

Appel à partenariat pour la numérisation et la valorisation des collections de la Bibliothèque nationale de France (BnF)

Mercredi 6 juillet 2011

Contacts presse

Département de l'information et
de la communication
01 40 15 74 71
service-presse@culture.gouv.fr

Commissariat général à l'Investissement
01 42 75 64 43
coralie.mercier@pm.gouv.fr

Bibliothèque nationale de France
01 53 79 41 18
claudine.hermabessiere@bnf.fr

www.culture.gouv.fr



Sommaire

Appel à partenariat pour la numérisation et la valorisation des collections de la Bibliothèque nationale de France (BnF)

Communiqué de presse

Fiche 1 : Présentation de l'appel à partenariats pour la numérisation et la valorisation des collections de la BnF

Fiche 2 : Les Corpus

Fiche 3 : Le Système de Préservation et d'Archivage Réparti (SPAR)

Fiche 4 : Gallica, la bibliothèque numérique de la BnF

ANNEXE

L'appel à partenariats

Communiqué de presse

Communiqué de presse



Paris, le 6 juillet 2011

**INVESTISSEMENTS D'AVENIR :
LANCEMENT D'UN APPEL À PARTENARIAT POUR LA NUMÉRISATION ET LA VALORISATION DES COLLECTIONS BnF**

Frédéric Mitterrand, ministre de la Culture et de la Communication, René Ricol, commissaire général à l'Investissement, et Bruno Racine, président de la Bibliothèque nationale de France (BnF) ont lancé mercredi 6 juillet 2011, un appel à partenariat dans le cadre des Investissements d'Avenir pour la numérisation et la valorisation des collections de la BnF.

Depuis une dizaine d'années, la BnF s'est engagée dans une politique de numérisation de ses collections qui permet d'enrichir sa bibliothèque numérique Gallica, composée actuellement d'un million et demi de documents. Ceux-ci sont également accessibles via Europeana, la bibliothèque numérique européenne dont la BnF est l'un des tout premiers contributeurs. Ces programmes de numérisation sont financés par une subvention du Centre national du Livre (CNL) et par le budget propre de la BnF. L'enveloppe annuelle est d'environ 7M€ depuis 2007.

Pour renforcer et accélérer la numérisation de ses collections, la BnF souhaite s'inscrire dans le programme des Investissements d'Avenir, voulu par le Président de la République. À cette fin, elle a identifié les 12 corpus suivants, qui sont susceptibles d'être numérisés en partenariat avec des acteurs privés qui souhaiteraient les valoriser commercialement : le livre ancien de 1470 à 1700 ; un ensemble de 3 à 500 000 imprimés français du domaine public ; la presse française de 1780 à 1940 ; les disques 78 tours et microsillons ; les manuscrits médiévaux ou modernes ; les collections de portraits ; la généalogie et l'histoire des familles ; les territoires : cartographie et représentations de la France et des pays étrangers ; l'histoire locale ; les partitions musicales ; la photographie ; le cinéma français des origines à la deuxième guerre mondiale.

Certains de ces corpus sont susceptibles d'une extension nationale, ce qui permettra de faire bénéficier d'autres bibliothèques, à Paris et en région, notamment municipales ou universitaires, du dispositif des Investissements d'Avenir.

Le coût global de ces programmes de numérisation s'élèverait à quelque 150 millions d'euros, soit l'équivalent de plus de vingt années au rythme actuel.

La numérisation de tous ces ensembles avec des partenaires privés qui, en échange d'une possibilité de valorisation commerciale, contribuerait au financement de la numérisation, permettrait de tripler le volume actuel de Gallica en quelques années.

Avec l'appui du Fonds national pour la Société Numérique (FSN), fonds mis en place par l'Etat et géré par la Caisse des Dépôts dans le cadre du programme « développement de l'économie numérique des « Investissements d'avenir », la BnF sera amenée à créer une filiale détenue à 100 %, « BnF-Partenariats ». Sa mission sera de conclure des accords de numérisation et de valorisation commerciale de ses collections en tenant compte des recommandations du Comité des Sages de l'Union Européenne. Cette filiale sera mise en place avant la fin de l'année 2011. Les projets qui en découleraient sont susceptibles d'être financés par le FSN.

Enfin, un appel à partenariat distinct porte sur la valorisation du Système de Préservation et d'Archivage Réparti (SPAR), développé et utilisé par la BnF depuis 2007 pour la conservation pérenne de ses ressources numériques.

Pour en savoir plus sur les investissements d'avenir, consultez le site internet :

<http://investissement-avenir.gouvernement.fr>

Contacts presse :

- Ministère de la Culture et de la Communication : 01 40 15 74 71, service-presse@culture.gouv.fr
- Commissariat général à l'Investissement : 01 42 75 64 43, coralie.mercier@pm.gouv.fr
- Bibliothèque nationale de France : 01 53 79 41 18, claudine.hermabessiere@bnf.fr

Fiche 1 : Présentation de l'appel à partenariat pour la numérisation et la valorisation des collections de la BnF

Le 6 juillet 2011, la BnF lance un appel à partenariat pour proposer, dans le cadre du programme des investissements d'avenir, la numérisation et la valorisation de 12 corpus essentiels parmi ses collections présentant un intérêt manifeste pour la BnF, le public et le secteur privé.

Les projets déposés par les partenaires seront sélectionnés à partir des critères d'évaluation suivants :

- Qualité et ampleur de la numérisation et durée de réalisation ;
- Originalité, intérêt scientifique et culturel et dimension du projet de valorisation ;
- Expérience, capacité du partenaire à porter le projet ;
- Modalités d'accès aux collections sur Gallica et Europeana ;
- Intérêt et réalisme financier du projet

Dans ce cadre, la BnF :

- rend accessibles ses collections non numérisées ;
- autorise la numérisation de ses collections ;
- autorise le partenaire à exploiter et valoriser, dans son secteur d'activité, la version numérisée du corpus complétée, le cas échéant, par des collections déjà numérisées ;
- assure la mise en ligne des collections dans ses emprises et sur Gallica, selon les modalités à définir avec le partenaire et signale l'ensemble de ces collections dans Europeana.

Les caractéristiques de la filiale et ses missions

Pour bénéficier des apports du Fonds pour la société numérique (FSN) dans le cadre des investissements d'avenir, la BnF est amenée à créer une filiale « BnF -Partenariats ». Sa mission sera de conduire les opérations partenariales de numérisation et de valorisation de ses fonds.

Détenue à 100% par la BnF, la filiale « BnF -Partenariats » sera mise en place avant la fin de l'année 2011.

Cette filiale, sous le contrôle de la BnF, mettra en œuvre des projets de développement et de valorisation de ses ressources numériques dans le cadre des investissements d'avenir, conformément aux orientations stratégiques et scientifiques de l'établissement, et en partenariat avec d'autres structures publiques ou privées.

Le calendrier

Date de limite de réponse : le 20 octobre 2011

Analyse approfondie des dossiers pendant l'automne 2011

Finalisation de l'ensemble des paramètres techniques, juridiques et financiers : premier trimestre 2012.

Fiche 2 : Les Corpus

Corpus 1 : Le livre ancien de 1470 à 1700

Corpus 2 : Un ensemble de 300 000 à 500 000 imprimés français du domaine public

Corpus 3 : La presse française de 1780 à 1940

Corpus 4 : Les disques 78 tours et les microsillons

Corpus 5 : Les manuscrits médiévaux ou modernes

Corpus 6 : Les collections de portraits

Corpus 7 : La généalogie et l'histoire des familles

Corpus 8 : Les territoires : cartographie et représentations de la France et des pays étrangers

Corpus 9 : L'histoire locale

Corpus 10 : Les partitions musicales

Corpus 11 : La photographie

Corpus 12 : Le cinéma français des origines à la deuxième guerre mondiale

Corpus 1 : Le livre ancien français de 1470 à 1700

Le patrimoine imprimé le plus rare et le plus précieux.

L'intégralité des collections exceptionnelles de livres anciens conservés à la BnF, des origines de l'imprimerie à 1700, dans toutes les disciplines (histoire, philosophie, religion, droit, sciences, littérature, art...)

- 2 500 incunables (livres imprimés avant 1500)
- 25 000 ouvrages de la Renaissance
- environ 40 000 ouvrages du XVIIe siècle

Un projet susceptible d'extension aux collections des bibliothèques publiques et universitaires de France sous réserve de leur accord pour atteindre :

- 4 200 incunables
- 65 000 ouvrages de la Renaissance
- Environ 100 000 ouvrages du XVIIe siècle

Un projet pour tous les chercheurs français et étrangers, pour les bibliophiles et pour les amateurs de livres, d'éditions originales, de reliures, d'illustrations.

Des potentiels forts pour les enseignants, les scolaires et les étudiants, et le commerce du livre ancien.



François Rabelais

Pantagrueline prognostication, certaine, véritable et infalible ["sic"] pour l'an mil DXXXIII. Nouvellement composé, au prouffit et advisement de gens estourdis et musars de nature, par Maistre Alcofribas, architriclin dudidict Pantagruel

BnF, Réserve des Livres rares

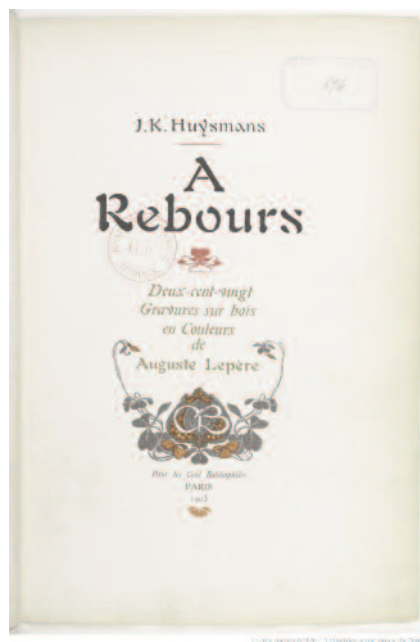
Corpus 2 : Un ensemble de 300 000 à 500 000 imprimés français du domaine public

Une contribution majeure à l'extension des collections numériques de la BnF par la numérisation de 300 000 à 500 000 ouvrages et revues de 1701 à 1930 et dans toutes les disciplines (histoire, philosophie, religion, droit, sciences, littérature, art...) afin d'atteindre un jour le tout numérique pour le patrimoine imprimé français. Ce corpus moderne complète celui qui est proposé sur la période 1470-1700.

Un principe clair : numériser des documents qui existent à la BnF en plusieurs exemplaires de sorte que le travail des chercheurs en salles de lecture soit parfaitement préservé.

Des possibilités de privilégier une période (par exemple le siècle des Lumières) ou des thématiques (par exemple l'histoire de France).

Un projet pour tous les publics de Gallica : chercheurs institutionnels ou individuels, enseignants, étudiants, scolaires, passionnés, curieux ...



Jors-Karl Huysmans
A rebours

Corpus 3 : La presse française de 1780 à 1940

Une des sources majeures de la mémoire nationale, des documents fragiles très difficiles à conserver dans leur intégrité en particulier sur la période 1780-1945, une des grandes priorités de la BnF.

Le projet le plus ambitieux de sauvegarde de la presse jamais proposé par la BnF : de 3 à 16 millions de pages, selon les ambitions du partenaire.

Une grande diversité documentaire : presse quotidienne nationale, presse quotidienne régionale, presse hebdomadaire, presse d'opinion, presse féminine, presse satirique, presse enfantine, presse ouvrière...

La presse figure parmi les documents les plus consultés dans Gallica comme dans la plupart des bibliothèques numériques étrangères (Grande-Bretagne, Mexique, Allemagne, Autriche, pays scandinaves...). Elle intéresse les chercheurs comme tout un chacun, notamment pour les recherches généalogiques ou personnelles.



Le Petit journal illustré (La France en danger ; Guerre d'Espagne)

Corpus 4 : Les disques 78 tours et les microsillons

L'intégralité du patrimoine sonore édité en France de la fin du XIXe siècle à 1960.

300 000 titres de 78 tours (1887-1957) et 70 000 titres de microsillons (1949-1960).

Une très grande diversité de répertoires : airs d'opéra, musique classique, chansons de tous les pays, jazz, musique de danse, archives sonores ... feront de cette collection un grand voyage dans le temps, de Bach aux yéyés.

La consultation gratuite d'extraits sur Gallica est un apport majeur à l'offre légale et payante de musique en ligne, chez soi, comme de façon nomade, pour tous les publics, les nostalgiques et les branchés.

En complément, la numérisation des pochettes et de leurs notes pour les collectionneurs et les fans clubs.



La Marseillaise / Rouget de L'Isle, comp. ; F. I. Salâpin, B.
BnF, département de l'Audiovisuel

Corpus 5 : Les manuscrits médiévaux ou modernes

Les manuscrits enluminés

La numérisation d'une collection majeure pour l'histoire des idées et l'histoire de l'art en Europe, du II^e au XVIII^e siècle, et de tous les pays occidentaux : le Pentateuque de Tours, l'Evangélaire de Charlemagne, les deux Bibles de Charles le Chauve, le psautier de saint Louis, *les Grandes chroniques de France* de Charles V, *les Petites et Grandes Heures du duc de Berry*, *le Livre des Merveilles*, *les Grandes Heures d'Anne de Bretagne*, par les plus grands maîtres de la peinture dans les livres...

Un projet pour les chercheurs, mais aussi un très fort potentiel pour tous les publics qui aiment la littérature courtoise, les chroniques historiques, les romans de chevalerie, l'imagerie du Moyen-Âge, et notamment les bestiaires, les images de métiers, les scènes religieuses, les témoignages sur la vie de Cour...

Des possibilités de choisir des corpus thématiques (par exemple les romans de Chrétien de Troyes, les Bibles ou les livres d'heures, les bibliothèques royales ou princières...) ou de numériser l'intégralité des collections (15 000 manuscrits, 500 000 images).

Une possibilité d'extension du projet à tous les manuscrits médiévaux de la BnF ou aux grandes collections françaises des bibliothèques publiques ou universitaires sous réserve de leur accord, soit 70 000 manuscrits et 1 million d'images.

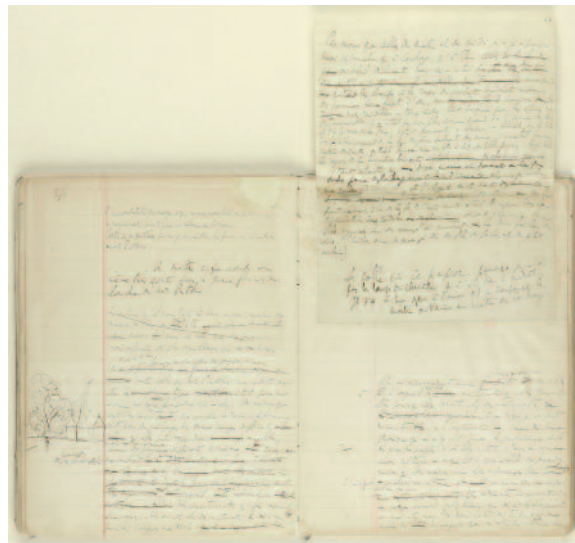


Petites heures de la reine Anne de Bretagne (XV^e s.)

BnF, département des Manuscrits

Les manuscrits littéraires

Une opportunité de bâtir la bibliothèque des manuscrits de plus de mille chefs d'œuvre de la littérature et de la pensée française, grands romans (*La Chartreuse de Parme* de Stendhal, *Les Liaisons dangereuses* de Choderlos de Laclos, *Les Femmes galantes* de Brantôme...) indispensables de la poésie (*Alcools* d'Apollinaire, *Une saison en enfer* de Rimbaud, *Stèles* de Segalen...) mémoires et correspondances célèbres (Saint-Simon, Benjamin Constant, Gide, les frères Goncourt ...) soit environ 1 million de pages.



Source : gallica.bnf.fr - Bibliothèque nationale de France

Marcel Proust, *A la recherche du temps perdu - Le côté de Guermantes*
BnF, département des Manuscrits

Corpus 6 : les collections de portraits

Une collection sans équivalent de 340 000 portraits de 150 000 personnes françaises ou étrangères du Moyen-Âge au début du XXe siècle : souverains, présidents, ministres et autres personnalités politiques, écrivains, penseurs, acteurs, musiciens, chanteurs, hommes d'affaires, inventeurs, sportifs, héros de faits divers, venus de fonds d'ateliers de graveurs ou de photographes (Disdéri, Reutlinger...) ou signés de grands artistes (Clouet, Robert Nanteuil, Gauguin, Daumier, Man Ray, Nadar...).

Un projet pour tous les publics et pour tout type de recherches.

Une encyclopédie sociologique extraordinaire : au-delà des personnes, mille lectures possibles des groupes sociaux (nationalités, groupes ethniques, classes sociales, classes d'âge, religions, professions...)



Source gallica.bnf.fr / Bibliothèque nationale de France

R.J. Bingham, *Recueil. Célébrités du 19e siècle: la reine Victoria*
BnF, département des Estampes et de la photographie

Corpus 7 : La généalogie et l'histoire des familles

Des ressources très connues mais encore peu valorisées au service de la nouvelle passion internationale pour la généalogie et l'histoire des familles.

Le « Cabinet des titres » : 6 800 volumes – 1,6 million d'images : un ensemble exceptionnel de documents relatifs à toute la noblesse de France, contenant de nombreuses pièces autographes de personnages célèbres.

La collection Clairambault : 1 354 volumes – 340 000 images, du nom du généalogiste des Ordres du roi sous Louis XIV et Louis XV, qui contient notamment une partie des archives du Cabinet de l'Ordre du Saint-Esprit et de la Chambre des comptes, toutes deux détruites pendant la Révolution française.

Des sources nouvelles et de première main pour les historiens, les chercheurs et toutes les personnes investies dans la recherche généalogique.

Un projet à rapprocher de l'histoire locale (corpus 9).



D'Hozier
Armorial général de France
BnF, département des Manuscrits

Corpus 8 : Les territoires, cartographie et représentations de la France et des pays étrangers

Une sélection exceptionnelle de cartes et plans (100 000 pièces) et de 550 000 images représentant la France et certains pays étrangers, de la Renaissance à aujourd'hui : plans de villes, cartes routières, cartes militaires, territoires de France mais aussi d'Outre-mer, cartographie marine...

Dans cet ensemble, une collection de photographies à redécouvrir, celle du Touring Club de France, dont la plupart des Français ont le souvenir, et d'innombrables cartes postales.

Un projet au service de l'histoire, de la géographie, de la généalogie, mais aussi des préoccupations environnementales, de l'urbanisme, de l'aménagement du territoire, un voyage au cœur des paysages et des villes.

Des potentiels forts pour les usages scientifiques, pédagogiques, professionnels, mais aussi pour les usages nomades (géolocalisation), le tourisme.



Source : gallica.bnf.fr / Bibliothèque nationale de France

Agnus aîné, *Plan-omnibus de Paris*
BnF, département des Cartes et plans

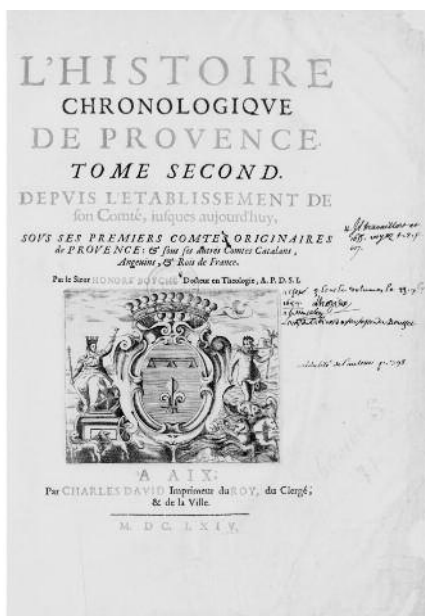
Corpus 9 : L'histoire locale

Deux ensembles exceptionnels dont la numérisation faciliterait la large diffusion :

- L'Histoire des provinces de France : 2585 volumes – 800 000 images : collection de documents et mémoires sur l'histoire des provinces de France réunie par des congrégations bénédictines : chartes, archives de châteaux, registres de Cours des comptes ...
- La collection Moreau : 1 834 volumes – 550 000 images : réunie par Jacob-Nicolas Moreau, historiographe du roi (1717-1804), elle rassemble des copies des chartes disséminées dans les innombrables chartiers du royaume : l'histoire administrative de la France en train de se faire. Egalement de très nombreux documents copiés en Angleterre, dans les Pays-Bas autrichiens, à Rome, en Bourgogne, en Franche-Comté et au parlement de Paris.

Un projet scientifique pour les chercheurs et les sociétés savantes mais aussi au service des historiens amateurs, des généalogistes, de tous ceux qui s'intéressent à l'histoire de leur territoire ou de leur localité.

Un fort potentiel auprès des collectivités territoriales, des centres d'archives privés ou publics.



Honoré Bouche

La Chorographie ou Description de Provence, et l'histoire chronologique du même pays, par le sieur Honoré Bouche...

Corpus 10 : Les partitions musicales

La musique dans sa version écrite, partition manuscrite ou publiée, prestigieuse ou courante.

Un choix très ouvert de trésors de la musique (100 000 pages de Bach, Rameau, Schubert, Berlioz, Massenet, Fauré, Debussy, Ravel, Satie...) mais aussi près d'un million de pièces musicales de toute nature et notamment de chansons, dont beaucoup à redécouvrir : chansons politiques, chansons militaires, comptines, chansons d'amour, chansons licencieuses, chansons à boire... ainsi que des répertoires célèbres ou oubliés.

Un potentiel fort auprès des musicologues, des mélomanes, des musiciens professionnels ou amateurs, des sociétés d'édition de musique, pour découvrir ou redécouvrir la musique savante ou populaire, mais aussi des historiens et des sociologues pour les nombreuses illustrations des partitions.

Un potentiel technologique certain (océrisation de la musique, exécution automatisée...).

Le projet sur les partitions musicales pourrait être rapproché du projet sur les disques 78 tours et les microsillons (corpus 4). L'écoute du son pourrait s'appuyer simultanément sur la musique écrite.



C.W. Gluck, *La clemenza di Tito*
BnF, département de la Musique

Corpus 11 : La photographie

Dans l'immense collection de photographies de la BnF, trois ensembles cohérents et passionnants :

- la photographie de mode à travers le fonds des frères Séeberger : toutes les modes de 1900 à 1975, de Deauville à New-York, de Saint-Moritz à Londres (290 000 images) ;
- des fonds intégraux d'agences de presse (Rol, Meurisse, Mondial, SAFARA) actives à partir de 1900 et disparues après la deuxième guerre (200 000 images) : toute l'actualité, inscrite dans les mémoires ou oubliée, de la vie politique et des événements internationaux aux faits divers et aux caprices de la météo ;
- une histoire technique, plastique et sociale de la photographie des origines à 1920 (30 000 images) : les grandes inventions, les grands mouvements, les grandes signatures et les premiers usages commerciaux.

Un projet pour tous les publics et pour toutes les recherches.



Gustave Le Gray, *Effet de soleil sur l'océan*

BnF, département des Estampes et de la Photographie

Corpus 12 : Le cinéma français des origines à la deuxième guerre mondiale

Un ensemble exceptionnel de 30 000 documents (630 000 images) dont beaucoup de documents inédits ou peu connus sur le premier âge du cinéma, réuni par un collectionneur, Auguste Rondel, banquier passionné de spectacles.

Constituée au moment même de l'émergence du septième art, la collection rassemble des livres, des journaux, des photographies, mais aussi d'innombrables documents d'archives (affichettes, programmes, brochures publicitaires, tickets...) dont beaucoup n'ont été conservés que par Rondel.

Un projet pour la recherche sur cette industrie naissante devenue un art, et pour tous les passionnés de vedettes, de réalisateurs, de genres de films, d'anecdotes sur le monde du cinéma.

Un matériel de premier ordre pour des bonus de DVD ou de VOD et toute édition scientifique.



Pichon, *Parisiana le roi des cinémas*
BnF, département des Estampes et de la Photographie

Fiche 3 : Le Système de Préservation et d'Archivage Réparti (SPAR)

Afin de préserver son patrimoine culturel numérisé et de le conserver à très long terme, la BnF a conçu et développé un système de pérennisation de ses ressources numériques appelé SPAR (Système de Préservation et d'Archivage Réparti) qui s'appuie sur des normes internationales faisant autorité dans le monde de la pérennisation des informations numériques (notamment OAIS).

Bien plus perfectionné qu'un simple entrepôt de données sécurisé, le système effectue de multiples copies sur disques et sur bandes des données numériques stockées et assure une surveillance continue de l'état des équipements et des médias d'enregistrement des fichiers permettant d'anticiper les recopies. Il permet également, grâce à une reconnaissance précise 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.

Mis en service en mars 2010, le système archive désormais la production numérique courante de la BnF. Il gère une collection numérique de près de 1000 téraoctets et qui ne cesse de croître. Indépendamment des coûts liés à l'infrastructure matérielle du système, l'investissement consenti jusqu'ici pour la réalisation logicielle s'est élevée à 3,3 M€.

L'appel à partenariat vise à valoriser l'investissement opéré en permettant à des acteurs privés ou publics de bénéficier de ces développements afin qu'eux-mêmes puisse développer une offre commerciale sur cette base. Que ce soit pour une offre de type SaaS (Software as a Service) ou pour répondre à des appels d'offre relatifs à la préservation numérique, la BnF a la conviction que les acteurs économiques pourront y trouver un intérêt.

Fiche 4 : Gallica, la bibliothèque numérique de la BnF

Gallica, la bibliothèque numérique de la Bibliothèque nationale de France, offre près d'un million et demi de documents numérisés dont 300 000 livres, 812 000 périodiques, 340 000 images, 12 000 manuscrits et 23 000 cartes issus des collections de la BnF ou de bibliothèques partenaires.

On y trouve aussi bien des livres, des périodiques, des estampes, des cartes postales et des affiches que des partitions musicales, des documents sonores, des manuscrits ou des cartes.

De nombreux documents sont plébiscités par les internautes comme les manuscrits de Casanova entrés récemment dans les collections de la BnF ou la presse quotidienne numérisée (Le Figaro, l'Humanité, la Croix, ...). De même, le zoom ou le téléchargement font partie des fonctionnalités les plus appréciées par les Gallicanautes.

Aujourd'hui consultée en moyenne par 25 000 visiteurs par jour, Gallica a connu une progression de 85 % de son audience en 2010 passant de 4 à 7,5 millions de visites.

Pour favoriser la réappropriation des contenus au-delà du site Gallica, la BnF a mis au point un lecteur exportable qui permet d'implanter et de feuilleter les documents sur les sites web et les blogs et, depuis peu, sur les pages Facebook. Cette initiative a été récompensée par le prix i-expo de l'innovation 2011 décerné à la BnF par le Groupement français de l'industrie de l'information (GFII).

Plusieurs actions permettent à la BnF de renforcer sa politique de diffusion des contenus. Le partenariat signé en octobre 2010 entre Microsoft et la BnF vise à mieux indexer et mettre en valeur les contenus de Gallica, facilitant ainsi l'accès au patrimoine numérisé. La BnF a signé récemment un marché avec la société EXALEAD pour remplacer, d'ici à 2012, l'actuel moteur de recherche de Gallica.

<http://gallica.bnf.fr/>

Gallica sur Twitter : <http://twitter.com/#!/GallicaBnF>

Gallica sur Facebook: <http://www.facebook.com/GallicaBnF>

ANNEXE

L'appel à partenariats

APPEL À PARTENARIAT

POUR LA NUMÉRISATION ET LA VALORISATION DES COLLECTIONS DE LA BnF ET

POUR LA VALORISATION

DU SYSTEME DE PRESERVATION ET D'ARCHIVAGE REPARTI DEVELOPPE ET UTILISÉ PAR LA BnF

POUR SES DONNÉES NUMÉRIQUES

I – Appel à partenariat pour la numérisation et la valorisation des collections BnF

1. Cadre général

1.1. Contexte et enjeux de la numérisation et de la valorisation des contenus BnF

La BnF dispose de collections d'une ampleur et d'une richesse exceptionnelles. Elle s'est engagée depuis une dizaine d'années dans une politique de numérisation et de mise à disposition de ses collections, à travers sa bibliothèque numérique Gallica (gallica.bnf.fr).

- livres imprimés (13 millions dont 300 000 numérisés),
- presse (plusieurs dizaines de millions de pages dont 3 millions numérisées),
- collections audiovisuelles (qui comprennent notamment 200.000 disques 78 tours dont 1.000 numérisés),
- collections dites spécialisées (manuscrits, estampes, photographies, monnaies et médailles, cartes et plans, partitions, etc.).

Les programmes de numérisation et de valorisation déjà engagés présentent un intérêt manifeste pour la BnF, le public et le secteur privé.

Pour la BnF, ces programmes permettent de garantir la conservation patrimoniale de collections dont le support d'origine peut se dégrader.

Ils entraînent également une diffusion croissante des collections de la BnF favorisant l'accès de tous à la culture. En effet, ils ont permis de constituer une bibliothèque numérique composée aujourd'hui de près d' 1 500 000 documents. Gallica a désormais atteint une taille critique qui suscite l'intérêt croissant du public puisque le site a reçu plus de 7,5 millions de visiteurs en 2010 (+85% par rapport à 2009). Ces programmes sont en outre une des sources principales d'enrichissement d'Europeana, la bibliothèque numérique européenne, qui constitue une priorité de la BnF.

Enfin, ces collections numérisées ont très largement contribué à l'émergence d'un marché permettant la commercialisation de ressources numériques et la valorisation des actifs immatériels de la BnF.

Pour des partenaires (entreprises privées, fondations...), cette offre large de documents numérisés présente de multiples intérêts commerciaux : elle permet notamment l'enrichissement de sites spécialisés (généalogie, musique), la création de bouquets documentaires à forte valeur ajoutée, la création de nouveaux services dont les modèles économiques sont, pour certains, d'ores et déjà stabilisés.

L'ensemble de ces programmes de numérisation ont été conduits jusqu'à maintenant à l'aide de crédits publics, qu'ils proviennent du Centre national du Livre ou directement du budget de la Bibliothèque nationale de France.

Pour renforcer cette politique de numérisation, la BnF a identifié parmi ses collections, différents corpus susceptibles d'intéresser des partenaires pour une numérisation et une exploitation commerciale directe.

Ainsi, en complément des programmes qui entrent dans le circuit habituel de financement de numérisation sur crédits publics, la BnF lance un appel à partenariat pour proposer la numérisation et la valorisation de ces corpus à des partenaires.

Une partie du financement de la numérisation de ces corpus sera portée par un ou plusieurs partenaires, accompagné d'un droit d'exploitation commerciale des fichiers numériques.

En parallèle, une filiale sera créée par la BnF, chargée de piloter les opérations partenariales de numérisation et de valorisation dans le cadre du programme Investissements d'Avenir. La mise en place de cette filiale est en cours.

La BnF invite les partenaires potentiels à tenir compte des recommandations du « comité des Sages » de l'Union Européenne.

1.2. Le dispositif des investissements d'avenir lancé par le Gouvernement

Dans le cadre du programme Investissements d'Avenir, le Gouvernement a décidé la création du Fonds national pour la Société Numérique (FSN), destiné à soutenir le développement de l'économie numérique tant au niveau des réseaux à très haut débit que des services tirant profit de ces réseaux.

Il intervient avec l'objectif de participer au financement des projets rentables notamment pour la « valorisation et numérisation des contenus scientifiques, éducatifs et culturels ». Qu'il s'agisse du livre, de la presse, de la musique, de la photographie, des contenus techniques et scientifiques, il est jugé pertinent par le FSN d'accompagner le développement d'acteurs innovants sur les technologies de la numérisation, création, gestion et valorisation des contenus numériques

Une attention particulière sera portée aux modèles économiques proposés par les partenaires intéressés et à leur capacité à concilier l'accès le plus large aux contenus numérisés et un mode de valorisation adapté, tout en assurant le respect des droits d'auteur et des droits de propriété intellectuelle.

Les projets de numérisation et de valorisation décrits ci-dessous peuvent être éligibles à un soutien financier du FSN.

Dans ce cas, le FSN dotera la filiale BnF de fonds qui lui permettront de procéder à un apport financier pour accompagner l'investissement du ou des partenaires, en fonction du type et de l'importance des projets envisagés.

Ces fonds serviront à régler le coût de la mise à disposition physique des collections, de leur préparation éventuelle, de l'élaboration de leurs métadonnées (catalogage et exemplarisation) et des capacités de stockage supplémentaires ainsi qu'éventuellement une partie du coût de la numérisation, l'autre partie étant prise en charge par le partenaire selon des modalités à définir.

2. Champ de l'appel à partenariat

2.1 Type de projets

L'appel à partenariat propose douze corpus à numériser et à valoriser qui feront chacun l'objet d'une annexe détaillée.

Chaque partenaire peut répondre à un ou plusieurs projets. Pour chaque corpus, la proposition du partenaire doit intégrer un dispositif de numérisation et de valorisation des collections une fois celles-ci numérisées.

L'intérêt de ces projets de numérisation se fonde sur « l'effet longue traîne » de l'univers numérique, où la valeur des contenus est liée à la complétude de l'offre. La valorisation des contenus numérisés doit pouvoir permettre de diffuser largement les collections et de générer ainsi, en fonction du modèle économique et du secteur d'activité de chaque partenaire, un retour sur investissement à un horizon de temps adapté.

Certains projets proposent des lots qui peuvent correspondre à des phases et visent à donner aux partenaires potentiels des échelles d'ambition. Ce phasage éventuel peut être objet de discussions et les partenaires devront se prononcer sur leur opportunité et leur faisabilité.

Certains projets comportent une possibilité d'extension à d'autres collections que celles de la BnF afin de leur donner une dimension nationale. De telles extensions impliquent l'accord des institutions évoquées et de leurs tutelles respectives (collectivités territoriales, Ministère de la Culture et de la Communication, Ministère de l'Enseignement Supérieur et de la Recherche, ...). Ces deux derniers ministères appuient cette initiative. Ces élargissements éventuels sont cependant donnés sous toute réserve.

De leur côté, les partenaires peuvent également proposer d'autres pistes. Encyclopédique et enracinée dans l'histoire, la BnF est en mesure de constituer des corpus dans de très nombreux domaines, du Moyen âge à la période contemporaine. On citera par exemple : la documentation et l'imagerie scientifique, l'architecture, les arts plastiques, les arts décoratifs, les arts du spectacle et en particulier le théâtre, l'histoire de France. Il serait également possible d'élaborer des projets autour de certains types de documents (les monnaies, l'imagerie populaire, l'imagerie publicitaire...). Les suggestions des partenaires sont bienvenues et seront étudiées avec soin.

Enfin, pour la plupart de ces corpus, la logique thématique et les besoins de la valorisation peuvent amener les partenaires potentiels à souhaiter ajouter aux collections à numériser des ressources d'ores et déjà numérisées et disponibles dans Gallica. Il en va ainsi par exemple de l'histoire locale, de la généalogie, de la presse, de la cartographie où des ressources numériques disponibles pourraient venir compléter les futurs ensembles.

2.2. Collections concernées

Corpus 1 : Le **livre ancien français** de 1470 à 1700

Corpus 2 : Un ensemble de **documents imprimés français du domaine public** (300.000 à 500 000 documents).

Corpus 3 : La **presse française du Journal de Paris (1777-1827) à Paris-Soir (1923-1943)**

Corpus 4 : Les **disques 78 tours et microsillons**

Corpus 5 : Les **manuscrits**

Corpus 6 : Célébrités et Anonymes : les **collections de portraits**

Corpus 7 : La **généalogie et l'histoire des familles**

Corpus 8 : Les **territoires : cartographie et représentations de la France et des pays étrangers**

Corpus 9 : L'**histoire locale**

Corpus 10 : Les **partitions musicales**

Corpus 11 : La **photographie**

Corpus 12 : Le **cinéma français des origines à la deuxième guerre mondiale**

Les informations détaillées concernant les corpus (contenu, paliers, aspects techniques) sont consultables dans les annexes I.1 à I.12. L'annexe I.14 précise les consignes techniques

3. Dispositions générales

La BnF :

- rend accessibles ses collections non numérisées.
- autorise la numérisation de ses collections.
- numérise le corpus pour le compte du partenaire si celui-ci désire passer par la filière de numérisation de la BnF.
- autorise le partenaire à exploiter et valoriser, dans son secteur d'activité, la version ainsi numérisée du corpus complétée, le cas échéant, par des collections déjà numérisées, pour favoriser une complétude de l'offre.
La durée et les conditions d'exploitation sont à définir.
- assure la mise en ligne des collections dans ses emprises
- assure la mise en ligne des collections sur le site Gallica, selon des modalités à définir avec le partenaire, et signale l'ensemble de ces collections dans Europeana.
- autorise le partenaire à utiliser la marque BnF pour la valorisation des collections pendant la période d'exploitation.

Le partenaire :

- propose un modèle de prise en charge par lui-même ou de partage des coûts de numérisation entre lui-même et la filiale BnF.
- fait effectuer la numérisation de tout ou partie du corpus, dans un délai contractuel qui dépend de l'ampleur du projet.
La numérisation pourra s'effectuer sur tout site en France à l'exception de la numérisation des collections rares et fragiles qui se fera sur place, à la BnF.
Le partenaire peut faire numériser par ses propres équipes ou par un prestataire agréé par la BnF. Il peut également demander à la BnF de numériser pour son compte. Dans ce cas, il peut bénéficier des tarifs négociés par la BnF auprès de ses prestataires.

- remet à la BnF la version numérisée de chaque document du corpus.
Cette version numérisée devra être conforme aux normes et formats demandés pour permettre à la BnF une conservation pérenne. Voir Annexe I.14 « Consignes techniques de numérisation ».
- exploite et valorise cette version numérisée pendant une durée et dans des conditions qu'il précisera.
- verse une redevance sur les produits d'exploitation dont il indiquera le montant.

4. Processus d'instruction des dossiers

4.1. Contenu du dossier

Le dossier de réponse à l'appel à partenariat doit au minimum comporter les éléments suivants :

- la description du marché, de l'offre et de l'objectif visé à terme (corpus, paliers, volume, etc.)
- la présentation de l'équipe, des compétences indispensables à la réussite du projet, de l'organisation
- la description des modalités de numérisation
- la présentation du modèle économique envisagé (plan d'affaires à court et moyen terme, modalités d'exploitation commerciale, retour financier pour la BnF).

N.B. : la BnF sera tenue à des obligations de confidentialité protégeant les porteurs de projet dans le cadre de l'appel à partenariat. En règle générale, ces dispositions rendent inutile la signature d'engagement supplémentaire relatif à la confidentialité des éléments transmis (« NDA »).

La date limite de réponse à cet appel à partenariat est fixée au **20 octobre 2011**.

4.2. Phases ultérieures

Les projets font l'objet d'une **analyse approfondie pendant l'automne 2011**. Au cours de cette phase, des rencontres sont organisées avec les porteurs du projet, auxquels il pourra être demandé de remettre des informations complémentaires. Il s'agit de valider la compréhension du projet, l'intérêt du dispositif proposé et de sélectionner le ou les partenaires.

La mise au point de l'ensemble des paramètres techniques, juridiques et financiers devra s'achever **au premier trimestre 2012**.

4.3. Critères d'évaluation

La sélection des partenaires s'appuiera sur les critères suivants :

- qualité de la numérisation et conformité aux objectifs, normes et formats techniques demandés par la BnF ;
- ampleur des projets de numérisation ;
- durée de réalisation ;
- originalité, intérêt scientifique et culturel et dimension du projet de valorisation ;
- expérience, capacité du partenaire à porter le projet ;
- modalités d'accès aux collections sur Gallica et Europeana :
plus l'accès aux documents numériques sur Gallica et Europeana sera large, plus l'appréciation de la proposition sera favorable.
- Intérêt et réalisme financiers du projet :
 - part des coûts de numérisation pris en charge par le partenaire :
plus la part sera importante, plus l'appréciation de la proposition sera favorable.

- La durée d'exploitation et le champ d'une exclusivité éventuelle sur la version numérisée du corpus :
plus la durée et le champ de l'exclusivité seront limités, plus l'appréciation de la proposition sera favorable¹.
- Niveau de retour financier pour la BnF :
plus le taux sera élevé, plus l'appréciation de la proposition sera favorable.

¹ Pour indication, le « Comité des Sages » recommande une période d'exclusivité d'un maximum de 7 ans suivant l'achèvement de la numérisation.

II - Appel à Partenariat pour la valorisation du système de préservation et d'archivage réparti (SPAR) développé et utilisé par la BnF pour ses données numériques

1. Cadre général

1.1. Contexte et enjeux de la valorisation de SPAR

La BnF a conçu dès 2007 et fait développer un système de pérennisation de ses ressources numériques dénommé SPAR (Système de Préservation et d'Archivage Réparti).

Il s'appuie sur des normes internationales faisant autorité dans le monde de la pérennisation des informations numériques. En particulier, SPAR respecte la norme OAIS (ISO-14721:2003), modèle de référence pour un système ouvert d'archivage d'information.

SPAR est bien plus qu'un simple entrepôt de données sécurisé :

- ➔ Il effectue de multiples copies des objets numériques (sur disques, sur bandes) et assure une surveillance continue de l'état des équipements et des médias d'enregistrement des fichiers permettant d'anticiper les recopies avant la perte définitive.
- ➔ 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. Ainsi, par exemple, lorsque le format d'image JPEG deviendra obsolète, SPAR sera en mesure de transformer les images concernées dans un nouveau format plus performant.
- ➔ Cette garantie implique un travail permanent de veille technologique sur les formats, de prototypage et de test des outils. L'ensemble de cet arsenal est intrinsèquement prévu dans la conception de SPAR.
- ➔ SPAR permet à tout moment de revenir en arrière pour restituer les objets dans leur format d'origine.

Le logiciel SPAR a été réalisé dans le cadre d'un marché par la société Atos Origin. Ce marché, d'une durée de quatre ans, a permis de développer un système modulaire, essentiellement en utilisant des briques de bases Open Source et dont l'intégration a été réalisée grâce à la plate-forme J2EE et utilise pleinement des technologies novatrices de gestion de métadonnées (RDF).

Le système SPAR a été mis en production en mars 2010 et archive désormais la production numérique courante de la BnF.

La BnF prépare actuellement la certification de son système auprès des Archives de France (SIAF).

Indépendamment des coûts liés à l'infrastructure matérielle du système, l'investissement consenti jusqu'ici pour la réalisation logicielle a été de 3,3 M€.

Pour obtenir un retour sur son investissement et faire bénéficier des acteurs privés ou publics de cet outil, modèle de référence dans son domaine, la BnF lance un appel à partenariat pour la valorisation et la commercialisation de SPAR.

Le ou les partenaires pourront, en contrepartie d'un retour financier, obtenir la mise à disposition de l'outil SPAR pour une utilisation commerciale.

Une filiale sera créée par la BnF, chargée de piloter l'ensemble des opérations partenariales de valorisation dans le cadre du programme Investissements d'Avenir. La mise en place de cette filiale est en cours.

Le financement des coûts d'investissements liés à la réutilisation de l'outil et à la mise en place d'une offre commercialisable seront portés par le ou les partenaires.

1.2. Le dispositif des investissements d'avenir mis en place par le Gouvernement

Dans le cadre du programme Investissements d'Avenir, le Gouvernement a décidé la création du Fonds national pour la Société Numérique (FSN), destiné à soutenir le développement de l'économie numérique tant au niveau des réseaux à très haut débit que des services tirant profit de ces réseaux.

Il intervient avec l'objectif de participer au financement des projets rentables qui accompagnent le développement d'acteurs innovants sur les technologies de la numérisation, création, gestion et valorisation des contenus numériques.

Le projet de valorisation de SPAR décrit ci-dessous peut être éligible à un soutien financier du FSN.

Dans ce cas, le FSN dotera la filiale BnF de fonds qui lui permettront de procéder à un apport financier pour accompagner l'investissement du ou des partenaires, en fonction du type et de l'importance des projets envisagés.

2. Champ de l'appel à partenariat

La BnF souhaite mettre à disposition son logiciel SPAR pour en permettre la commercialisation par un ou plusieurs partenaires spécialisés. Cette mise à disposition pourrait se faire selon deux modalités :

1. la réutilisation commerciale des sources logicielles du système SPAR par une société de type intégrateur informatique ou éditeur, lui permettant d'utiliser tout ou partie de ses composants logiciels afin de répondre à des appels d'offres, nationaux comme internationaux autour de la préservation numérique.
2. L'utilisation du logiciel SPAR par une société dans le cadre d'une offre type SaaS (Software as a Service) qui permettrait à une société commerciale d'offrir un service hébergé sur ses plateformes.

2.1. Réutilisation commerciale des sources logicielles de SPAR

La première utilisation commerciale possible du logiciel SPAR est sa réutilisation dans le cadre d'autres projets de préservation de documents numériques par une société de services souhaitant baser sa solution sur tout ou partie du logiciel.

Cette solution permettrait à cette société de services de limiter son coût de développement en partant de la souche logicielle SPAR et en réutilisant et en adaptant les modules de SPAR.

En plus de la conception du système en modules², garante de son évolutivité, SPAR utilise un mécanisme de filières spécifiques à une typologie de documents et indépendantes du cœur du système. Cela lui permet une évolutivité importante et une capacité forte à s'adapter à des besoins métiers différents de

² Au sens des entités de la norme OAIS.

ceux de la bibliothéconomie. Il sera d'ailleurs bientôt utilisé par la BnF pour ses propres besoins d'archivage de documentation administrative.

Une autre possibilité serait qu'un éditeur de logiciel spécialisé décide de créer un produit à partir de SPAR. Cet éditeur pourrait en effet, après un travail sur le logiciel, à partir de SPAR créer une solution « clé en main » lui permettant de viser un marché ciblé (par exemple un module de gestion d'archives pérennes du dossier patient pour les hôpitaux, etc.).

2.2. Réutilisation commerciale de SPAR en mode SaaS

Il existe sur le web plusieurs offres de services d'archivages de documents numériques (offre DigiCoffre de la Poste, etc.) permettant à des usagers de bénéficier d'un espace de sauvegarde de leurs fichiers numériques, accessible de n'importe où sur le web.

Il serait tout à fait possible, pour un prestataire de services, d'utiliser SPAR selon le même principe pour offrir un service plus ambitieux de préservation en ligne des données numériques de ses clients. Ce service pourrait en effet bénéficier de la capacité de SPAR à faire de l'archivage à long terme et faire bénéficier à un large public de l'expertise capitalisée dans le cadre du projet en termes de gestion des formats et de gestion de risques.

3. Dispositions complémentaires

Le code source de SPAR sera fourni en l'état et sans garantie de corrections de la part de la BnF par rapport à des dysfonctionnements qui pourraient apparaître lors de son utilisation par le titulaire.

La BnF utilise SPAR pour son propre compte et se réserve la possibilité d'offrir des services d'hébergement à ses partenaires à travers une offre « BnF-Archivages numérique » en cours de définition.

La BnF pourra, sous réserve de l'intérêt du projet, consentir à une exclusivité, pour un usage ou une cible donnée.

La BnF est intéressée par une mutualisation des coûts de maintenance corrective, adaptative et évolutive du logiciel SPAR. Les propositions des partenaires pourront prévoir un dispositif sur ce sujet.

4. Processus d'instruction des dossiers

4.1. Contenu du dossier

Le soumissionnaire devra détailler en particulier dans sa réponse, les éléments suivants :

- la description du marché, de l'offre et de l'objectif visé à terme.
- la présentation de l'équipe, des compétences indispensables à la réussite du projet, de l'organisation
- le modèle économique envisagé dans le cadre de la réutilisation commerciale de SPAR
- les gains directs pour la Bibliothèque (rétribution financière liée au point 2)
- les gains indirects pour la Bibliothèque (mutualisation des coûts de maintenance corrective, adaptative et évolutive du logiciel).
- le mode de collaboration envisagée ainsi que le type de contractualisation (partenariat, accord cadre, co-propriété du code, exclusivité, etc.).

N.B. : la BnF sera tenue à des obligations de confidentialité protégeant les porteurs de projet dans le cadre de l'appel à partenariat. En règle générale, ces dispositions rendent inutile la signature d'engagement supplémentaire relatif à la confidentialité des éléments transmis (« NDA »).

Une journée de présentation du logiciel SPAR sera organisée début septembre sur le site François Mitterrand de la Bibliothèque pour tous les candidats intéressés pour répondre au présent appel à partenariats. Cette journée aura comme objectif de présenter le système SPAR, la partie infrastructure (baies de disques, bandothèques, etc. etc.) mais également la partie logicielle : les fonctionnalités du système, son architecture et sa conception technique. A l'issue de cette journée, les candidats intéressés par un audit des codes sources se verront offrir un accès, dans les locaux de la BnF et autant que de besoin, à une « data-room » leur permettant de consulter le code source de SPAR.

La date limite de réponse à cet appel à partenariat est fixée au **20 octobre 2011**.

4.2. Phases ultérieures

Les projets font l'objet d'une **analyse approfondie pendant l'automne 2011**. Au cours de cette phase, des rencontres sont organisées avec les porteurs du projet, auxquels il peut être demandé de remettre des informations complémentaires. Il s'agit de valider la compréhension du projet, l'intérêt du dispositif proposé, et de sélectionner le ou les partenaires.

La phase de finalisation de l'ensemble des paramètres techniques, juridiques et financiers est prévue **au premier trimestre 2012**.

4.3. Critères d'évaluation

La sélection des réponses s'appuiera sur l'intérêt et réalisme financiers du projet :

- pertinence et crédibilité du projet de valorisation.
- Expérience, capacité du partenaire à porter le projet.
- Niveau de retour financier pour la BnF.
Plus le taux de retour financier sera élevé, plus l'appréciation de la proposition sera favorable.
- Gains indirects pour la Bibliothèque (mutualisation des coûts de maintenance corrective et évolutive du logiciel).
Plus les gains indirects seront élevés, plus l'appréciation de la proposition sera favorable.