

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 numérisés),
- 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 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.