



ÉTUDE SUR LA NUMÉRISATION DE RESSOURCES DOCUMENTAIRES AU SERVICE DE LA RECHERCHE, DESTINÉE À ACCOMPAGNER LA DÉFINITION D'UNE POLITIQUE NATIONALE DE NUMÉRISATION DANS LE CADRE DE L'INFRASTRUCTURE DE RECHERCHE COLLEX-PERSÉE

Synthèse de l'étude :

- 1) État des lieux des programmes de numérisation au service de la recherche, mis en œuvre depuis 2006 et bilan de la diffusion des ressources numérisées
 - 2) Retours d'expérience des projets et besoins en termes de services et d'accompagnement
-

Sommaire

1	Introduction.....	4
1.1	Rappel des objectifs de l'étude.....	4
1.2	Rappel des questions posées dans le CCTP	4
1.2.1	<i>Pour l'objectif n°1 : état des lieux des programmes de numérisation au service de la recherche, mis en œuvre depuis 2006, et bilan de la diffusion des ressources numérisées.....</i>	<i>4</i>
1.2.2	<i>Pour l'objectif n°2 : retours d'expérience des projets et besoins en termes de services et d'accompagnement.....</i>	<i>4</i>
2	Synthèse de l'étude.....	5
2.1	Synthèse de l'enquête (objectif n°1).....	5
2.2	Synthèse des retours d'expérience (objectif n°2)	7
2.3	Synthèse des besoins (objectif n°2).....	8
2.4	Orientations pour CollEx-Persée.....	9
3	Synthèse de l'objectif n°1 : état des lieux des projets de numérisation au service de la recherche, mis en œuvre depuis 2006 et bilan de la diffusion des ressources numérisées	11
3.1	L'enquête et le questionnaire.....	11
3.1.1	<i>Les structures sollicitées.....</i>	<i>11</i>
3.1.2	<i>Les programmes de numérisation concernés.....</i>	<i>11</i>
3.1.3	<i>Le questionnaire en ligne.....</i>	<i>12</i>
3.2	Les acteurs des programmes de numérisation	13
3.2.1	<i>Les structures ayant répondu</i>	<i>13</i>
3.2.2	<i>Les programmes de numérisation déclarés</i>	<i>13</i>
3.2.3	<i>L'évolution du nombre de programmes menés chaque année</i>	<i>13</i>
3.2.4	<i>Les principaux acteurs de la numérisation.....</i>	<i>14</i>
3.2.5	<i>La BnF.....</i>	<i>16</i>
3.3	Les finalités des programmes.....	17
3.3.1	<i>Les disciplines des programmes.....</i>	<i>17</i>
3.3.2	<i>Les objectifs déclarés des programmes</i>	<i>17</i>
3.3.3	<i>L'implication des chercheurs dans les programmes</i>	<i>17</i>
3.4	Les résultats des programmes	18
3.4.1	<i>Unités matérielles numérisées.....</i>	<i>18</i>
3.4.2	<i>Vues numérisées</i>	<i>18</i>
3.4.3	<i>Formats des données.....</i>	<i>18</i>
3.4.4	<i>Formats des fichiers.....</i>	<i>19</i>
3.5	Les ressources mobilisées sur les programmes	19
3.5.1	<i>Ressources financières.....</i>	<i>19</i>
3.5.2	<i>Ressources humaines</i>	<i>20</i>
3.6	Les traitements réalisés dans les programmes.....	20
3.6.1	<i>Traitements réalisés.....</i>	<i>20</i>
3.6.2	<i>Partenaires des programmes.....</i>	<i>20</i>
3.6.3	<i>Recours éventuel à l'externalisation.....</i>	<i>20</i>
3.6.4	<i>Sauvegarde et archivage pérenne des données.....</i>	<i>21</i>
3.7	La diffusion des ressources	22
3.7.1	<i>Les modalités de diffusion.....</i>	<i>22</i>
3.7.2	<i>Les solutions de diffusion</i>	<i>22</i>

3.7.3	<i>Autres dispositifs de diffusion et valorisation des ressources</i>	24
3.7.4	<i>L'évaluation de la diffusion par les structures</i>	24
3.8	Les projets à venir.....	25
4	Synthèse de l'objectif n°2 : retours d'expérience des projets et besoins en termes de services et d'accompagnement	26
4.1	Rappel des objectifs.....	26
4.2	Structures sollicitées.....	26
4.3	Retours d'expérience	27
4.3.1	<i>Synthèse des retours d'expérience</i>	27
4.3.2	<i>L'aboutissement des projets</i>	28
4.3.3	<i>L'utilité pour la recherche et la satisfaction des chercheurs</i>	28
4.3.4	<i>La coopération entre recherche et documentation</i>	28
4.3.5	<i>L'acquisition de compétences sur les projets de numérisation</i>	28
4.3.6	<i>Le financement des projets</i>	28
4.3.7	<i>Le fonctionnement des appels à projets</i>	29
4.3.8	<i>Les partenariats</i>	29
4.3.9	<i>Traitement des ressources avec des dispositifs internes</i>	29
4.3.10	<i>Traitement des ressources par des prestataires externes</i>	30
4.3.11	<i>Solutions de référencement et de diffusion des ressources</i>	30
4.3.12	<i>Droit de numérisation pour diffusion des ressources</i>	30
4.3.13	<i>Solutions d'archivage pérenne</i>	31
4.4	Perspectives et besoins.....	32
4.4.1	<i>Synthèse des besoins</i>	32
4.4.2	<i>Assurer un soutien financier et humain durable</i>	32
4.4.3	<i>Adapter le fonctionnement des appels à projets (notamment CollEx-Persée)</i>	33
4.4.4	<i>Soutenir des programmes de numérisation menés avec des chercheurs autant qu'avec des structures de recherche</i>	33
4.4.5	<i>Définir des orientations ou des priorités de numérisation à l'échelle nationale</i>	33
4.4.6	<i>Rendre visible les acteurs et les actions pour favoriser la collaboration</i>	34
4.4.7	<i>Mutualiser les connaissances et les formations</i>	34
4.4.8	<i>Mutualiser et partager les expertises juridiques</i>	35
4.4.9	<i>Mutualiser et partager les expertises techniques</i>	35
4.4.10	<i>Définir et partager des formats et des référentiels pour assurer l'interopérabilité</i>	35
4.4.11	<i>Mutualiser les numérisations (opérations et solutions de gestion)</i>	36
4.4.12	<i>Mutualiser les solutions de bibliothèques numériques pour les besoins des structures</i> ...	36
4.4.13	<i>Mutualiser les bibliothèques numériques à l'échelle nationale ou sur des thématiques</i> ..	36
4.4.14	<i>Mutualiser les solutions d'archivage pérenne</i>	37
4.4.15	<i>Autres besoins exprimés</i>	37
5	Orientations pour CollEx-Persée	38
6	Annexes	39
6.1	Annexe 1 : synthèse du contact avec la BnF.....	39
6.2	Annexe 2 : bibliothèques numériques en fonctionnement	40
6.3	Annexe 3 : établissements avec lesquels ont été menés les entretiens	48

1 Introduction

1.1 Rappel des objectifs de l'étude

Deux objectifs ont été fixés pour cette étude :

1. réaliser un **état des lieux des programmes de numérisation de ressources documentaires au service de la recherche menés entre 2006 et 2016**, ainsi que le **bilan de la diffusion des ressources numérisées**, via une enquête en ligne auprès des principales structures de recherche ou de documentation d'ESR susceptibles d'avoir mené des programmes de numérisation pour la recherche ;
2. identifier les **besoins des acteurs, actuels et futurs, en services et en accompagnement à la numérisation**, via le recueil, en entretien, des **retours d'expériences** et des besoins d'un échantillon de porteurs de programmes de numérisation et de chercheurs.

Cette étude servira de **base à la définition d'une politique nationale de numérisation dans le cadre de l'infrastructure de recherche CollEx-Persée.**

1.2 Rappel des questions posées dans le CCTP

1.2.1 Pour l'objectif n°1 : état des lieux des programmes de numérisation au service de la recherche, mis en œuvre depuis 2006, et bilan de la diffusion des ressources numérisées

- Quelle proportion des acteurs de l'ESR a mené des programmes de numérisation ces dix dernières années ?
- Quels sont les objectifs visés dans ces différents programmes, en fonction des acteurs ?
- Quelles sont les grandes caractéristiques de ces programmes (contenus, disciplines, caractéristiques techniques, moyens humains et financiers, partenaires, licences) ?
- Quels en sont les usages ?
- Quelle proportion des documents numérisés ces dix dernières années a fait l'objet d'une mise en ligne effective ?

1.2.2 Pour l'objectif n°2 : Retours d'expérience des projets et besoins en termes de services et d'accompagnement

- Quels sont les freins, au sein des établissements, à la mise en œuvre de projets de numérisation, à la mise en ligne des ressources numérisées et au développement de services autour de la numérisation ?
- Quelles sont les priorités en matière de numérisation et de diffusion pour la recherche dans les années à venir, par disciplines et types de structures de recherche ?
- Quels sont les besoins des porteurs de projets de numérisation en matière de services et d'accompagnement ?
- Comment le GIS CollEx-Persée peut concourir à leur mise en œuvre ? Quels types de services peut-il rendre aux établissements dans ce cadre ?

2 Synthèse de l'étude

2.1 Synthèse de l'enquête (objectif n°1)

Sur les 500 structures d'enseignement supérieur et de recherche sollicitées par courriel pour répondre au questionnaire en ligne, 181 ont répondu, dont 55% de structures de recherche (laboratoires, UMR, UMS, USR, MSH...) et 45% de structures documentaires (SCD, BIU, bibliothèques, centres de documentation). On peut supposer, sans pouvoir le chiffrer, que de nombreuses structures n'ont pas répondu car elles ne se sont pas senties concernées par l'objet de l'enquête.

Sur ces 181 structures d'ESR, 131 ont mené au moins 1 programme, pour un total de 405 programmes de numérisation menés sur 2006 -2016, avec un pic de programmes en 2015 (218). Sur les 405 programmes signalés dans le questionnaire, 386 ont effectivement été décrits.

151 programmes de numérisation restent en cours en 2018.

En majorité, peu de programmes sont menés par structure (moins de 3 programmes sur 10 ans dans 2/3 des cas), mais quelques structures sont très actives (de 7 à 8 programmes)¹.

Toutefois, le nombre de programmes réalisés est un indicateur à considérer avec prudence, car il n'est pas forcément représentatif de l'activité réelle de numérisation d'une structure. En effet, un programme peut concerner des volumes et des délais très différents d'une structure à l'autre. Certains établissements peuvent avoir mené un seul programme thématique concernant un grand nombre de volumes sur une longue période, tandis que d'autres peuvent avoir réalisé beaucoup de programmes concernant des corpus plus ciblés et plus modestes, sur des périodes plus courtes. Ce n'est pas parce qu'une structure a mené peu de programmes qu'elle n'a pas une activité de numérisation importante.

Principaux acteurs de la numérisation

Les principaux acteurs de la numérisation sont, en financement sur 10 ans : l'Infrastructure de recherche RECOLNAT (16 M€), Persée (5,2 M€), la BIU Santé (1,45 M€), le Collège de France (1,2 M€), la MSH de Dijon (1 M€), la COMUE PSL (1 M€).

En nombre de vues numérisées, les principaux acteurs sont : l'Infrastructure de recherche RECOLNAT (8,2 millions), Persée (6,2 millions) la BIU Santé (3,7 millions), la MSH de Dijon (1,6 million), PSL (1,4 million), Sciences Po Paris (1,2 million), Lille (1,2 million), Mathdoc (1,2 million).

Il est à noter que les acteurs RECOLNAT et Persée font figure d'exceptions, de par l'ampleur de leur activité de numérisation, et qu'ils ne sont pas représentatifs de l'activité moyenne de numérisation menée par les établissements de l'ESR. L'UMS Persée notamment est un opérateur dont l'activité est quasi intégralement tournée vers la numérisation, ce qui lui donne une place à part dans le paysage de la numérisation de l'ESR. Pour cette raison, les chiffres de ces deux acteurs ont été retirés des moyennes de volumes numérisés et d'ETP mobilisés par les structures de l'ESR dans le cadre de leurs programmes.

Si l'on rapproche les ordres de grandeur des financements des structures de l'ESR et de la BnF, sur l'ensemble des programmes menés sur 10 ans, respectivement :

¹ Le questionnaire permettait à chaque structure de signaler au maximum 8 programmes.

- pour l'ESR : 41 M€ pour 39 millions de vues
- pour la BnF : 45 M€ pour 144 millions de vues numérisées.

Les disciplines dominantes des programmes de numérisation des structures d'ESR sont les Lettres et Sciences humaines (71%) et ensuite les Sciences et Techniques (26%).

Objectifs des programmes de numérisation

Les objectifs des programmes de numérisation signalés sont essentiellement axés sur la valorisation du patrimoine (82%). Un peu plus de la moitié des programmes ont été réalisés en soutien ou en contribution à un projet de recherche (53%), mais il est à noter que cet objectif de contribution à un programme de recherche sera, d'après l'enquête, en progression dans les projets de numérisation à venir (67%).

Cependant, un programme de recherche peut être réalisé dans un objectif de service à la recherche, même s'il ne s'inscrit pas directement dans le cadre d'un programme de recherche précis.

Une grande majorité des programmes ont ainsi été réalisés en association avec des chercheurs (72%), particulièrement pour la sélection des documents à numériser (75%).

Profil-type des programmes de numérisation

Il existe des écarts types très importants entre les chiffres clés des différents programmes mais, hors Persée et Reclon, on peut définir ainsi le profil-type des programmes de numérisation et des structures qui les ont menés :

- en moyenne par structure (tous programmes de la structure sur 10 ans confondus) :
 - 210 K€ de financement, avec des sources très éclatées (essentiellement issues de l'établissement, puis de l'ANR, de la BnF, avec de très nombreuses autres sources....),
 - 47.000 unités matérielles manipulées,
 - 250.000 vues numérisées.
- en moyenne par programme :
 - une charge de travail de 1,4 ETP mobilisé par programme et par an, avec un recours à l'externalisation de tout ou partie des traitements dans 70% des programmes, surtout pour la numérisation et l'OCR,
 - une durée de 4,8 ans par programme,
 - des partenariats de tous types sur 50% des programmes.

Choix techniques

Concernant les choix techniques, les formats de données utilisés dans le cadre des programmes sont plutôtinteropérables (en majorité Dublin Core voire EAD, MARC XML et METS, pas de REFNUM qui est le format de la BnF²) mais cette situation peut être améliorée (encore un certain nombre de formats de données utilisés sont spécifiques à la structure).

Les formats de fichiers utilisés sont généralement pérennes (en majorité TIFF + JPEG + PDF/A), sauf exception.

² Le choix d'un nouveau format (METS) a été fait par la BnF en 2014. Il est aujourd'hui appliqué à tous les marchés de la BnF.

La distinction entre des sauvegardes sécurisées et l'archivage pérenne serait à revoir, pour assurer au final à la fois une réelle sauvegarde efficace et un archivage réellement pérenne des données et fichiers.

Diffusion des ressources numérisées

On note un bon taux de diffusion des ressources numérisées, puisque celles-ci font l'objet d'une diffusion totale dans 76% des programmes signalés et partielle dans 15% des programmes.

L'accès aux ressources est (sauf exception) ouvert à tous, gratuit et en ligne à distance.

La non-diffusion (qui concerne 9% des programmes) et une partie de la diffusion partielle s'expliquent souvent par le fait que de nombreux programmes sont encore en cours de réalisation. Il est à noter que des blocages juridiques sont responsables de 27% des cas de non-diffusion.

Les solutions de diffusion reposent majoritairement (60%) sur des solutions locales, éclatées, avec une majorité relative de solutions Omeka. Gallica est ensuite la solution nationale la plus utilisée (31%).

80% des structures répondantes ont indiqué avoir mis en place un entrepôt et un moissonneur OAI-PMH. La description des données selon les principes du web sémantique commence à émerger (17% des structures indiquent l'avoir mise en œuvre). Mais le référencement avec les moteurs du Web reste partiel (41% des répondants déclarent assurer le référencement de leurs ressources numérisées sur le Web).

L'évaluation de la diffusion reste encore très partielle (52% des répondants ont mis en place un outil d'évaluation, 14% ont réalisé des enquêtes auprès du public).

Projets de numérisation à venir

107 structures (sur les 181 ayant répondu au questionnaire) annoncent des projets de numérisation à venir. Dans 89% des cas, ces projets impliquent des partenariats avec d'autres structures, notamment la BnF et Persée.

Plusieurs structures ont déjà exprimé des attentes vis à vis d'un futur dispositif national de soutien aux programmes de numérisation, tant en termes de renfort en ressources que de contenus à privilégier. Certaines nomment explicitement CollEx-Persée comme acteur potentiel de ce dispositif.

2.2 Synthèse des retours d'expérience (objectif n°2)

Ces retours d'expérience ont été recueillis sur la base de 28 entretiens avec des structures ayant mené des programmes de numérisation et ayant répondu à l'enquête.

Satisfactions

Les principales satisfactions sont les suivantes :

- **l'aboutissement des projets** -numérisation et diffusion-, surtout après un processus souvent long. Il représente la satisfaction majeure ;
- l'utilité pour la recherche apparaît au moins de façon informelle (peu d'enquêtes ont été menées ou peu d'outils mis en place pour évaluer les usages réels) ;

- en matière de référencement et de diffusion, **la mise en œuvre réussie de solutions locales** (ex. Omeka), **nationales** (Nakalona, Gallica, Persée) ou autres (Internet Archive), sachant que des limites fonctionnelles sont apparues dans certains outils (Omeka-Classique ou Internet Archive notamment) ;
- **l'acquisition de compétences pour la numérisation**, en particulier dans le cadre de projets menés en interne, ce qui est une garantie de qualité pour les documents complexes ou précieux ;
- **des partenariats réussis** - avec la BnF, Huma-Num, Persée, Biblissima, des MSH, des partenaires en Région, des COMUE – parfois plus délicats à l'international ou dans le cas d'un nombre de partenaires trop important ;
- **une coopération entre recherche et documentation** faite de relations naturelles, souvent quotidiennes, sachant que la connaissance croisée des ressources reste à améliorer.

Difficultés

Les principales difficultés rencontrées sont :

- **d'abord le financement des projets et en particulier le manque de maintien durable des contrats des compétences internes**, avec des équipes réduites (1 à 2 personnes) ;
- **le fonctionnement complexe des appels à projets** : ils impliquent une mobilisation importante pour répondre, et reposent sur des critères parfois difficiles à respecter (exigence Open Access, association avec une structure de recherche) ;
- **le traitement des questions de droit de numérisation pour diffusion** : temps nécessaire, complexité juridique, etc ;
- **le coût trop élevé de la solution d'archivage pérenne** (solution PAC du CINES).

2.3 Synthèse des besoins (objectif n°2)

Ces besoins ont également été recueillis sur la base des 28 entretiens avec des structures ayant mené des programmes de numérisation et ayant répondu à l'enquête.

On peut les résumer ainsi (les besoins les plus fréquemment exprimés sont en caractères gras) :

- **Assurer un soutien financier et humain durable** (*besoin le plus cité*) : pour pérenniser le financement des ressources humaines (personnels contractuels, compétences scientifiques et techniques), et ainsi venir à bout des programmes, au-delà de la durée des appels à projets ;
- Adapter le fonctionnement des appels à projets (AAP) notamment dans le cadre de CollEx-Persée : alléger le travail de réponse, privilégier le temps long, assurer la cohérence avec les appels à projets BnF et ANR... ;
- Soutenir des programmes de numérisation menés avec les chercheurs autant qu'avec les structures de recherche, en clarifiant la nature de la relation souhaitée avec les projets de recherche (avec des chercheurs, avec des structures ?) ;
- **Clarifier / définir les orientations et les priorités à l'échelle nationale** : soutenir explicitement des pôles de compétences en région, de « petites structures », des axes thématiques, des types de documents (ex. sources originales, récentes, non libres de droits, nativement numériques), des techniques (3D, traitement d'image...) ;

- **Rendre visibles les acteurs et les actions pour favoriser la collaboration** : disposer d'une plate-forme collaborative des « acteurs et des actions de numérisation au service de la recherche » pour valoriser l'existant et développer les partenariats ;
- **Mutualiser les connaissances et les formations** : mettre en commun (via la plate-forme ?) des guides pratiques, des manuels de formation ;
- **Mutualiser et partager les expertises juridiques et techniques** : mettre en commun (via la plate-forme ?) en s'appuyant sur les compétences actuelles des acteurs, des expertises juridiques (sources originales, droit international...) ou techniques (3D, HTR, OLR, traitement images 2D, RDF) ;
- Définir et partager des formats (de données, de fichiers) et des référentiels : imposer l'interopérabilité des données dans les appels à projets ; assurer la cohérence des recommandations entre acteurs (ex. BnF) (et les diffuser via la plate-forme ?) ;
- Mutualiser les numérisations (opérations et solutions matérielles et logicielles) : mettre en commun des marchés (prestations) ou des moyens de numérisation, à l'échelle métropolitaine ou régionale (matériels, logiciels : ex. Numa-Hop), en s'appuyant sur des structures existantes ;
- Mutualiser les bibliothèques numériques pour les structures : disposer d'une solution de référence de bibliothèque numérique, pour pallier les limites de solutions actuelles, en s'appuyant sur des acteurs existants (notamment Huma-Num Nakalona) ;
- Mutualiser les bibliothèques numériques locales à l'échelle nationale et internationale : mettre en place un dispositif national pour valoriser et pérenniser les contenus des bibliothèques numériques locales, par exemple via des plates-formes thématiques nationales, en s'appuyant sur les réalisations actuelles (Persée, Bibliissima, Reclonat...) ;
- Mutualiser les solutions d'archivage pérenne / réduire les coûts du CINES : bâtir pour l'offre PAC du CINES un modèle économique avec des prix plus attractifs, en s'appuyant sur les volumes prévisionnels des établissements ;

2.4 Orientations pour CollEx-Persée

Sur la base des besoins exprimés, 6 grandes orientations et pistes d'actions peuvent être définies pour CollEx-Persée, dans le cadre de son rôle de soutien aux porteurs de projets de numérisation :

1. Renforcer et cibler le **soutien financier et humain** : objet/durée, portage des emplois durables... ;
2. Préciser la **cible des projets à soutenir** : critères d'appels à projets (mettre la priorité sur le soutien à la recherche plus que sur le soutien au patrimoine), priorités de numérisation (types de documents, techniques innovantes...) ;
3. Définir et mettre en œuvre une **plateforme de partage et de coopération entre acteurs de la numérisation** : mise en réseau des acteurs et des structures (ESR, CNRS, BnF...), partage de guides, de formations, d'expertises juridiques, d'expertises techniques, de formats et référentiels, de programmes de numérisation (à venir, en cours), d'appels à projets... ;
4. Définir et mettre en œuvre une politique de **mutualisation des opérations de numérisation** (plates-formes locales/régionales ?) **et de leur gestion** (ex. déploiement Numa-Hop ?) ;

5. Définir et mettre en œuvre une politique en matière de **mutualisation des solutions de bibliothèque numérique pour le niveau local et pour le niveau national / thématique**, en concertation avec les opérateurs actuels (Huma-Num, Persée...);
6. Définir et mettre en œuvre une politique en matière de **mutualisation des solutions d'archivage pérenne** (en lien avec le CINES).

3 Synthèse de l'objectif n°1 : état des lieux des projets de numérisation au service de la recherche, mis en œuvre depuis 2006, et bilan de la diffusion des ressources numérisées

3.1 L'enquête et le questionnaire

3.1.1 Les structures sollicitées

Pour établir un état des lieux des programmes de numérisation menés entre 2006 et 2016, et un bilan de la diffusion des ressources numérisées, une enquête a été menée auprès des principales structures assumant des missions d'enseignement supérieur et de recherche.

Environ **500 structures** ont été sollicitées dans cette l'enquête, dont :

- environ **310 structures de recherche** (laboratoires, unités mixtes de recherche (UMR), unités mixtes de services (UMR), unités de services et de recherche (USR), maisons des sciences de l'homme (MSH), ...),
- environ **190 structures documentaires** (bibliothèques interuniversitaires, services communs ou directions de la documentation, bibliothèques ou centres de documentation...).

3.1.2 Les programmes de numérisation concernés

Les structures ont été invitées à décrire les programmes de numérisation de ressources documentaires qu'elles ont menés entre 2006 et 2016 **dans une perspective de services à la recherche**.

Un programme de numérisation est compris comme une démarche scientifique coordonnée de reproduction numérique de documents sur support (par numérisation ou transcription), concernant un ou plusieurs corpus, dotée de moyens humains et financiers dédiés, et ayant produit des résultats.

Les programmes de numérisation ont pu être dotés d'une logique thématique (ex : « Les sources du droit ») ou être réalisés dans le cadre d'un projet de recherche (ex : opération de numérisation menée par un laboratoire dans le cadre d'un projet ANR). Ils ont pu concerner des éditions critiques en ligne.

En revanche, **ne sont pas entrés dans le périmètre de cette enquête** :

- **les services de numérisation à la demande,**
- **la fourniture de collections pour des opérations de numérisation menées par d'autres acteurs** (ex. fourniture de collections pour numérisation et diffusion par Persée),
- **les programmes dont les objectifs exclusifs étaient la préservation et la valorisation du patrimoine documentaire**, même si comme on le verra des réponses ont été faites en ce sens.

Le questionnaire a pris en compte les programmes menés en totalité ou en partie sur la période 2006-2016. Ainsi, les programmes ayant débuté avant 2006 ou ceux qui se prolongent au-delà de 2016 sont également rentrés dans le périmètre de l'étude.

3.1.3 Le questionnaire en ligne

Cette enquête a reposé sur un questionnaire en ligne, construit avec le logiciel Sphinx, et envoyé par courriel aux responsables des structures.

Les questions étaient organisées de la manière suivante :

- la fiche d'identité de la structure,
- les questions spécifiques à chaque programme de numérisation mené par la structure (8 programmes pouvaient être décrits au maximum³) :
 - o description, objectifs, disciplines, corpus, périodes,
 - o sources et montants des financements,
 - o moyens humains,
 - o travail avec les chercheurs ou avec d'autres partenaires,
 - o modalités de diffusion des ressources,
 - o types et volumes de documents numérisés (unités matérielles et vues),
 - o traitements réalisés et externalisation éventuelle,
 - o formats des données et des fichiers,
- les questions globales concernant l'ensemble des ressources numérisées par la structure :
 - o solutions techniques de gestion et de diffusion des ressources,
 - o évaluation de l'utilisation des ressources,
 - o projets à venir et partenaires éventuels,
 - o observations libres⁴.

Ce questionnaire représente un maximum de 220 questions traitant 473 « variables de choix » pour une structure qui décrirait 8 programmes.

Le calendrier du questionnaire a été le suivant :

- lancement du questionnaire le 06/12/2017,
- relance de tous les répondants le 09/01/2018,
- relance de tous les répondants en cours de réponse le 16/01/2018,
- clôture du questionnaire le 19/01/2018 au soir,
- extraction des résultats pour l'analyse figurant dans ce document le 30/01/2018 à 18h30.

A noter que le questionnaire est resté ouvert tout au long de l'étude, ce qui a rendu possible des corrections éventuelles de réponses, ou l'ajout de réponses de nouvelles structures⁵.

³ Ce maximum a été rendu nécessaire pour rester en deçà de 500 variables, tant pour limiter le temps de réponse, que pour rester sur une architecture standard du logiciel utilisé.

⁴ Voir en annexe la synthèse des observations libres (Annexe 5, p. 33).

⁵ 2 nouvelles réponses ainsi été enregistrées entre le 30/01/18 et le 14/02/18.

3.2 Les acteurs des programmes de numérisation

3.2.1 Les structures ayant répondu

Sur les 500 destinataires du questionnaire, 181 structures ont répondu au 30/01/18 (159 réponses terminées + 22 réponses en cours)⁶, soit **37% environ des destinataires**.

Parmi ces réponses, on trouve **une majorité de structures de recherche** :

- **100 laboratoires / UMR / UMS / USR / MSH et autres structures de recherche (55%)**
(sur plus de 310 structures consultées, soit environ 33% de répondants),
- **81 bibliothèques / SCD / directions de la documentation (45%)**
(sur plus de 190 structures consultées, soit environ 43% de répondants).

La légère prépondérance des structures de recherche dans le panel des répondants s'explique par le fait qu'elles ont été sollicitées en plus grand nombre que les structures documentaires (310 laboratoires sollicités contre 190 structures documentaires).

Les répondants se sont réellement impliqués dans ces réponses :

- dans **50%** des cas, c'est le ou la directeur/directrice, responsable, secrétaire général(e), de la structure qui a validé le questionnaire,
- le temps consacré à la saisie a été en moyenne de **1h50**⁷.

3.2.2 Les programmes de numérisation déclarés

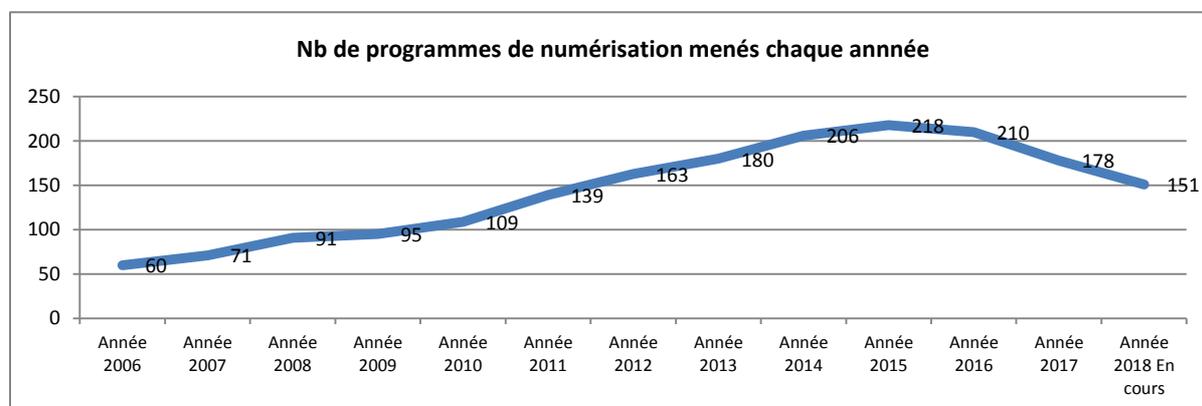
Sur les 181 structures ayant répondu, **131 ont déclaré au moins un programme (72%)**.

Au total, **405 programmes de numérisation ont été signalés et 386 ont été effectivement décrits**.

67% des structures ont mené moins de 3 programmes. Mais 16 structures en ont mené 7 ou 8⁸.

Toutefois, le nombre de programmes ne constitue pas un indicateur pertinent de l'activité réelle de numérisation : en effet, les 2 structures ayant mené les programmes les plus importants en termes de volumes numérisés et de moyens consacrés (RECOLNAT et Persée) ont déclaré respectivement 1 et 2 programmes (cf. aussi tableau des acteurs ci-dessous).

3.2.3 L'évolution du nombre de programmes menés chaque année



⁶ Environ 4 autres structures ont fait connaître par mël qu'elles n'étaient pas concernées.

⁷ Durée entre la connexion et l'enregistrement, avec 20' de time-out.

⁸ La liste de ces structures figure en annexe.

On remarque une forte progression du nombre de programmes de numérisation menés, en particulier entre 2011 et 2015. La mise en place de certains dispositifs publics de soutien à la numérisation a pu encourager ce mouvement, à l'instar des deux appels à projets de numérisation lancés par la Bibliothèque scientifique numérique en 2013 et 2014, qui ont permis de soutenir 29 programmes de numérisation⁹.

Début 2018, 151 programmes de numérisation sont en cours.

3.2.4 Les principaux acteurs de la numérisation

Tout d'abord, deux acteurs nécessitent d'être distingués dans les analyses, eu égard aux volumes de ressources mises en jeu (unités matérielles, vues numérisées, ressources humaines, financements) : **l'UMS Persée et l'Infrastructure de recherche RECOLNAT.**

3.2.4.1 UMS Persée

Initié en 2003, le portail Persée est un programme national d'édition électronique de collections patrimoniales de publications scientifiques (revues, actes de colloques, publications en série, livres, dictionnaires, etc.) et de diffusion ouverte (www.persee.fr).

Ces publications scientifiques relèvent principalement des sciences humaines et sociales mais aussi des sciences de la terre et de l'environnement. Les collections s'étendent du 19^{ème} au 21^{ème} siècle.

Le label "Perséides" identifie les collaborations entre Persée et des équipes de recherche avec pour objectif de produire, diffuser et exploiter des corpus numérisés.

Les chiffres clés de Persée sont les suivants :

	Persée
Nb de programmes	2 (portail Persée, label Perséides)
Nb d'unités matérielles numérisées	35.000
Nb de vues numérisées	6,2 millions
Financements	5,24 M€
Nb d'ETP moyen / an / programme	Portail Persée : 15 ETP / an Label Perséides 2 : 5 ETP / an

3.2.4.2 Infrastructure de recherche RECOLNAT

L'infrastructure de recherche RECOLNAT a mis en œuvre la plate-forme e-ReColNat qui a pour objectif de présenter à terme sur une plateforme unique les images et les données de la totalité des collections naturalistes françaises. Ces données proviennent de bases existantes et interopérables et des programmes de numérisation qui ont lieu à un niveau national.

⁹ 16 projets de numérisation ont été financés dans le cadre de l'appel à projets 2013, puis 17 dans le cadre de l'appel à projets 2014 (dont 4 s'inscrivaient dans la continuité de projets déjà retenus en 2013).

Les chiffres clés de RECOLNAT sont les suivants :

	RECOLNAT
Nb de programmes	1
Nb d'unités matérielles	7,2 millions
Nb de vues numérisées	8,2 millions
Financements	16 M€ (ANR)
Nb d'ETP moyen / an / programme	82,4 ETP / an

3.2.4.3 Les autres acteurs clés

Certains acteurs méritent d'être signalés en raison de l'importance des volumes qu'ils ont numérisés et des moyens qu'ils ont consacrés à la numérisation (> 500K€ en financement ou > 100.000 unités matérielles numérisées ou > 500.000 vues numérisées).

M Millions K Milliers	Nb Prog.	Financement	Nb Unités mat. num.	Nb vues num.	Nb ETP moy / an / prog.
BIU Santé	6	1,45 M€	237 K	3,7 M	1,1
Collège de France (Réseaux documentaires)	1	1,20 M€	16 K	131 K	3
MSH Dijon USR 3516	1	1,06 M€	30 K	1,6 M	5
Paris Sciences et Lettres (Doc)	3	1,06 M€	63 K	1,4 M	1,4
Institut recherche et histoire des textes	7	970 K€	14 K	NC	1,1
Muséum national d'histoire naturelle (Bibliothèque)	3	804 K€	3,5 K	686 K	1,1
Laboratoire HiSoMA	7	739 K€	1,3 M	87 K	2,1
Sciences Po Paris (Bibliothèque)	5	674 K€	33 K	1,2 M	1,1
Laboratoire CIHAM-UMR 5648	7	619 K€	647	214 K	2,2 (dont 8 P8)
Université de Lille (SCD)	1	610 K€	3,2 K	1,2 M	1,5
BDIC	7	571 K€	8,2 K	823 K	1,4
CNAM (SCD)	1	540 K€	3,4 K	1,0 M	2
UMR 7270 Laboratoire Ligérien de Linguistique	1	540 K€	1,3 K	1,5 K	8
Musée du Quai-Branly Jacques Chirac	5	500 K€	457 K	760 K	0,9
INHA (DBD)	8	399 K€	6,9 K	1,0 M	2
Paris 3 Sorbonne nouvelle (DBU)	6	354 K€	106 K	247 K	0,42
UniVegE - Herbarium universitaire	2	345 K€	293 K	293 K	0,6
Mathdoc UMS 5638	1	323 k€	6,7 K	1,2 M	3
UMR 7306 CRD Océanie	1	90 K€	260 K	260 K	2,5
Université Fédérale Toulouse (SICD)	8	90 K€	9,2 K	719 K	0,53

M Millions K Milliers	Nb Prog.	Financement	Nb Unités mat. num.	Nb vues num.	Nb ETP moy / an / prog.
Université Paris-Dauphine (SCD)	1	83 K€	3,1 K	825 K	1
Aix-Marseille Université (SCD)	5	80 K€	253	506 K	1,2
Centre Camille Jullian (UMR7299)	2	65 K€	101 K	7,5 K	0,7
Paris Jourdan Sciences Eco - UMR8545	1	NC	NC	625 K	NC
USR 3123 CEDEJS (Le Caire)	3	52 K€	1,5 M	1,5	19 (dont 50 P1)

On constate qu'il est assez difficile de dégager des ratios entre financements, nombre d'unités matérielles et nombre de vues numérisées, car il faut tenir compte de l'avancement des projets, et des types de ressources.

3.2.5 La BnF

Des contacts ont également été pris avec la BnF.

Une rencontre a été organisée le 15/12/2017 et la BnF a fourni une description de ses programmes le 30/01/2018.

Une synthèse de cette réponse figure en annexe de ce document.

Si l'on rapproche les ordres de grandeur des chiffres clés de la BnF avec le total des structures de l'ESR ayant répondu :

	Structures de l'ESR ayant répondu (131)	BnF
Programmes	405	122
Financements	Environ 41 M€	Plus de 45 M€
Vues numérisées	Environ 39 millions de vues	Près de 144 millions de vues ¹⁰

¹⁰ Hors production de BnF Partenariats.

3.3 Les finalités des programmes

3.3.1 Les disciplines des programmes

Les programmes déclarés couvrent principalement les Lettres et Sciences Humaines :

- **Lettres et Sciences Humaines : 71%**
- Sciences et Techniques : 26%
- Droit-Sciences Politiques : 19%
- Économie et AES : 10%
- Santé : 10%
- STAPS et Autres : 7%

197 programmes (soit 49% du total des programmes signalés) sont axés uniquement sur les lettres et sciences humaines. Cette prédominance des lettres et sciences humaines dans les disciplines des programmes signalés dans cette étude peut s'expliquer en partie par le fait que les structures de recherche sollicitées étaient essentiellement spécialisées en sciences humaines et sociales, et par la centralité de la constitution de corpus et d'exploitation des sources pour ces disciplines.

En fonction des priorités en matière de recherche, le dispositif CollEx-Persée pourrait à l'avenir apporter un soutien particulier sur certaines disciplines, pour faire évoluer cet état des lieux.

3.3.2 Les objectifs déclarés des programmes

Les programmes de numérisation déclarés restent encore partiellement au service de projets de recherche.

- 82% des programmes ont pour objectif la valorisation du patrimoine,
- 62% des programmes ont pour objectif la préservation du patrimoine,
- **53% des programmes ont pour objectif le soutien et la contribution à un projet de recherche.**

À noter que 28 programmes ont pour objectif unique le soutien à un projet de recherche. Il est à souligner que le terme de « soutien ou contribution à un projet de recherche » réfère à un contexte de numérisation plus précis que la notion de « service à la recherche », qui peut être comprise comme englobant un plus grand nombre de cas de figure. Un programme de numérisation peut ainsi être considéré comme étant mené au service de la recherche sans s'insérer formellement dans le cadre d'un programme de recherche. Cela peut ainsi expliquer en partie que seuls 53% des programmes signalés dans le cadre de cette étude déclarent s'inscrire dans le cadre du soutien à la recherche.

CollEx-Persée a annoncé sa volonté de soutenir en priorité les programmes de numérisation contribuant à des projets de recherche, cet objectif devrait donc progresser à l'avenir.

3.3.3 L'implication des chercheurs dans les programmes

Malgré le caractère partiel de l'objectif de soutien à des projets de recherche, **les chercheurs sont impliqués dans 72% des programmes signalés.**

Lorsque les chercheurs sont impliqués, ils interviennent en priorité sur les choix de contenus :

- dans **75% des programmes, sur la sélection des documents à numériser,**
- dans 52% des programmes, sur la définition du type de documents à numériser,
- dans 47% des programmes, sur les besoins de structuration et d'enrichissement.

Cette coopération des chercheurs dans le cadre des programmes de numérisation sera particulièrement analysée au cours des entretiens qui seront menés dans la seconde partie de l'étude.

3.4 Les résultats des programmes

3.4.1 Unités matérielles numérisées

Un total de **12,5 millions d'unités matérielles numérisées a été déclaré dans le cadre de cette enquête¹¹**:

- dont **5,3 millions d'unités matérielles hors Persée et RECOLNAT,**
- soit une moyenne de 47.000 unités matérielles par structure.

Total Unités matérielles (hors P&R)	Imprimés	Archives et Manuscrits	Docs icono.	Cartes et plans	Son et Audiovisuels	Objets 3D	Autres types
5 296 941	389 870	386 367	1 373 612	57 469	63 267	8 364	3 026 044
100%	7%	7%	26%	1%	1%	0%	57%

À noter la proportion importante de documents iconographiques (26%) et d'autres types hors ceux cités (57%).

3.4.2 Vues numérisées

39 millions de vues numérisées ont été déclarées¹² :

- dont **24,5 millions de vues numérisées hors Persée et RECOLNAT,**
- soit une moyenne de 250.000 vues numérisées par structure.

Total Vues (hors P&R)	Imprimés	Archives et Manuscrits	Docs icono.	Cartes et plans	Son et Audiovisuels	Objets 3D	Autres types
24 523 212	16 140 995	4 088 588	1 647 021	34 836	54 179	12 435	2 548 156
100%	66%	17%	7%	0%	0%	0%	10%

À noter la proportion importante d'imprimés (66%).

3.4.3 Formats des données

Les réponses indiquent que l'interopérabilité des données est assurée mais pas totalement :

- **en majorité du Dublin Core (41%),**
- ensuite de l'EAD (12%), du Marc XML (9%), METS (9%), TEI (7%), MODS (2%),
- **dans 20% des cas, des formats autres, dont il conviendrait de garantir l'interopérabilité :**
 - o des formats a priori interopérables : Marc 21, Unimarc, XML et modèles associés (ALTO, EXODIC, OLAC, MADS, EDM, SKOS....), IPTC/EXIF, CSV, ISAD/G, Darwin Core,

¹¹ Par 114 des 131 structures : la réponse n'était pas obligatoire

¹² Par 97 des 131 structures : la réponse n'était pas obligatoire

- **d'autres a priori moins, ou pas directement** (pour 20 à 25 structures) : formats « Maison », « Local », « Propriétaire ».

À noter l'absence du format REFNUM utilisé par la BnF¹³.

Un des enjeux du dispositif CollEx-Persée sera de déterminer quels formats de données privilégier et surtout quelles actions entreprendre pour les formats a priori non interopérables. Outre la pérennité des données, la diffusion est bien sûr conditionnée par l'interopérabilité des données.

3.4.4 Formats des fichiers

Les réponses indiquent que la pérennité des fichiers est assurée sauf exception¹⁴ :

- **en majorité du TIFF (41%)** ensuite du JPEG (31%) et PDF/A (14%),
- **dans 10% des cas, des formats autres mais standards sauf exception** : WAV, PDF, MP3, MPEG4, XML (TEI, ALTO), TXT.

Un des enjeux du dispositif CollEx-Persée sera de déterminer quels formats de fichiers privilégier. Outre la pérennité des fichiers, la diffusion serait facilitée par une certaine standardisation des formats de fichier.

3.5 Les ressources mobilisées sur les programmes

3.5.1 Ressources financières

Un total de 41 M€ a été consacré à la numérisation de ressources documentaires¹⁵ :

- dont **19,5 M€ de dépenses hors Persée et RECOLNAT**,
- soit une moyenne de 210 K€ par structure.

Total financements (hors P&R)	Établissement	ANR	Collectivités	Europe	BnF	BSN	Autres sources
19 494 349 €	7 583 519 €	2 580 798 €	1 178 580 €	372 500 €	1 914 534 €	891 026 €	4 980 392 €
100%	39%	13%	6%	2%	10%	5%	26%

Les financements des programmes sont très éclatés, mais c'est l'établissement qui assure la majorité relative des financements (39%), puis l'ANR (13%) et la BnF (10%), avec toutefois de très nombreuses autres sources (26%).

Les autres sources les plus fréquemment citées sont d'abord le CNRS et la TGIR Humanum (plus de 15 programmes), les Fondations, les Idex, les DRAC, des Ministères.

Il est à noter que les réponses au questionnaire n'offrent pas une vision exhaustive des dépenses réalisées pour la numérisation de ressources documentaires. Sur la période 2009-2016, le soutien de la BnF aux structures de l'ESR s'élève ainsi à 3,4M€ (pour 9,6 millions de pages numérisées). Il se traduit par des subventions à la numérisation (2,1M€ pour 5,9 millions de pages numérisées) et par l'intégration de documents dans les marchés de numérisation (imprimés) de la BnF (1,3M€ pour 3,7 millions de pages).

¹³ Voir note 1 page 6.

¹⁴ Cf. recommandation du SIAF des Archives de France sur les formats de fichiers pérennes.

¹⁵ Par 96 des 131 structures : la réponse n'était pas obligatoire.

3.5.2 Ressources humaines

En moyenne, les structures répondantes ont mobilisé 1,4 ETP par programme de numérisation et par an (hors Persée et Reclnat)¹⁶.

Ces moyens annuels sont à rapprocher de la **durée moyenne des programmes de 4,8 années.**

12 structures (hors Persée et Reclnat) ont consacré plus de 4 ETP en moyenne par programme et par an¹⁷.

3.6 Les traitements réalisés dans les programmes

3.6.1 Traitements réalisés

Les principaux traitements réalisés dans les programmes sont (dans l'ordre décroissant de fréquence) :

- **la numérisation (83%),**
- **la saisie manuelle des métadonnées (67%),**
- l'indexation (54%),
- l'OCR (37%),
- la création d'index (28%),
- la structuration manuelle du document (28%),
- l'intégration de métadonnées de bases de données existantes (28%).

A noter très peu d'OLR (optical lay-out recognition), de transcriptions, d'annotations.

3.6.2 Partenaires des programmes

En dehors des partenariats financiers, en moyenne **50% des programmes renseignés ont été menés en partenariat avec une ou plusieurs autres structures.**

- les partenaires sont variés : internationaux, nationaux, locaux, territoriaux, autres laboratoires, etc. ;
- les partenariats sont de différentes natures : fourniture, sélection, prêt de documents, numérisation (le plus fréquent), choix éditoriaux, mise à disposition de plateforme, etc. ;
- Persée et la BnF sont cités comme partenaires d'au moins 10 programmes ;
- il existe quelques cas où des programmes ont été menés uniquement par des chercheurs.

Il conviendrait donc de réfléchir au rôle que pourrait jouer le dispositif CollEx-Persée pour organiser cette diversité.

3.6.3 Recours éventuel à l'externalisation

Pour 70% des programmes, les traitements sont au moins partiellement externalisés.

Les principaux traitements réalisés de manière externalisée sur les programmes sont (dans l'ordre décroissant de fréquence) :

- **la numérisation (61%),**

¹⁶ Sur 379 des 405 programmes : la réponse n'était pas obligatoire.

¹⁷ Voir tableau en annexe.

- **l'OCR (29%),**
- la génération automatique des métadonnées (16%),
- la saisie manuelle des métadonnées (11%).

Parmi les prestataires (et partenaires) externes :

- **Arkhênum est de loin le plus cité car impliqué dans 60 programmes,**
- Jouve / Safig et Azentis, sont prestataires dans 15 programmes,
- Diadeis, Puce et Plume dans 8 programmes,
- Tribun, Numen dans moins de 5 programmes,
- Persée et la BNF sont cités à 10 reprises.

Il conviendrait peut être de réfléchir au rôle que pourrait jouer le dispositif CollEx-Persée pour faciliter le recours aux marchés de prestations externalisées (ex. recommandations du SIAF).

Par ailleurs, il est à noter que **55% des structures disposent de matériel en propre pour mener des programmes de numérisation.**

3.6.4 Sauvegarde et archivage pérenne des données

Les réponses indiquent que les structures assurent :

- des sauvegardes internes dans 85% des cas,
- des sauvegardes externes dans 33% des cas.

Les solutions de sauvegarde décrites les plus courantes sont Huma-Num/Nakala/Nakalona, BNF¹⁸, mais aussi le CINES, Persée, Numdam.... (Pour le CINES il s'agit plutôt d'archivage pérenne que de sauvegarde).

Concernant l'archivage pérenne :

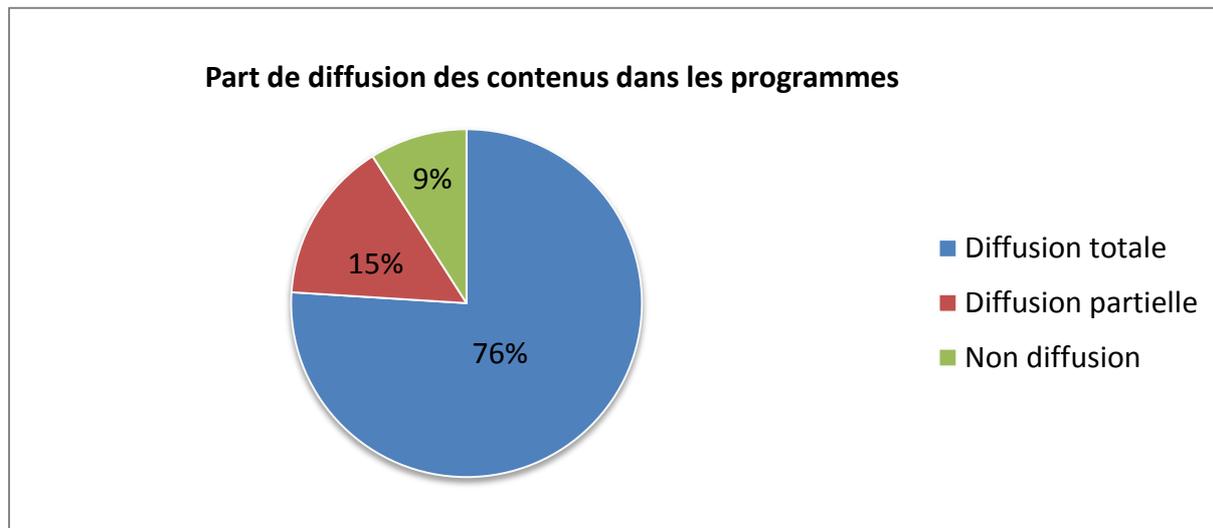
- **l'archivage pérenne est mis en œuvre dans 51% des cas,**
- **la solution décrite la plus courante est bien sûr celle du CINES,**
- d'autres solutions ont été citées :
 - o des solutions nationales (BnF SPAR, HAL-TEL, PERSEE, Huma-Num Nakala- Nakalona),
 - o des solutions internes à l'établissement (ex. Observatoire, MAE, MNHN, FMSH), qui ne constituent peut être pas réellement des solutions d'archivage pérenne).

La distinction entre une sauvegarde sécurisée et une archive pérenne est sans doute un concept à mieux expliciter dans le cadre des actions de CollEx-Persée, pour assurer au final à la fois une réelle sauvegarde efficace (interne + externe) et un archivage réellement pérenne des données et fichiers.

¹⁸ Les documents des partenaires numérisés par la BnF sont sauvegardés, car le stockage des documents numériques sur serveurs conditionne techniquement la diffusion dans Gallica. Cette sauvegarde est également appliquée aux fichiers numériques « intégrés », dans le cas par exemple de la diffusion en Gallica marque blanche. La BnF a également une offre de tiers-archivage (SPAR) qui elle est payante.

3.7 La diffusion des ressources

3.7.1 Les modalités de diffusion



Les ressources numérisées font l'objet d'une diffusion dans la grande majorité des programmes signalés. Les motifs de non diffusion, y compris partielle, ont été les suivants :

- **dans 27% des cas des problèmes de droit,**
- dans 13 % des cas des problèmes techniques,
- dans 12 % des cas des moyens financiers insuffisants.
- **mais dans 41% des cas, d'autres raisons que celles indiquées : essentiellement des projets encore en cours car s'étalant sur plusieurs années (plus de 20 programmes dans ce cas) ou ayant pris du retard, faute de temps ou de ressources humaines).**

Dans la base des réponses, il serait possible d'identifier les programmes terminés mais pas ou partiellement diffusés, pour évaluer si une aide spécifique de CollEx-Persée pourrait être apportée (soutien technique, financier, juridique pour les problèmes de droits).

Concernant les modalités de diffusion, l'accès est :

- **ouvert à tous à 70%** (vs. sur authentification à 9,8%),
- **gratuit à 83%** (vs. payant à 2,8%¹⁹),
- **en ligne à distance à 80%** (vs. accès dans les locaux à 9,8%).

3.7.2 Les solutions de diffusion²⁰

Une solution locale a été mise en œuvre par 60% des structures répondantes.

Les répondants ont connu des difficultés à réaliser des distinctions entre les différents types de solutions (SIGB, GED « Bibliothèque numérique », « Portail », site web ...), mais parmi les solutions citées, **Omeka est la plus fréquente.**

¹⁹ Manuscrits médiévaux, archives théologiques : registres pontificaux, fonds Ricoeur, archives archéologiques, etc.

²⁰ Les questions relatives à la diffusion, à la valorisation des ressources ainsi qu'à l'évaluation de la diffusion concernaient l'ensemble des ressources numérisées par chaque structure, et non les ressources numérisées dans chaque programme.

Les solutions nationales ou internationales de diffusion mises en œuvre sont les suivantes :

- **Gallica (31%)**
- Internet Archive (9,4%)
- Calames (7,6%)
- Persée (4,2%)
- Nakalona (2,4%)
- Numdam (0,5%)

Ces réponses au questionnaire sont confortées par l'analyse des Bibliothèques numériques identifiées (cf. annexe).

Sur les 126 réalisations de bibliothèques numériques identifiées :

- 14 ne sont pas en fonctionnement²¹,
- 30 fonctionnent avec une solution qui n'a pas été identifiée,
- 36 fonctionnent avec OMEKA,
- 2 fonctionnent avec OMEKA HUMA NUM,
- 9 fonctionnent avec Gallica (dont 1 avec Gallica Marque Blanche : BNU Strasbourg)²²,
- 8 fonctionnent avec des développements internes (dont il conviendrait de garantir la pérennité),
- 3 solutions sont utilisées 3 fois : FLORA, PLEADE, LIMB Gallery (I2S),
- 3 solutions sont utilisées 2 fois LODEL, Mnesys (NAONED), WordPress,
- 10 solutions sont utilisées 1 fois : ALEXANDRIE, ARMADILLO, Content DM (OCLC), DIGITOOL (Ex-Libris), DRUPAL, DSPACE, E-Critur, GREENSTONE, ICEBERG, Internet Archive.

Un des enjeux du dispositif CollEx-Persée sera donc de trouver des solutions optimisant l'hétérogénéité de ce panorama, tout en préservant l'indépendance de choix des structures.

Les principales fonctionnalités du site de diffusion des ressources numérisées sont les suivantes :

- **la recherche par métadonnées (94,4%),**
- **le téléchargement des images (73,2%),**
- **la recherche en texte intégral (69%),**
- **les identifiants des documents (67,6%),**
- les URL pérennes (60,6%),
- le feuilletage dans les index, les référentiels, les glossaires (56,3%),
- les outils avancés de visualisation et l'impression des images (54,9%).

À noter, la très faible présence de fonctionnalités d'annotation.

²¹ Test réalisé au 30/11/2017.

²² Gallica assure la diffusion des documents des partenaires qui ont été soit numérisés par la BnF, soit intégrés après numérisation. 47 partenaires de l'ESR contribuent ainsi à Gallica par intégration de documents physiques, et 3 par intégration de fichiers numériques. 42 bibliothèques numériques de l'ESR sont également moissonnées par Gallica (données fin 2017).

3.7.3 Autres dispositifs de diffusion et valorisation des ressources

Les autres dispositifs mis en œuvre liés à la diffusion sont les suivants :

- **un entrepôt et un moissonneur OAI-PMH ont été mis en place par 79% des répondants** (la solution est en général liée à la plateforme locale, parfois à une solution nationale. Il y a quelques rares cas de développement spécifique), donc comment exploiter ce potentiel ?
- **le référencement des ressources numérisées dans les moteurs de recherche du web a été mis en place par 41% des répondants**, donc des actions pour optimiser ce référencement seraient sans doute à envisager,
- **la description et l'exposition des données selon les principes du web sémantique ont été mises en œuvre par 17% des répondants**

3.7.4 L'évaluation de la diffusion par les structures

Un outil de mesure d'usage est mis en œuvre par 52% des structures répondantes :

- Google Analytics le plus souvent, ensuite Piwik et AWStats,
- le suivi concerne d'abord le nombre de vues, puis le nombre d'unités téléchargées et de recherches effectuées.

Cela signifie en creux que, dans près de la moitié des cas, l'utilisation réelle des ressources diffusées n'est pas mesurée.

Des enquêtes sont réalisées auprès des publics par 14% des répondants (suivi des usages, des publics, et des attentes).

Dans la suite de l'étude, il serait sans doute intéressant de demander aux structures ce qu'elles ont tiré comme enseignement de ces évaluations, qu'il s'agisse des résultats de mesures ou des enquêtes.

Un des enjeux du dispositif CollEx-Persée serait aussi de favoriser des mesures et bilans d'usages.

3.8 Les projets à venir

Des projets de programmes ont été annoncés par 66% des répondants.

Les objectifs annoncés pour ces projets sont les suivants,

- 92% des projets visent à valoriser les ressources,
- 77% visent à préserver les ressources,
- **67% visent à soutenir des programmes de recherche (ce qui montre un progrès par rapport aux programmes menés sur 2006-2016 où ce taux était de 53%).**

Comme pour les programmes menés sur 2006-2016, les partenariats prévus pour la réalisation des projets de numérisation sont nombreux :

- **89% des projets signalés sont prévus en partenariat – technique, financier, etc,**
- **dont 60% en partenariats multiples et 29% avec un seul partenaire.**
- **19 projets sont prévus en partenariat avec la BNF ;**
- Humanum est concerné formellement par 6 projets à venir et Persée par 4 projets ;
- 5 ou 6 projets seront lancés avec les mêmes partenaires que ceux mobilisés pour les précédents programmes des structures ;
- quelques rares structures mèneront leurs projets en interne sans partenaires ;
- enfin, 8 à 10 structures n'ont pas encore réfléchi aux partenariats possibles ou n'ont pas encore trouvé tous les partenaires souhaités. CollEx-Persée est cité à 2 reprises comme partenaire souhaité.

Dans les observations libres fournies en fin de questionnaire, on trouve déjà l'expression d'attentes vis-à-vis d'un futur dispositif national :

- souhait d'une politique nationale pour la numérisation de la recherche qui s'affranchisse de l'actualité des projets de recherche nécessairement temporaires, pour préserver et numériser des gisements non encore exploités mais porteurs de potentiels d'usages scientifiques variés et sur le long terme ;
- besoin d'une structure pour mettre en œuvre un dispositif d'appel à projets. CollEx-Persée est explicitement cité pour répondre à ce besoin ;
- besoin d'aide aux petites structures qui souhaitent numériser ;
- besoin de soutiens : postes statutaires, vacations, post doc ;
- besoin d'attribution d'identifiants pérennes d'accès aux ressources de l'ESR.

4 Synthèse de l'objectif n°2 : retours d'expérience des projets et besoins en termes de services et d'accompagnement

4.1 Rappel des objectifs

Cette phase de l'étude avait pour objectif, en s'appuyant sur les résultats de l'enquête en ligne, et sur la base d'entretiens en présentiel ou à distance :

- **d'établir un retour d'expérience sur les projets menés afin d'identifier les satisfactions et difficultés rencontrées lors de ces projets ;**
- **de faire le point sur les perspectives en matière de numérisation pour la recherche et les besoins des porteurs de projets de numérisation**, en particulier ceux auxquels le GIS CollEx-Persée pourrait contribuer à l'avenir.

4.2 Structures sollicitées

Parmi les répondants au questionnaire, des structures représentatives de la diversité des situations ont été sélectionnées :

- type de structures : documentaires / de recherche,
- nombre de programmes réalisés,
- importance des programmes réalisés,
- réussites et difficultés rencontrées,
- panel des disciplines concernées,
- panel des types de documents,
- projets à venir,
- observations apportées, à creuser.

Au total, 28 structures ont été consultées :

Structures documentaires	Structures de recherche
1. Bibliothèque Nationale et Universitaire de Strasbourg	1. Cellule Mathdoc – Université Grenoble Alpes
2. Bibliothèque interuniversitaire de la Sorbonne	2. OSUG Université Grenoble Alpes
3. Bibliothèque interuniversitaire de Santé	3. Université Grenoble Alpes - Projet Fonte Gaia
4. École nationale des Chartes	4. CRAHAM Caen
5. SCD Université Claude Bernard Lyon 1	5. UMR7041 Archéologies et Sciences de l'Antiquité (ArScAn)
6. SCD CNAM	6. Laboratoire HisCant-MA (Université de Lorraine)
7. Bibliothèque de documentation internationale contemporaine – La contemporaine	7. Maison des Sciences de l'Homme Ange-Guépin Nantes
8. Bibliothèque de Sciences Po Paris	8. CITERES - Laboratoire Archéologie et Territoires Tours
9. Bibliothèque Mazarine	9. Photothèque CESCМ Poitiers
10. SCD Université Bordeaux Montaigne	10. Institut de recherche et d'histoire des textes
11. Bibliothèque de l'Observatoire de Paris	11. Institut international de la marionnette
12. BUSIM (Bibliothèque de l'Université et de la Société Industrielle de Mulhouse)	12. HiSoMA (Université Lyon 2)

Structures documentaires	Structures de recherche
13. Bibliothèque Sainte-Geneviève 14. SCD Université Lille : SHS (ex L3) + LILLIAD (ex L1) 15. Bibliothèque de l'école française d'Extrême-Orient	13. Infrastructure Reolnat

Une présentation plus détaillée des structures avec lesquelles ont été menés les entretiens figure en annexe.

4.3 Retours d'expérience

4.3.1 Synthèse des retours d'expérience

Les thèmes des retours d'expérience peuvent être représentés selon qu'il s'agisse plutôt de satisfactions et/ou de difficultés, exprimées plus ou moins fréquemment par les établissements.

Thèmes du retour d'expérience	Fréquence des satisfactions	Fréquence des difficultés
Aboutissement des projets	16	
Utilité pour la recherche et satisfaction des chercheurs	6	
Coopération entre Recherche et Documentation	13	6
Acquisition de compétences pour la numérisation	8	
Financement des projets	7	22
Fonctionnement des appels à projet		3
Partenariats	13	3
Traitement des ressources avec dispositifs internes	7	5
Traitement des ressources avec prestataires externes	6	8
Solutions de référencement et de diffusion	17	12
Droit de numérisation pour diffusion		15
Solutions d'archivage pérenne	2	8

Légende : Nb d'établissements ayant exprimé le thème

Plus de 21 x
Plus de 14 x
Plus de 7 x
Moins de 7 x

Les comptes rendus d'entretien regroupés par thèmes ont été fournis sous forme d'une annexe au rapport d'étude. On précise ce bilan ci-dessous, thème par thème.

4.3.2 L'aboutissement des projets

L'aboutissement des projets est une satisfaction qui s'exprime :

- par la complétude et les volumes des ressources numérisées par rapport aux collections disponibles ;
- par la complétude et les volumes des ressources diffusées, en lien avec la mise en œuvre de solutions de diffusion ;
- par la capacité à mener à bien un processus de long terme (souvent plus de 10 ans).

4.3.3 L'utilité pour la recherche et la satisfaction des chercheurs

Les retours des chercheurs sur l'usage des bibliothèques numériques sont plutôt positifs, mais les mesures d'usage outillées et les enquêtes de satisfaction sont rares.

2 facteurs de succès sont mis en évidence :

- la capacité de la numérisation à prendre en compte diverses qualités typographiques et des structures de texte complexes ;
- la richesse de l'indexation.

L'utilité pour les chercheurs de la numérisation à la demande a également été soulignée.

4.3.4 La coopération entre recherche et documentation

Cette coopération est source de satisfactions :

- pour beaucoup de structures documentaires, la coopération avec les chercheurs est naturelle voire quotidienne (pour la consultation sur place, la numérisation à la demande), ce qui facilite la « co-construction » des projets de numérisation ;
- les MSH jouent par nature un rôle de pivot dans l'articulation recherche et documentation.

Mais elle rencontre aussi quelques difficultés :

- pour faire connaître auprès des chercheurs le potentiel des ressources documentaires ;
- pour les structures documentaires, pour identifier les chercheurs potentiellement intéressés, car complexité des structures de recherche ;
- pour les lauréats d'appels à projets, pour obtenir des financements dédiés à la numérisation.

4.3.5 L'acquisition de compétences sur les projets de numérisation

Certaines structures ont acquis de réelles compétences sur la numérisation des ressources sur support : numérisation, contrôle qualité, référencement, diffusion, archivage....

Certaines structures ont également développé des compétences propres à certains types de documents, par ex. :

- la numérisation de ressources récentes, nativement numériques et sous droit ;
- la numérisation 3D ;
- l'acquisition et la structuration du texte, notamment en TEI.

4.3.6 Le financement des projets

Le financement des projets apparaît comme une satisfaction :

- lorsqu'il est pérenne et régulier en provenance notamment de l'établissement, de la BnF, du Ministère de la culture ;
- lorsqu'il est important en provenance par ex. de la COMUE (ex. PSL pour l'Observatoire et les Chartes).

Mais l'absence de régularité et de pérennité des financements constitue la principale difficulté pour plusieurs raisons :

- il est difficile de maintenir dans la durée les contrats des ressources internes, ayant un niveau de compétences scientifiques et techniques élevé ;
- les équipes actuelles dédiées à la numérisation sont en nombre réduit (1 à 2 p. max) et donc vulnérables à la mobilité des personnels (retraite, départs...) ;
- il existe parfois une incertitude sur la pérennité des financements des COMUE (ex. PSL) ;
- les financements des appels à projets (ex. BSN, projets ANR) sont par nature ponctuels.

4.3.7 Le fonctionnement des appels à projets

Outre un financement sur des périodes courtes, les appels à projets :

- demandent un travail important à des ressources réduites et souvent déjà mobilisées sur les programmes en cours, y compris pour des structures importantes ;
- peuvent reposer sur des critères difficiles à respecter par les bibliothèques dans le cadre du soutien à la recherche : intérêt durable pour la recherche, relations structurées entre bibliothécaires et chercheurs ;
- contiennent souvent une exigence de diffusion en Open Access, alors que les sources intéressantes à traiter portent souvent des restrictions de droit de diffusion.

4.3.8 Les partenariats

Travailler en partenariat constitue globalement une source de satisfaction (financement, partage d'outils et de compétences), avec

- l'Etablissement,
- d'autres bibliothèques en France,
- la BnF,
- Huma-Num,
- Persée,
- l'Equipex Biblissima,
- des acteurs en région (ex. Bibliothèque municipale),
- une MSH, notamment pour les outils,
- la COMUE (ex. PSL).

Mais il existe parfois quelques difficultés, par exemple avec des acteurs internationaux ou lorsque la pérennité de la fonction documentaire dans les COMUE n'est pas assurée.

4.3.9 Traitement des ressources avec des dispositifs internes

Disposer de dispositifs de numérisation en interne de la structure (scanners, matériels photographiques, logiciels de gestion des traitements) est une source de satisfaction, qui garantit la qualité des traitements sur les documents les plus précieux ou complexes.

Mais ces dispositifs ne sont pas toujours adaptés aux traitements de grandes quantités.

4.3.10 Traitement des ressources par des prestataires externes

Le recours à un prestataire extérieur donne souvent satisfaction, notamment pour traiter des volumes importants de ressources ou des documents hors formats standard.

Mais les compétences des équipes des prestataires sont parfois insuffisantes. En effet, ceux-ci rencontrent parfois des difficultés à maintenir les compétences de leurs équipes dans la durée ou à traiter certaines ressources complexes, ce qui peut générer un travail important de contrôle de la qualité de la numérisation et de saisie des métadonnées par les équipes en interne.

4.3.11 Solutions de référencement et de diffusion des ressources

La mise en œuvre d'une solution de référencement et de diffusion constitue globalement une source de satisfaction, qu'il s'agisse :

- de solutions de bibliothèques numériques locales (développements internes, Omeka, un SIGB, en complément ou pas des solutions de bibliothèques numériques locales, d'autres solutions locales) ;
- de solutions de bibliothèques numériques nationales, le plus souvent en complément des solutions locales (Omeka via Huma-Num (Nakalona), Gallica, avec le cas de Gallica Marque blanche (BNU), Persée, le SUDOC, Calames....) ;
- d'autres solutions (ex. Internet Archive).

Mais des difficultés dans la mise en œuvre ne sont pas rares :

- des saisies multiples sont nécessaires entre les bibliothèques numériques locales et des catalogues nationaux (Calames, SUDOC) ;
- des compétences sont à acquérir sur les formats de description ;
- des solutions de sauvegardes sécurisées sont à mettre en œuvre ;
- les solutions logicielles (Omeka-Classique, Gallica Marque Blanche, Internet Archive...) possèdent des limites, notamment en ce qui concerne l'intégration de certains types de documents non imprimés, l'éditorialisation des contenus, la valorisation d'expositions virtuelles, ou encore la différenciation des accès en fonction des ressources (accès ouvert/accès dans les locaux de l'établissement) ;
- la pérennité des solutions locales ou des développements internes n'est pas toujours assurée ;
- les fonctions statistiques des bibliothèques numériques restent insuffisamment utilisées.

4.3.12 Droit de numérisation pour diffusion des ressources

Connaître les conditions juridiques de la numérisation pour la diffusion des ressources est, après les financements, la 2ème difficulté la plus fréquemment citée.

Cela concerne en particulier les documents sources, récents, non édités, sur support ou nativement numériques (archives privées, manuscrits, photographies) qui constituent parfois les gisements les plus intéressants pour la recherche.

Cela concerne également les contextes internationaux (ex. fouilles archéologiques à l'étranger d'équipes françaises).

Indépendamment des conditions juridiques réelles, les chercheurs demandent parfois un « délai d'embargo » avant la publication des ressources numérisées (pour mener à leur terme leur travail de recherche).

4.3.13 Solutions d'archivage pérenne

La nécessité d'une solution d'archivage pérenne est exprimée par des structures de toutes tailles.

L'utilité de la solution PAC du CINES est reconnue, en utilisation directe (ex. BSG) ou indirecte (via Huma-Num, ou Persée).

Mais le coût de la solution PAC du CINES constitue une réelle difficulté. De plus, l'absence de pérennité du financement (paiements annuels sur des budgets non pérennes) est problématique.

4.4 Perspectives et besoins

Les besoins de services et d'outils mutualisés découlent naturellement du bilan précédent.

4.4.1 Synthèse des besoins

Les besoins peuvent être représentés selon leur fréquence d'expression par les établissements.

Besoins de services et d'outils mutualisés	Fréquence du besoin
Assurer un soutien financier et humain durable	12
Adapter le fonctionnement des appels à projets (notamment CollEx-Persée)	5
Soutenir des programmes de numérisation menés avec les chercheurs autant qu'avec les structures de recherche	5
Définir des orientations et des priorités à l'échelle nationale	10
Rendre visibles les acteurs et les actions pour favoriser la collaboration	14
Mutualiser les connaissances et les formations	11
Mutualiser et partager les expertises juridiques	12
Mutualiser et partager les expertises techniques	3
Définir et partager des formats et des référentiels	6
Mutualiser les numérisations (moyens et solutions)	5
Mutualiser les bibliothèques numériques pour les structures	6
Mutualiser les bibliothèques numériques à l'échelle nationale	9
Mutualiser les solutions d'archivage pérenne / réduire les coûts du CINES	13

Légende : Nb d'établissements ayant exprimé le besoin.

Plus de 21 x
Plus de 14 x
Plus de 7 x
Moins de 7 x

On précise cette expression des besoins ci-dessous, thème par thème.

4.4.2 Assurer un soutien financier et humain durable

Il s'agit essentiellement d'un besoin de pérennité de financement des ressources humaines (*):

- pour des personnels contractuels, en majorité à un niveau élevé de compétences scientifiques ou techniques ;
- pour venir à bout des traitements (numérisation, référencement) et pour assurer *in fine* la diffusion des ressources ;
- parfois pour des dépenses externes (matériels de numérisation ou de sauvegardes, prestataires de numérisation).

(*) Rappel des réponses au questionnaire : 1,4 ETP en moyenne par programme et par an et 4,8 années en moyenne de durée d'un programme.

La question est donc posée des modèles économiques à envisager pour porter des emplois durables au service des structures (faut-il mutualiser des emplois au niveau régional ou métropolitain ? faut-il envisager des marchés de personnels en régie ?).

4.4.3 Adapter le fonctionnement des appels à projets (notamment CollEx-Persée)

Il s'agit aussi d'une difficulté souvent citée.

Les adaptations proposées sont les suivantes :

- alléger le travail de réponse ;
- privilégier le temps long des financements ;
- clarifier les critères, avec notamment :
 - o des programmes en soutien à la recherche, mais qui ne reposeraient pas forcément sur des relations organisées entre structures (ex. SCD / Labo) mais sur des relations et coopérations directes avec les chercheurs (voir ci-dessous) ?
 - o des programmes dont l'intérêt peut être durable ou pas ?
- assurer la cohérence avec les appels à projets de l'ANR, de la BnF...

Il serait imaginable de diffuser les appels à projets dans la « plate-forme des acteurs et des actions de numérisation pour la recherche » (voir ci-dessous).

4.4.4 Soutenir des programmes de numérisation menés avec des chercheurs autant qu'avec des structures de recherche

L'orientation de CollEx-Persée de soutenir en priorité les programmes de numérisation au service de la recherche a bien été comprise : il est attendu du GIS CollEx-Persée qu'il soutienne la coopération entre recherche et documentation.

Toutefois des structures documentaires à vocation patrimoniale (ex. BNUS, BIS), ont exprimé la volonté de continuer à développer une « offre » de ressources patrimoniales numérisées, potentiellement au service de futurs projets de recherche, sans que des projets de labos ne soient obligatoirement à l'initiative des programmes de numérisation.

Rappel des réponses au questionnaire : « seulement » 53% des programmes ont pour objectif le soutien à un projet de recherche, mais 72% des programmes impliquent des chercheurs.

4.4.5 Définir des orientations ou des priorités de numérisation à l'échelle nationale

Ces orientations pourraient porter :

- sur le type d'établissements :
 - o pour favoriser les pôles de compétence existant en région,
 - o pour ne pas exclure les « petites » structures,
- sur des axes thématiques (ex. en Archéologie) ;
- sur des types de documents, par exemple sur des sources originales (archives privées, manuscrits, photos...) et/ou des ressources récentes et nativement numériques, souvent non libres de droits et non cataloguées ;

- sur des techniques de numérisation ou des services innovants (ex. numérisation 3D, retranscription automatique du texte des documents audios/vidéos, traitement de l'image....).

Enfin, la coordination des politiques de numérisation et de conservation partagée est souhaitée (quelles collections déposer en conservation partagée en fonction de ce qui est numérisé et qui reste accessible ?).

4.4.6 **Rendre visible les acteurs et les actions pour favoriser la collaboration**

La mise en place d'une « plate-forme collaborative des acteurs et des actions de numérisation pour la recherche » valorisant à l'échelle nationale est fréquemment souhaitée.

Cette plate-forme pourrait valoriser :

- les acteurs de la numérisation, pour favoriser le partage d'expériences et d'expertises, et pour faciliter les partenariats de projets (structures de recherche, structures documentaires, dispositifs nationaux (ex. CollEx-Persée, BSN/Comité pour la Science Ouverte, Huma-Num, Biblissima) structures nationales (ex. BnF, CINES, Ministère de la culture) et acteurs internationaux) ;
- les disciplines et les thématiques, pour favoriser les synergies de projets ;
- les expertises transversales (techniques, juridiques...) pour favoriser le partage des compétences (ressources en ligne et rencontres) ;
- les programmes de numérisation, pour favoriser l'organisation et la mise en cohérence des actions ;
- une cartographie ou un annuaire des bibliothèques numériques.

Il existe plusieurs exemples de fonctionnement en réseau dans le domaine de la recherche :

- en France :
 - o consortiums thématiques Huma-Num (ex. MASA - Archéologie, CAHIER, IMAGEO, 3D-SHS....) : <https://www.huma-num.fr/consortiums>
 - o réseau des MSH : <http://www.msh-reseau.fr>
 - o réseau des bases de données (CCSD) : <http://rdbd.cnrs.fr>
- à l'international :
 - o à l'échelle d'un pays : Fondation allemande pour la recherche (DFG : Deutsche Forschungsgemeinschaft) : <http://www.dfg.de>
 - o à l'échelle d'un sujet : HOPE (Héritage des peuples européens - Heritage of the People's Europe) : <http://www.peoplesheritage.eu/fr/index.htm>

Il conviendrait donc de préciser les articulations entre ces réseaux et la plate-forme des acteurs et des actions de numérisation pour la recherche.

4.4.7 **Mutualiser les connaissances et les formations**

Il s'agirait de développer les compétences par la mise en commun :

- des guides et des manuels de numérisation, des recommandations sur les formats de données et de fichiers, sur les règles de nommage des fichiers (en complément avec ce que propose déjà la BnF) ;
- des fiches de poste types (compétences requises...) nécessaires à la mise en œuvre de projets de numérisation ;
- des formations, y compris des formations communes chercheurs - bibliothécaires.

Ces contenus pourraient être valorisés dans la plate-forme des acteurs et des actions.

4.4.8 Mutualiser et partager les expertises juridiques

Un soutien est attendu sur les conditions juridiques de la numérisation et la diffusion (*):

- notamment pour les sources originales (ex. archives privées, manuscrits, photos...), contemporaines, sur support ou nativement numériques,
- dans le cadre du droit français et du droit international.

Il serait opportun de s'appuyer sur les compétences existantes (BIS, Contemporaine)

Ces contenus pourraient être valorisés dans la plate-forme des acteurs et des actions.

(*) Rappel des réponses au questionnaire : 27% des programmes n'ont pas été diffusés en raison de problèmes de droits.

4.4.9 Mutualiser et partager les expertises techniques

Parmi les expertises à développer ont été citées :

- la numérisation 3D ;
- l'acquisition du texte, avec des techniques de HandWritten Text recognition (HTR) ou d'Optical Lay-out Recognition (OLR) ;
- le traitement de l'image 2D (reconnaissance d'items dans les images) ;
- le RDF, le web sémantique.

Ces contenus pourraient être valorisés dans la plate-forme des acteurs et des actions.

4.4.10 Définir et partager des formats et des référentiels pour assurer l'interopérabilité

En lien avec tous les acteurs, il serait souhaitable de prendre une position commune sur :

- les formats de données (ex. METS, TEI...) ;
- les formats d'OCR (ex. ALTO) ;
- les formats de fichiers ;
- les référentiels (ex. avec OpenTheso soutenu par Huma-Num) ;
- les règles de nommage et d'identification des fichiers (ex. ARK).

Il ne s'agirait pas d'imposer l'unicité des formats, mais d'imposer, par exemple dans les appels à projets CollEx-Persée, l'interopérabilité des données.

Ces contenus pourraient être valorisés dans la plate-forme des acteurs et des actions.

4.4.11 Mutualiser les numérisations (opérations et solutions de gestion)

Il s'agirait de mutualiser les opérations de numérisation, avec :

- des marchés communs à plusieurs établissements ;
- la mise en commun de moyens de numérisation à l'échelle métropolitaine ou régionale.

En Île-de-France, la question a été posée du rôle que pourrait jouer le CTLes dans ce cadre.

Il s'agirait aussi d'envisager de mutualiser les solutions de gestion des numérisations :

- des outils développés en local pour le contrôle qualité de la numérisation ;
- exemple de la solution Numa-Hop développée par la Bibliothèque Sainte-Geneviève (avec Sciences Po Paris et la BULAC), dans le cadre de la COMUE Université Sorbonne Paris Cité (et souhaitée par la BIS), avec la question du potentiel d'extension de cette solution à l'échelle nationale.

4.4.12 Mutualiser les solutions de bibliothèques numériques pour les besoins des structures

Il s'agirait de disposer d'une solution « de référence » de bibliothèque numérique (pour le référencement et la diffusion des ressources), répondant aux besoins des chercheurs, et permettant de pallier les limites fonctionnelles des solutions actuelles.

Il est rappelé :

- que la mise en œuvre d'une bibliothèque numérique constitue la 3ème difficulté la plus citée ;
- que 60% des structures ayant répondu au questionnaire ont mis en œuvre une solution locale de diffusion (avec une grande diversité de solutions, dont Omeka qui est la solution la plus présente en majorité relative).

Des souhaits d'évolution ayant été clairement exprimés par plusieurs structures, faudrait-il envisager des projets communs à plusieurs structures ?

La question est enfin posée de privilégier, au niveau national, la solution Omeka proposée par Humanum dans Nakalona (une évolution est prévue vers Omeka-S en 2019).

4.4.13 Mutualiser les bibliothèques numériques à l'échelle nationale ou sur des thématiques

En complément des bibliothèques numériques au niveau des établissements, il s'agirait de mettre en place un dispositif national pour valoriser les contenus des bibliothèques numériques :

- pour favoriser la dissémination des ressources, à l'échelle nationale et internationale (via des mécanismes d'interopérabilité de types API, web services, OAI, IIIF...) ;
- pour favoriser la pérennisation des bibliothèques numériques.

Ce dispositif national serait alimenté en interopérabilité avec les solutions locales.

Les scénarios d'architecture restent évidemment à définir :

- Faut-il préférer une plate-forme nationale, avec une « marque » unique et reconnue (comme Gallica) ?

- Faut-il préférer plusieurs plates-formes nationales spécialisées sur certaines disciplines et/ou types de document en s'appuyant sur les réalisations actuelles (Persée, Biblissima, Reconnat, MOM/Archéologie, Cartomundi...) qu'il conviendrait de renforcer ?
- Dans les deux cas, est-il préférable de mettre en œuvre de nouvelles solutions ou d'adapter et de faire évoluer des solutions actuelles (ex. Isidore, Gallica) ?

4.4.14 Mutualiser les solutions d'archivage pérenne

Les besoins d'archivage pérenne sont exprimés par tous types de structures.

Il est rappelé que 49% des structures ayant répondu au questionnaire n'ont pas mis en œuvre des solutions d'archivage pérenne.

Pour l'offre PAC du CINES, Il serait indispensable de trouver un nouveau modèle économique permettant de réduire les prix actuels, en utilisant le recensement des besoins des structures.

La question peut être posée de savoir si le modèle actuel de refacturation entre structures de l'ESR reste le plus adapté pour développer ce service.

4.4.15 Autres besoins exprimés

Certains besoins ont été exprimés plus ponctuellement, comme :

- mutualiser des services de numérisation à la demande, et définir une position commune vis à vis de la plateforme européenne (EOD Ebook On Demand : <https://books2ebooks.eu/fr>),
- soutenir le référencement des données des bibliothèques numériques sur le web,
- mettre à disposition des chercheurs des outils de datamining.

5 Orientations pour CollEx-Persée

Sur la base des besoins exprimés, 6 grandes orientations et pistes d'actions peuvent être définies pour CollEx-Persée, dans le cadre de son rôle de soutien aux porteurs de projets de numérisation :

1. Renforcer et cibler le **soutien financier et humain** : objet/durée, portage des emplois durables...
2. Préciser la **cible des projets à soutenir** : critères d'appels à projet (soutien à la recherche plutôt qu'au patrimoine), priorités de numérisation (types de documents, techniques innovantes...).
3. Définir et mettre en œuvre une **plateforme de partage et de coopération entre acteurs de la numérisation** : mise en réseau des acteurs et des structures (ESR, CNRS, BnF...), partage de guides, de formations, d'expertises juridiques, d'expertises techniques, de formats et référentiels, de programmes de numérisation (à venir, en cours), d'appels à projet...
4. Définir et mettre en œuvre une politique de **mutualisation des opérations de numérisation** (plates-formes locales/régionales ?) **et de leur gestion** (ex. déploiement Numa-Hop ?).
5. Définir et mettre en œuvre une politique en matière de **mutualisation des solutions de bibliothèque numérique pour le niveau local et pour le niveau national / thématique**, en concertation avec les opérateurs actuels (Huma-Num, Persée...).
6. Définir et mettre en œuvre une politique en matière de **mutualisation des solutions d'archivage pérenne** (en lien avec le CINES).

6 Annexes

6.1 Annexe 1 : synthèse du contact avec la BnF

La réponse de la BnF a été apportée le 30/01/2018 (pas de saisie en ligne dans le questionnaire mais réponse fournie dans un fichier Excel).

Plus de 122 programmes ont été décrits, plus ou moins complètement, selon la trame des questions du questionnaire.

Les tendances qui ressortent de la réponse sont les suivantes :

- des fonds multiples : spécialisés, presse, réserve, indisponibles, fonds partenaires,
- les types de documents : imprimés, manuscrits, documents iconographiques, documents cartographiques, musique, objets et monnaies,
- plus de 160.000 monographies, 5.800 titres de périodiques, 45.000 documents d'archives et manuscrits, 230.000 documents iconographiques, ont été numérisés,
- au total près de 144 millions de vues numérisées ont été produites,
- le financement depuis 2006 a été supérieur à 45 M€ (provenant de la BnF bien sûr, mais aussi du CNL, d'Europeana 14-18),
- le budget a été de 5M€ en 2017,
- la stabilité des formats de conservation : Format Refnum pour les données²³ et TIFF (parfois JPEG) pour la conservation des documents.

Pour la période 2017-2021, 10 grands programmes de numérisation sont envisagés :

- numérisation de masse et par corpus des livres, revues et magazines courants : 75000 monographies pour 12 millions de vues,
- numérisation des collections spécialisées: 4 millions d'images,
- numérisation de la presse : 7,6 millions d'images,
- numérisation de l'audiovisuel : 400 000 supports sonores, vidéo, électroniques et diapositives,
- numérisation des Réserves des départements : 4 millions d'images,
- numérisation des documents imprimés déclarés hors d'usage : 80 000 documents et 14 millions d'images,
- numérisation des microformes : 19 millions d'images,
- numérisation des documents les plus demandés par les lecteurs,
- numérisation en lien avec la vie culturelle nationale et internationale,
- programmes parallèles à la numérisation de ses collections par la BnF, dans le cadre de BnF partenariats : 16,5 millions de vues.

²³ Le choix d'un nouveau format (METS) a été fait par la BnF en 2014. Il est aujourd'hui appliqué à tous les marchés de la BnF.

6.2 Annexe 2 : bibliothèques numériques en fonctionnement

Source : http://www.bibliopedia.fr/wiki/Bibliothèques_numériques + compléments Six et Dix

Structures	Bibliothèque Numérique en fonctionnement	Note : cette liste recense des bibliothèques numériques et des collections numérisées accessibles uniquement sur Gallica.
Aix en Provence - MMSH	OK (Alexandrie)	Phonothèque de la MMSH d'Aix-en-Provence http://phonotheque.mmsh.huma-num.fr/
Aix-Marseille	OK (Omeka)	Odysée: bibliothèque patrimoniale https://odyssee.univ-amu.fr/
Antilles. Guyane. Saint Claude. SCD	OK (Greenstone)	MANIOC : Bibliothèque numérique Caraïbe Amazonie Plateau des Guyanes http://www.manioc.org/
Artois	OK (Omeka)	Bibliothèque patrimoniale http://bibnum-bu.univ-artois.fr/
Arts et Métiers ParisTech	OK (Limb Gallery)	Bibliothèque patrimoniale numérique http://patrimoine.ensam.eu/
BDIC	OK (Mnesys Naoned)	Argonaute http://argonaute.u-paris10.fr/
Besançon. Université de Franche Comté	OK	LibrIsta Bibliothèque numérique (Institut des Sciences et techniques de l'Antiquité/Presses Universitaires de France Comté) http://ista.univ-fcomte.fr/catalogue_biblio_numerique.html
Bibliothèque interuniversitaire de la Sorbonne	OK (Omeka)	NuBIS https://nubis.univ-paris1.fr/
Bibliothèque interuniversitaire Santé	OK	Bibliothèque numérique Medica http://www.biusante.parisdescartes.fr/histoire/medica/index.php Banque d'images et de portraits : http://www.biusante.parisdescartes.fr/histoire/images/index.php
Bibliothèque Jacques Doucet	OK (Mnesys Naoned)	http://bljd.sorbonne.fr/Bibliotheques-numeriques/p31/Les-bibliotheques-numeriques
Bibliothèque Mazarine	OK (Limb Gallery)	Mazarinum http://mazarinum.bibliotheque-mazarine.fr/
Bibliothèque Sainte-Geneviève	OK Archive.org	Pas de bibliothèque numérique mais documents numérisés accessibles sur archive.org http://www.archive.org/details/bibliothequesaintegenevieve
Bibliothèque Sainte-Geneviève	OK (Dev. Interne)	Reliures estampées à froid http://bsg-reliures.univ-paris1.fr/fr/accueil/
Bibliothèque Sainte-Geneviève	OK	Expositions Bibliothèque nordique/Expositions virtuelles http://www.bsg.univ-paris3.fr/nordique/expositions.htm

Structures	Bibliothèque Numérique en fonctionnement	Note : cette liste recense des bibliothèques numériques et des collections numérisées accessibles uniquement sur Gallica.
Bibliothèque Sainte-Geneviève	KO	Réserve/Expositions en ligne http://www-bsg.univ-paris1.fr/la_reserve/home.html
BNU Strasbourg	OK. Gallica Marque blanche	Collections numérisées http://www.bnu.fr/collections/la-bibliotheque-numerique
Bordeaux 3 Montaigne	OK (Omeka)	http://1886.u-bordeaux3.fr
Bordeaux 3 Montaigne	OK (Omeka)	Mauriac en ligne : http://mauriac-en-ligne.u-bordeaux-montaigne.fr
Bordeaux Université	OK (Omeka)	http://www.babordnum.fr
Brest. Centre de recherche bretonne et celtique	OK (Omeka Huma Num)	http://bibnumcbbc.huma-num.fr/
Caen Université	OK (Dev. Interne)	Corpus Descartes https://prodescartes.greyc.fr/
Caen. SCD de l'Université de Caen. Fonds de la société Linnéenne de Normandie	OK Gallica-Flickr	Pas de bibliothèque numérique mais un fonds numérisé (dépôt sur Gallica et Flickr Commons) http://scd.unicaen.fr/parteneriats/bibliotheque-nationale-de-france/
Catho de Paris	OK	http://bibliotheque-numerique.icp.fr
Cité des Sciences et de l'Industrie	KO	Scientifica : bibliothèque numérique de la BSI http://www.cite-sciences.fr/cs/Satellite?c=Page&cid=1195217022070&pagename=Portail%2FMED%2FPortailLayout&pid=1195217022149 (HS janvier 2017)
Clermont Université	OK (Omeka)	Carnets géologiques de Philippe Glangeaud, http://bibliotheque.clermont-universite.fr/glangeaud/
Clermont Université BCU	OK	http://bibliotheque.clermont-universite.fr/collections-numerisees
UMS 3726 Centre pour les humanités numériques et l'histoire de la justice	OK	Criminocorpus/Bibliothèque numérique (Centre A.-KOYRE/CRHST. UMR n°8560) http://www.criminocorpus.cnrs.fr/rubrique4.html
Centre de recherche en histoire des sciences et des techniques (CRHST) de la Cité des Sciences et de l'Industrie	OK (Plate-forme interne)	Jean-Baptiste Lamarck (CHRST/CNRS) http://www.lamarck.cnrs.fr/index.php?lang=fr
CNRS / CRHST	OK (Plate-forme interne)	Oeuvres de Lavoisier (CRHST/CNRS) http://www.lavoisier.cnrs.fr/
CNRS / CRHST	OK (Iceberg)	Ampère et l'histoire de l'électricité (CRHST/CNRS) http://www.ampere.cnrs.fr/?lang=fr
CNRS-Maison des Sciences de l'Homme Paris Nord	OK (e-Critur)	Bibliothèque virtuelle Claude Bernard http://claudobernard.in2p3.fr/index.php?name=paratexte&page=Le_Projet

Structures	Bibliothèque Numérique en fonctionnement	Note : cette liste recense des bibliothèques numériques et des collections numérisées accessibles uniquement sur Gallica.
CNRS/CERAM	KO	ColiSciences (CNRS/CERAM) http://colisciences.in2p3.fr/index.php
CNRS/EHESS/Muséum National d'Histoire Naturelle/Cité des Sciences et de l'Industrie	OK (Dev. Interne)	Buffon et l'histoire naturelle : l'édition en ligne (UMR 8560 du CNRS/EHESS/Muséum National d'Histoire Naturelle/Cité des Sciences et de l'Industrie, pôle HSTL) http://www.buffon.cnrs.fr/
Collège de France	OK (Pleade)	Salamandre https://salamandre.college-de-france.fr/
Conservatoire national des Arts et Métiers. Paris	OK	CNUM-Conservatoire numérique des Arts et Métiers http://cnum.cnam.fr/
Corse	OK Gallica	Pas de bibliothèque numérique, mais documents numérisés accessibles via Gallica http://gallica.bnf.fr/services/engine/search/sru?operation=searchRetrieve&exactSearch=false&version=1.2&query=((bibliotheque%20adj%20%22Biblioth%C3%A8que%20de%20l%27Universit%C3%A9%20de%20Corse%22))
Corse	OK (Omeka)	M3C Médiathèque culturelle de la Corse et des Corses (LISA) http://m3c.univ-corse.fr
Cujas	OK (Digitool)	Expositions virtuelles http://bcujas-digitool.univ-paris1.fr/R/8V9DRMHLE2I85GIL33PCTQIT75QBSPI33LLN6BYYUNBQ3X4A4X-01946?func=collections-result&collection_id=8873
Cujas	OK	Ouvrages en ligne http://cujasweb.univ-paris1.fr/search
Dijon. SCD	OK (Dev. Interne)	Fonds Queneau http://www.queneau.fr/
Dijon. SCD	KO	Cabinet des sciences http://collections-rares.u-bourgogne.fr/pleade/ead.html?id=D0002 (ne fonctionne plus en 2017)
Dijon. SCD	-	Pas de bibliothèque numérique globale mais plusieurs fonds numérisés
Ecole centrale Lyon	OK (Lodel)	Histoire de l'Ecole Centrale de Lyon http://histoire.ec-lyon.fr
Ecole des Chartes	OK (Omeka)	Bibliothèque numérique de l'Ecole des Chartes http://bibnum.enc.sorbonne.fr/
Ecole des Chartes	OK	Elec : Editions électroniques en ligne de l'Ecole des Chartes http://elec.enc.sorbonne.fr/
Ecole des Ponts ParisTech	OK (Omeka)	Bibliothèque numérique patrimoniale http://patrimoine.enpc.fr/
Ecole des Ponts ParisTech	OK (Omeka)	Bibliothèque des phares : http://bibliothequedesphares.fr/
Ecole Française d'Extrême Orient. Paris	KO	Photothèque http://collection.efeo.fr/
Ecole polytechnique	OK	Numix, bibliothèque numérique http://numix.sabix.org/

Structures	Bibliothèque Numérique en fonctionnement	Note : cette liste recense des bibliothèques numériques et des collections numérisées accessibles uniquement sur Gallica.
ENAP- Médiathèque Gabriel Tarde. Agen.	KO	Bibliothèque numérique http://enap-mediathèque.paprika.net/enap1/
ENS Cachan	OK (Omeka)	Documents d'archives récents liés à l'enseignement des sciences et disciplines technologiques. http://artheque.ens-cachan.fr
ENS Lyon	OK (Omeka)	Le roman des Morand, histoire d'une correspondance. http://morand.ens-lyon.fr
ENS Lyon	OK	L'Echo de la fabrique et la petite presse ouvrière lyonnaise des années 1831-1835 http://echo-fabrique.ens-lyon.fr/
ENS Lyon	OK	Corpus des Textes Linguistiques Fondamentaux http://ctlf.ens-lyon.fr/
ENSSIB Lyon	OK	Corpus de revues et d'ouvrages anciens numérisés http://www.enssib.fr/bibliotheque-numerique/
Fondation des MSH	OK	Bibnum (FMSH): Textes fondateurs de la science analysés par les scientifiques d'aujourd'hui http://www.bibnum.education.fr/
Grenoble 2 et 3	OK (Omeka)	Hydraulica. Bibliothèque numérique d'ouvrages d'arts en hydraulique http://hydraulica.grenet.fr
Grenoble 2 et 3	KO	http://bibnum-stendhal.upmf-grenoble.fr/
Grenoble. MSH	OK (Omeka)	Fonte Gaia : échange d'information et de débat d'idées autour de l'italianisme contemporain et international. http://www.fontegaia.eu
Paris 7-Groupe International de Recherches Balzaciennes	OK	Balzac. La Comédie humaine. Edition en ligne critique (Groupe International de Recherches Balzaciennes, Mairie de Paris, Université de Chicago) http://www.v1.paris.fr/commun/v2asp/musees/balzac/furne/presentation.htm
Ifremer. Brest	OK	Photographies anciennes du littoral http://www.ifremer.fr/sextant/fr/web/photos-anciennes-littoral/
Institut de France. Académie des Sciences	KO	Histoire et Mémoires de l'Académie royale des sciences http://ads.ccsd.cnrs.fr/
Institut national d'histoire de l'art	OK (Limb Gallery)	Bibliothèque numérique http://bibliotheque-numerique.inha.fr/
Institut National du patrimoine. La Plaine Saint-Denis	OK	Médiathèque numérique http://mediathèque-numerique.inp.fr/
Institut Pasteur	OK (Armadillo)	Patrimoine numérique de l'Institut Pasteur http://bibnum.pasteur.fr/
ITEM	OK ((Huma Num)	E-Man Plateforme d'édition électronique (et génétique) de manuscrits modernes (ITEM) http://eman.hypotheses.org/
Le Mans. SCD	KO	Pas de bibliothèque numérique globale mais un fonds numérisé. Planches mycologiques http://planches-mycologiques.univ-lemans.fr/ (non accessible en 2017)
Lille 3	OK (Dev. Interne)	Pôlib : Bibliothèque numérique patrimoniale http://polib.univ-lille3.fr/

Structures	Bibliothèque Numérique en fonctionnement	Note : cette liste recense des bibliothèques numériques et des collections numérisées accessibles uniquement sur Gallica.
Lille 1. SCD	OK (Dspace)	IRIS: Bibliothèque numérique en histoire des sciences http://iris.univ-lille1.fr
Lille 3. SCD	OK (PLEADE)	Nordnum : Bibliothèque numérique d'Histoire régionale du Nord Pas-de-Calais http://nordnum.univ-lille3.fr
Limoges	OK Flickr	Herbier numérisé : accès sur Flickr http://www.flickr.com/photos/scdlimoges et le Museum d'Histoire Naturelle https://science.mnhn.fr/institution/bibliimo/collection/limo/item/search
Limoges	OK (Lodel)	Album de cartes postales de Simon Jeanjean : http://epublications.unilim.fr/jeanjean/
Limoges. ESPE	OK (Omeka)	Fonds Histoire de l'Education. http://www.unilim.fr/histoire-education/
Lorraine Université. DDE	OK (Omeka)	Pulsar : Bibliothèque numérique patrimoniale http://pulsar-bu.univ-lorraine.fr/
Lorraine. MSH	OK (Omeka)	Henri Poincaré : http://henri-poincare.ahp-numerique.fr
Lorraine. MSH	OK (Omeka)	Bourbaki. http://archives-bourbaki.ahp-numerique.fr
Lyon 1	OK Gallica	Pas de Bibliothèque numérique mais documents numérisés accessible via le catalogue et Gallica http://gallica.bnf.fr/services/engine/search/sru?operation=searchRetrieve&exactSearch=false&version=1.2&query=(provenance%20adj%20%22bulyon1%22)
Lyon 1	KO	Œuvres Complètes de d'Alembert (CNRS/Université Lyon 1) http://dalembert.univ-lyon1.fr/
Lyon 2. Maison de l'Orient et de la Méditerranée	OK	Bibliothèque numérique Digimom http://www.mom.fr/digimom/index.php
Lyon Bibliothèque Diderot	OK Gallica	Accès via Gallica http://gallica.bnf.fr/services/engine/search/sru?operation=searchRetrieve&exactSearch=false&collapsing=true&version=1.2&query=dc.source%20all%20%22Biblioth%C3%A8que%20Diderot%20de%20Lyon%22
Lyon Bibliothèque Diderot	OK	Bibliothèque numérique de la BD http://www.bibliotheque-diderot.fr/collections-numerisees-236974.kjsp?RH=bdl&RF=bdl-0204 Accès via le catalogue: http://www.bibliotheque-diderot.fr/collections-numerisees-bis/
Mathdoc	OK	Numdam Recherche et téléchargement d'archives de revues mathématiques numérisées (Cellule Mathdoc) http://www.numdam.org/
Mines ParisTech. Paris	OK (Omeka)	Bibliothèque numérique patrimoniale https://patrimoine.mines-paristech.fr/
Montpellier. Bibliothèque interuniversitaire	OK (Flora)	Bibliothèque numérique Foli@ http://www.biu-montpellier.fr/redir/fofia
Montpellier. Bibliothèque interuniversitaire	OK (Flora)	Exposition virtuelle "Manuscrits musicaux de la bibliothèque universitaire de Médecine de Montpellier" http://manuscrits.biu-montpellier.fr/Enluminures des manuscrits de la BU de Médecine http://www.biu-montpellier.fr/florabium/jsp/redir.jsp?redir=form_enluminures.

Structures	Bibliothèque Numérique en fonctionnement	Note : cette liste recense des bibliothèques numériques et des collections numérisées accessibles uniquement sur Gallica.
Montpellier. Bibliothèque interuniversitaire	OK (Dev. Interne)	Exposition virtuelle "Belles lettres en lumière" http://manuscrits.biu-montpellier.fr/expo_initiales/ .
Mulhouse. Université de Haute Alsace	OK (Flora)	Photothèque http://www.phototheque-scd.uha.fr/
Museum d'Histoire Naturelle	OK (Ermes Archimed)	Bibliothèque centrale. Textes numérisés / Base d'images http://mussi.mnhn.fr/cda/?INSTANCE=INCIPIO
Nantes	OK Gallica	Documents numérisés accessibles via Gallica http://gallica.bnf.fr/services/engine/search/sru?operation=searchRetrieve&version=1.2&maximumRecords=50&page=1&query=%28bibliotheque+adj+%22Biblioth%C3%A8que+universitaire+de+Nantes%22%29
Observatoire. Paris	OK (Omeka)	https://bibnum.obspm.fr Alidade -Accès en Ligne aux Instruments, Documents et Archives De l'astronomie- est une plate-forme donnant accès aux différents inventaires décrivant les collections patrimoniales conservées à l'Observatoire de Paris: https://alidade.obspm.fr/cdc.html
Orsay - Paris Sud	OK Gallica	Pas de bibliothèque numérique globale mais documents numérisés accessibles sur Gallica
Orsay - Paris Sud - Maison de la Chimie	OK Gallica	http://www.bibliotheques.u-psud.fr/fr/collections/collections-specialisees/fonds-patrimonial-maison-de-la-chimie.html accès en ligne sur Gallica http://gallica.bnf.fr/services/engine/search/sru?operation=searchRetrieve&exactSearch=false&version=1.2&query=((bibliotheque%20adj%20%22Universit%C3%A9%20Paris-Sud%22))
Paris 1 Université	OK	Bibliothèque virtuelle PHARE (Pôle d'Histoire de l'Analyse et des Représentations Economiques-Université Paris 1) http://phare.univ-paris1.fr/
Paris 4	OK	Observatoire de la vie littéraire (Labex Obvil) http://obvil.paris-sorbonne.fr
Paris 6 et Paris 7	KO	Journal Electronique d'Histoire des Probabilités et de la Statistique (EHESS, Université Paris Vi et VII) http://www.jehps.nethttp://ista.univ-fcomte.fr/catalogue_biblio_numerique.html
Paris 6 Université Pierre et marie Curie	OK (Pleade)	Jubilothèque http://jubilotheque.upmc.fr/
Paris 8 SCD Saint Denis	OK (Omeka)	http://octaviana.fr
UMR 5478 IKER	OK (Omeka)	Anpersana bibliothèque numérique recevant et diffusant des sources primaires de la recherche dans le domaine des études basques. https://anpersana.iker.univ-pau.fr
Perpignan. BU	OK (Omeka)	Estudi https://estudi.univ-perp.fr/
Poitiers	OK (WordPress)	Exposition virtuelle L'état de l'Irlande http://etat-irlande.edel.univ-poitiers.fr/
Poitiers	OK (WordPress)	Exposition virtuelle Regards sur l'enfance http://regards-enfance.edel.univ-poitiers.fr/

Structures	Bibliothèque Numérique en fonctionnement	Note : cette liste recense des bibliothèques numériques et des collections numérisées accessibles uniquement sur Gallica.
Poitiers	OK (Omeka)	SOCIÉTÉS RÉELLES, SOCIÉTÉS RÊVÉES : UNE HISTOIRE DE L'UTOPIE http://une-histoire-de-lutopie.edel.univ-poitiers.fr
Poitiers	OK (Omeka)	Bibliothèque virtuelle des premiers socialismes http://premierssocialismes.edel.univ-poitiers.fr/
Poitiers	OK (Omeka)	Coutumiers du Centre-Ouest http://coutumiers.edel.univ-poitiers.fr/
Poitiers	OK (Omeka)	Biblioteca Virtual Cordel http://cordel.edel.univ-poitiers.fr/
Poitiers	OK	Biblioteca virtual CIBeLES http://cibeles.edel.univ-poitiers.fr/
Poitiers	OK	Exposition virtuelle La Rose et l'imprimé http://la-rose-et-limprime.edel.univ-poitiers.fr/
Poitiers	OK	Exposition virtuelle Traversées: récits de voyages des Lumières http://expo-recits-de-voyage.edel.univ-poitiers.fr/index.html
PSL	OK (Omeka)	Explore https://explore.univ-psl.fr/fr/collections
PSL	OK (Omeka)	Expositions virtuelles: https://explore.univ-psl.fr/fr/virtual-exhibitions
Reims. SCD	OK	Eaux minérales, eaux médicales http://www.univ-reims.fr/site/autre/eaux-minerales-eaux-medicales/preface,15662,26933.html
Rennes 1. SCD	OK Gallica	Documents accessibles via Gallica http://gallica.bnf.fr/services/engine/search/sru?operation=searchRetrieve&exactSearch=false&version=1.2&query=dc.source%20adj%20%22universit%C3%A9%20de%20Rennes%201%22
Rennes 2. SCD	OK (Omeka)	Bibliothèque numérique http://bibnum.univ-rennes2.fr/
Rouen Université	OK	Les manuscrits de Madame Bovary http://bovary.univ-rouen.fr/
Rouen Université	OK	Gustave Flaubert http://flaubert.univ-rouen.fr/
Sorbonne Universités	OK	http://documentation.sorbonne-universites.fr/ressources/collections-numerisees.html
Strasbourg SCD	OK (Content DM)	Bibliothèque numérique patrimoniale http://docnum.unistra.fr/
Toulouse Université SICD	OK (Drupal)	Tolosana, la bibliothèque virtuelle des fonds anciens http://tolosana.univ-toulouse.fr/
Toulouse. MSH.	OK (Omeka)	Portail d'archives numérisées de chercheurs. http://architoul.huma-num.fr
Tours. MSH Val de Loire	OK (Omeka)	Alter Prop, Alternatives de propriétés pour l'habitat (programme de recherche de l'Université François-Rabelais de Tours et de l'Université de Bretagne Occidentale de Brest)
Tours. MSH Val de Loire	KO (Omeka)	Mame et Fils, collection de l'A.N.R. "Mame (1796-1975) : deux siècles d'édition pour la jeunesse"
Tours. MSH Val de Loire	KO (Omeka)	Mutual Heritage, projet sur le patrimoine architectural et urbain récent dans le monde méditerranéen
Tours. Université François	OK	Architectura http://architectura.cesr.univ-tours.fr/index.asp

Structures	Bibliothèque Numérique en fonctionnement	Note : cette liste recense des bibliothèques numériques et des collections numérisées accessibles uniquement sur Gallica.
Rabelais.		
Tours. Université François Rabelais. Centre d'Etudes Supérieures sur la Renaissance.	OK	Bibliothèques Virtuelles Humanistes http://www.bvh.univ-tours.fr/index.htm
Transcrire (consortium des archives des ethnologues, Paris & Marseille)	OK (Omeka)	http://transcrire.huma-num.fr

6.3 Annexe 3 : établissements avec lesquels ont été menés les entretiens

Structures documentaires	Structures de recherche
<ul style="list-style-type: none"> • Bibliothèque Nationale et Universitaire de Strasbourg • Bibliothèque interuniversitaire de la Sorbonne • Bibliothèque interuniversitaire de Santé • École nationale des Chartes • SCD Université Claude Bernard Lyon 1 • SCD du Conservatoire national des arts et métiers (Cnam) • Bibliothèque de documentation internationale contemporaine – La contemporaine • Bibliothèque de Sciences Po Paris • Bibliothèque Mazarine • SCD Université Bordeaux Montaigne • Bibliothèque de l’Observatoire de Paris • BUSIM (Bibliothèque de l'Université et de la Société Industrielle de Mulhouse) • Bibliothèque interuniversitaire Sainte-Geneviève • SCD Université de Lille : SHS (ex-Lille 3) + Lilliad (ex-Lille 1) • Bibliothèque de l’École française d’Extrême-Orient 	<ul style="list-style-type: none"> • Cellule Mathdoc – Université Grenoble Alpes • OSUG (Observatoire des sciences de l’univers de Grenoble) - Université Grenoble Alpes • Université Grenoble Alpes - Projet Fonte Gaia • CRAHAM Caen • UMR7041 Archéologies et Sciences de l'Antiquité (ArScAn) • Laboratoire HisCAnt-MA - Université de Lorraine • Maison des Sciences de l'Homme Ange-Guépin Nantes • CITERES - Laboratoire Archéologie et Territoires Tours • Photothèque CESCOM Poitiers • Institut de recherche et d'histoire des textes • Institut international de la marionnette • Laboratoire HiSoMA - Université Lumière Lyon 2 • Infrastructure ReColnat

