dans le budget des établissements. C'est également le cas pour la numérisation des collections patrimoniales conservées par les établissements d'enseignement supérieur et de recherche.

Dans ce cadre, le soutien aux projets de numérisation participant à la construction de la bibliothèque numérique européenne devrait être inscrite dans les programmes de soutien mise en place par la Commission européenne, comme cela a été initié dans l'appel à projets en cours du programme eContent+.

Le plan national de numérisation inscrit un budget de 3 millions d'euros annuellement, qui finance des projets sur la base d'un appel à projets. La majorité de ces projets sont menés en partenariat avec les collectivités locales qui apportent un financement public à hauteur de 50%. En 2007, des crédits spécifiques (10 millions d'euros) pour la BnUE ont été inscrits au budget de la BnF. Ils proviennent du Centre National du Livre (CNL), qui tire ses ressources de deux taxes, l'une sur le chiffre d'affaires des éditeurs les plus importants, la seconde sur la vente des appareils de reprographie sur le territoire national ("redevance sur l'emploi de la reprographie").

1) Un accord entre BnF et SNE pour expérimenter un accès commun aux oeuvres sous droits et aux oeuvres du domaine public

La nouvelle dimension de la diffusion du patrimoine écrit dans le cadre de bibliothèques en ligne nécessite la recherche d'un partenariat nouveau avec les éditeurs. Celui-ci doit permettre la mise à disposition, aux côtés des documents libres de droits, des documents soumis au droit d'auteur. La BnF, forte de son engagement en faveur du développement de bibliothèques numériques, travaille étroitement depuis l'été 2006 avec le Syndicat national de l'Edition (SNE) pour définir les conditions d'une mise en ligne d'ouvrages protégés, dans le strict respect des législations nationales, communautaires et internationales relatives à la propriété intellectuelle.

Un groupe de travail BnF / SNE comprenant des représentants de plusieurs maisons d'édition françaises a été constitué à cette fin. Il a commandé l'élaboration d'un rapport permettant de créer un modèle économique à un consultant spécialiste des bibliothèques numériques (Numilog). Numilog a considéré que pour l'accès aux ouvrages, il fallait un modèle payant pour garantir l'équilibre économique de la filière livre. Il a proposé un modèle de « librairie numérique », reposant sur des licences individuelles d'accès aux livres, reproduisant le modèle de vente des livres imprimés. Les bibliothèques numériques, qu'elles se nomment Gallica 2, BnUE ou Bibliothèque numérique francophone, pourraient ainsi offrir des fonctions de recherche plein texte tant pour les ouvrages du domaine public que pour les ouvrages soumis au droit d'auteur. L'internaute pourrait alors consulter librement et gratuitement les œuvres libres de droit dans leur intégralité, ainsi que les notices, voire de courts extraits, des œuvres protégées, avec l'accord des ayants droit.

Le feuilletage virtuel des œuvres protégées se ferait par l'intermédiaire de diffuseurs agréés, choisis par les éditeurs. Le feuilletage serait payant. Des formules d'abonnements à des collections de livres ou de droits temporaires de lecture pourraient être proposées. Le diffuseur aurait également le choix de vendre le téléchargement des titres à l'exemplaire, au prix fixé par l'éditeur. La commercialisation des exemplaires papier serait ainsi possible, par l'intermédiaire des libraires.

Les recommandations de l'étude ont été approuvées par la BnF et le bureau du SNE. A la suite de l'adoption du modèle économique, une charte technique

a été élaborée, qui prescrit les choix de normalisation nécessaires pour l'inclusion des données concernant les livres sous droits.

Suite à ces travaux, la BnF et les éditeurs réalisent une expérimentation du modèle économique : le site <http://gallica2.bnf.fr> réalisé par la BnF qui sera présenté au public au prochain Salon du livre de Paris (14-19 mars 2008), à la fois portail et bibliothèque numériques, contiendra environ 50 000 livres du domaine public, et au moins 10 000 titres sous droit, mis à disposition par des éditeurs qui ont souhaité participer à l'expérience.

Ce projet constitue une expérimentation de l'offre numérique légale ; il est ouvert à tous les éditeurs et les e-distributeurs selon les modalités suivantes. Le modèle prévoit l'accessibilité, pour le public à des documents numériques sous droits par l'intermédiaire de plates-formes d'e-distributeurs agréés par les éditeurs des ouvrages considérés. Les éditeurs peuvent aussi jouer le rôle de e-distributeur et assurer eux-mêmes la communication en ligne des ouvrages de leurs fonds. La durée de l'expérimentation est de un an, à partir du Salon du Livre de Paris 2008. A cette date, ce sont plusieurs milliers de documents de l'édition contemporaine qui seront référencés dans Gallica 2. Au terme de la phase expérimentale, qui servira à mieux cerner les usages, une évaluation sera conduite par les différents partenaires. Gallica 2 dispose d'un moteur de recherche fédérée permettant à l'utilisateur de lancer une requête unique sur les fonds patrimoniaux et les éditions contemporaines. Pour les documents appartenant aux collections numériques de la BnF, la recherche portera sur les métadonnées et le plein-texte, s'il existe. La visualisation et le téléchargement s'effectuent également au sein de Gallica 2. L'accès est libre et gratuit. Pour les éditions contemporaines référencées dans Gallica 2, le modèle retenu prévoit une recherche sur les métadonnées et sur le plein-texte. Le titre de l'ouvrage apparaît dans la liste de résultats, accompagné d'un court extrait qui situe les termes de la recherche de l'utilisateur dans leur contexte. L'utilisateur peut choisir d'en savoir plus en cliquant sur le titre. S'affiche alors la notice enrichie de l'ouvrage (table des matières, visuel de la première de couverture, texte de la quatrième de couverture, description détaillée de l'ouvrage). L'utilisateur pourra alors cliquer sur un lien vers la plate-forme du e-distributeur où lui seront

proposés, selon les cas, différents services gratuits ou payants : feuilletage d'extraits, accès au document intégral, téléchargement, abonnement en ligne, location, prêt de la version numérique pour les abonnés de certaines bibliothèques. Les tarifs et les prestations diffèrent selon les e-distributeurs.

Le graphique présentant la plate-forme d'accès aux documents numériques de Gallica 2 est disponible à l'adresse : http://www.bnf.fr/pages/catalog/pdf/gallica2_schema.pdf.

Les e-distributeurs désirant participer à l'expérimentation doivent signer au préalable une convention avec la BnF. Ils doivent en outre pouvoir se prévaloir d'accords conclus avec les éditeurs dont ils distribuent les ouvrages. Ils doivent être en mesure de mettre en place un entrepôt OAI dans lequel la BnF viendra moissonner les métadonnées des documents selon le protocole technique OAI-PMH : notice descriptive, table des matières, visuel grand format de la première de couverture, texte de la quatrième de couverture, description longue de l'œuvre, indicateur de gratuité du document... Un lien fin (URL) conduisant vers l'emplacement du document numérique doit également être fourni. Au lancement de l'expérimentation, un même document ne peut être accessible que par le biais d'un seul e-distributeur. Enfin, l'e-distributeur doit être en mesure de fournir le contenu du document selon un formalisme XML simple. L'échange des données cryptées se fera à l'aide du protocole FTP. Ce contenu ne sera utilisé que pour indexer le document et aider l'internaute à le trouver.

Une charte technique a été élaborée par la BnF. Elle prescrit les choix de normalisation nécessaires pour l'inclusion des données concernant les livres sous droits dans le portail d'accès Gallica 2 (formalisme de transmission des contenus des documents, format des métadonnées, protocoles de fourniture des données etc.). Afin d'encourager les éditeurs à participer à cette expérimentation, le Centre national du Livre propose, sous la forme de subvention, une aide pour la diffusion numérique de documents soumis au droit d'auteur dans le cadre de la mise en œuvre de la phase de test du prototype français pour une bibliothèque numérique européenne. La Commission "Politique Numérique" du CNL examine régulièrement les demandes. L'un des critères de sélection mis en œuvre par le CNL réside dans la compatibilité de l'offre des e-distributeurs et éditeurs candidats avec la politique documentaire de Gallica 2.

2) Un partenariat technologique BnF/France Télécom

En janvier 2007, la BnF et France Télécom ont signé un protocole préliminaire à un futur accord visant à promouvoir le patrimoine culturel numérique français et européen, et en faciliter l'accès au plus grand nombre, en particulier au travers du projet de Bibliothèque numérique européenne. Une partie de ce partenariat portait sur un apport de compétence technologique, de la part de France Télécom R&D, en matière d'organisation des documents, de recherche d'information et d'accès aux résultats, ou des techniques de traitement des langues naturelles appliquées au texte.

Le thème « réseaux de documents » a été traité en priorité par la BnF, avec l'élaboration d'un rapport synthétisant l'ensemble des données et métadonnées disponibles. Au sein de ce thème, les travaux se focaliseront sur la géolocalisation (reconnaissance d'entités nommées de type géographique, i.e. entités toponymiques, pour association de documents et de lieux).

L'objectif des traitements sur les données fournies à France Telecom par la BnF sera principalement la création de liens entre différents documents et ressources. Ces liens peuvent être des liens directs d'un document à un autre (par exemple, reconnaissance automatique de citation d'un document par un autre), des liens entre documents passant par des référentiels (par exemple notices d'autorités géographiques pour la géolocalisation), ceux-ci pouvant être également fournis par la BnF.

L'objectif pour 2008 est de tester les résultats des premiers traitements et pour permettre de démontrer leur potentiel d'innovation en terme de services, de réaliser dans le cadre du partenariat des interfaces de démonstration sur des points précis jugés particulièrement intéressants, comme par exemple : la géolocalisation, ou accès cartographique aux collections numériques et le parcours d'un utilisateur dans le réseau de documents conçu autour d'une perception géographique.

La prochaine phase du travail BnF/France Télécom portera sur les formats de données possibles pour échanger les données, stocker les nouveaux liens issus des traitements et construire les interfaces, ainsi que sur les questions de propriété intellectuelle et de la possibilité d'intégrer des outils développés dans le cadre de ce partenariat au sein de la bibliothèque numérique.

3) Un accord signé avec l'INA par la chaîne de télévision TF1 pour la commercialisation de ses archives d'actualités en exclusivité mondiale à destination des professionnels

TF1 a décidé de confier à l'Ina l'exploitation de ses archives d'actualités de la période 1989-2006 en exclusivité mondiale, à destination des professionnels. L'Ina les numérise et les commercialisera progressivement dans le monde entier pour les 12 prochaines années. Ce mandat d'exploitation concerne tous les journaux télévisés et magazines d'actualités produits par TF1 sur cette période. Cet accord vient en complément du mandat qu'elle avait déjà signé en 2006 avec l'Ina sur son fonds d'actualité 1982-1989.

Ce fonds TF1 enrichit les fonds de l'Ina et permettra de rendre accessibles sur www.inamediapro.com environ 20 000 heures supplémentaires de programmes d'actualité de TF1. C'est donc au total un fonds d'actualité TF1 de plus de 30 000 heures de programmes qui devient disponible grâce à la numérisation de ces archives.

Ce fonds deviendra également visible pour les particuliers puisqu'il sera mis en ligne, au fur et à mesure de sa numérisation, parallèlement sur les sites Internet grand public du groupe TF1 et de l'Ina.

Cet accord confirme la politique de commercialisation de fonds "tiers" (autres que ceux constitués par les archives publiques de la télévision et de la radio) dans laquelle l'Ina est engagé : depuis 3 ans, de nombreux producteurs audiovisuels (comme CAPA, AFP vidéo...) ont confié à l'Ina un mandat de commercialisation de leurs archives.

Inamediapro.com est le service BtoB de l'Ina qui s'adresse aux professionnels de l'audiovisuel (journalistes, producteurs d'émissions de stock ou de flux, éditeurs de DVD, ...) et leur permet de visionner en ligne le contenu d'une archive, d'en segmenter un extrait et d'en passer commande. Ce service totalisait à la fin de l'année 2007 plus de 450 000 heures de programmes de télévision et de radio, ce qui en fait le leader mondial dans le domaine.

4) Un groupe de travail de haut niveau sur la diffusion et la réutilisation numérique des données publiques culturelles mis en place en 2008 par le ministère de la culture et de la communication

A la confluence du politique, de l'économique et du droit, les questions relatives à la diffusion numérique des données publiques donnent lieu depuis longtemps à des réflexions de fonds, à d'importants travaux

réglementaires, tant au plan communautaire que national, à des actions judiciaires, à des travaux scientifiques et à des programmes européens. La création en 2007 de l'Agence pour le patrimoine immatériel de l'Etat (APIE) a marqué le début d'une réflexion interministérielle, principalement tournée vers la commercialisation des données publiques. Compte tenu de la nature spécifique de son patrimoine, le ministère de la culture et de la communication (MCC) souhaite mener en interne une réflexion propre sur la diffusion numérique et la réutilisation des données publiques culturelles, prenant en compte ses missions de service public culturel, qui sera une traduction sectorielle et une contribution à ce chantier interministériel. Ainsi, a-t-il été décidé de constituer un groupe de travail de haut niveau afin de définir la doctrine et les modes opératoires du ministère en matière de diffusion numérique et de réutilisation des données publiques culturelles. Ce groupe est composé de représentants du ministère et des opérateurs et de personnalités qualifiées. Il partira des recommandations du rapport Ory-Lavollée de 2002 et les éclairera des évolutions technologiques, juridiques et économiques des années ultérieures en s'inspirant des comparaisons internationales et en tenant compte de l'existant recensé.

Ce groupe de travail proposera courant 2008 une « charte » de diffusion et de réutilisation visant à harmoniser les actions du ministère de la culture et de la communication et de ses opérateurs. Une fois validée par la ministre, les grands principes de cette charte se déclineront en autant d'outils de mise en œuvre de la stratégie du ministère.

A4. Progress on the set-up and sustaining of large scale digitisation facilities.

Les opérations de numérisation sont essentiellement techniques et sont sous-traitées par les institutions culturelles dans le cadre de marchés publics à des prestataires privés.

1) Numérisation de masse du livre (BNF)

Une opération de numérisation de masse de ses collections patrimoniales (livres) est menée par la BnF dans le cadre d'un marché lancé avec la société SAFIG en septembre 2007 (voir question (2)) *Deux marchés, qui s'achèveront en 2008, permettent déjà la numérisation de 30 000 documents imprimés et la conversion en mode texte de 60 000 documents*

déjà numérisés en mode image et consultables sur Gallica. En septembre 2007, un nouveau marché de numérisation « de masse » a été notifié par la BnF à la Société Safig. Il porte sur la numérisation et la conversion en mode texte de 300 000 documents sur trois ans. Les premiers ouvrages traités par la société Safig seront consultables dès avril 2008, l'accroissement se fera ensuite au rythme de 8 000 documents par mois. Fin 2010, la BnF mettra à disposition du public via sa bibliothèque en ligne (Gallica 2) près de 400 000 documents imprimés, simultanément disponibles en mode image et texte, ce qui représente plus de 45 millions de pages. A ces 400 000 documents imprimés, s'ajouteront les fascicules de presse issus du plan quinquennal de numérisation de la presse (3,5 millions de pages), ainsi que les documents issus des chaînes de sauvegarde numérique de la BnF : collections spécialisées, imprimées et audiovisuelles.

Par ailleurs, dans le cadre de sa politique de conservation préventive, la BnF abandonne progressivement les supports de substitution argentiques au profit du numérique : des chaînes internes sont en cours de montée en charge, dont la production, si les droits le permettent, alimentera Gallica.

2) Fonds audiovisuels (INA)

Pour la numérisation des fonds audiovisuels, l'Ina s'est également engagé dans la numérisation de masse ("Plan de Sauvegarde et de numérisation", voir question (2)). L'INA est un laboratoire unique de recherche qui développe de nouvelles technologies telles que « Signature », procédé technique de gestion légale et de protection des contenus sur l'Internet. Premier pôle européen de formation aux métiers de l'image et du son, l'Ina est aujourd'hui un des acteurs européens majeurs de l'enseignement professionnel à l'audiovisuel. Depuis octobre 2007, Ina'Sup, l'école supérieure de l'audiovisuel et du numérique, propose deux masters : Production-édition audiovisuelle et numérique et Gestion et conservation des patrimoines audiovisuels et numériques.

A5. Progress of cultural institutions, as well as publishers and other rightholders to make their digitised material searchable through the European Digital Library. Describe in particular the progress on the application of common digitisation criteria by cultural institutions and, where

relevant, private companies, in order to achieve interoperability with the European Digital Library and to facilitate cross-language searchability.

1) Des objectifs communs de numérisation définis en terme de contenus au niveau national

Les priorités en terme de contenus présentées ci-dessous sont articulées afin de numériser en cohérence l'ensemble des supports (imprimés, archives sonores, collections photographiques, cartes et plan, manuscrits...).

– le plan national de numérisation du ministère de la culture et de la communication concerne l'ensemble des domaines (archives, bibliothèques, musées, monuments, archéologie...) et des supports . Il est structuré autour de six programmes thématiques visant à organiser les contenus numérisés selon des priorités nationales : les territoires ; les personnes ; le français et les langues de France ; l'art et l'archéologie ; l'architecture ; la création contemporaine.

Domaine	Collections
Programme 1 : Les territoires	
Géographie	Plans du cadastre ancien Cartes et plans Patrimoine maritime, lagunaire et fluvial
Histoire des communes	Registres de délibérations de plus de 100 ans
Histoire locale	Presse régionale et locale Instruments de recherche Dossiers documentaires Archives sonores
Histoire nationale	Documents emblématiques de l'histoire de France Première et deuxième guerre mondiale Mémoire de l'esclavage Manuscrits littéraires Manuscrits médiévaux
Programme 2 : Les personnes	
Généalogie	Tables décennales de l'état civil (exclus registres) Registres paroissiaux Listes nominatives (dépouillement généalogie)
Société	Mémoires des groupes sociaux Mémoire de l'immigration Archives du monde du travail
Patrimoine immatériel	Savoir faire et artisanat Chants, danses traditionnelles, contes, tradition orale

Programme 3 : Le français et les langues de France	
Linguistique	Corpus oraux
Programme 4 : L'art et l'archéologie	
Beaux arts	Reproductions d'œuvres : vidéos, peintures, sculptures, affiches, objets mobiliers... Cabinets d'arts graphiques Arts décoratifs (papiers peints, ...) Photographies Arts premiers Portraits (écrivains, personnages, caricatures...) Instruments de musique Jouets et jeux
Archéologie	Archéologie sous-marine Art rupestre et grottes ornées
Sciences et techniques	Histoire des sciences et techniques
Arts du spectacle	Manuscrits musicaux, ...
Domaine	Collections
Programme 5 : L'architecture	
Monuments historiques	Patrimoine architectural (sites du patrimoine mondial, label européen du patrimoine, châteaux, cathédrales, monuments antiques...)
Architecture	Patrimoine du XX ^{ème} siècle
Jardins	Dossiers documentaires, fonds photographiques....
Programme 6 : La création contemporaine	
Musique	Musique contemporaine Musiques du Monde Musique orchestrale
Art contemporain	Œuvres et créateurs
Design	Œuvres et créateurs
Mode et textile	Collections patrimoniales et contemporaines
Livres d'artistes	Collections contemporaines

Domaine	Collections
Cinéma expérimental	Oeuvres et créateurs
Théâtre	Photographies, vidéos, archives
Cirque, arts forains, arts de la rue	Collections patrimoniales et contemporaines
– la numérisation de l'imprimé par la BnF dans le cadre de Gallica est opérée selon une charte documentaire <http://www.bnf.fr/pages/infopro/numerisation/pdf/Charte_Gallica.pdf> qui définit des critères, des priorités et des objectifs	en matière de numérisation pour assurer une couverture géographique, chronologique, linguistique et par domaine des contenus et garantir une complémentarité avec les autres opérations de numérisation de l'imprimé existant au niveau national (numérisation en réseau). Dans le cadre du marché des 100 000 documents imprimés
Programmes sur le patrimoine national	Histoire de France : histoire régionale, histoire par période, biographies historiques L'institution de la sociologie comme champ du savoir Histoire de l'enseignement Emergence de la question sociale et fondation de la protection sociale Histoire du protestantisme français et son influence en Europe. Revue en sciences humaines et sociales, artistiques et littéraires Les sciences naturelles au XIXe siècle : contribution française au développement et à la diffusion des connaissances dans ces domaines. La médecine française au XIXe siècle : la maladie et son traitement. Poésie et théâtre français Enfantina (corpus de littérature enfantine) Salons artistiques Ouvrages de référence en art Catalogues d'éditeurs et de libraires
Programmes à vocation internationale	Les langues: • Histoire et diversité des langues • Apprentissage de la langue tant orale qu'écrite • Traductions des chefs-d'œuvre de différentes littératures nationales Traductions des chefs-d'œuvre de l'Antiquité Presse quotidienne Documentation juridique La généalogie
Une bibliothèque française sur une Europe ouverte sur le monde (œuvres et documents français consacrés spécifiquement à l'Europe)	Réception de la philosophie européenne en France Les expositions universelles au XIXe, vitrines du progrès et instruments de diffusion des connaissances. Sinica (documents écrits ou traduits en français sur la Chine) Questions coloniales

annuels, sont privilégiées les thématiques suivantes :
 – en matière de contenus audiovisuels, l'INA s'est engagé dans une opération de numérisation exhaustive de ses fonds, présentés de manière structurés au grand public via le site "Archives pour tous" sur Ina.fr : thèmes (architecture, art, arts du spectacle, cinéma, économie, environnement, géographie, histoire, justice, littérature, médias, musique, politique intérieure, politique internationale, sciences, sciences humaines, société, sport, télévision), personnalités, époques et programmes de radio-télévision.

2) L'accès au patrimoine numérisé : 3 outils complémentaires prêts à être intégrés dans Europeana

Ressources des archives, bibliothèques, musées, monuments historiques et archéologie : le portail Collections sur Culture.fr (Ministère de la culture et de la communication)

Le portail "Collections", service d'accès aux ressources des archives, bibliothèques, musées et services du patrimoine via le portail Culture.fr, a été mis en ligne en août 2007. Cette version intègre 29 ressources, ce qui porte à 3 millions le nombre de données patrimoniales accessibles, illustrées par plus de 2 millions d'images. Elle inclut des fonctions de recherche avancées, basées sur les fonctionnalités sémantiques, et le multilinguisme (anglais, français, espagnol). Le ministère de la culture et de la communication en association avec ses établissements publics (Louvre, Centre Pompidou, musée du Quai Branly, Réunion des musées nationaux...) et plusieurs collectivités territoriales, poursuit ainsi l'ouverture de ses ressources numériques au grand public par l'intermédiaire de son moteur de recherche « Collections » inclus dans le portail <<http://www.culture.fr>>. Ce "guichet unique" couvre les domaines les plus divers : Architecture, archéologie, objets mobiliers, peinture, sculpture, gravure, dessin, enluminures, arts du spectacles, ethnologie, plans, cartes, archives, généalogie, photographie, événements... Par le biais d'une interrogation unique, « Collections » recherche dans 29 sources documentaires différentes et affiche l'ensemble des résultats par ordre de pertinence.

Collections imprimées : bibliothèque numérique Gallica de la BnF

Le site Gallica <<http://gallica.bnf.fr>>, bibliothèque numérique de la BnF, a été réalisé en 2000 et donne aujourd'hui accès à 250 000 ouvrages numérisés,

collections de presse comprises, et 100 000 images. Une refonte s'est avérée nécessaire pour l'inscrire dans les évolutions actuelles du Web. Existant en version beta depuis octobre 2007, l'interface Gallica 2 <<http://gallica2.bnf.fr>> coexistera quelques mois avec l'ancienne version de Gallica, avant de prendre définitivement sa place en 2008 lorsque toutes les fonctionnalités seront opérationnelles et que l'ensemble des documents de l'ancienne version auront été versés dans la nouvelle. Les documents issus des programmes de numérisation en cours enrichissent également progressivement le corpus disponible (100 000 documents imprimés numérisés en mode texte chaque année). Egalement seront intégrées progressivement des données relatives aux livres récents accessibles au format numérique via des sites de e-distributeurs (accès libre ou sous conditions, voir question (A3-1) et des documents des bibliothèques numériques partenaires de la BnF. Fin 2008, les "Gallicanauts" bénéficieront aussi de nouvelles fonctionnalités de personnalisation et de partage dans l'esprit du Web 2.0. Fin 2010, près de 400 000 documents imprimés libres de droit seront simultanément disponibles en mode image et texte. A ces 400 000 documents imprimés, s'ajouteront les fascicules de presse issus du plan quinquennal de numérisation de la presse (3,5 millions de pages), ainsi que les documents issus des chaînes de sauvegarde numérique de la BnF : collections spécialisées, imprimées et audiovisuelles.

La conception de Gallica 2 a tenu compte des remarques des usagers collectées grâce à deux études réalisées par la BnF (voir : *Rapport de bilan sur les usages et les attentes des utilisateurs d'Europeana*, Noémie Lesquins, juin 2007 disponible à l'adresse : <http://bibnum.bnf.fr/usages/BnF_Europeana_EtudeUsages2007.pdf>).

La version actuelle du site Gallica 2 ouvert en octobre 2007 permet l'accès aux livres et à l'iconographie déjà présents dans Gallica ; elle dispose d'un moteur de recherche plein-texte et offre aux lecteurs la possibilité de se créer un espace personnel où ils peuvent stocker leurs documents et apposer des marque-pages.

Collections audiovisuelles : portail "Archives pour tous" de l'INA

Depuis 2006, l'Ina ouvre ses archives au grand public sur ina.fr portail "Archives pour tous". Ina.fr est ainsi aujourd'hui en tête de l'offre de Vidéo à la Demande (Video On Demand) en termes de volume et de qualité

des contenus : plus de 100 000 documents sont consultables en ligne, ce qui correspond à plus de 18 000 heures de programmes (14 000 heures de télévision et 4 000 heures de radio). Cette offre s'enrichit de 50 000 émissions par an et s'est étendue au domaine pédagogique avec Apprendre, un outil éducatif audiovisuel en ligne. Plus d'un million de visiteurs uniques par mois naviguent sur le site.

3) Des standards et des recommandations pour garantir l'interopérabilité

Les opérations de numérisation menées au niveau national s'intègrent dans les recommandations interministérielles d'interopérabilité (RGI - référentiel général d'interopérabilité) pilotées par la Direction générale de la modernisation de l'Etat (DGME).

Le plan national de numérisation du ministère de la culture et de la communication vise à soutenir des initiatives facilitant l'accès des fonds patrimoniaux à un large public. Il privilégie les technologies ouvertes et les protocoles standards. Outre la mise en ligne obligatoire des collections numérisées, les porteurs de projets sont invités à rendre accessibles les fonds à travers les outils fédératifs nationaux, régionaux ou thématiques. C'est ainsi l'accès par le portail agrégatif national "Collections" (voir ci-dessus) du Ministère de la Culture et de la Communication qui est privilégié, dans la perspective de l'intégration de ce portail dans la bibliothèque numérique européenne.

Par ailleurs, des actions de coordination et de diffusion des bonnes pratiques sont également menées pour favoriser la prise en compte des standards d'interopérabilité. Dans ce cadre, une réunion nationale sur le protocole OAI a été organisée le 8 octobre 2007. De plus, la BnF, dans le cadre de son réseau des pôles associés, offre des formations spécifiques à l'OAI ; elle met son expertise dans le domaine au service de ses partenaires en validant la conformité des entrepôts OAI aux standards et recommandations sur le sujet.

Le ministère de la culture et de la communication (direction des Archives de France, avec la participation du comité de pilotage numérisation) a mis en ligne en février 2008 un guide de rédaction des cahiers des charges de numérisation. Il s'agit notamment de prendre en compte les expériences récentes et les dernières évolutions techniques, afin de favoriser la qualité des ressources numérisées et leur intégration dans des portails d'accès.

En ce qui concerne le livre, une charte technique a été élaborée par la BnF dans le cadre de Gallica 2. Elle prescrit les choix de normalisation nécessaires pour l'inclusion des données concernant les livres sous droits dans le portail d'accès Gallica 2 (formalisme de transmission des contenus des documents, format des métadonnées, protocoles de fourniture des données etc.). Dans les projets des établissements qu'il finance au travers des contrats quadriennaux comme dans les programmes dont il assume le pilotage, le MESR recommande l'utilisation de technologies ouvertes : l'emploi, en amont, des normes informatiques et documentaires reconnues au niveau internationales (XML et modèles de données) et en aval, du protocole international d'échanges de données OAI-PMH.

A6a. Progress on mechanisms to facilitate the use of orphan works.

A6b. Progress on mechanisms to facilitate the use of works that are out of print or out of distribution.

A6c. Progress on the availability of lists of known orphan works and works in the public domain.

A6d. Progress on the identification of barriers in your legislation to the online accessibility and subsequent use of cultural material that is in the public domain – and the steps taken to remove them.

Le Conseil supérieur de la propriété littéraire et artistique, organe consultatif placé auprès du ministre chargé de la culture, doit examiner prochainement, un projet d'avis ayant pour objet la recherche de dispositifs permettant d'améliorer la situation des œuvres orphelines. Le rapport s'attachera à définir les œuvres orphelines, à quantifier le phénomène secteur par secteur, à établir un diagnostic par secteur et dégager si des initiatives doivent être prises tenant compte des différences de situations possibles entre les catégories d'œuvres concernant les situations en cours et, dans une perspective de prévention notamment, par des mécanismes d'identification des œuvres. D'ores et déjà, il a été relevé que les œuvres orphelines sont marginales dans les secteurs du cinéma, de la musique et, dans une moindre mesure, de l'audiovisuel et qu'elles sont plus nombreuses dans les secteurs de l'écrit et de l'image fixe ce qui pourrait conduire à des propositions diversifiées selon les secteurs. Par ailleurs, le code de la propriété intellectuelle prévoit des dispositions législatives de nature à résoudre nombre de situations. L'article L. 122-9 du code de la propriété

intellectuelle dispose : « En cas d'abus notoire dans l'usage ou le non-usage des droits d'exploitation de la part des représentants de l'auteur décédé visés à l'article L. 121-2, le tribunal de grande instance peut ordonner toute mesure appropriée. Il en est de même s'il y a conflit entre lesdits représentants, s'il n'y a pas d'ayant droit connu ou en cas de vacance ou de déshérence. Le tribunal peut être saisi par le ministre chargé de la culture [Ce cas est assez courant notamment pour les fonds de cartes postales.]. L'article L. 211-2 du même code prévoit un mécanisme comparable pour les droits voisins : « Outre toute personne justifiant d'un intérêt pour agir, le ministre chargé de la culture peut saisir l'autorité judiciaire, notamment s'il n'y a pas d'ayant droit connu, ou en cas de vacance ou de déshérence ». Il convient de ne pas assimiler l'œuvre anonyme à l'œuvre orpheline. A cet égard, le droit français prévoit des règles qui sont de nature à faciliter la gestion de ces œuvres. En vertu de l'article L. 123-3 du code de la propriété intellectuelle, les œuvres anonymes tombent dans le domaine public soixante-dix ans suivant l'année de leur publication et peuvent alors être exploitées sans soulever les difficultés posées par les œuvres orphelines. Si elles devaient entrer dans la catégorie des œuvres orphelines, elles perdraient définitivement leur statut d'œuvres anonymes. L'article L. 113-6 du même code prévoit que les auteurs d'œuvres anonymes sont représentés dans l'exercice de leurs droits « par l'éditeur ou le publicateur originaire, tant qu'ils n'ont pas fait connaître leur identité civile et justifié de leur qualité ». Les droits attachés à ces œuvres sont donc exercés par un représentant de l'auteur ou de ses ayants droit.

B. DIGITAL PRESERVATION

B1. Progress on national strategies for the long-term preservation of and access to digital material. Describe the organisational approach, indicating the roles and responsibilities of the parties involved as well as the allocated resources. Describe the specific action plans outlining the objectives and a time-table for the specific targets to be met.

1) Un référentiel national en cours de constitution

Dans le cadre du comité de pilotage national "numérisation", un travail spécifique est en cours depuis début 2007 pour coordonner l'action du ministère concernant la préservation

à long terme du patrimoine numérisé. Il vise à mettre en place un référentiel national fédérant de manière transversale les préconisations existantes dans les différents secteurs et les différents acteurs.

Ce référentiel national sera constitué sous la forme d'un site Internet de référence, facilement mis à jour, accessible sur le site "numérisation" du ministère de la culture et de la communication.

2) Le "Système de préservation et d'archivage réparti" (SPAR) de la BnF

La récente accélération de l'accroissement des collections numériques (supérieure à 100 Téraoctets par an aujourd'hui) et la diversité de leurs formats placent la BnF devant le défi de leur conservation : il est essentiel d'entreposer de manière sécurisée et pérenne les objets numériques en s'appuyant sur un socle solide et performant. Après plus d'une année d'étude, la BnF a lancé le projet SPAR, véritable magasin numérique. Sa conception s'appuie sur des normes internationales faisant autorité dans le monde de la pérennisation des informations numériques (modèle OAIS). SPAR est bien plus qu'un simple entrepôt de données sécurisé. Il effectue de multiples copies des objets numériques et assure une surveillance continue de l'état des fichiers d'enregistrement permettant d'anticiper les recopies avant la perte définitive. Mais il permet également, grâce à une reconnaissance précise et complète des formats de données versées, de garantir la continuité d'accès en procédant aux transformations nécessaires en cas d'obsolescence technologique des outils informatiques de restitution. Apporter cette garantie implique un travail permanent de veille technologique sur les formats, de prototypage et de tests des outils. L'ensemble de cet arsenal est intrinsèquement prévu dans la conception de SPAR. De plus, SPAR permet à tout moment de revenir en arrière pour restituer les objets dans leur format d'origine.

SPAR est un système au service d'une communauté.

Il doit garantir que les documents restitués n'ont pas été altérés. Pour ce faire, SPAR estampille chaque objet versé d'une empreinte numérique. De plus, afin de garantir les droits de « communicabilité » des objets numériques diffusés, SPAR s'adosse à un système expert de gestion de droits qui lui permet de calculer les licences d'utilisation des objets numériques et d'appliquer les restrictions nécessaires en fonction du profil de l'utilisateur (chercheur situé à la BnF, internaute...)

SPAR n'a pas vocation à être uniquement un outil interne à la BnF. Dans une volonté de mutualisation des expertises et des coûts, une étude copilotée par la BnF et la Caisse des Dépôts et Consignations est en cours pour préciser la nature de l'offre que pourra proposer la BnF et offrir ainsi un service de « tiers-archiviste ».

SPAR est un projet de grande ampleur qui s'étend sur plusieurs années. Après avoir acquis l'infrastructure en 2005, la BnF a lancé en juin dernier un appel d'offres pour la réalisation de la partie logicielle et a adopté une orientation en faveur du logiciel libre afin de s'assurer une indépendance maximale. SPAR sera opérationnel dès le début 2009.

3) Le développement de l'archivage électronique

Le projet pilote de plate-forme PILAE porté par la direction des Archives de France (DAF).

La DAF, avec l'appui de la direction générale de la modernisation de l'Etat (DGME), est en train de réaliser une plate-forme pilote qui fonctionnera durant une période transitoire 2008-2012 pour les Archives nationales, pour recevoir, conserver et communiquer les archives nativement numériques produites par les services centraux de l'Etat (données extraites de bases de données, documents bureautiques ou produits à partir d'une gestion électronique de documents, messageries électroniques, flux de données dématérialisés comme les marchés publics...). Le pilote implémente le standard d'échange de données pour l'archivage, permet de recevoir suivant ce format les archives transférées, de les contrôler en partie automatiquement selon les conditions fixées par des conventions passées entre les différents acteurs du processus d'archivage, en partie manuellement par les archivistes des archives nationales, de convertir en entrée automatiquement vers des formats pérennes des documents dont les formats d'entrée ne sont pas des formats cibles d'archivage, de les conserver d'une manière sécurisée grâce à des répliques sur deux sites distants et enfin de les communiquer par téléchargement. Dans un premier temps, le pilote fonctionnera pour les producteurs dans les ministères des documents archivés, pour les archivistes en place dans les différents ministères (en tant que services versants) et les archivistes des archives nationales. Il pourra être réutilisé comme modèle par d'autres services qui souhaiteraient mettre en œuvre une plate forme d'archivage dans leur propre environnement.

Dans le cadre du programme pour le développement de l'administration électronique et de la mise en place du référentiel général d'interopérabilité au niveau national, lancement et suivi de l'initiative "IF-08 archivage numérique" (DAF et DGME) : élaboration de référentiels (politique d'archivage sécurisé, standard d'échange de données pour l'archivage).

La DAF est également partie prenante (en tant que chef de projet, les Archives de l'Etat de Victoria en Australie étant éditeurs du projet) du projet eArchiving, qui vise à définir un standard d'échange entre producteurs et services d'archives, qui a été accepté par l'UN/CEFACT en octobre 2006. Il avait précédemment reçu le soutien de l'ICA (International Council on Archives).

Un questionnaire à l'attention des services d'archives du monde entier a été diffusé en décembre 2006 par la voie de l'UN/CEFACT, de l'ISO et de l'ICA. Sur cette base, Andrew Waugh, éditeur du projet, a rédigé une première version du standard en septembre 2007. Est actuellement en cours le lancement d'un appel à commentaires international

Les actions du ministère de l'enseignement supérieur et de la recherche (MESR)

Le Centre Informatique National de l'Enseignement Supérieur (CINES) a été chargé par le MESR de l'archivage pérenne de ses données numériques, et notamment des données issues d'opérations de numérisation. La Plate-forme d'Archivage du CINES (PAC) a été désignée, fin 2007, comme l'outil susceptible d'assurer cette mission d'archivage pérenne pour les documents numériques de la recherche.

B2. Progress on exchange of information with other Member States on your strategies and action plans.

La concertation avec les autres Etats membres de l'Union européenne s'opère à la fois dans le cadre des procédures mises en place par la Commission européenne (participation du MCC et de la BnF au groupe d'experts des états membres, participation de la direction des archives de France au groupe des archivistes européens EBNA), dans le cadre des réseaux professionnels actifs au niveau européen (CENL pour les bibliothèques, EURBICA pour les archives...) et par la participation aux projets européens soutenus par la Commission européenne.

Ainsi, des participants français (du ministère de la culture et de la communication, de la Bibliothèque nationale de France, de l'Institut national de l'audiovisuel) participent depuis septembre 2007 au réseau thématique EDLnet (réseau d'interopérabilité bibliothèques, musées, archives, audiovisuel), en charge de développer le prototype baptisé EUROPEANA sur proposition de la France.

Deux membres français (Martine de Boisdeffre -qui est la vice-présidente de la Fondation-, pour la branche européenne du Conseil international des Archives, et Christophe Dessaux, pour l'association Michael) sont membres de la Fondation "EDL", destinée à gérer la future BNuE, et qui regroupe les associations et fédérations européennes d'institutions patrimoniales (bibliothèques, musées, archives, audiovisuel). L'INA ainsi que la BnF ont également rejoint la fondation à la fin 2007. Par ailleurs, les équipes françaises ont participé aux projets en cours contribuant à la Bibliothèque numérique européenne :

- MICHAEL, portail donnant accès à l'inventaire des collections numérisées en Europe, dont l'élargissement en cours permettra de compléter le service existant pour donner accès aux collections de 18 pays au total à partir de mai 2008 (participation française : ministère de la culture et de la communication),
- les équipes de la BnF participent activement au projet TELplus, qui s'attache, depuis octobre 2007, à l'amélioration des fonctionnalités d'accessibilité, de recherche et de services collaboratifs pour la future BNuE. La BnF est coordonnatrice pour le WP3 (amélioration de l'accessibilité) et a une participation importante dans le WP1 (OCRisation de contenus numériques) auquel elle contribuera à hauteur de 7 millions de pages avec un OCR de haute qualité,
- dans le domaine de l'audiovisuel, l'INA coordonne ou participe à de nombreux projets européens et mondiaux destinés à innover dans le domaine de la conservation et de la valorisation du patrimoine audiovisuel. Parmi ces projets, certains comme K-Space, Mesh ou Caspar regroupent plusieurs dizaines de partenaires internationaux (instituts, universités, établissements privés, etc.). L'INA a coordonné le projet PRESTOSPACE (34 partenaires issus de 9 pays), visant à développer des solutions concrètes pour la sauvegarde et la numérisation des archives audiovisuelles à grande échelle selon les critères de qualité, de rapidité, et de réduction des coûts.

Enfin, les partenaires français se sont impliqués dans les propositions suivantes :

- dans le cadre du 7ème PCRD et des projets relatifs aux centres de compétence, la BnF participe au projet IMPACT, proposé par la KB (Pays-Bas) pour l'amélioration des procédés d'OCR.
- projet de création d'un portail européen des archives : la France a contribué au projet de création d'un portail européen des archives (APENET), en cours de négociation suite à l'appel à projets eContentplus clos le 4 octobre dernier. La nécessité d'un tel portail avait été posée par la recommandation du Conseil du 14 novembre 2005. Elle a été renforcée par le lancement du projet de Bibliothèque numérique européenne. Le portail européen des archives devrait permettre, si le projet est accepté, de rendre accessibles des instruments de recherche et des documents numérisés conservés par des services d'archives de divers pays d'Europe. Il s'agit de rendre plus lisible, pour des chercheurs ou des citoyens à la recherche de leurs racines, une histoire européenne complexe, marquée notamment par de nombreux conflits et déplacements de populations. Les ressources ainsi disponibles seront également consultables par le portail Bibliothèque numérique européenne, dont le portail des archives pourra constituer un maillon intermédiaire.
- enfin, toujours dans le même appel à projet eContent+, clos le 4 octobre dernier, et dans le cadre de la mise en place de la bibliothèque numérique européenne, la France participe aux projets ARROW (participation BnF, développement de l'intégration des contenus sous-droits) et ATHENA (participation ministère de la culture et de la communication, développement notamment de l'intégration du secteur des musées), en cours de négociation. Il faut enfin citer la participation de l'INA au projet VideoActive, portail d'accès européen aux collections des archives de radio-télévision.

La mise en place du dépôt légal des sites Internet par la BnF se fait en étroite collaboration avec d'autres bibliothèques et archives nationales européennes regroupées au sein du Consortium international pour la préservation de l'Internet (IIPC). Un grand nombre de bibliothèques nationales de l'Union ont en effet rejoint le consortium, participant à des développements logiciels et à un effort de normalisation qui favorisent la prise en compte de l'interopérabilité et de la complémentarité des stratégies adoptées par chaque Etat membre associé à IIPC. La BnF, qui assure le pilotage technique

du consortium veille tout particulièrement à la cohérence et à l'interopérabilité des solutions retenues dans ce secteur, au niveau européen et international. Elle participe plus globalement à la promotion et à la valorisation de l'archivage du Web dans les instances et associations professionnelles européennes et internationales.

B3. Progress on legal provisions for multiple copying and migration of digital cultural material by public institutions for preservation purposes.

La loi « Droit d'Auteur et Droits Voisins dans la Société de l'Information » (loi n° 2006-961 du 1er août 2006 dite DADVSI) transpose dans le droit français la directive européenne 2001/29/CE sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information.

Le Chapitre I du Titre I traite des "exceptions aux droits d'auteur et aux droits voisins" et amende le code de la propriété intellectuelle (CPI). En matière de conservation à long terme du patrimoine numérique, la loi autorise désormais "la reproduction d'une œuvre, effectuée à des fins de conservation ou destinée à préserver les conditions de sa consultation sur place par des bibliothèques accessibles au public, par des musées ou par des services d'archives, sous réserve que ceux-ci ne recherchent aucun avantage économique ou commercial".

Sont également prévues les exceptions suivantes :

- la représentation ou la reproduction d'extraits d'œuvres (sous réserve de certaines catégories d'œuvres) dans le cadre de l'enseignement et de la recherche, en échange d'une rémunération forfaitaire (à compter du 1er janvier 2009),
- la reproduction (stockage) provisoire des œuvres dans le cadre de procédés techniques permettant l'utilisation ou la transmission de l'œuvre par la voie d'un réseau : cela concerne la mise en cache temporaire effectuée par les fournisseurs d'accès à Internet afin d'optimiser l'accès aux ressources sur le réseau,
- la reproduction et la représentation (par les personnes morales ou par les établissements ouverts au public, tels que bibliothèques, musées, archives...) en vue d'une consultation strictement personnelle par les personnes handicapées,
- "la reproduction ou la représentation, intégrale ou partielle, d'une œuvre d'art graphique, plastique ou architectural par voie de presse écrite, audiovisuelle

ou en ligne dans le but exclusif d'information immédiate et en relation directe avec cette dernière, sous réserve d'indiquer clairement le nom de l'auteur".

La même loi DADVSI a introduit dans le Code du patrimoine une exception au droit de reproduction et de représentation en faveur des organismes en charge du dépôt légal. En vertu de ce texte, l'auteur ne peut interdire aux organismes dépositaires :

1° La consultation de l'œuvre sur place par des chercheurs dûment accrédités par chaque organisme dépositaire sur des postes individuels de consultation dont l'usage est exclusivement réservé à ces chercheurs

2° La reproduction sur tout support et par tout procédé d'une œuvre, nécessaire à la collecte, à la conservation et à la consultation sur place dans les conditions prévues au 1°.

Il en est de même pour les titulaires de droits voisins :

« L'artiste interprète, le producteur de phonogrammes ou de vidéogrammes ou l'entreprise de communication audiovisuelle ne peuvent interdire la reproduction et la communication au public des documents mentionnés à l'article L. 131-2 dans les conditions prévues à l'article L. 132-4 »

Le texte prévoit également que le producteur d'une base de données ne peut interdire l'extraction et la réutilisation par mise à disposition de la totalité ou d'une partie de la base dans les mêmes conditions.

B4. Progress on policies and procedures for the deposit of born-digital material. Please note how you are taking into account developments in other Member States in order to prevent a wide divergence in depositing arrangements.

Les dispositions du code du patrimoine concernant le dépôt légal (article L131-2) incluent les documents numériques : « les documents imprimés, graphiques, photographiques, sonores, audiovisuels, multimédias, quel que soit leur procédé technique de production, d'édition ou de diffusion, font l'objet d'un dépôt obligatoire, dénommé dépôt légal, dès lors qu'ils sont mis à la disposition d'un public. Les logiciels et les bases de données sont soumis à l'obligation de dépôt légal dès lors qu'ils sont mis à disposition d'un public par la diffusion d'un support matériel, quelle que soit la nature de ce support. »

A la BnF, c'est le département de l'Audiovisuel qui est plus particulièrement chargé de recevoir le dépôt légal de tout document multimédia ou base de données diffusés

sur support. L'INA est en charge du dépôt légal de la radio-télévision, le CNC (Centre national de la cinématographie) du film.

En outre, le décret 2006-696 du 13 juin 2006 prévoit que la BnF puisse désormais "demander le dépôt d'un fichier numérique se substituant au dépôt du document imprimé, graphique ou photographique". Enfin, le titre IV de la loi DADVSI précédemment citée (loi n° 2006-961 du 1er août 2006) a ouvert le cadre légal de l'archivage des sites Internet du domaine français :

« Sont également soumis au dépôt légal les signes, signaux, écrits, images, sons ou messages de toute nature faisant l'objet d'une communication au public par voie électronique. ». Cette disposition est reprise dans le code du patrimoine (article L.131-2). Le loi prévoit la « collecte selon des procédures automatiques » (aspiration de sites Internet), mais ouvre également la possibilité de déterminer d'autres modalités en accord avec les éditeurs de sites. Les organismes chargés de ce dépôt légal sont la Bibliothèque nationale de France et l'Institut national de l'audiovisuel. La loi prévoit que la mise en œuvre d'un code ou d'une restriction d'accès ne pourra pas faire obstacle à la collecte par les organismes dépositaires. La BnF envisage de réaliser ce dépôt légal en recourant cumulativement aux trois procédures suivantes :

- des captures massives et automatiques du domaine français .fr;
- des collectes ciblées également automatiques de sites sélectionnés selon des critères garantissant la continuité et l'enrichissement des collections du dépôt légal. Ces collectes automatiques seront effectuées selon une périodicité déterminée par la BnF et qui variera pour les collectes ciblées en fonction des fréquences de mise à jour des publications. En cas d'échec des procédures de collectes automatiques, la remise d'une copie ou extraction de la publication électronique par l'éditeur ou le producteur pourra être envisagée selon des modalités à déterminer en accord avec celui-ci.

La réalisation du projet d'archivage de la Toile s'inscrit également dans un mouvement international de coopération. La BnF assure le pilotage technique du consortium IIPC (Consortium international pour la préservation de l'Internet), créé par onze bibliothèques nationales en 2003 à son initiative et qui fédère les réalisations les plus innovantes dans le monde. En février 2008, le consortium international pour la préservation de l'Internet (IIPC) compte 38 membres, parmi lesquels des représentants des Etats-membres suivants : France, Royaume Uni,

Espagne, Italie, Allemagne, Pays-Bas, Danemark, Norvège, Suède, Finlande, Autriche, Pologne, République Tchèque, Lettonie, Croatie, Slovaquie. Cette configuration assure une forte convergence des intérêts et des politiques des pays de l'Union en matière d'archivage du Web, qui rejoignent également ceux des autres membres, principalement implantés en Amérique du Nord, Asie et Australasie. La diversité des membres d'IIPC favorise la collaboration entre des structures publiques différentes (bibliothèques, archives, universités...) ainsi que des partenariats avec des sociétés ou fondations privées, comme Internet Archive, la société britannique Hanzo Ltd. ou la fondation European Archive. La concertation porte d'abord sur des partages documentaires, chaque institution (bibliothèque ou archive nationale) se concentrant sur la collecte de sites web propres à sa production nationale. Elle implique ensuite une mise en commun de moyens financiers et d'expertises pour le développement de normes et d'outils destinés à l'ensemble de la communauté internationale. Ces applications concernent principalement la collecte automatique de sites, les logiciels d'indexation, les interfaces de consultation et les stratégies de préservation. Ainsi, tous les logiciels développés par les membres du consortium, tels le robot de moissonnage Heritrix, sont des logiciels libres.

Par ailleurs, le nouveau format des archives du Web, le format WARC, a été essentiellement conçu par des membres d'IIPC. Il est en cours de normalisation à l'ISO. Enfin, le consortium constitue un forum d'échange d'expériences et de bonnes pratiques particulièrement utile, au niveau technique, organisationnel, documentaire ou juridique. Tant pas son rôle de pilotage technique et administratif du consortium que par ses propres choix techniques, la BnF s'inscrit pleinement dans ce mouvement de coopération. Au sein de l'Union, elle a en outre développé des partenariats privilégiés avec la British Library (cofinancement d'un projet de "smart crawler") et la Bibliothèque Royale du Danemark (la BnF projette d'utiliser le logiciel libre de gestion des archives Netarkive Suite).

B5. Progress on legal provisions for the preservation of web-content by mandated institutions.

La loi sur les Droits d'auteur et droits voisins dans la société de l'information (DADVSI, titre IV)

du 1er août 2006 étend le dépôt légal aux « signes, signaux, écrits, images, sons ou messages faisant l'objet d'une communication au public par voie électronique ». La BnF se doit donc d'intégrer à ses fonds les nouvelles formes éditoriales « nées numériques », dans l'esprit du dépôt légal depuis cinq siècles : constituer un patrimoine représentatif de la dynamique éditoriale nationale, sans préjuger de la qualité ni de la forme de ses contenus.

Des collectes automatiques de l'Internet, de grande échelle (instantanés annuels), sont réalisées depuis 2004 en partenariat avec la fondation américaine Internet Archive au moyen du robot Heritrix. La BnF a par ailleurs construit en 2006 l'infrastructure informatique lui permettant de gérer de manière autonome des collectes thématiques de taille moyenne. Le projet pilote de collecte de sites électoraux « Internet en campagne » (réalisé d'octobre 2006 à juillet 2007, avec la collaboration de Sciences Po, le Forum des droits sur Internet, mais aussi huit bibliothèques municipales de dépôt légal imprimeur) a ainsi servi de laboratoire à la mise en place d'une filière complète de traitement des sites (collecte, indexation, sauvegarde).

L'archivage de la Toile est ainsi devenu une activité quotidienne de la BnF. Les collections réunies à ce jour représentent 110 Téraoctets de données dont certaines

remontent jusqu'à 1996, soit 10 milliards de documents. Face à cette nouvelle mission, les priorités de l'établissement pour les trois prochaines années concerneront :

- L'ouverture des accès au public : un décret doit prochainement préciser les conditions de communication au public des archives de l'Internet. C'est à partir de 2008 que les chercheurs devraient pouvoir commencer à explorer en salle de lecture ce nouveau matériau scientifique et expérimenter de nouveaux services de recherche (recherche par URL grâce à la Wayback Machine, indexation plein texte...).
- La montée en charge des capacités de production de la BnF, afin que l'établissement réalise lui-même des « instantanés » de grande échelle et soit en mesure de les préserver durablement à partir de 2009.
- De nouvelles campagnes de collectes thématiques
- L'extension à d'autres filières de collecte numérique qui permettront d'archiver des fichiers que les collectes automatiques, pour des raisons techniques ou juridiques, ne permettent pas de capturer.
- La recherche, l'amélioration et la création d'outils de collecte, d'indexation et d'accès aux archives qui répondront aux attentes des chercheurs de demain.

