

La numérisation des collections patrimoniales imprimées et iconographiques du SCD de Montpellier 2

Dossier d'aide à la décision

Cyril Czernielewski

Sous la direction de Christian Ollès
Enseignant à l'ENSSIB

Conseiller
professionnel
Vial Mireille
Responsable du service des livres anciens
BIU Montpellier section Médecine

Correspondant
d'établissement
Galceran Mireille
Directeur du SCD Montpellier 2
BIU Montpellier section Sciences

Résumé : Le Service Commun de Documentation de Montpellier 2 souhaite valoriser ses collections patrimoniales en élaborant un projet de numérisation. Ce dossier d'aide à la décision présente les différentes possibilités techniques et normatives à mettre en œuvre, et interroge la faisabilité d'une telle démarche au sein du service.

Descripteurs :

Bibliothèque interuniversitaire section sciences (Montpellier).

Numérisation.

Bibliothèque ** Fonds spéciaux ** Sciences.

Images, traitements des ** Techniques numériques.

Toute reproduction sans accord express de l'auteur à des fins autres que strictement personnelles est prohibée.

Sommaire

SOMMAIRE.....	2
INTRODUCTION.....	4
I° LE CONTEXTE DU PROJET DE NUMÉRISATION.....	5
1. L’UNIVERSITÉ DE MONTPELLIER 2.....	5
2. LA BIBLIOTHÈQUE INTERUNIVERSITAIRE DE MONTPELLIER.....	6
3. LE SCD DE MONTPELLIER 2.....	6
II° LA DÉFINITION DU PROJET.....	8
1. LES OBJECTIFS DE RÉALISATION.....	8
2. LES ENJEUX DU PROJET.....	9
3. LES COLLECTIONS À NUMÉRISER.....	10
4. LES RESSOURCES EXISTANTES.....	17
5. LES CONTRAINTES DU PROJET.....	19
6. LES PUBLICS VISÉS.....	20
III° LES RÉFÉRENCES EXTÉRIEURES.....	23
1. LA BIBLIOTHÈQUE CENTRALE DU MUSÉUM D’HISTOIRE NATURELLE.....	23
2. CNUM, LA BIBLIOTHÈQUE NUMÉRIQUE DU CNAM.....	26
3. NORDNUM ET LIBRIS : DEUX PROJETS NUMÉRIQUES AUTOUR DE L’UNIVERSITÉ DE LILLE 3.....	27
4. SYNTHÈSE.....	30
IV° LE CYCLE DE NUMÉRISATION : SPÉCIFICATIONS TECHNIQUES ET ORGANISATION GÉNÉRALE.....	31
1. LA PRÉPARATION À LA NUMÉRISATION.....	31
2. LA DÉMATÉRIALISATION.....	34
3. LE STOCKAGE.....	46

4. LE CONTRÔLE QUALITÉ.	50
5. L'EXPLOITATION DES DOCUMENTS.....	52
6. LE TRAITEMENT DOCUMENTAIRE.....	59
7. LA MAINTENANCE.....	61
V° QUELLES SOLUTIONS POUR LE SCD DE MONTPELLIER 2.	63
1. TROIS INVARIANTS : LA CONSTITUTION D'UNE ÉQUIPE, LA FORMATION DES PERSONNELS ET L'INVESTISSEMENT MINIMAL.	63
2. INTERNALISATION VERSUS EXTERNALISATION.	67
3. TROIS SCÉNARIOS.	69
BIBLIOGRAPHIE	76
TABLE DES ANNEXES	81
TABLE DES MATIÈRES	105

Introduction

L'Université et le SCD de Montpellier 2 disposent de collections patrimoniales imprimées et iconographiques en nombre et en qualité. Ces collections sont cependant très peu signalées et, par bien des aspects, fort peu valorisées, alors même que l'engouement des publics pour le patrimoine et les questions, voire les inquiétudes, sur la « Science » apparaissent plus pressants. Les développements technologiques permettent aujourd'hui d'apporter une réponse positive à ces attentes légitimes, tout en conciliant, grâce à la numérisation, la mission de conservation des bibliothèques. Ces technologies arrivent à maturité et nombre d'expériences, en France ou à l'étranger, permettent d'avoir maintenant un certain recul.

Dans le cadre du contrat quadriennal 2003-2006, le responsable du SCD nous a demandé, afin de lui soumettre un dossier d'aide à la décision, de mener une étude relative au projet de numérisation des collections patrimoniales imprimées et iconographiques du SCD de Montpellier 2. Elle a pour objectifs d'établir un état des lieux en matière de numérisation, en termes juridique, normatif et technique ; de déterminer, pour les collections patrimoniales, les solutions envisageables de dématérialisation et de communication ; d'apprécier, enfin, comment elles pourraient s'articuler dans des scénarios vraisemblables en relation avec le contexte du SCD.

Selon ce schéma, on présentera dans un premier temps le contexte (partie I), puis, on définira plus précisément le projet (partie II) avant d'aborder quatre expériences extérieures au SCD (partie III). Par la suite, on s'attachera à décrire le cycle de numérisation appliqué aux collections patrimoniales du SCD et les différentes opérations à mener (partie IV). Pour terminer, nous discuterons des solutions envisageables et proposerons trois scénarios (partie V).

I° Le contexte du projet de numérisation.

1. L'Université de Montpellier 2.

Créée en 1809 par décret du 17 mars 1808, la Faculté des Sciences de Montpellier devient en 1964 l'Université des sciences et techniques du Languedoc¹. L'Université de Montpellier 2 accueille aujourd'hui quelque 17 000 étudiants du 1^{er} au 3^{ème} cycle et 1000 chercheurs environ. Ils sont répartis entre le site central et les sites distants, dont, pour la pédagogie, les IUT de Montpellier, de Sète, Nîmes et Béziers². Les principales filières d'enseignement figurent à l'annexe n°1. Les laboratoires de recherche sont regroupés au sein de 7 départements. L'enseignement à la recherche est structuré selon 8 écoles doctorales³.

En 2004, l'Université s'est dotée d'un Service Commun des Collections, destiné à organiser et à valoriser le patrimoine scientifique et technique des collections physiques ; le Service Commun de Documentation étant, pour sa part, chargé des collections patrimoniales imprimées et iconographiques.

Par ailleurs, l'Université porte, sous l'égide du Pôle Universitaire Européen de Montpellier et du Languedoc-Roussillon, le projet MuseUM, en partenariat avec les Universités de Montpellier 1 et 3. Ce musée des Universités de Montpellier se voulant à la fois musée et centre de culture scientifique, s'est établi dans un périmètre comprenant le Jardin des Plantes, les locaux historiques de la Faculté de Médecine, l'Institut de Botanique et certains locaux de l'hôpital St Charles. Doté d'un budget de 50 millions d'euros, MuseUM vise 250 000 visiteurs par an⁴. Dans ce cadre, l'Université de Montpellier 2 présente une thématique « Trésors des sciences » autour de ses collections d'objets scientifiques témoignant de l'histoire du savoir scientifique et de sa transmission.

¹ Dulieu Louis. La Faculté des sciences de Montpellier de ses origines à nos jours. 1981. p.9 et p.34.

² L'annexe n°1 détaille la population étudiante de l'Université en 2004.

³ L'annexe n°2 présente les départements et les formations doctorales.

2. La bibliothèque interuniversitaire de Montpellier.

La documentation universitaire de Montpellier est structurée autour de la Bibliothèque Inter-Universitaire, SICD, et des 3 SCD de Montpellier 1, 2 et 3, qui fédère 8 bibliothèques universitaires. Les principales missions de la BIU concernent la politique générale de l'établissement en accord avec les universités, la coordination des BU et de la politique documentaire, ainsi que la gestion administrative et financière de l'ensemble. Parmi ses domaines de compétence on notera, aussi, la définition et la mise en place de la politique patrimoniale et de la politique en matière informatique de la BIU, qui s'ancrent, respectivement, autour de la mission transversale du patrimoine et du service informatique. Ce dernier assure les fonctions d'informatisation et de réinformatisation, de maintenance et de développement au niveau matériel et logiciel, et participe à l'administration de la documentation électronique. Pour l'ensemble des bibliothèques universitaires, le système intégré de gestion de bibliothèque est ALEPH 500 de la société Ex-Libris. La mission transversale du patrimoine assure des fonctions de coordination et de financement, les SCD disposant d'une certaine autonomie politique en matière patrimoniale. La BIU dispose, de plus, de deux ateliers, l'un de restauration où 3 techniciens d'art formés à la BnF sont affectés, l'autre de photographie avec un photographe chargé des reproductions.

3. Le SCD de Montpellier 2.

Le SCD de Montpellier 2 fédère les bibliothèques des composantes de l'Université et coopère avec les bibliothèques des départements de recherche. Ce réseau documentaire comprend la bibliothèque universitaire du campus, les 5 bibliothèques des sites distants (IUT de Nîmes et de Montpellier ; UFR des sciences aux Carmes à Nîmes) et les bibliothèques de recherche de l'ISTEEM, de BEE et de Mathématiques. 2 personnels des bibliothèques sont en poste à la bibliothèque de l'IUT de Nîmes et à la bibliothèque de l'antenne à Béziers de l'IUT de Montpellier. Les bibliothèques des départements de recherche sont

⁴ Pellequer Bernard. Projet MuseUM, la science en partage. 2004.

pourvues de documentalistes recrutés par l'Université. La BU sciences dispose de 29 personnels titulaires ou stagiaires dont 3 conservateurs, 4 bibliothécaires, 3 bibliothécaires adjoints spécialisés, 2 assistants de bibliothèque, 3 adjoints administratifs et 10 magasiniers. 3 agents en Contrat Emploi Solidarité viennent compléter les effectifs, l'un pour renforcer l'équipe de catalogage, les deux autres assurant des fonctions de magasinage.

Le chantier d'extension de la BU de 1999 à 2001 a débouché sur la mise à disposition d'un bâtiment câblé, sonorisé et conforme aux normes d'accessibilité, grâce auquel se sont développés de nouveaux services (Référence, Patrimoine) et le système d'information documentaire (SID).

Les services du SCD sont actuellement organisés autour de deux pôles que sont les Collections et les Nouvelles Technologies de l'Information et de la Communication. Ce dernier assure les fonctions d'administration du réseau, du SID et de mise en œuvre de la politique documentaire en matière électronique. Le service du Patrimoine, en charge de la conservation, de la communication et de la valorisation des documents anciens, est rattaché au pôle Collections. N'ayant pu bénéficier d'une équipe stable et en effectif suffisant ces deux dernières années, il est aujourd'hui opérationnel avec l'arrivée d'un conservateur. Pour autant, il n'offre qu'un espace de consultation des documents réduit. Le fonds imprimé ancien et les documents iconographiques sont très peu signalés. Sur le site Web de la BIU, le patrimoine du SCD est absent.

Parmi les cinq grands axes de développement prévus par le contrat de plan quadriennal 2003-2006, on retiendra les trois suivants auxquels le projet de numérisation est, directement ou indirectement, lié :

- Accroître l'offre documentaire et faciliter l'accès aux documents, notamment grâce à la numérisation des thèses, de la littérature grise et de textes libres de droit.
- Conserver et valoriser le patrimoine, avec notamment, le développement du service de patrimoine, l'inventaire et la signalisation des fonds, la restauration d'ouvrages et la numérisation des textes et de l'iconographie.

- Politique culturelle, avec l'association aux différentes célébrations nationales (Mois du Patrimoine, Fête de la Science, etc.) et locales.

Le projet de contrat de plan prévoit une dotation de 26000 € au titre de la conservation et de 21200 € au titre de la valorisation pour les collections patrimoniales du SCD.

II° La définition du projet.

1. Les objectifs de réalisation.

L'objectif final de réalisation est la mise à disposition des publics, à partir des fonds patrimoniaux du SCD, d'un fonds numérique thématique en relation avec l'histoire de la botanique à Montpellier. Les objectifs intermédiaires de réalisation sont de deux ordres: d'une part, la production des données : la dématérialisation, le traitement documentaire et le stockage des enregistrements de référence ; d'autre part, la communication des données avec la mise en place d'un outil de consultation et la production de données et métadonnées compatibles avec cet outil. Ces deux objectifs intermédiaires ne sont pas indépendants l'un de l'autre dans la mesure où, la solution technique de communication conditionne en partie le travail de production, en particulier du point de vue du traitement documentaire. Inversement, le travail de production des données conditionne les possibilités techniques de valorisation et une qualité technique maximale des enregistrements de référence accroît l'étendue des actions. Ils demeurent cependant distincts dans la mesure où l'un correspond à un travail de numérisation à des fins de stockage dont l'enjeu est la conservation de l'information, et où l'autre correspond à un travail de numérisation à des fins de transmission dont les enjeux sont la communication et la valorisation des collections patrimoniales. Si la publication des documents numériques répond à ces derniers, elle n'épuise cependant pas la totalité des destinations possibles des enregistrements de référence : expositions virtuelles, besoins éditoriaux internes ou externes, publics ou privés...etc. Par

ailleurs, les actions de communication et de valorisation peuvent faire l'objet d'une programmation dans le temps, fixant, par exemple, une première étape avec une communication intranet, puis une seconde avec une diffusion Internet, tout en participant à des actions culturelles en relation avec l'Université ou la bibliothèque interuniversitaire de Montpellier. En terme opérationnel enfin, cette distinction souligne le caractère central de l'enregistrement de référence à partir duquel, après adaptation, les actions de communication et de valorisation seront menées, et dont il est nécessaire d'assurer la conservation à plus long terme.

2. Les enjeux du projet.

On distinguera les enjeux propres aux missions des bibliothèques auxquelles le SCD est assujetti et les enjeux propres au SCD.

Au niveau des missions générales des bibliothèques :

Il s'agit d'articuler dans une même opération les missions de conservation, de communication et de valorisation des collections patrimoniales du SCD⁵ :

- En accroissant la visibilité des fonds patrimoniaux scientifiques de l'Université et en assurant leur diffusion selon de multiples modalités.
- En améliorant les conditions de conservation des collections, ce qui suppose la restauration éventuelle des documents mais aussi la mise à disposition de documents de substitution dans des conditions telles qu'elles n'accroissent pas la pression sur les originaux.

Au niveau du service :

Il s'agit de développer les capacités de service du SCD en matière documentaire :

- En poursuivant le développement du service du patrimoine du fait de l'intégration des collections écrites, imprimées et iconographiques de l'Université, prévu par le contrat de plan 2003-2006 en complémentarité avec le projet de recensement et de valorisation de l'ensemble des collections de l'Université. Au quotidien, la mise à

⁵ Bechedelièvre Danielle. Pour une valorisation du patrimoine scientifique écrit de l'Université : évaluation des collections du 19^{ème} siècle de la BU Sciences du SCD de Rennes 1. 2002. p. 8 à 27.

Seta Frédérique. Le rôle de la numérisation dans la mise en valeur des fonds patrimoniaux : l'exemple de la bibliothèque interuniversitaire Cujas. 1999. p. 21 à 23.

disposition de reproductions et de fac-similés numériques pallierait le problème de surface de l'espace de consultation.

- En dotant le SCD de Montpellier 2 d'outils, de méthodes et de compétences susceptibles de répondre à un large panel de besoins en matière de numérisation mentionnés dans le contrat de plan 2003-2006 : travaux universitaires, collections patrimoniales, politique culturelle. Et, donc, de développer, à terme, un service de numérisation.

3. Les collections à numériser.

3.1. La collection des vélins de Node-Véran

La collection des vélins de Node-Véran est conservée à l'Herbier de l'Institut de Botanique de Montpellier. Sont représentées des espèces botaniques, des plantes vasculaires, principalement exotiques, mais aussi, quoique en nombre moins important, des vignes et des champignons. La technique est celle de « l'aquarelle à la colle ». La qualité d'exécution est indéniable et l'expression d'ensemble est remarquable. L'état de fraîcheur des pigments ne l'est pas moins. Michel Denizot, René Marie et Guy Privat témoignent de la qualité des vélins, *dessins d'art, mais aussi documents scientifiques apportant un complément substantiel aux exsiccata*⁶.

Node-Véran, élève des Ponts et Chaussées, puis instituteur, devient le peintre officiel du Jardin des plantes de Montpellier en 1813. Les prémices de la collaboration de Node-Véran avec le Jardin remonteraient selon Charles Flahault à l'année 1810. Il le restera jusqu'à sa mort en 1852. Selon Michel Denizot, René Marie et Guy Privat⁷, De Candolle aurait fait état de 200 dessins à son départ, Dunal en 1842 évoque le nombre de 800. Node-Véran est resté 39 ans au service

⁶ Rioux Jean Antoine. Le Jardin des Plantes de Montpellier. Quatre siècles d'histoire. 1994. p. 175.

⁷ Id. p. 175.

du Jardin des plantes. La collection compte 978 dessins, ce qui fait une moyenne de 25 par an, chiffre que relate déjà Dunal en 1842⁸.

Matériellement, la collection des aquarelles est rangée dans un meuble métallique fermé à clé. Elle se présente sous la forme de 10 classeurs en cuir dont l'état de conservation est préoccupant. Certaines coiffes sont en déliquescence. Un ou plusieurs onglets en carton explicitent pour chacun des classeurs la famille à laquelle appartiennent les taxons représentés. Les intercalaires en papier sont jaunis et certains souffrent de roussissure. L'état général des classeurs et du mobilier ne correspond pas aux meilleures conditions de conservation. Si les dessins semblent avoir peu souffert du temps, il n'en semble pas de même des supports, encore moins du matériel de conservation.

Les aquarelles sont présentées selon la classification en vigueur au tournant du XX^{ème} siècle ; Charles Flahault est l'instigateur du changement de classification. Ainsi, chaque aquarelle comporte un numéro en haut à gauche du recto, de 1 à 978. Les traces d'une ancienne classification sont parfois présentes au verso, avec une numérotation différente, de même qu'un tampon, accompagné de la date et de la signature du recteur d'académie.

Les aquarelles disposent pour la plupart d'une légende originale des mains de l'artiste (seules 2 aquarelles sur l'échantillon ne présentent pas de légende originale), parfois d'une autre, bien plus récente, ajoutée au crayon à papier, qui témoigne de l'évolution de la classification en botanique. Ces légendes correspondent aux noms scientifiques des taxons représentés. Sur une même aquarelle, on peut ainsi observer trois marques qui témoignent de l'évolution de la discipline : au recto, la légende originale correspondant à la date du verso, le numéro de la classification du temps de Charles Flahault, et la légende au crayon à papier correspondant à la dénomination scientifique actuelle.

Sur 978 vélins, 245 ont été collationnés à titre d'échantillon, soit 25%. 158 sont de véritables vélins, soit 65% ; 87 ont un support papier, soit 35%. Ce résultat est en accord avec les estimations du responsable de l'Herbier de l'Institut de Botanique. On peut donc estimer qu'il y aurait 636 supports vélins et 342 supports papier. Il existe deux formats, le plus courant est d'environ 32x47 cm, l'autre, un peu plus

⁸ Id. p. 175.

grand, de 42x55 cm et ne s'applique vraisemblablement qu'au support papier. L'état des vélin est variable. L'annexe n°3 reprend les principaux résultats. Par problème de planéité, il faut entendre que les vélin présentent des déformations localisées, témoins, sans doute, des variations d'hygrométrie et de température au cours des années. Sur les 100 premières aquarelles consultées, le plus petit objet considéré comme significatif a été mesuré ; il témoigne régulièrement d'une surface d'1 mm carré. Des détails de texture, telle la pilosité des tiges sont au demeurant bien plus petits. En extrapolant ces chiffres à l'ensemble de la collection, on peut s'attendre, en première approximation, à ce que 28 aquarelles sur vélin souffrent d'un brunissement rédhibitoire et 310 d'un défaut de planéité ; que 8 aquarelles sur papier présentent une roussissure très importante. Cette collection a fait l'objet une édition partielle de 48 reproductions⁹.

3.2. Le fonds patrimonial photographique

Localisées à la bibliothèque universitaire section Sciences de Montpellier, les collections photographiques de l'Institut de Botanique sont très mal connues. Il s'agit majoritairement d'originaux noir et blanc sur plaques de verre grand (13x18 ou 9x12) ou moyen format (4,5x6) et de tirages d'époque ne dépassant pas le format 18x24. Ils représentent des paysages, des campagnes d'herborisation, des vues du Jardin des plantes, et des portraits, parfois en groupe. Ces collections ont été créées à l'initiative de Charles Flahault, directeur de l'Institut de Botanique de 1889 à 1927. On y observe la vie universitaire de l'Institut, du Jardin des plantes de Montpellier, des paysages régionaux, français et étrangers au tournant du siècle dernier (entre 1883 et 1917 sous réserve), ainsi que des vues au microscope. Un inventaire établi par Flahault accompagne la collection, qui indique parfois l'auteur des clichés. Un nombre très faible de plaques a, semble-t-il, fait l'objet d'une édition.

Une première évaluation des collections a été menée en 2004. Matériellement entreposés dans une armoire métallique fermée à clé, les documents photographiques ont fait l'objet d'une première action de conservation lors des travaux d'extension de la bibliothèque avec leur mise en boîte papier neutre. Cette

⁹ Id. Vélin 1 à Vélin 48, p. 176.

action n'a concerné que les plaques de verre grand format et certains tirages papier, le reste des collections, plaques de verre moyen format et certains tirages papier ne dispose pas de matériel de conservation adéquat. Une composante essentielle des bonnes conditions de conservation des films photographiques ancien est cependant absente : le froid. Si la pièce dans laquelle ils sont entreposés ne présente pas a priori une température ambiante trop élevée, ni d'écart trop importants, il n'en demeure pas moins que ces conditions idéales pour la conservation des documents papier ne le sont pas réellement pour le film photographique.

Il est possible de distinguer 6 sous-ensembles dans ces collections :

- Les plaques de verre grand format, mises sous boîte, au format 13x18, 650 environ.
- Les plaques de verre grand format, mises sous boîte, au format 9x12, 30 environ.
- Les « biotomes », au format 9x12, 60 environ.
- Les stéréotypes, 130 environ.
- Les tirages papier et des gravures mis sous boîte, de divers formats, 200 environ, dont certains sont les épreuves des plaques de verre.
- Les plaques de verre moyen format, en carton, en nombre indéterminé.

Au total, il y aurait au moins 1000 documents (dont certains cependant sont les reproductions des autres).

Les plaques de verre sont numérotées. Cette numérotation n'est pas d'origine et semble récente. Un échantillon de 67 plaques 13x18 et 9x12 a été analysé en détail afin d'en évaluer l'état de conservation. Elles ont été classées en 4 catégories, très bon état, bon état, état moyen, et mauvais état. Les critères retenus pour effectuer cette classification ont été la présence d'une coloration anormale partielle ou totale, l'existence d'un vignelage dû au vieillissement sur la partie périphérique de l'image, le décollement périphérique de l'émulsion du support en verre, la présence de tâches sur l'émulsion, la détérioration du support en verre, ou bien encore la perte significative de contraste. Les résultats sont reproduits par l'annexe n°3. Du point de vue de la numérisation, on peut considérer que 40 plaques sur 67

peuvent faire l'objet de l'opération ; les 27 autres ne donneront pas a priori de bons résultats sans traitement.

Les tirages papier n'ont pas été analysés en détail. L'état général de conservation semble bon voire très bon pour certains. Quelques uns souffrent cependant d'une détérioration, notamment d'une perte de contraste. Une partie des tirages présentent, par ailleurs, un problème de planéité. Une évaluation plus approfondie doit être menée sur cette partie du fonds.

3.3. Les ouvrages anciens du SCD de Montpellier 2.

Les fonds patrimoniaux du SCD regroupent des ouvrages anciens, antérieurs à 1810, et des ouvrages du 19^{ème} siècle. Ces fonds sont évalués au nombre de 22000 volumes. Ils procèdent de la réunion des collections propres à la BU et de l'intégration récente des collections de l'Institut de Botanique. Une campagne de restauration a déjà été menée pour les ouvrages les plus abîmés. Une ligne budgétaire est par ailleurs prévue à cet effet dans le contrat quadriennal 2003-2006. Un inventaire est en cours de réalisation, qui devrait permettre de cerner le contenu exact de ces fonds. De l'avis du responsable du SCD et du service du patrimoine, les fonds correspondent assez bien à l'histoire de l'enseignement des Sciences à Montpellier avec une forte dominante en botanique qui représenterait environ 50% des volumes. La plus grande majorité de ces documents n'est pas signalé dans le SUDOC, 200 ouvrages ayant fait pour l'instant l'objet d'un catalogage informatisé, essentiellement des ouvrages antérieurs à 1810. La recherche des autres documents s'effectue grâce au catalogue sur fiches. Une fois l'inventaire effectué, le service du patrimoine compte établir un plan d'actions relatif au catalogage. Dans le cadre du projet de numérisation, une première sélection de 6 ouvrages, pressentis comme candidats à la numérisation pour leur contenu, avait été faite à titre d'échantillon. Une recherche menée dans le cadre de ce travail a révélé que parmi ces 6 ouvrages 5 avaient déjà fait l'objet d'une numérisation et étaient disponibles sur Gallica, et ne sont donc plus candidats au projet. Le choix du corpus reste donc à faire et ne peut être envisagé avant la fin de l'inventaire.

3.4. Les droits.

3.4.1 Dispositif général

Le droit français repose sur la distinction et l'indépendance du droit de propriété intellectuelle et du droit de propriété matérielle. Parmi les droits de propriétés intellectuelles, les droits d'auteurs perdurent jusqu'à la mort de l'auteur, pendant 70 ans à partir de la date de décès au profit de ses ayants droits. L'œuvre tombe ensuite dans le domaine public. Seront donc libres de droits en 2005 les œuvres dont l'auteur est décédé avant 1935. On distingue parmi les droits d'auteur les droits moraux (droit de divulgation ; droit à la paternité ; droit au respect de l'œuvre), inaliénables, et les droits patrimoniaux (droit de représentation et droit de reproduction), cessibles contractuellement par l'auteur ou ses ayants droits.

De la diffusion d'œuvres numérisées naît une autre catégorie de droits, les droits voisins, dont il est nécessaire de tenir compte. Parmi ceux-ci, celui du droit à l'image concerne particulièrement le projet. Le droit à l'image, composant des droits de la personnalité, en regard du respect à la vie privée défini par l'article 9 du code civil, soumet le diffuseur à une autorisation contractuelle de la personne dont l'image est reproduite. Lorsque celle-ci est décédée, la jurisprudence interdit la diffusion d'une image portant atteinte à la dignité de la personne humaine. Cette autorisation préalable est, en outre, nécessaire pour la diffusion d'une image comportant une œuvre protégée par le droit d'auteur. Elle l'est aussi pour la diffusion de l'image d'un bien dont la protection s'établit sur la base des droits de propriété, ou du respect de la vie privée.

En anticipant, il faut ajouter que *toute copie numérique d'une œuvre du domaine public crée un nouvel objet juridique également protégé*, que les bases de données sont aussi protégées par les droits de propriété intellectuelle, et que *l'administration est investie d'un droit sui generis qui empêche toute extraction substantielle*¹⁰. A ce titre, les éventuels droits d'exploitation, suivant le dispositif adopté, devraient faire l'objet d'une convention entre le SCD et l'Université. Si

¹⁰ Battisti Michèle. Assises internationales du PNER Questions juridiques liés à la numérisation à des fins d'enseignement et de recherche. 2002. p.61.

redevance d'utilisation il y a, elle devrait, au même titre que le signalement nécessaire des droits¹¹, faire l'objet d'une mention lors de la diffusion.

3.4.2 Situation des documents soumis au projet.

La situation paraît claire pour les vélins et les ouvrages imprimés ; elle l'est moins pour le fonds particulier des documents iconographiques noir et blanc.

Les ouvrages imprimés anciens du SCD sont a priori libres de droits. Une attention particulière devra cependant être portée aux éditions critiques et à la qualification juridique des œuvres de collaboration, pour les ouvrages édités après 1875¹².

La collection des vélins Node-véran est la propriété de l'Université de Montpellier 2. Les vélins sont libres de droits dans la mesure où l'auteur est décédé en 1852. Leur numérisation et leur diffusion nécessiteront une autorisation formelle de l'Université.

Le fonds patrimonial photographique de l'Institut de Botanique pose des problèmes juridiques plus pointus et un examen détaillé des droits devra être accompli, qui comportera l'étude des droits à l'image déjà signalés, et des droits d'auteur des photographes dont l'histoire juridique complexifie un peu la tâche. En effet, la loi de 1957 restreignait la protection des photographies aux catégories artistique et documentaire jusqu'à l'extension du champ d'application avec la loi de 1985. Il s'en est suivi l'absence de signature des œuvres, surtout au début du siècle, qui pénalise l'identification. Les documents photographiques soumis au projet relèvent pour certains de ce cas de figure puisque les plus récents sont datés de l'année 1917. Si l'identification demeure possible grâce, notamment, à l'inventaire de Charles Flahault, des recherches seront poursuivies afin de déterminer la date de décès de l'auteur. Si l'identification de l'auteur demeure impossible, les documents photographiques peuvent être considérées comme des œuvres anonymes. La durée du droit exclusif est alors de 70 ans à partir du 1^{er} janvier de l'année civile suivant l'année de publication ; si l'auteur ou ses ayants droit se font connaître, la durée de protection redevient celle de la règle commune, 70 ans après la mort de l'auteur. Un autre cas de figure peut aussi se présenter : le statut du photographe peut relever du droit d'auteur, s'il s'agit par exemple d'un

¹¹ Nguyen Duc Long Christine. La numérisation des œuvres Aspects de droits d'auteur et de droits voisins. 2001. p.136.

photographe de métier et que l'on ne dispose en toute rigueur d'aucun acte de cession des droits ; ou bien, s'il s'agit d'un fonctionnaire, des dispositions spécifiques s'appliquent relevant de la qualification juridique de l'œuvre collective et de l'avis OFRATEME du Conseil d'Etat¹³, auquel cas, l'administration est investie des droits d'auteur. Cet examen juridique est complexe et demandera l'avis du service juridique de l'Université.

4. Les ressources existantes.

4.1. Les ressources humaines.

4.1.1 Le service du patrimoine

Le service comprend un conservateur, chartiste (0,5 ETP), un magasinier spécialisé (1 ETP) et un magasinier en chef principal (1 ETP), tous deux formés à la conservation des ouvrages anciens.

4.1.2 Le service des nouvelles technologies de l'information et de la communication

Le service comprend à l'heure actuelle, 1 conservateur (0,5 ETP) et 2 bibliothécaires (2 ETP) dont un partirait à la retraite en 2005 ; un magasinier spécialisé (1 ETP) a intégré le service en septembre 2004.

4.2. Les ressources matérielles.

4.2.1 Le système d'information.

Matériellement le réseau de la bibliothèque universitaire est un sous réseau de l'Université ; il est câblé en 10/100 et ne souffre d'aucun problème de débit. La liste complète et détaillée des matériels informatiques du SCD est reproduite dans l'annexe n°4.

4.2.1.1 Architecture client serveur à l'intérieur du bâtiment

Le système d'information documentaire (SID) est basé sur une architecture client serveur Microsoft 2000. Le serveur héberge le dispositif logiciel de la

¹³Game Valérie, Aspects juridiques de la numérisation, In Conduire un projet de numérisation, 2002, p. 53.
Battisti Michèle, Assises internationales du PNER Questions juridiques liées à la numérisation à des fins d'enseignement et de recherche, 2002, p.59.

société Archimed, en particulier les applications CD Line et le navigateur sécurisé Archimed Internet Explorer. Faisant fonction de serveur de fichiers pour le SID et de serveur d'impression pour les imprimantes professionnelles, il dessert 28 stations clientes pour l'accès aux bases de données, aux ouvrages de références et aux périodiques électroniques, 10 stations spécialisées pour l'OPAC et 6 dédiées à l'enseignement à distance. Les postes clients sont répartis sur les deux étages de la bibliothèque universitaire. En sus, un poste par étage est un serveur d'impression équipé de Watchdoc d'Archimed. Les postes clients sont des matériels acquis il y a quatre ans déjà, ne disposant pas d'un circuit vidéo indépendant et suffisamment performant pour consulter des images en haute résolution.

4.2.1.2 Architecture client léger pour les 6 postes distants

Basé sur l'application Métaframe de Citrix, ce dispositif permet de desservir 6 bibliothèques distantes du réseau documentaire du SCD. Le serveur spécifique à l'application Métaframe est un client du serveur « Archimed ». En terme de performances vidéo, les PC mis à la disposition des bibliothèques appellent la même remarque que ceux de la bibliothèque centrale.

4.2.2 Les locaux.

La mise en œuvre du projet de numérisation, quelle que soit la solution adoptée, nécessite un local adapté pour entreposer le matériel informatique et mener les différentes opérations dévolues au personnel du SCD. La bibliothèque universitaire a bénéficié d'un réaménagement dû à l'extension de sa surface à partir de l'année 2000. Ce réaménagement continue aujourd'hui avec la nouvelle occupation de locaux contigus au bâtiment principal. Les locaux pressentis pour le projet sont les suivants :

- La salle climatisée des serveurs (12m carré) :

Cette pièce climatisée qui abrite le matériel du service des NTIC est totalement occupée. Elle ne peut recevoir une station de travail supplémentaire mais constituerait avec un mobilier adapté un lieu de stockage idéal pour les enregistrements de référence.

- L'ancien laboratoire photographique (18m carré) :

Il sert de magasin de stockage des entrées ; une photocopieuse est installée pour le PEB. Pièce aveugle, il n'y a pas de lumière naturelle. En outre, il ne dispose pas de câblage réseau.

- Deux emplacements sont de plus actuellement vacants :

Les nouveaux bureaux du niveau des magasins. Lumineux, ils ne sont cependant pas, à l'heure actuelle, équipés d'un câblage réseau. A partir de septembre 2004, des travaux de réhabilitation sont prévus qui rendront ces surfaces prochainement opérationnelles.

Le bureau du niveau du fonds professionnel, anciennement bureau de la direction, est câblé et dispose de lumière naturelle. Il renferme une partie des archives de la direction du SCD.

5. Les contraintes du projet.

Les contraintes relatives au projet sont de deux ordres : les unes concernent les documents candidats à la numérisation, les autres sont propres au service.

5.1. Les contraintes liées aux documents.

On en distingue quatre :

On aura noté la multiplicité des documents, imprimés, aquarelles sur vélin, plaques photographiques, tirages papier et gravures, qui impose des solutions techniques de dématérialisation différentes, de par leur valeur, parfois précieuse, leur format et leur support physique. Cette multiplicité des documents soulève, de plus, des questions relatives aux modalités de consultation dans la mesure où celles-ci peuvent apparaître spécifiques aux documents, notamment en matière iconographique. En terme d'usage, la consultation d'une collection de documents imprimés numérisés se révèle différente d'une collection d'images numérisées, et les moyens techniques et ergonomiques mis en œuvre pourraient s'en ressentir.

L'état de conservation de certains documents iconographiques ne permet en l'état d'espérer un niveau qualitatif de dématérialisation satisfaisant. Si leur nombre peut paraître, suivant l'estimation, relativement faible, une numérisation exhaustive des collections semble apporter une réelle valeur supplémentaire au projet. Dans ce

cas, il devrait être procédé à une campagne préalable de restauration dans la mesure où elle reste physiquement et financièrement possible.

Il faut, de plus, considérer le travail d'inventaire indispensable à une bonne connaissance des fonds. A une collation précise des vélins s'ajoute un véritable travail d'archiviste concernant le fonds photographique patrimonial.

Enfin, le statut juridique de certains documents photographiques sur plaques de verre, peut représenter une contrainte qu'on lèvera soit par négociation des droits, soit par leur retrait de la sélection.

5.2. Les contraintes liées au service

Pour l'heure, on insistera sur deux types de contraintes liées au service. Le premier concerne la disponibilité et la formation des personnels. Comme le souligne Valérie Néouze pour le projet de numérisation de la bibliothèque centrale du Muséum d'histoire naturelle, le travail à accomplir *s'ajoute – et là réside une difficulté du projet – aux activités quotidiennes du personnel de la bibliothèque*¹⁴. La réalisation des tâches liées au projet implique la décharge de certaines autres pour les personnels concernés. En ce sens, la réalisation du projet serait celle de l'ensemble des personnels de la bibliothèque et devrait induire des modifications dans l'organisation du service qu'il faudra faire accepter. D'autre part, il faut prendre en considération le niveau de compétence des personnels face à une « nouvelle » technologie et des types de documents peu traités par le service, qui appellera des actions de formation. Le second type de contraintes est d'ordre technique. Il s'agit, d'abord, de la faible capacité d'affichage des postes du SID, et ensuite, de l'absence de site Web propre au SCD qui demandera donc, si on opte pour cette solution immédiatement, d'accorder le dispositif technique à celui des hébergeurs potentiels.

6. Les publics visés.

La palette des usages des documents numérisés est aussi large que l'éventail des publics. Thierry Delcourt en présente une typologie dont on ne retiendra dans

¹⁴ Néouze Valérie. Quelle bibliothèque numérique pour une bibliothèque patrimoniale ? 2000. p.12.

le contexte du projet que les usages éditoriaux, les usages savants, les usages culturels et les usages pédagogiques¹⁵. Toute proportion gardée, les résultats du projet de recherche BibUsages réalisé dans le cadre de Gallica sont, à cet égard, instructifs¹⁶. Les publics de Gallica sont composés à 75% d'enseignants, de chercheurs et d'érudits. 60% des consultations sont effectuées à partir de la France, 85% à partir d'Internet et 15% à partir de l'intranet. 75% des entrants proviennent d'un site non Bnf (moteur de recherche, fournisseur d'accès, site offrant un lien vers Gallica). L'étude montre que l'audience des autres sites du même type est non négligeable et que Gallica se révèle une source de textes numérisés parmi d'autres. Elle souligne l'importance de la bibliophilie et de la recherche en tant que contexte d'accès à Gallica.

Les publics potentiels du projet du SCD sont un public de chercheurs, lié aux usages savants, et un public plus élargi, lié plus directement aux usages pédagogiques et culturels.

6.1. Un public des chercheurs :

Face à des collections iconographiques encore mal signalées, on ne peut que partager le souhait de Claude Collard, d'Isabelle Giannattasio et de Michel Melot lorsqu'ils affirment qu'*une autre variable qui va profondément modifier les données des usages sera, espérons-le, l'accessibilité croissante des collections aujourd'hui mal identifiées par les chercheurs, que nous vivons une période transitoire marquée encore par la disproportion entre une offre limitée, ne permettant guère que des usages occasionnels, et une demande savante et professionnelle qui ne peut que s'accroître, et qu'à cet égard, la responsabilité des bibliothécaires est entière : à eux de donner matière à de nouvelles études*¹⁷. D'où une politique d'offre que l'on doit cependant concevoir avec soin. Les chercheurs, systématiciens, botanistes, historiens des sciences, etc..., expriment des besoins spécifiques qu'il faudra nécessairement prendre en compte. Ces derniers ont des implications en matière d'indexation, d'interface de consultation, d'accès à distance, de services de reproduction et de caractéristiques techniques de

¹⁵ Delcourt Thierry. Les publics. In Conduire un projet de numérisation. 2002. p. 15 à 31.

¹⁶ Lupovici Catherine, Cloarec Thierry et de Charentenay France. Les usages de Gallica. 2003. p. 40 à 44.

¹⁷ Collard Claude, Giannattasio Isabelle et Melot Michel. Les Images dans les bibliothèques. 1995. p. 91.

dématérialisation. Viser ce type d'usagers induit aussi, que l'on tende, si possible, vers une numérisation exhaustive des collections iconographiques. A l'Université, on notera l'existence d'une formation doctorale en histoire des sciences avec le DEA Construction des Savoirs scientifiques : didactique, histoire et épistémologie. On signalera, en outre, l'existence d'une demande préalable des chercheurs de l'Institut de Botanique relative à la numérisation d'une dizaine de documents patrimoniaux imprimés dont ils font grand usage.

6.2. Un public élargi :

Par ce terme, on entend ici le public étudiant, et le grand public. Toucher ce dernier dans le contexte du SCD revient immédiatement à proposer un accès grâce à l'Internet ou à participer à des manifestations culturelles, expositions temporaires ou permanentes comme La Comédie du Livre ou le projet MuseUM. Cela demande une autre médiation et le développement d'un outil de communication adapté faisant une place plus large au contexte, aide en ligne, commentaires, liens vers d'autres ressources patrimoniales et scientifiques, etc... Le public étudiant dispose d'un accès plus évident à la bibliothèque et la solution d'une diffusion intranet suffirait. Ce public est essentiellement composé d'étudiants du 1^{er} et 2^{ème} cycle de l'UFR des Sciences et d'étudiants inscrits à l'IUFM. Les sciences de la vie représentent de loin le premier effectif des inscrits avec 1112 inscrits sur 2689 inscrits à l'UFR en 1^{er} cycle. Les effectifs en 2^{ème} cycle et les volumes horaires d'enseignements font apparaître la même tendance¹⁸. Cette caractéristique de l'Université de Montpellier 2 se retrouve aussi au niveau des laboratoires de recherche où les sciences de la vie sont très bien représentées. La botanique et la systématique sont, dans ce domaine, des disciplines fondamentales au premier sens du terme. L'existence d'un fonds numérisé relatif à ces dernières permet d'envisager des usages pédagogiques, mais aussi des usages culturels individuels. Par ailleurs, Il existe une longue tradition d'enseignement de l'histoire des sciences en 1^{er} et 2^{ème} cycle à l'Université de Montpellier 2 et Montpellier 3 qui rencontre un succès évident auprès des étudiants puisqu'un tiers

¹⁸ L'annexe n°5 reprend l'ensemble des statistiques de l'UFR en 2004.

des étudiants biologistes en DEUG étaient inscrits dans cet enseignement optionnel¹⁹.

Viser des publics aussi hétérogènes du point de vue des attentes ne va pas sans soulever des difficultés. Pour les résoudre, on peut dans une première approche considérer que les chercheurs et les étudiants bénéficient d'une même solution de communication dont la consultation s'effectuerait sur deux niveaux comme la recherche « simple » et « avancée » tandis que les actions de communication en direction du grand public prendraient une forme différente, suivant les opportunités, à partir des enregistrements de référence, leur utilisation, en tant que matière première, pouvant répondre à de multiples fins.

III° Les références extérieures.

1. La bibliothèque centrale du Muséum d'histoire naturelle.

1.1. La bibliothèque centrale

La bibliothèque centrale du Muséum (BCM) a été créée par le décret de la Convention du 10 juin 1793 qui transforma le Jardin royal des Plantes médicinales en Muséum d'Histoire Naturelle. Spécialisée dans le domaine des sciences naturelles, elle appartient au réseau documentaire du Muséum qui comprend 27 bibliothèques. Elle est, de plus, Centre d'Acquisition et de Diffusion de l'Information Scientifique et Technique pour la botanique, la zoologie et la biologie végétale et animale depuis 1982. Bibliothèque scientifique, patrimoniale et publique, elle dispose de collections d'environ 300 000 ouvrages dont plus du tiers, 105 000 environ, sont des imprimés anciens. Par ailleurs, elle abrite 7000 manuscrits et la célèbre collection des vélins du Muséum, au nombre de 6984,

¹⁹ Lecourt Dominique. L'enseignement de la philosophie des sciences. 1999. p. 90.

initiée par Gaston d'Orléans ainsi qu'environ 800 œuvres d'art. On recense de plus quelque 12 000 titres de périodiques. La photothèque rassemble quelque 23 000 images fixes. Les collections sont pour l'essentiel en accès indirect ; la collection des vélins du Muséum n'est pas communicable. Le vidéodisque des vélins du Muséum, un temps disponible à la médiathèque, ne l'est plus aujourd'hui, pour des raisons techniques. La salle de lecture offre des postes de consultation basés sur le logiciel AIE de la société Archimed. Le système intégré de gestion de bibliothèque (SIGB) de la BCM est Horizon de la société Dynix.

Dans le cadre du projet Grand Chantier des Collections, la BCM participe à la valorisation des collections patrimoniales du Muséum avec la restauration, le microfilmage et la numérisation des documents patrimoniaux imprimés et iconographiques en sa possession. Une synthèse des besoins et des coûts concernant ces opérations a été publiée par Valérie Néouze²⁰.

1.2. La photothèque et les travaux de reproductions argentiques et numériques

Le service de la photothèque de la bibliothèque centrale du muséum dépend du service du patrimoine. Le personnel de la photothèque est composé d'un bibliothécaire, de deux photographes et de deux vacataires. Sa principale mission est la commercialisation des reproductions argentiques et numériques du patrimoine iconographique et imprimé de la bibliothèque auprès d'une clientèle de particuliers et de professionnels. Elle entreprend, en outre, des campagnes préventives de sauvegarde d'ouvrages anciens et de manuscrits. Ces reproductions représentent des planches d'imprimés du fonds patrimonial dans les domaines de la botanique, de la zoologie, de la paléontologie, des voyages exploratoires du XIXème siècle, mais aussi des portraits de scientifiques du Muséum, des vues historiques du Jardin des Plantes et du Muséum, ainsi que des œuvres d'art. La recherche documentaire des images fixes s'effectue grâce à des fichiers papier, les bases numériques dont il sera question plus loin n'étant, pour l'instant, pas à la disposition du public. Les fichiers papier sont de trois types : fichier classé par cotes des documents originaux ; fichiers systématiques classés par noms

²⁰ Néouze Valérie. Quelle bibliothèque numérique pour une bibliothèque patrimoniale ? 2000. p.XIV

vernaculaires de plantes et d'animaux ; fichiers matières traitant les autres thèmes, les portraits et les œuvres d'art.

Les travaux photographiques, argentiques et numériques, se font sur commande. Tout usage public des reproductions fait l'objet d'une autorisation préalable de la directrice de la bibliothèque et nécessite, en sus, l'acquiescement d'une redevance d'utilisation. Par ailleurs, si le document n'est pas tombé dans le domaine public, l'utilisateur doit obtenir l'autorisation de l'auteur ou de ses ayants droits. Le service photographique n'exécutera les travaux que sur présentation de cette autorisation.

L'équipement du service de la photothèque est le suivant : un appareil reflex 24x36 Nikon F3 équipé d'un verre de visée quadrillé et d'un objectif macro AIS F2.8 de 55mm ; une chambre Sinar 4x5 inch muni d'un dos numérique Phase One ; un scanner 24x36 LS 4000 Nikon ; un scanner à plat Mikrotek; un scanner à tête zénithale Minolta; deux stations de travail Apple munies d'écran de 21 pouces ; une imprimante couleur Epson 2100 aux encres pigmentaires spéciales ; un serveur Dell sur lequel sont stockées les bases de données d'images numérisées ;

L'inventaire du fonds iconographique en 2004 fait apparaître que le service dispose de 29 654 images fixes²¹. Depuis la fin de l'année 2002, la photothèque s'est équipée de matériel numérique. Des campagnes de numérisation avaient cependant été réalisées depuis 1998 par des prestataires extérieurs. La création de ces reproductions numériques vient alimenter les trois bases de données de la bibliothèque que sont Botanica, Zoologica et Histoire du Muséum. Pour Botanica, 5500 images ont été numérisées à la BCM et 2000 au service des cultures du Muséum ; deux lots supplémentaires de 645 images sont en cours de réalisation. Pour Zoologica, 3391 images ont été numérisées en partenariat avec la Bibliothèque nationale de France ; 13 lots de 300 images ont été gravés sur cédérom et sont en cours d'intégration et d'indexation dans la base. Pour Histoire du Muséum, un premier lot de 304 images sur la Ménagerie a été gravé ; il est en cours d'intégration et d'indexation. Après dématérialisation par le laboratoire photographique, les images sont stockées sur CDRW avant leur transfert sur le serveur Dell pour les opérations de traitement documentaire. Le cédérom

« Ménagerie-01 », par exemple, contient 304 fichiers au format JPEG qualité maximum. La taille des fichiers est comprise entre 2500x1900 et 4200x3200 pixels et pèsent entre 700 Ko et 4 Mo. Chaque pixel est codé soit sur 24 bits en RVB pour la couleur, soit sur 8 bits en niveaux de gris pour le noir et blanc.

Lors de la mise en chantier des campagnes de numérisation, le SIGB Horizon ne disposant pas d'un module spécialisé dans le traitement bibliothéconomique des images fixes numérisées, le service de la photothèque s'est orienté vers l'achat du SIGB Taurus+ de la société Ever interfacé avec le module Gesco-Média permettant de constituer une base de données d'images fixes numérisées. La version installée de Taurus+ est la version 3.6.0. Gesco-Média permet l'organisation, l'indexation, la consultation, l'acquisition et l'impression des données. L'indexation des images numérisées a été confiée à deux vacataires doctorants, l'un, en zoologie, pour Zoologica, et l'autre, en botanique, pour Botanica. Les champs indexés renseignent précisément sur le document comme le montre l'annexe n°6. Pour la recherche, Gesco-Média permet d'utiliser les opérateurs booléens et les comparateurs ; Gesco-Média supporte la troncature qui remplace un ou plusieurs caractères. La recherche s'effectue selon plusieurs modalités qui sont décrites dans l'annexe n°6 : elle peut être globale, guidée par rubrique, par équation sur une rubrique ou séquentielle.

2. CNUM, la bibliothèque numérique du CNAM.

<http://cnum.cnam.fr>

Le CNUM est un projet démarré en 1998, sur budget propre, avec trois services du CNAM, la bibliothèque centrale, le centre d'histoire des techniques et le centre d'études et de recherches en informatique²². Il s'est développé avec rigueur à l'aide d'une méthode et d'une organisation du travail que l'annexe n°7 reproduit, et qui mettent en évidence les dimensions scientifique et patrimoniale du projet. En particulier, on notera les difficultés rencontrées en terme de pagination

²¹ Dont 5988 négatifs noir et blanc au format 9x12, 183 négatifs noir et blanc au format 13x18, 23 666 diapositives couleur au format 24x36, 76 ektachromes couleur au format 6x7, 1451 ektachromes couleur au format 10x12,5 et, 285 ektachromes couleur au format 13x18.

du fait du caractère ancien des éditions et les solutions pour y remédier, ainsi que le soin apporté à la sélection des ouvrages grâce à la constitution d'une instance de validation scientifique. Les fonds numérisés constituent un corpus francophone extrait des collections du CNAM sur l'histoire des techniques dans toutes ses dimensions : constitution des savoirs ; histoire des institutions ; médiation vers le public. Les ouvrages sélectionnés l'ont aussi été en fonction des formats (12°, 8° et 4°). Ils sont libres de droits pour la plus part.

La dématérialisation est effectuée par des prestataires extérieurs ; les premiers lots l'on été en mode bitonal à 400 dpi, puis de 300 à 400 dpi en niveaux de gris. Les formats des enregistrements de référence sont TIFF Gr 4 pour le bitonal, JPEG ou PNG pour les niveaux de gris. L'opération de dégradation des fichiers ramène les enregistrements de communication à une résolution de 100 à 133 dpi. Une opération de conversion pour les fac-similés de textes et de planches gravées est réalisée au format PNG ou GIF. Le système de nommage des fichiers permet de s'assurer d'un stockage et d'un archivage rigoureux en conformité avec la nature de la pagination.

Pour la communication, un site Web est proposé, comprenant une rubrique de présentation, une rubrique d'aide, une rubrique outils, un moteur de recherche et un catalogue alphabétique des documents numérisés. L'accès se fait grâce au catalogue qui renvoie sur une page où figure(nt) la(les) notice(s) et table(s) du document, soit à l'aide du moteur de recherche. La consultation séquentielle à l'intérieur du document s'effectue à l'aide d'une barre d'outils.

Le traitement documentaire comprend l'établissement des notices bibliographiques et l'indexation RAMEAU. L'établissement des tables des matières et d'illustrations est effectué par le prestataire chargé de la dématérialisation.

3. NordNum et Libris : deux projets numériques autour de l'Université de Lille 3

Ces deux projets sont indépendants mais présentent comme partenaire commun le Centre de recherches sur l'histoire de l'Europe du Nord-Ouest

²² Cubaud Pierre, Deblock Geneviève et Radix Marie-Christine. Le conservatoire numérique des arts et métiers :

(CRHEN-O) de l'Université de Lille 3. Ils sont, de plus, par bien des aspects complémentaires. L'un, NordNum, est une bibliothèque numérique d'histoire régionale, l'autre, Libris, est une base de données multimédia du patrimoine régional.

3.1. Libris

<http://libris.univ-lille3.fr>

Trois partenaires sont à l'origine de la base de données multimédia Libris, soutenue par le Pôle Universitaire européen Lille-Nord-Pas-de-Calais: le CRHEN-O, la bibliothèque municipale de Lille et la bibliothèque centrale de l'Université catholique de Lille²³.

Les publics visés par le projet sont les enseignants, les chercheurs et le grand public. Les fonds numérisés, quelque 6000 documents, sont composés d'iconographies anciennes, photographies, illustrations de livres anciens, cartes géographiques et cartes postales. Ils émanent de fonds publics et de fonds privés. Pour ces derniers, le mode d'acquisition est original, en faisant appel aux particuliers, collectionneurs, etc. Des prêts de documents sont consentis par les particuliers pour être dématérialisés assorties d'une autorisation de diffusion sur la base de données à titre gracieux. Toute autre exploitation nécessite une nouvelle autorisation.

La dématérialisation est effectuée en interne à l'aide d'un scanner à plat et d'un scanner Leaf Lumina ; les logiciels utilisés sont Photoshop, Illustrator et Acrobat.

Le traitement documentaire est réalisé sur le logiciel Phrasea Suite. La structuration des données est fonction des besoins avec la définition des champs obligatoires et indexés. Le niveau de description est souhaité le plus fin possible. L'indexation suit les besoins ressentis à partir d'adaptations de divers thésaurus. L'outil de consultation offre la possibilité d'interroger en texte intégral sur l'ensemble des champs.

La communication est basée sur une diffusion internet grâce à Phrasea Server (TCP/IP) hébergé par l'Université de Lille 3. L'application Phrasea Client permet

historique du projet et organisation du site. In La numérisation des textes et des images... 2002. p.13-22.

²³ Aubry Martine. Libris : images régionales et recherche. In La numérisation des textes et des images... 2002. p.149-162.

la consultation en local, tandis que Phrasea e-Publisher autorise la publication sur le Web des bases ouvertes par Phrasea Server. Dans ce dispositif, Phrasea Viewer, téléchargeable gratuitement, assure la consultation à distance.

3.2. NordNum

<http://nordmun.univ-lille3.fr>

Les partenaires du projet sont des composantes du SCD : la bibliothèque du CRHEN-O et la bibliothèque centrale du SCD de Lille ²⁴. NordNum est financé par le contrat quadriennal de l'Université. Le public cible se compose de chercheurs, enseignants, étudiants et grand public. Les fonds publics numérisés rassemblent un corpus d'ouvrages du XIX^{ème} siècle, monographies d'histoire régionale et ouvrages techniques relatifs à l'histoire économique régionale.

La dématérialisation est réalisée en externe à l'Atelier National de Reproduction des Thèses de Lille. En mode image à 600 dpi au format TIFF Gr4 pour les imprimés, au format JPEG pour les illustrations en couleur. Le nommage des fichiers suit la concaténation du code RBCCN, du numéro d'inventaire, du numéro de volume et du numéro séquentiel du fichier, suivie de l'extension qui indique le format. Le rythme des opérations est fixé à 100 ouvrages par an.

Le traitement documentaire est effectué à l'aide du SIGB Horizon et de la plate-forme XML SDX. Le signalement est réalisé dans le catalogue du SCD Lille 3 et le SUDOC en exploitant la zone 856 de la notice ; sur la plate-forme SDX, il est obtenu en convertissant les notices USMARC en UNIMARC, puis vers BiblioML. Une saisie des tables des matières et des index en XML (DTD BiblioML) est réalisée par des moniteurs-étudiants. Les champs d'indexation des documents XML sont les titres principaux et secondaires ; les auteurs ; les dates de publication ; les sujets ; le texte des tables et index.

La communication des documents XML s'appuie sur la plate-forme SDX. L'outil de recherche propre à NordNum a été développé par la société AJLSM. Les formats de diffusion sont, selon la configuration du poste client, TIFF, JPEG, PNG, PDF. L'interface de consultation permet une interrogation à partir des champs indexés et des opérateurs booléens ET et OU, avec deux modes de

²⁴ Sévigny Martin. NordNum : rapport technique. In La numérisation des textes et des images... 2002. p.27- 46.

recherche, simple ou avancée. Une autre interrogation est possible à partir des listes de titres, auteurs, sujets. L'affichage des résultats est obtenu sous la forme d'une notice brève à partir de laquelle, en cliquant sur le titre, on consulte l'ouvrage numérisé. La navigation à l'intérieur de ce dernier peut s'effectuer page par page ou grâce aux tables et index. La possibilité de télécharger tout ou partie de l'ouvrage en PDF est offerte.

4. Synthèse

On peut constater à travers ces références extérieures :

- Qu'en terme de dématérialisation, la production en interne nécessite une équipe formée et compétente ; que les outils sont de haut niveau (chambre et dos numérique);
- Que différentes approches sont possibles en matière de communication et qu'en particulier, les solutions retenues font la part belle soit à l'image soit à l'imprimé ;
- Les difficultés rencontrées avec la pagination des ouvrages anciens ;
- La rigueur de la démarche en terme d'organisation avec notamment les règles de nommage des fichiers ;
- Le choix de standards techniques.

IV° Le cycle de numérisation : spécifications techniques et organisation générale.

Les spécifications techniques dont il sera question plus loin représentent la matière d'un futur cahier des charges. Dans ce cadre, nous les détaillerons afin de justifier les choix.

1. La préparation à la numérisation.

1.1. La collection des vélins.

Il n'existe pas à notre connaissance un inventaire officiel des vélins. Dans ce cas, il faudra procéder à l'inventaire complet de la collection. La table des illustrations, succincte et en mauvais état, présente au début des classeurs pourra servir de base de départ. Il comprendra entre autres : les dimensions physiques, la nature du support et son état, le numéro du classeur, le relevé des noms des familles de la classification, le numéro de la classification actuelle, le relevé des légendes, le relevé des mentions manuscrites au recto et au verso, la date au verso lorsqu'elle existe, et la signature ou l'absence de signature.

Un examen physique de la collection permettra de déterminer les vélins qui doivent être restaurés. Une expertise semble nécessaire qui permettra de déterminer finement le budget nécessaire de conservation. Conjointement, il pourrait être fait appel aux prestataires de service de numérisation (s'il y a lieu) afin d'évaluer l'état minimal requis pour une numérisation de qualité. Une campagne de restauration des vélins sera alors programmée.

1.2. Le fonds des documents photographiques.

Ces collections présentent plus de difficultés. Les documents sont quasiment inconnus et n'ont jamais fait l'objet d'une étude des personnels de la

bibliothèque. Un document manuscrit décrit incomplètement l'inventaire des plaques photographiques avec les dates, la légende et parfois le nom du photographe d'une série. Pour les autres types aucun document d'accompagnement ne semble disponible. Claude Collard propose une méthode de prise en main d'une collection photographique patrimoniale que l'on pourrait ici appliquer²⁵. On y ajoutera, si les conditions le permettent, une action de conservation en procédant au nettoyage des plaques de verre selon les méthodes préconisées par la DLL : nettoyage du coté verre avec de l'alcool éthylique appliqué rapidement avec un chiffon non pelucheux ; essuyage de l'émulsion avec une brosse douce en poils de martre²⁶. Le programme se déclinerait selon les opérations suivantes :

Nettoyer. Inventorier et répertorier. Sélectionner et vérifier les légendes. Classifier en respectant le classement s'il existe ou en créant de grandes rubriques. Etablir un catalogue pour la consultation. A partir de là, il sera possible de mener des investigations relatives aux droits d'auteurs et aux droits à l'image, d'examiner les restaurations éventuelles, puis de procéder à la campagne de dématérialisation.

1.3. Le traitement documentaire des collections iconographiques.

Le catalogage s'effectue dans le SUDOC et ALEPH. Cette opération est particulièrement importante dans la mesure où elle constitue un moyen de signalement des collections iconographiques et la base à partir de laquelle seront produites les notices bibliographiques des documents numériques. La description des documents iconographiques devrait suivre la norme AFNOR Z 44-077 : règles pour la rédaction de la description bibliographique des images fixes, en deux dimensions et sur support mobile). La norme Z 44-077 ne traite ni la description iconographique, ni les vedettes pour lesquelles on fera alors appel aux normes suivantes Norme Z 44-059 : Catalogage – Choix des accès à la description bibliographique ;

Norme Z 44-060 : Catalogues d'auteurs et d'anonymes -- ...vedettes de collectivités-auteurs ;

Norme Z 44-061 : Catalogage--...vedettes noms de personnes... ;

²⁵ Collard Claude, Giannattasio Isabelle et Melot Michel. Les Images dans les bibliothèques. 1995. p. 160.

²⁶ Ministère de la culture et de la communication. Protection et mise en valeur du patrimoine des bibliothèques. 1998. p.73.

Norme Z 44-070 : Indexation analytique par matière²⁷.

L'indexation des documents suivra le vocabulaire contrôlé RAMEAU. Ce choix est déterminé par sa simplicité de mise en œuvre, mais il n'est pas définitif. L'utilisation d'un thésaurus spécialisé pourrait être mise à l'étude ; c'est une problématique difficile qui demande cependant une forte mobilisation des moyens humains et dont l'efficacité demeure incertaine. Elle s'articule avec l'indexation et la production de métadonnées des reproductions numériques, et donc indirectement avec les solutions techniques du dispositif de communication. Certains produits disponibles sur le marché disposent d'une fonction de recherche « plein texte ».

D'un point de vue pratique, cette opération soulève deux problèmes : celui de la qualification particulière du personnel de la bibliothèque en matière de catalogage des images fixes auquel on répondrait par une action de formation. Celui, plus technique, du niveau de catalogage. Prenons la collection des vélins pour exemple. Doit-on procéder au catalogage des documents au niveau du document ou au niveau de la collection ? Autrement dit doit-on établir : une « notice mère » dans laquelle il faudra mentionner un titre (lequel ?) au niveau de la zone du titre et des responsabilités, quitte à dresser un inventaire plus détaillé dans une note de dépouillement ; une notice par document au risque de diluer les documents dans les catalogues et de produire du bruit à la recherche ? Ce questionnement n'est pas non plus indépendant des solutions techniques de communication des reproductions numériques et des moyens humains disponibles pour ce traitement. Les réponses ne peuvent, dès lors, être apportées qu'en ayant une vision plus globale du projet. En anticipant quelque peu, on soulignera, ici, qu'un choix est possible entre un traitement archivistique des collections iconographiques, en effectuant directement la saisie de l'inventaire dans un éditeur XML suivant la définition type de document de la description archivistique encodée (EAD), et un traitement bibliographique traditionnel.

1.4. La préparation des ouvrages imprimés.

L'inventaire des collections patrimoniales imprimées en cours nécessite 12 mois pour arriver à son terme. Il permettra de choisir le corpus à numériser. La

²⁷Collard Claude, Giannattasio Isabelle et Melot Michel. Les images dans les bibliothèques. 1995. p. 210.

sélection des ouvrages effectuée, on consignera l'état des documents et on engagera une vérification des notices en terme d'existence et de complétude dans le catalogue de la BIU. S'il s'avère que la notice est absente, on se reportera aux fichiers papier du fonds ancien et on procédera au catalogage dans le SUDOC, en application des normes en vigueur de catalogage des livres anciens. La responsabilité de cette opération devrait échoir au service de patrimoine de la BU dont le personnel est compétent en matière de catalogage des livres anciens. On pourra alors envisager la constitution des lots et la campagne de dématérialisation.

1.5. La constitution des lots de numérisation.

Une campagne de dématérialisation réussie suppose un niveau minimum de formalisation dans l'organisation des flux entrants et sortants, afin de garantir son efficacité, et la sécurité des documents. Sa programmation doit prendre en compte les caractéristiques propres aux originaux et aux reproductions numériques. La constitution de lots homogènes répond à ces objectifs. En cas d'externalisation, la taille des lots devra être déterminée en relation avec le prestataire. En production interne, il devra l'être avec le responsable du service. Un conditionnement adapté aux lots sera requis pour le transport. Ce dernier nécessitera que le prestataire (s'il y a lieu) soit assuré. Chaque lot sera identifié. Les informations suivantes devront figurées sur la fiche d'identification du lot : nom de la bibliothèque expéditrice ; numéro unique identifiant le lot ; date d'expédition ; nombre de documents contenus dans le lot ; liste des documents et de leurs caractéristiques²⁸.

2. La dématérialisation.

2.1. La dématérialisation des vélins

2.1.1 Analyse du document

La valeur et l'état des documents impliquent une numérisation sans contact²⁹ : intermédiation photographique argentique et scanner de films, scanner couleur à tête zénithale (rare et très cher) ou appareil photo numérique. Avec le

²⁸Sitruk Alfred. Cahier des charges de numérisation. In Conduire un projet de numérisation. 2002. p. 277

moins de manipulations possibles, l'idéal étant un processus de numérisation et un seul ; ce qui implique que les fichiers numériques obtenus puissent répondre ultérieurement aux plus grands nombre d'applications possibles.

La qualité des vélin nécessite que l'on numérise en couleur haute résolution :

- par respect pour l'œuvre et le travail du miniaturiste.
- pour disposer d'enregistrements de référence permettant au maximum une sortie offset au format A3 en qualité édition, la norme est au minimum de 300 dpi (quasi standard et recommandations multiples). Au demeurant, l'encodage doit s'effectuer au minimum sur 24 bits (3x8 bits en RVB). Le couple résolution de 300 dpi / encodage en 24 bits apparaît dans la littérature comme une norme de fait en matière de numérisation des documents iconographiques patrimoniaux. Ce niveau d'exigence correspond à l'hypothèse la plus haute que l'on puisse soutenir car, compte tenu des dimensions des originaux, le volume théorique des fichiers numériques s'accroît en fonction du carré de la résolution.

2.1.2 Mode d'acquisition

Il existe deux modes d'acquisition, le mode image et le mode texte. Les documents iconographiques nécessitent évidemment une numérisation en mode image. Il reste cependant à choisir le niveau de codage. Chaque pixel est codé selon un certain nombre de bits. Plus ce dernier est élevé, plus l'image numérique restituera de nuances de couleur et plus la reproduction sera fidèle à l'original. En couleur, chaque pixel est codé sur 8 bits minimum par couleur. En RVB, le codage s'effectue donc sur 3x8 bits, soit 24 bits. Ce mode d'acquisition couleur en 24 bits minimum apparaît dans les programmes de numérisation en couleur comme une norme de fait. Il est aujourd'hui possible de codé chaque pixel sur plus de 8 bits par couleur ; ce qui permet alors de disposer d'un codage de la couleur sur 48 bits. Il faut cependant être prudent sur ces possibilités car si certains fabricants annoncent un tel niveau d'échantillonnage, il demeure que les fichiers obtenus ne

²⁹ Stuart D Lee. Digital imaging a practical handbook. 2001. p. 51

sont parfois codés qu'en 24 bits³⁰. De plus, plus on augmente le nombre de bits de codage plus le fichier numérique obtenu est volumineux.

2.1.3 Choix de la résolution

2.1.3.1 Acquisition numérique directe

Si on veut une résolution de restitution au moins égale à 300 dpi en A3, compte tenu que les dimensions de l'original sont plus grandes que les dimensions de l'image, il suffit de numériser à une résolution d'acquisition de 300 dpi ; il est alors nécessaire de disposer d'un outil d'acquisition susceptible de délivrer un fichier de 5499x3744 pixels minimum, qu'il s'agisse d'un scanner à tête de numérisation fixe ou d'un dos numérique. Cette valeur est significative à un point près : dès lors que les rapports d'homothétie des formats de l'original et du capteur ne sont pas égaux, il conviendrait de raisonner en terme de pixels utiles puisqu'une partie du capteur au rapport 1 :1 enregistre des informations inutiles (vide latéral). Il faut donc dans la pratique une taille supérieure de capteur numérique.

A l'heure actuelle, en photographie numérique, seuls les dos numériques grand format permettent d'obtenir ce niveau de définition. Le capteur pour dos numériques moyen format le plus performant à la fin 2003 ne pouvait délivrer une taille d'image supérieure à 5440x4080 pixels. Certains fabricants, Phase One et Dicomed, proposent des dos numériques grand format (de type à balayage) pour chambre photographique, couleur 24 bits ou plus, d'une définition de 6000x8000 pixels³¹. Certains modèles comme le PowerPhase FX de Phase One, 48 bits, permettent une résolution d'acquisition de 300 dpi pour un document A0 (84,1x118,9 cm)³². Ces outils d'acquisition numérique, dont l'avantage est leur grande universalité, ont pour inconvénient d'être chers ; il faut, de plus, une chambre photographique et un objectif adapté aux travaux de reproduction, ainsi qu'une station graphique sur laquelle les images sont enregistrées directement, sans compter les accessoires indispensables de prises de vue : photomètre, éclairage artificiel et système de support du document. Il demeurerait cependant possible d'accéder à moindre coût à ce type de matériel grâce à la location, mais il

³⁰ Niemetzky Gérard. Gestion de la couleur Calibration & profils ICC pour l'image numérique et la chaîne graphique. 2002. p.56.

³¹ Jacquesson Alain et Rivier Alexis. Bibliothèques et documents numériques. 1999. p.81.

nécessite aussi un opérateur formé compétent. En cas d'externalisation, ce type de matériel peut être considéré comme le premier choix. Quant aux scanners spécialisés, très onéreux aussi, ils se révèlent moins universels et l'investissement apparaît déraisonnable eu égard au volume de documents à numériser. Ils peuvent cependant être aussi un critère de choix en cas d'externalisation chez un prestataire de service.

2.1.3.2 Intermédiation argentique

La technique de l'intermédiation argentique permet de disposer aussi d'un minimum de 300 dpi en résolution de restitution sur la reproduction argentique³³. Cette technique est une solution non dépourvue d'intérêts encore de nos jours (avantages : support de conservation connu et qui a fait ses preuves ; souplesse du processus de numérisation ; sécurité des originaux ; coûts; préférence encore actuelle des documents de substitution argentiques chez les professionnels de l'édition ; moindre technicité des processus de production). L'intermédiation argentique implique cependant la gestion d'un support supplémentaire.

Pour obtenir une résolution de restitution de 300 dpi en format A3 (30x40), il faut à partir d'un original de 10x12cm (4x5inch) une résolution d'acquisition de 900 dpi. $RA=RR \times I/O$ ou I/O est de coefficient d'agrandissement (ici égal à 3 environ), c'est-à-dire le rapport entre la taille physique de l'image en sortie et la taille physique du document numérisé³⁴. Pour un document au format 6x7, il faut pourvoir numériser à 1500 dpi ($I/O=5$) pour obtenir un A3 à 300 dpi. Pour obtenir, avec ce même format 6x7, un tirage au format A4 à 300 dpi, ($I/O=3$), la résolution d'acquisition est de 900 dpi. Cependant, au-delà de la résolution d'acquisition, la qualité de l'image numérique dépendra aussi de la qualité du document intermédiaire. A ce titre, le choix du film et la compétence du photographe en matière de reproduction sont primordiaux.

En théorie, toute reproduction sur un support analogique représente une perte d'information due aux imperfections optiques, mécaniques et physico-chimiques. En outre, à définition égale de film, cette perte d'information est fonction du

³³ Stuart D Lee. Digital imaging a practical handbook. 2001. p.60.
Id. p.64.

³⁴ Fernandez Edmond. Les techniques de numérisation In La numérisation des textes et des images. 2002. p.92. RA est la résolution d'acquisition ; RR, la résolution de restitution.

rapport entre le format de l'original et la surface du film. En pratique, il faut être plus prudent. Les tests effectués à la BNF montrent qu'au final les résultats peuvent être meilleurs que la numérisation directe selon le format de l'original et la nature des supports. Schématiquement, les résultats seraient meilleurs pour les documents originaux de taille moyenne au plus (A3 max), avec du film noir et blanc plutôt qu'avec de l'ektachrome³⁵.

Pour dématérialiser ces phototypes, on peut utiliser un scanner de film ou un scanner à plat de qualité photo. Suivant les modèles, la vitesse du scanning, la densité maximale et les spécifications de la connectique sont variables. La densité maximale est une norme destinée à mesurer l'étendue maximale de densité du film que le scanner peut restituer ; plus elle est élevée, meilleur est l'outil. Les normes connectiques se traduisent par des taux de transfert différents ; le Firewire IEEE 1394 est plus rapide que l'USB 2.0.

Pour exemples, le modèle Epson P4870 Photo est un scanner à plat de qualité photo dont la résolution atteint 4800x9600 dpi sur 48 bits pour une surface de film de 14,4x23,1 cm, avec une Dmax de 3,8, disposant des ports USB2.0 et Firewire IEEE 1394 ; le modèle Nikon Coolscan 9000 ED est un scanner de film moyen format (6x9 max.) numérisant à 4000 dpi sur 48 bits, avec une Dmax 4,8, et des ports Firewire IEEE 1394 et USB2.0.

2.1.4 Le format des enregistrements de référence.

Il existe une multitude de formats image, 70 d'après la DLL³⁶. Selon le point de vue adopté, il est possible d'établir des typologies : format propriétaire versus format libre de droit, format non compressé versus format compressé, format norme officielle versus format norme de fait, etc.

La Sous Direction des Bibliothèques (SDB), pour qui *le format constitue un élément clé pour la pérennité du document numérisé car il conditionne sa lisibilité à moyen et long terme*, préconise de limiter le *choix des formats à ceux qui*

³⁵ Duchemin Pierre-Yves. La numérisation des documents graphiques : enjeux techniques et financiers. 1999. p.12.

³⁶ Ministère de la culture et de la communication. Protection et mise en valeur du patrimoine des bibliothèques. 1998. p. 110.

présentent le caractère de stabilité et de normalisation nécessaires –préférer un produit standard du marché.³⁷

Des projets français ou étrangers de numérisation on peut voir émerger deux formats particulièrement utilisés pour les documents iconographiques : le TIFF (Tagged Image File Format) et le JPEG (Joint Photographic Expert Group). Ces formats ne s'excluent pas l'un et l'autre, mais sont complémentaires. Ils répondent à des besoins différents d'un projet de numérisation. Le TIFF est le format des enregistrements de référence ; il est non-destructeur d'informations ; il peut être compressé ou non ; il apparaît comme une norme de fait ; il permet d'échanger des fichiers images entre différentes plates-formes ; il est recommandé par la SDB ; il correspond aux besoins de conservation. Mais il est volumineux. Le JPEG est le format de diffusion sur les réseaux : il est le plus répandu ; il est fortement compressible ; il est une norme officielle ISO et AFNOR. Mais il est destructeur d'informations. Les fichiers d'images numériques peuvent être enregistrés en TIFF puis convertis en JPEG.

Les fichiers d'images numériques peuvent comporter d'autres types d'informations dits « formats dans les fichiers³⁸ », tels les dimensions de l'original, le modèle de couleur, les profils ICC (International Color Consortium) et les métadonnées EXIF et IPTC. Pour le modèle de couleur, il s'agit de distinguer le CMJN (impression) du RVB (écran). Pour les métadonnées EXIF, il s'agit des métadonnées techniques relatives au mode d'acquisition. Pour les métadonnées IPTC, il s'agit des métadonnées en relation avec la sémantique de l'image, tels légendes, mots-clés, etc... Le dernier élément, les profils ICC, fera l'objet du chapitre suivant. Certaines de ces métadonnées, au sens large, ont absolument besoin d'être connues en cas de sous-traitance : le modèle de couleur et le profil ICC de l'appareil de numérisation. Les métadonnées EXIF sont d'un intérêt moindre mais leur présence ne serait pas inutile, en terme d'informations pour la conservation et de contrôle qualité. De même une légende sommaire dans les champs IPTC permettrait d'identifier grossièrement le contenu iconographique.

³⁷ France. Sous direction des bibliothèques et de la documentation. A propos de la numérisation. Notions et conseils techniques élémentaires, p.9.
³⁸ Lopez Jean-Gabriel. Images numériques : quels formats, quelles caractéristiques, pour quels usages ? 2003.

2.1.5 Le profil ICC de l'outil d'acquisition numérique.

Dans une chaîne de traitement numérique couleur, il est fondamental de disposer d'outils de gestion de la couleur. Chaque périphérique interprète en fonction des réglages qui sont les siens (la calibration) les données colorimétriques. Afin de pouvoir assurer une interprétation homogène, il est nécessaire de calibrer et de caractériser chaque périphérique, appareil d'acquisition, écran et imprimante. En 1998, la DLL préconisait les recommandations suivantes : *Le prestataire doit pouvoir fournir une image test créée par la numérisation des mires de contrôle qui sont des phototypes obtenus en photographiant des éléments suivants :*

- *une gamme de gris allant de la densité 0,20 à la densité 2 par pas de 0,20 ;*
- *une gamme de couleurs comportant 10 plages uniformes répartis sur le spectre visible (Rouge, Vert, Bleu et mélanges) ;*
- *un quadrillage noir et blanc.*

La prise de vue de ces deux gammes et du quadrillage est faite simultanément de manière à être sur le même phototype. Si les valeurs retenues après un test de numérisation conviennent à la bibliothèque, elles seront considérées pour toute la suite de la prestation comme valeurs étalon³⁹.

Depuis, l'utilisation des profils ICC s'est répandue. Le profil ICC de chaque périphérique répond à cette obligation d'étalonnage des périphériques. Sans entrer dans les détails techniques, cette opération consiste à numériser une charte de couleur dont les valeurs sont connues, l'étalon (charte IT8 par exemple), et à les comparer aux valeurs restituées dans le fichier numérique, ce que « devrait voir » l'outil d'acquisition et ce qu'il « voit » réellement, ce qui permet de déterminer son profil ICC par la mesure des écarts. Ce dernier est ensuite attaché au fichier numérique de telle manière que les périphériques suivant de la chaîne numérique, un fois eux aussi calibrés et caractérisés, interprètent le plus fidèlement possible les valeurs du document primaire.

Dans le cas de la numérisation des vélins, quelle que soit la solution retenue, interne ou externe, il est indispensable de disposer du profil ICC de l'outil

³⁹ Ministère de la culture et de la communication. Protection et mise en valeur du patrimoine des bibliothèques. 1998. p.109

d'acquisition car il permettra d'accroître l'efficacité du travail en aval et participera à l'intégrité des données. Il ne dispense cependant pas de la fourniture de l'image test.

2.2. La dématérialisation des plaques de verre et des documents iconographiques annexes.

2.2.1 Mode d'acquisition.

S'agissant de documents iconographiques en noir et blanc, il est recommandé de numériser en mode image et en niveaux de gris sur 8 bits. Pour les tirages photographiques et les documents iconographiques imprimés, cela devrait suffire. Cependant, pour les plaques de verre, la DLL recommande de numériser en couleur RVB 24 bits, puis de les faire passer en niveaux de gris lors du traitement par un logiciel de retouche d'images⁴⁰.

2.2.2 Choix de la résolution

Dans le chapitre précédent, il a été établi que le choix de la résolution d'acquisition dépendait du produit du coefficient d'agrandissement par la résolution de restitution souhaitée. Partant d'un document original de 13x18 cm, la résolution d'acquisition minimum pour obtenir un fac-similé papier au format A3 est donc de 700 dpi ; elle est de 500 dpi pour obtenir un fac-similé papier au format A4. Pour les documents originaux au format 9x12, une résolution d'acquisition de 750 dpi et de 1050 dpi permettront d'obtenir respectivement un fac-similé au format A4 ou A3.

2.2.3 Le format des enregistrements de référence.

Les fichiers numériques de références seront enregistrés au format TIFF.

2.2.4 Numérisation directe ou intermédiation argentique.

Cette question est ici moins essentielle que pour les vélins dans la mesure où la taille des documents originaux est plus petite. La mise en œuvre de solutions numériques adaptées à ces formats est plus courante. Cependant, l'état de

⁴⁰Ministère de la culture et de la communication. Protection et mise en valeur du patrimoine des bibliothèques. 1998. p.110

conservation n'ayant pas été établi de manière définitive, il convient ici de poser la question d'une numérisation avec ou sans contact.

S'il advient qu'il faille numériser sans contact, les solutions techniques de dématérialisation des plaques de verre sont les mêmes que celles des vélins : appareils d'acquisition photo-numérique ; scanners à tête zénithale ; intermédiation argentique. Pour numériser directement un document au format 13x18 à 500 dpi ou de 700 dpi, il faut disposer en numérisation fixe d'un capteur pouvant générer respectivement un fichier de 2535x3510 pixels ou de 3550x4915 pixels; pour un document au format 9x12, la taille du fichier doit être de 2635x3510 pixels pour 750 dpi et de 3685x4915 pour 1050 dpi. La taille minimum envisagée ici implique l'utilisation d'un appareil photographique équipé d'un dos numérique moyen format.

S'il demeure possible de numériser avec contact, d'autres solutions techniques de dématérialisation sont possibles. Ainsi, Edmond Fernandez envisage la possibilité de dématérialiser des plaques de verre à l'aide d'un *scanner « bureautique »*⁴¹. Il décrit ici un scanner à plat, et l'usage de guillemets suppose qu'il s'agisse d'un scanner à plat de qualité, capable, selon nous, de délivrer la résolution appropriée et disposant d'une densité maximale (Dmax) élevée. Il faudrait, en outre, qu'il accepte les documents transparents d'un format au moins égal au 13x18 et que sa manipulation soit compatible sans dommage avec l'épaisseur des plaques. Pour les trois premiers critères, l'offre actuelle des constructeurs permet de répondre favorablement à un coût relativement faible. Le modèle P4870 photo d'Espon, par exemple, correspondrait. Par contre, il est nécessaire de poursuivre plus en avant les investigations quant à l'épaisseur compatible ou non des plaques.

2.3. La dématérialisation des ouvrages imprimés

Pour des raisons déjà évoquées, le choix du corpus imprimé n'a pas été arrêté. Les principes de dématérialisation suivants sont donc donnés à titre provisoire dans la mesure où chaque document, en fonction de son format, de sa typographie, de sa mise en page et de son état de conservation appelle des

⁴¹Fernandez Edmond. Les techniques de numérisation In La numérisation des textes et des images. 2002. p. 101

solutions appropriées. Dans les lignes qui suivent, nous établirons les principes généraux et les recommandations techniques des organismes de tutelle.

2.3.1 Mode d'acquisition.

Le premier choix possible en matière d'acquisition numérique de documents imprimés concerne le mode de codage du pixel. Il dépend en particulier de la présence ou non d'une iconographie. Pour les documents textuels, la SDB recommande de dématérialiser en mode bitonal⁴². Le codage du pixel répond seulement à deux valeurs, soit noir, soit blanc. Si l'ouvrage candidat à la numérisation comporte des éléments iconographiques, il sera possible de numériser ces derniers en niveaux de gris.

Le second choix possible s'effectue entre le mode image ou le mode texte. Le mode image restitue l'original tel qu'il est avec sa typographie et sa mise en page ; il s'agit alors d'un fac-similé. Il respecte l'apparence exacte du document, mais il n'est pas possible d'effectuer une recherche en texte intégral dans le document numérisé. Le mode texte lui permet ce type de recherche mais la typographie et la mise en page du document se trouvent fortement altérées. Le mode texte nécessite l'emploi d'un logiciel de reconnaissance de caractères (OCR) dont les résultats ne sont pas fiables à 100%. L'utilisation du mode texte oblige à un contrôle supplémentaire suivi de corrections manuelles. De plus, la DLL signale que *la reconnaissance optique de caractère n'est possible que sur certains imprimés*⁴³.

La numérisation en mode texte nécessite l'usage d'un standard de codage en machine des systèmes d'écriture, la table de caractères. Parmi ces standards de codage, on trouve l'ASCII et l'Unicode. Le premier permet de coder sur 8 bits (256 valeurs différentes) l'écriture latine. Il ne convient pas pour les documents textuels présentant des symboles et caractères spéciaux comme les textes anciens ou scientifiques. L'Unicode ouvre depuis 1991 d'autres possibilités dans la mesure où le codage s'effectue sur 16 bits (65536 valeurs différentes) permettant ainsi de coder des systèmes d'écriture non latine et en particulier les symboles

⁴² France. Sous direction des bibliothèques et de la documentation. A propos de la numérisation. Notions et conseils techniques élémentaires. p.1

⁴³ Ministère de la culture et de la communication. Protection et mise en valeur du patrimoine des bibliothèques. 1998. p.110

mathématiques et scientifiques⁴⁴. L'Unicode n'est cependant pas encore supporté par tous les logiciels et suppose de disposer des polices de caractères adéquates⁴⁵. Le mode texte ne permet pas a priori de respecter la présentation du document. Cependant, il existe des solutions techniques permettant de se rapprocher du mode image à partir d'une numérisation dans ce mode. Parmi ces dernières, on trouve en particulier le format PDF qui permet de conserver la structure physique. Grâce à Acrobat Capture de la société Adobe, il est permis de convertir en mode texte un fichier préalablement numérisé en mode image. Ce logiciel permet, en outre, *d'encapsuler des images et en particulier des conversions en format PDF de fichiers TIFF*⁴⁶. On trouve une application de cette solution sur Gallica. Pour pouvoir conserver la structure logique du document, il est nécessaire de mettre en œuvre un autre type de solutions techniques basées sur la norme SGML (ISO 8879, 1996) dont HTML et XML sont dérivés et constituent des DTD (définition de type de document) définies comme des syntaxes particulières de SGML. Nous n'entrerons pas ici dans les détails techniques relatifs à ces normes. Il importe néanmoins d'avoir présent à l'esprit ces possibilités car elles permettent de mettre à disposition des publics, en particulier des chercheurs et de leurs attentes, des outils particulièrement performants. Elles supposent cependant la mise en œuvre d'outils complexes nécessitant des compétences de haut niveau. Par ailleurs, ces solutions de conservation des structures physique et/ou logique ne sont pas sans conséquence sur le coût de revient de la mise à disposition sous forme électronique des documents imprimés. Nous y reviendrons plus loin lorsque nous évoquerons les différentes modalités de communication.

2.3.2 Choix de la résolution.

La bibliothèque de Cornell aux Etats-Unis, spécialiste des programmes de numérisation, a mis au point une méthode permettant de déterminer avec exactitude la résolution d'acquisition des documents imprimés en mode bitonal⁴⁷. Partant d'un indice de qualité représentatif du nombre de pixels pour représenter la

⁴⁴Jacquesson Alain et Rivier Alexis. Bibliothèques et documents numériques. 1999. p. 34

⁴⁵Lupovici Catherine. Les choix techniques de la numérisation des documents imprimés. In Conduire un projet de numérisation. 2002. p. 139-140.

⁴⁶Id. p.140-141.

⁴⁷Jacquesson Alain et Rivier Alexis. Bibliothèques et documents numériques. 1999. p.83-84.

hauteur d'un caractère, une formule a été mise au point pour déterminer la résolution.

$RA=3 \times QI / 0,039 \times h$ où QI représente l'indice de qualité et h la hauteur en millimètre des caractères. Un indice de qualité égal à 8 assure une parfaite lisibilité tandis qu'un indice de qualité égal à 3,6 assure une visibilité minimum. En mode niveaux de gris, la formule devient la suivante : $RA=1,5 \times QI / 0,039 \times h$. Il s'en suit une série de recommandations de l'Université de Cornell par type de document : pour du texte et dessin au trait, la résolution d'acquisition sera de 600 dpi en mode bitonal ; pour du texte illustré en noir et blanc, la résolution d'acquisition sera de 600 dpi en mode bitonal ou de 300-400 dpi en mode niveaux de gris.

La SDB signale plusieurs niveaux de résolution selon le type d'exploitation désiré. Elle recommande pour une lecture et un travail à l'écran une résolution moyenne de 300 dpi afin d'assurer une bonne lisibilité et note qu'il s'agit de la résolution choisie par la BnF. Lorsque les caractères sont plus petits ou de mauvaise qualité, elle recommande 400 dpi. Si l'on souhaite numériser pour l'archivage, l'option sera de 600 dpi⁴⁸. La DLL recommande pour sa part une résolution d'acquisition minimale de 200 dpi pouvant aller jusqu'à 400 dpi, pour l'archivage⁴⁹.

A partir de ces considérations, on peut, suivant l'état et la nature des documents à numériser, émettre l'hypothèse selon laquelle la résolution d'acquisition des documents imprimés du SCD devrait se situer entre 300 et 600 dpi.

2.3.3 Format des enregistrements de référence.

Si l'on numérise en mode image, le format des enregistrements de référence sera le format TIFF suivant en cela les recommandations des organismes de tutelle. Ce format offre l'avantage de pouvoir être convertit en format PDF pour la consultation. Les éléments relatifs à ce dernier format seront développés dans le chapitre consacré à la diffusion.

⁴⁸ France. Sous direction des bibliothèques et de la documentation. A propos de la numérisation. Notions et conseils techniques élémentaires. p.2.

⁴⁹ Ministère de la culture et de la communication. Protection et mise en valeur du patrimoine des bibliothèques. 1998. p.110.

2.3.4 Des outils de dématérialisation des documents imprimés.

Les ouvrages imprimés à numériser seront des ouvrages patrimoniaux. Leur valeur et les nécessaires précautions de conservation impliquent que la dématérialisation s'effectue sans contact. Dès lors, il ne peut être envisagé que de recourir, soit aux scanners spécialisés, soit aux appareils photo-numériques. Le choix s'effectuera donc, une fois la résolution d'acquisition établie, en fonction du format des documents. La dématérialisation peut s'accompagner d'un post-traitement : redressement de la courbure des pages, retouche des imperfections, etc. que les prestataires de service proposent mais qui majore le tarif. La BIU dispose d'un scanner Minolta PS 7000 noir et blanc ; il autorise la numérisation d'un A3 à 600 dpi avec redressement de la courbure.

3. Le stockage.

3.1. Le volume de stockage des enregistrements de référence.

A partir des choix relatifs aux techniques de dématérialisation, il est possible de procéder à une estimation du volume de stockage. Le poids d'un fichier numérique est le produit de la définition par le nombre de bits. L'annexe n°8 présente les résultats des calculs pour les documents iconographiques.

3.1.1 Pour la collection des vélins

En acquisition numérique directe à 300 dpi : 62 mégaoctets, soit, en première approximation, à un volume total de 61 gigaoctets, les résultats ayant été arrondis au gigaoctet supérieur. Ces résultats ne tiennent pas compte des variations de formats. Or, les vélins présentent 2 formats différents. Le calcul se base sur le format le plus répandu (241 sur 245 de l'échantillon) de 32x47 cm, soit 12,48x18,33 pouces.

En intermédiation argentique 6x7 à 1500 dpi : 44 mégaoctets, soit, en première approximation, à un volume total de 44 gigaoctets.

3.1.2 Pour la collection des plaques photographiques

Le volume unitaire des fichiers dépendra du mode de numérisation :

En RVB 24 bits couleur, à 700 dpi, le volume unitaire est de 53 mégaoctets pour les plaques de verre au format 13x18. Au format 9x12, le volume unitaire, à 1050 dpi, est de 54 mégaoctets. Cela correspond à un volume total de 40 gigaoctets.

En niveaux de gris 8 bits, à 700 dpi, le volume unitaire est de 17,5 mégaoctets au format 13x18. Au format 9x12, il est de 18 mégaoctets à 1050 dpi. Cela correspond à un volume total de 13 gigaoctets.

3.1.3 Pour les documents iconographiques noir et blanc annexes

Une estimation grossière peut être effectuée en considérant que l'ensemble de ces documents est au format 18x24, format le plus grand rencontré, mais qui n'est pas représentatif. En niveaux de gris à 400 dpi, le volume unitaire est de 11 mégaoctets pour un volume total de 2,2 gigaoctets. Avec une résolution d'acquisition égale à 600 dpi, le volume unitaire est de 24 mégaoctets pour un volume total de 4,8 gigaoctets.

3.1.4 Pour les ouvrages imprimés du fonds patrimonial

3.1.4.1 *En mode image*

L'enregistrement au format TIFF compressé IUT Gr 4 permet de réduire le volume des fichiers dans un rapport de 1 à 10 selon la densité d'information contenue dans la page. Il n'est pas possible de déterminer a priori le volume exact d'un enregistrement TIFF IUT Gr 4 dans la mesure où la compression s'ajuste à la densité d'information. Supposons un ouvrage de 300 pages au format 15x23 numérisé en mode bitonal à 300 dpi. Son volume unitaire en format TIFF non compressé est de 0,6 Mo par page. Une compression maximale permettra d'obtenir un fichier de 60 Ko, une compression moyenne, un fichier de 120 Ko. Le volume de l'ouvrage est respectivement de 18 Mo ou de 36 Mo.

3.1.4.2 *En mode texte*

Le volume de stockage d'un document est déterminé par la formule : $v = n \times c \times o \times F$; où n représente le nombre de pages ; c, le nombre moyen de caractères par page ; o, le nombre d'octets par caractère ; F, le facteur d'accroissement lié à l'existence d'index.

Le facteur F dépend des besoins d'indexation et des systèmes informatiques ; il varie de 1,5 à 2,5 ; il est égal à 1 si aucun index n'accompagne le texte. Le nombre

d'octets par caractère dépend de la table des caractères utilisée, l'ASCII, 1 octet par caractère, ou l'Unicode, 2 octets par caractère⁵⁰.

Supposons à nouveau un ouvrage de 300 pages au format 15x23 dont le nombre moyen de caractères par page est de 3000, sans index, encodé en ASCII. Son volume unitaire sera de 0,9 Mo ; s'il est encodé en Unicode, 1,8 Mo ; s'il est encodé en Unicode et qu'il a été généré un index, facteur F moyen égal à 2, alors le volume unitaire sera de 3,6 Mo.

3.1.5 Le volume total

Le volume total estimé est de 107 Go. Il comprend les collections iconographiques et un volume estimé de 1 Go pour les collections imprimées.

3.2. Les règles de nommage et d'archivage des fichiers

Principe général : dénommer de manière univoque les fichiers en permettant d'identifier sommairement leur contenu. Il est donc nécessaire de définir un identifiant par document. La méthode adoptée par la bibliothèque numérique du CNAM et par le projet NordNum répond à cette préoccupation. Le volume d'un ouvrage imprimé ou la collection iconographique représente l'unité de base. L'identifiant attribué, soit à partir de la cote, soit à partir du numéro d'inventaire, sera le nom du répertoire. Deux types de fichiers numériques sont préparés : fichier fac-similé et fichier « métadonnées ».

Pour les lots de fac-similés :

Collections iconographiques : chaque vélin, chaque plaque photographique et chaque document iconographique annexe se verra attribuer un identifiant univoque de manière incrémentale à partir du nom de répertoire. Du fait de l'impossibilité de stocker quantitativement l'ensemble d'une collection iconographique sur un seul support optique de conservation, la notion de répertoire perd ici son utilité au profit du nom de la collection, à moins d'établir une notion de volume virtuel, propre à la collection numérisée. L'incrémental peut suivre ou non le classement de la collection originale. Dans ce dernier cas, il faut établir une table de concordance. Si la cote ou le numéro d'inventaire ne permet pas de déterminer

⁵⁰Jacquesson Alain et Rivier Alexis. Bibliothèques et documents numériques. 1999. p. 90.

le support original, il serait avantageux de le distinguer grâce au procédé de nommage. Dans ce cas, une nomenclature des signes attribués devra être établie.

Collections des ouvrages imprimés : chaque fichier correspond à une page de l'ouvrage, il s'agira alors d'identifier la page dans un volume d'un ouvrage. On partira soit de la cote, soit du numéro d'inventaire. Les projets NordNum et CNUM utilisent la même méthode de concaténation.

Pour les fichiers « métadonnées », on ajoutera à l'identifiant du répertoire un qualificatif permettant de renseigner sur la nature des métadonnées : notice bibliographique, table des matières, table des illustrations, etc. S'il est fait acquisition d'un logiciel d'archivage, il faudra le cas échéant adapter ces règles aux possibilités de l'application.

L'annexe n°7 illustre cette approche.

3.3. Les supports de stockage

Il existe une grande variété de supports de stockage répondant aux besoins d'un fonds numérisé. On citera en particulier le cédérom, le DVD, le disque dur (DD) et les bandes de sauvegarde. L'annexe n°9 décrit les caractéristiques, les avantages et les inconvénients de ces supports. Pour la conservation, les supports optiques apparaissent comme le meilleur choix possible. Parmi ceux-ci le cédérom (CDR), en plus d'être un support normalisé, est sans aucun doute le meilleur rapport coût/performances. Il ne peut cependant pas être un support de conservation à long terme, comme le DON ou le CD pressé, et nécessitera des migrations périodiques tous les cinq à dix ans, voire une migration vers d'autres types de supports comme le DVD une fois normalisé. Les CDR se distinguent du point de vue de la conservation par la nature du colorant sur lequel s'effectue la déformation thermique. La stabilité de l'information dépend, ceteris paribus, de la qualité des colorants. La norme AFNOR Z-42-011-2 décrit les spécifications de qualité des CDR et de qualité de gravure. Il est possible d'acquérir une interface de contrôle pour les CDR conforme à la norme AFNOR (ISO/DIS 12024 pour la DLL⁵¹) qui permettra d'accepter le CDR ou de le rejeter selon les taux d'erreur. Par ailleurs, la détérioration à plus long terme signifie le plus souvent la perte

complète des données enregistrées ; c'est tout ou rien ; d'où une interrogation quant à la densité d'information et au nombre de fichiers contenus sur chaque cédérom. Faut-il opter pour une logique de minimisation des coûts en enregistrant le plus de fichiers possibles par cédérom, ou bien opter pour une logique de minimisation des risques en limitant le nombre de fichiers par cédérom ?

Nous avons procédé à une évaluation du nombre de cédéroms dans la première hypothèse. L'annexe n°8 reprend les résultats en fonction des choix de dématérialisation.

A titre d'exemples, la collection des vélins, numérisée en mode image couleur 24bits à 300 dpi, représente 100 cédéroms ; la collection des plaques photographiques numérisée selon le même mode à 1500 dpi, 70 ; la collection des documents iconographiques annexes, en niveaux de gris à 600 dpi, 8 ; soit un total de 178 cédéroms. Cela correspond au nombre de cédéroms pour les enregistrements de 1^{ère} génération qui ne peuvent en aucun cas être considérés comme des supports de travail ; il conviendra dès lors de prévoir un jeu de substitution par duplication.

4. Le contrôle qualité.

L'objectif du contrôle qualité est d'assurer la sécurité et l'intégrité des données tout au long du cycle de numérisation. Par « données » il faut entendre aussi bien les documents originaux que leurs reproductions numériques. Ce contrôle qualité doit s'effectuer quelles que soient les options opératoires, internes ou externes ; il est à la charge des personnels de la bibliothèque. Il demande des compétences en matière de conservation des documents originaux et des compétences en matière de traitement numérique de l'image. Cette double compétence fait à l'heure actuelle défaut dans le service, mais la constitution d'une équipe « contrôle qualité » permet de pallier à cette situation en associant, par exemple, le responsable du service du patrimoine, conservateur formé à l'école des Chartes, et le bibliothécaire chargé des NTIC.

⁵¹Ministère de la culture et de la communication. Protection et mise en valeur du patrimoine des bibliothèques. 1998. p.114.

Pour la phase de dématérialisation et de stockage, une liste non exhaustive des contrôles a été dressée :

- Contrôle de la sécurité des originaux en conformité avec le cahier des charges : conditionnement ; identification du lot ; transport ; lieu de stockage.
- Contrôle des moyens techniques de dématérialisation mis en œuvre par le prestataire ou le service interne en conformité avec le cahier des charges.
- Contrôle qualité des documents originaux au retour : état et exhaustivité. Les personnels connaissant le mieux les collections semblent les plus à même d'effectuer cette opération. Elle nécessite de formaliser l'état des collections lors du dépôt chez le prestataire ou dans le service interne.
- Contrôle qualité des cédéroms par tests du taux de BLER conformes à la norme AFNOR. Ce contrôle nécessite un logiciel de tests que le prestataire de dématérialisation consent parfois à livrer, qui peut aussi être intégré dans une solution de GED, ou qu'il faudra acquérir indépendamment. Si le prestataire peut lui-même assurer cette tâche, la confiance n'exclut pas le contrôle.
- Contrôle qualité des documents numérisés : exhaustivité et conformité aux spécifications techniques de dématérialisation (mode d'acquisition, résolution, cadrage, formats). Cette opération nécessite un logiciel de traitement de l'image comme Photoshop CS. Elle demande un personnel formé à cet outil informatique. Le prestataire peut proposer d'assurer ce contrôle là aussi. A ce contrôle des valeurs s'ajoutera un contrôle visuel après impression du fichier sur un format A4 ou A5. Une imprimante de qualité photo sera donc nécessaire, voire deux, l'une pour la couleur et l'autre pour le noir et blanc, le surcoût étant rapidement amorti par l'usage. A défaut, on optera pour une très bonne imprimante noir et blanc capable d'effectuer de bonnes sorties couleur. D'autre part, eu égard au coût unitaire d'impression, il n'est pas souhaitable de contrôler sur papier la qualité de la totalité des fichiers ;

on utilisera donc une méthode d'échantillonnage et on ne testera qu'une partie de la collection (10% au minimum). Du fait des caractéristiques de fichiers à contrôler, une station de travail graphique est conseillée.

- Contrôle qualité du système de nommage des fichiers (nom des fichiers ; table de correspondance ; table des matières, tables des illustrations, index, etc...). Ce contrôle peut être effectué avec les mêmes outils en présence des documents originaux.

5. L'exploitation des documents.

Nous discuterons, ici, de l'exploitation documentaire seulement, conformément à l'objectif de réalisation : quel type de structuration des données et quelle interface d'accès ?

5.1. Une ou deux solutions ? Textuelle et/ou iconographique ?

Les documents soumis au projet sont de nature différente. Textes et images répondent à des usages, en terme de recherche documentaire, différents. Les options en terme de dématérialisation des documents pourraient amener à une collection de documents numérisés en mode image et à une collection de documents numérisés en mode texte. Pouvoir disposer d'une interface de consultation et de recherche commune ou proposer deux systèmes indépendants plus adaptés à la nature des documents est une alternative forte en terme de diffusion. C'est aussi une alternative en terme de compétences, de complexité de la mise en œuvre, d'agrément d'utilisation, et de coût. La question est ici de savoir si on développe une « iconothèque numérique » et une « bibliothèque numérique » ou une « bibliothèque numérique contenant des images ».

En matière de consultation, l'image ne fonctionne pas de manière identique au texte, car elle ne se réduit que difficilement à la description bibliographique conçue pour le document imprimé. L'image est un objet sériel, d'autant plus pour les images scientifiques, et n'est parfois pas indépendant de la collection ; la collection des vélins illustre ce propos. L'indexation, qui détermine en partie les performances des outils de recherche, demeure toujours délicate pour l'image.

C'est pourquoi *le recours à l'image elle-même est le plus souvent requis par les documentalistes et les chercheurs*⁵². La recherche dans un fonds iconographique, si elle nécessite la notice descriptive, n'en est que plus efficace s'il est possible de prévisualiser l'image, encore plus s'il est possible de prévisualiser une partie de la collection. La recherche et la visualisation par lot sont très appréciables pour qui veut consulter l'intégralité d'un fonds. Une étude spécifique pourrait être menée sur cet aspect du projet, afin de déterminer les fonctionnalités de l'interface de consultation. Elle s'avèrerait, de plus, déterminante pour le choix en matière de GED. Elle associerait idéalement professionnels de la bibliothèque et utilisateurs potentiels (enseignants, chercheurs, étudiants,...) par le biais d'un questionnaire et d'entretiens. Les études menées sur les usages des documents iconographiques par les chercheurs à la BnF montrent qu'ils préfèrent *l'interrogation en langage libre des légendes et des textes d'accompagnement, tant que les outils d'indexations personnalisés n'auront pas été développés*, et soulignent qu'ils s'accordent à ne pas vouloir de systèmes sophistiqués qui brident plus qu'ils ne facilitent la spécificité de leur recherche⁵³. Une autre approche, complémentaire, consisterait à partir des fonctionnalités des logiciels de gestion d'images présents sur le marché⁵⁴.

5.2. Les contraintes techniques de diffusion

Le débit des réseaux et les résolutions d'écran limitent le volume des fichiers. Pour une vignette destinée au Web le volume doit être de quelques dizaines de Ko ; pour une image plein écran de 100 à 200 ko. La résolution de restitution dépend de la taille de l'écran et de sa résolution d'affichage. Plus la taille de l'écran et sa résolution sont grands, plus il permet d'afficher une image définie et de taille confortable⁵⁵. Le volume des enregistrements de référence n'est pas compatible avec une exploitation réseau, qui rendrait les temps d'accès trop longs. Il convient dès lors de diminuer le volume des fichiers grâce aux techniques de dégradation et de conversion de format (soit d'avoir préalablement numériser le document dans une résolution et un format compatibles); cette opération conduit à

⁵² Collard Claude, Giannattasio Isabelle et Melot Michel. Les images dans les bibliothèques. 1995. p.221.

⁵³ Id. p.87.

⁵⁴ Schietse, Bérengère et Vandooren, Françoise. L'Iconothèque numérique Un nouveau service électronique pour l'université. 2004, p.93.

⁵⁵ Jacquesson Alain et Rivier Alexis. Bibliothèques et documents numériques. 1999. p.95.

la création d'un troisième jeu de fichiers numériques de faible volume enregistrés dans un format JPEG ou PNG pour le Web. Elle nécessite un logiciel de traitement de l'image comme Photoshop ou ImageMagick. En partant d'un volume moyen de 150 Ko par fichier au format JPEG, le volume total serait de 300 Mo pour une collection de 2000 images. Un tel volume ajouté à celui de l'application pourrait être aisément stocké sur les ressources informatiques existantes. Pour une meilleure qualité de restitution, en consultation locale sur une station spécialisée, il est possible d'utiliser un format non standardisé le JPEG 2000 et/ou d'augmenter la résolution. L'interface peut aussi proposer plusieurs résolutions si le paramétrage de l'écran le permet. La dernière étape de traitement des images numériques avant leur mise à disposition éventuelle sur le Web concerne « l'estampillage » du document. Des solutions existent et sont mises en œuvre par les éditeurs de GED, comme l'exploitation des entêtes et métadonnées IPTC des formats TIFF et JPEG, le watermarking, le tatouage, etc... mais elles ne sont pas sûres à 100%. La meilleure façon de contrôler les exploitations illégales qui pourraient en être faites, consiste à diffuser des documents dont la faible résolution interdit toute utilisation éditoriale.

Pour les ouvrages imprimés, les contraintes de volume et de résolution d'écran, certes plus lâches, demeurent. Une basse résolution d'écran pose des problèmes de lisibilité des documents PDF lorsqu'ils incluent des illustrations⁵⁶. De surcroît, il faut pouvoir communiquer un ensemble d'objets électroniques : des fichiers images de pages d'un document original structurés en un document numérique. Une solution consiste à convertir la série des fichiers TIFF en un document PDF à l'aide d'Acrobat 6.0 d'Adobe. On obtient alors un fac-similé électronique du document original en conservant la présentation. Une conversion en mode texte grâce à Acrobat Capture permet de disposer d'une recherche en texte intégral. Une autre consiste, si l'on veut conserver la structure logique, ou enrichir le document, à s'appuyer sur les possibilités qu'offre XML.

⁵⁶ Westeel Isabelle. NumNord : rapport technique. In La numérisation des textes et des images. 2003. p.43

5.3. Les solutions techniques de communication.

5.3.1 Une diffusion en intranet

5.3.1.1 D'une interface HTML simple...

Pour les documents imprimés cela semble être une solution intéressante pour le SCD. Dans un premier temps, on ne peut compter que quelques titres numérisés. Suivant un scénario intranet, on peut envisager la création d'un document HTML présentant d'une liste des ouvrages numérisés, par titre, par sujet, par auteur, qui permette d'effectuer un lien avec le document PDF. L'existant en matière informatique suffirait a priori. Cette technique semble particulièrement adaptée à de petits fonds numériques et les premiers besoins exprimés par les chercheurs de l'Institut de botanique font état d'une dizaine d'ouvrages du fonds patrimonial. Ce document HTML pourrait être enrichi de notices descriptives et d'un ensemble de signets relatifs aux ressources externes en relation avec les ouvrages numérisés, éventuellement d'un index. Il serait envisageable de développer l'arborescence en concevant un niveau supplémentaire de pages donnant accès à la table des matières, elle-même liée aux fichiers PDF contenant la partie désirée. Stockés sur le serveur, ou bien créés à la volée, ils seraient accessibles par l'interface du SID. Acrobat Reader étant installé sur les postes clients AIE permettrait d'ouvrir, de lire à l'écran et d'imprimer le document PDF. Le développement de ce document devrait être mis à l'étude en collaboration avec le prestataire de service qui assure la maintenance du système d'information de la bibliothèque. Du fait des particularités du fonds, du nombre de documents, et de l'ergonomie attendue d'une interface de consultation, la technique de la liste HTML n'est cependant pas envisageable pour la consultation des images numériques qui nécessite un outil plus structuré.

5.3.1.2 ...A une base de donnée SGBDR/SGBDO...

Le principe général est d'utiliser un système de gestion de base de données interfacé avec un document HTML, permettant de stocker les métadonnées et d'établir un lien entre ces dernières et les fichiers images. Une première solution serait d'utiliser un système de gestion de base de données standard comme Microsoft Access ou Filemaker Pro qui peuvent incorporer des objets multimédia,

interfacé avec un document HTML en utilisant par exemple Microsoft's Active Server Pages. La base serait interrogeable par les champs indexés. Ce document HTML pourrait être intégré au système d'information actuel. Son développement devrait s'effectuer avec le concours d'un informaticien en interne, ou avec un prestataire de service extérieur. Cette opération demandera la mise au point de l'architecture logique de la base de données, de son intégration au SID et le développement de l'interface d'interrogation.

Selon le même principe, il est aussi possible de faire appel au marché avec une solution SGBDR spécialisée image. Certains sont « full Web ». Ce type de logiciel offre des fonctions d'archivage, d'indexation, de recherche et de consultation. Selon leurs caractéristiques propres, ils intègrent des fonctions avancées de visualisation d'images, mais les fonctions de description, quoiqu'en développement, sont plus limitées⁵⁷. Leur inconvénient majeur est le manque d'interopérabilité. Spécialisés, ils nécessitent, de plus, que les personnels soient formés. Parmi les solutions SGBDR/SGBDO on peut citer Gesco-média, Armadillo Photo Média 6.20 et Phrasea Suite 1.3. Elles se présentent en version monoposte, intranet et Internet. Une évaluation plus complète de ces produits est proposée dans l'annexe n°12. Ce sont des outils « clés en main », dont une étude de faisabilité doit évaluer leur intégration possible au système d'information, en particulier avec le dispositif Citrix Métaframe ; à moins de ne mettre en place que des stations de consultation spécialisées à l'intérieur de la bibliothèque, ou d'acquérir une solution Internet.

5.3.1.3 ...A un document XML.

Utiliser XML pour représenter un fonds d'objets numériques, textes ou images, c'est apporter à l'utilisateur une forte valeur ajoutée documentaire, notamment pour les documents imprimés.

*XML [...] est une méthode universelle et standardisée de représentation textuelle des données structurées*⁵⁸. C'est aussi un langage à balise normalisé par le W3C qui supporte l'Unicode. Orienté vers l'échange de données, il est le standard en

⁵⁷ Schieste, Bérengère et Vandooren Françoise. L'Iconothèque numérique... 2004. p. 93

⁵⁸ Chahuneau François. XML : un langage universel pour la représentation textuelle de données structurées. In Bibliothèques numériques : cours INRIA... 2000. p. 119

passer de remplacer HTML sur le Web. Comme ce dernier, XML permet d'établir des liens internes et externes au document. La notion de « document XML » permet de décrire un fonds, une collection ou bien un ouvrage particulier. *Un document XML désigne une unité logique, une collection complète d'information qui s'interprète dans sa globalité, mais qui peut être décomposée en plusieurs fragments physiques (« entités »), chacun correspondant à un fichier (ou à une URI) différent*⁵⁹. De plus, *un document XML est une structure de données adressable, identifiable à une base de données : il est donc possible de lui appliquer une requête*⁶⁰. C'est ce que fait XQL (XML Query language) en cours de normalisation au W3C. XML utilise des modèles de documents appelés DTD (définition type de document) qui fixent les contraintes de validité propres au document. BibliML est une DTD qui représente la notice bibliographique Unimarc. Plus précisément, il s'agit *d'une spécification pour représenter les données bibliographiques et d'autorités*⁶¹. Des outils sont disponibles dans le domaine public pour convertir les notices Unimarc en BiblioML. Créer un document XML avec cette DTD revient à créer un catalogue dont les documents seront décrits par une notice Unimarc. Il est possible de développer en XML des index et des tables comme dans le cas du projet NumNord, ces outils permettant d'accroître l'ergonomie et donc l'efficacité de la recherche. Si cette DTD représente une solution pour les documents imprimés, elle implique un traitement bibliothéconomique des images. Mais il est possible d'opter pour un traitement archivistique des collections iconographiques. L'EAD est une DTD destinée aux inventaires d'archives et aux catalogues de manuscrits⁶². Elle paraît aussi pertinente pour décrire des fonds iconographiques car elle établit une hiérarchie et une structure entre les descriptions, permettant ainsi de gérer des « classes » d'objets numériques. Cette DTD est à la fois un outil de saisie informatique mais aussi une organisation intellectuelle qui s'adapte bien aux données sérielles comme les documents iconographiques. Elle admet jusqu'à 12 niveaux hiérarchiques. Dans le cas de la collection des vélins, il serait alors possible de structurer selon la

⁵⁹ Id. p. 125

⁶⁰ Id. p. 139

⁶¹ Bottin Michel. BiblioML : une application XML pour les données bibliographiques et d'autorités, compatible avec le format (UNI) MARC. In Numérisation des textes et des images. 2003. p. 69

classification et la présentation physique actuelles, en respectant les séries et la disposition en classeurs. Pour les collections iconographiques noir et blanc, l'organisation dépendra des premiers constats lors de la prise en main de ce fonds, mais pourrait là aussi s'appuyer sur l'inventaire manuscrit.

La communication des documents XML nécessite un système de gestion électronique de documents supportant XML. On a déjà noté d'existence dans le domaine public de la plate-forme SDX ; sur le marché des logiciels de GED, on notera la présence de SIM d'Archimed dont le système d'information actuel du SCD est un produit, et, celle de Digttool d'Ex-Libris dont le SIGB actuel est un produit interfaçable.

5.3.2 Une diffusion Internet

Elle demande, afin de s'intégrer techniquement dans l'existant, de développer une application Web en partenariat avec le service informatique de la BIU, ou le CRII de l'Université. Les solutions techniques de structuration des données et de leur diffusion demeurent les mêmes que la diffusion en intranet. Dans le cadre d'un partenariat les solutions techniques de diffusion doivent pouvoir s'intégrer au dispositif actuel ou en évolution (portail BIU, site de l'Université). Ainsi, le CR2I propose la possibilité d'héberger les sites Web des laboratoires de l'Université sur son serveur grâce à l'ouverture d'un compte auprès de ses services, possibilité dont pourrait bénéficier le SCD. Les caractéristiques techniques d'hébergement sont : Apache pour le serveur WWW, PHP pour la langage de script et MySQL comme langage de requêtes pour les bases de données. L'espace disque alloué suit la logique tarifaire suivante : les 100 premiers Mo sont mis à disposition gratuitement, puis, par tranche de 200 Mo, le service est facturé 150 € pour une durée illimitée. La mise à disposition d'un outil statistique de fréquentation du site par le CR2I introduirait un élément d'évaluation du projet. Dans la perspective d'une intégration au dispositif de la BIU, la solution de GED Digttool mériterait, parmi d'autres, d'être évaluée techniquement et financièrement, car elle s'interface avec Aleph 500. Basée sur plusieurs standards

⁶² Queyroux Fabienne. EAD, la description archivistique encodée. In Numérisation des textes et des images. 2003. p.75.

dont HTML, XML, EAD et Dublin Core, elle offre une certaine pérennité aux données et garantit leur indépendance vis-à-vis de la plate-forme logicielle.

6. Le traitement documentaire.

Numériser des documents anciens ne se réduit donc pas à en produire des versions électroniques. Il faut associer à ces versions numériques un riche ensemble de métadonnées qui restituent ce contexte indispensable à une bonne interprétation des documents. Par ailleurs, les documents anciens sont des documents rares, et des métadonnées sont donc nécessaires, qui en règlent l'accès et en fixent les modalités de reproduction. A ces métadonnées, dont l'existence est justifiée par le caractère ancien des documents, viennent s'en ajouter d'autres, d'ordre technique, applicables à tout document numérique (information sur les formats, sur les protocoles d'accès, etc.)⁶³.

6.1. Les principes généraux

A partir du moment où la solution technique de communication est déterminée et les principes de catalogage des documents originaux établis, on procédera au catalogage des documents numériques.

La description bibliographique suivra la norme AFNOR FD Z44-082 – Documentation – Catalogage des ressources électroniques – Rédaction de la description bibliographique. Cette norme Z44-082 s'appliquera en combinaison avec les normes spécifiques aux divers types de documents⁶⁴, en particulier avec la norme Z 44-77 pour les images numériques. L'indexation devrait suivre le vocabulaire contrôlé RAMEAU pour les raisons déjà évoquées.

Les accès aux documents électroniques sont potentiellement de deux types :

- les notices du catalogue ALEPH et par conséquent du SUDOC qu'il faudra créer;
- les interfaces de consultation intranet ou Internet propres à l'outil de recherche.

⁶³ François Role. Métadonnées et structuration des bibliothèques numériques. In Bibliothèques numériques : cours INRIA... 2000. p.163.

⁶⁴ Cazabon Marie-Renée. Le catalogage : méthodes et pratiques. Tome 2. 2003. p. 263.

Cette double approche est celle utilisée par le projet Nordnum. Elle permet un signalement optimum des ressources électroniques et correspond à la volonté d'accroître la visibilité des fonds patrimoniaux du SCD. Le signalement dans le catalogue de la bibliothèque et dans le SUDOC offrirait un accès aux documents numériques via la zone 856 de la notice Unimarc pour les ressources web. Dans un scénario intranet, la notice du document original et la notice du document numérique indiqueraient dans le catalogue les moyens d'accès aux documents grâce aux zones du bloc des notes 3xx pour la localisation et du bloc 4xx des liens, notamment les zones 452 (Autre édition sur un autre support) et 488 (Autres œuvre en liaison)⁶⁵. A ce niveau, l'interface de consultation des collections patrimoniales prendrait le relais, soit par le biais d'une liste en HTML ou d'un document XML pour les imprimés, soit par le biais d'une base de données interfacée avec un document HTML ou d'un document XML pour l'iconographie.

Si les ressources numériques étaient accessibles en ligne par Internet, il conviendrait outre l'exploitation de la zone 856 du format Unimarc de pouvoir signaler ces dernières grâce aux potentialités qu'offre l'usage des moteurs de recherche du Web. Ainsi, l'implantation dans la balise META de l'entête des pages HTML ou XML des métadonnées Dublin Core permettraient un référencement efficace, en décrivant le contenu des dites pages. Cette opération nécessite la saisie manuelle de ces métadonnées dans les pages statiques. Pour les pages dynamiques, cette implantation peut être effectuée à la volée si l'on dispose des outils appropriés. Le Dublin Core est aujourd'hui *le standard de métadonnées le plus répandu et le plus avancé pour la description des ressources Internet*⁶⁶. Les métadonnées Dublin Core sont au nombre de quinze réparties en trois groupes : les métadonnées relatives au contenu, les métadonnées relatives à la propriété intellectuelle et les métadonnées relatives à l'instance particulière. L'annexe n°11 détaille ce schéma.

⁶⁵ Cazabon, Marie-Renée et Witt, Maria. Unimarc Manuel de catalogage. 1993. p. 332-333

⁶⁶ Chung Nicolas et Cormenia Marie-Noelle. *Normes et standards de métadonnées*.
<http://www.educnet.education.fr/dossier/metadata/dc2.htm>

6.2. Les dispositifs particuliers.

Le questionnement relatif au catalogage des collections iconographiques trouve ici un écho particulier. Il était question de savoir s'il est utile de procéder à un catalogage au niveau des documents eux-mêmes ou au niveau de la collection, ou bien encore, au niveau de séries fictives ou non.

- S'il est fait appel à un document XML avec une DTD EAD pour la structuration des données et que sa consultation est possible sur l'internet, une notice bibliographique seulement est nécessaire, qui décrit une collection numérique et qui pointe grâce à la zone 856 vers l'URL de l'inventaire numérique.
- S'il est fait appel à un document XML avec une DTD BiblioML, l'exhaustivité des notices Unimarc par document est nécessaire afin d'alimenter après conversion le document à partir duquel sera effectué la recherche.
- S'il est fait appel à un système de gestion de base de données, il conviendra d'adapter les champs indexés de l'interface de recherche à la notice bibliographique Unimarc du document, afin de conserver autant que possible la cohérence bibliographique, et donc de cataloguer les ressources numériques au niveau le plus fin.

7. La maintenance.

Une fois le processus de numérisation accompli dans sa totalité, demeurent les questions de maintenance des outils de communication et de conservation des enregistrements de première et de deuxième génération. On aborde ici la problématique de la conservation des ressources électroniques : supports, informations, outils d'accès à l'information. On soulignera toute l'importance de l'indépendance des données vis-à-vis de l'application qui les a générées, et en particulier les performances d'XML en la matière, tout comme les avantages à disposer de standards tel TIFF et JPEG.

On distinguera les opérations à court, moyen et long terme.

7.1. La maintenance à court terme.

A court terme, il s'agira d'organiser la maintenance du système logiciel et matériel de communication : réseau, postes de consultation, imprimantes, interface de consultation et d'impression. Ces opérations, classiques, demandent un personnel formé apte à maintenir les outils de communication en état de manière à assurer la continuité du service. Ces fonctions sont, à l'heure actuelle, échues au service des nouvelles technologies de l'information et de la communication qui maintient le système d'information en collaboration avec le service informatique de la BIU et le prestataire de service l'ayant installé. Un relais du même type devrait être organisé avec le fournisseur éventuel du système informatique de consultation des fonds numérisés. De plus, afin d'assurer la sécurité des enregistrements de communication, il faut mettre en place un système de sauvegarde sur bande de type DAT. Un tel système, d'une capacité de 40 gigaoctets, est déjà présent pour les données relatives au système d'information documentaire stockées sur le serveur Dell.

7.2. La maintenance à moyen et long terme.

Eu égard au coût financier de l'opération de numérisation, il apparaît fondamental d'envisager la conservation à moyen et long terme des enregistrements de référence et la capacité future du service à pérenniser la communication des enregistrements de seconde génération au public. Cette approche permet d'accroître la durée d'amortissement et donc de relativiser l'investissement consenti. Elle s'appuie par ailleurs sur des choix en matière de formats (TIFF et JPEG) et de structuration éventuelle des données (XML) aptes à répondre à ces attentes légitimes. Concernant les outils de communication au public, il devra être apporté un soin particulier à la compatibilité ascendante des logiciels fournis sans pour autant se leurrer sur leur réelle capacité dans la mesure où *les programmes informatiques, les méthodes de traitement de l'information disparaissent au profit de nouvelles techniques selon un cycle de validité d'une durée de 2 à 5 ans*⁶⁷.

⁶⁷ Lupovici Catherine. Les stratégies de gestion et de conservation préventive des documents numériques. 2000. p. 45.

Afin d'assurer aux supports une durée de vie maximale, il est recommandé de les stocker dans des conditions adéquates de conservation. Les facteurs de dégradation de l'information qu'ils contiennent sont les suivants : altération du colorant, matière organique qui se dégrade avec le temps et la lumière ; excentricité du noyau ; planéité ; décollement des couches ; abrasions de la surface. L'annexe n°10 indique quelles sont, en terme de température et d'hygrométrie, les conditions adéquates de conservation.

La conservation à moyen et long terme passe en second lieu par le rafraîchissement du support. Cette opération doit être menée de manière périodique, tous les 5 ans au maximum. Elle consiste à transférer les contenus sur des supports neufs. Enfin, au regard de l'obsolescence rapide des technologies informatiques, il arrivera un moment où nécessité se fera sentir de mener des opérations conservatrices de migration ou d'émulation. Ces méthodes de conservation à long terme sont, à l'heure actuelle, expérimentales. Pour la migration, il s'agit de transférer *une ressource électronique d'un environnement matériel/logiciel à un autre, ou d'une génération de technologie informatique à une autre*⁶⁸. L'émulation, pour sa part, *consiste à émuler un environnement informatique différent de la plate-forme sur laquelle se fait l'émulation*⁶⁹.

Ces opérations révèlent l'obligation d'une veille technologique que devra assumer le service des nouvelles technologies de l'information et de la communication.

⁶⁸ Id. p. 45.

⁶⁹ Id. p. 46.

V° Quelles solutions pour le SCD de Montpellier 2.

1. Trois invariants : la constitution d'une équipe, la formation des personnels et l'investissement minimal.

1.1. La constitution d'une équipe projet

Dans la partie IV, ont été décrites les opérations nécessaires à la réalisation du projet : préparation à la numérisation (inventaire, catalogage des originaux, restauration et constitution des lots) ; dématérialisation (interne et/ou externe ; directe ou indirecte) ; stockage (duplication ; nommage des fichiers) ; communication (préparation des enregistrements de communication ; mise en place d'outils de communication) ; contrôle qualité ; traitement documentaire (catalogage et indexation des reproductions numériques) ; maintenance (à court, moyen et long terme). Ces opérations nécessitent des personnels et la constitution d'une équipe projet. A ce propos, on soulignera l'importance que revêt l'association de compétences multiples en citant Dominique Arot pour qui *les réalisations les plus exemplaires sont toujours le fruit d'un travail de conception autour d'un thème précis associant bibliothécaires et chercheurs*⁷⁰. On remarquera le choix des personnels chargés de l'indexation à la Bibliothèque Centrale du Muséum et la constitution d'une instance de validation scientifique au CNAM. On distinguera, ici, le noyau dur, composé essentiellement de personnels du SCD, des partenaires extérieurs.

1.1.1 L'équipe de base

Un conservateur de bibliothèque, responsable du service du Patrimoine, spécialisé en histoire des sciences, pour la partie politique documentaire (choix du

⁷⁰ Arot, Dominique. La bibliothèque numérique patrimoniale. In Bibliothèques numériques : cours INRIA... 2000. p. 20

corpus imprimé), le catalogage des ouvrages anciens, la constitution des lots et le contrôle qualité des ouvrages imprimés.

Un bibliothécaire, chargé de la maintenance du réseau informatique, pour la prise en main et l'inventaire du fonds particulier des plaques photographiques, la constitution des lots, le traitement numérique et documentaire des fichiers, la dématérialisation (en interne s'il y a lieu), le contrôle qualité de l'iconographie et la maintenance.

Un informaticien, extérieur au SCD, appartenant au service informatique de la BIU ou au CRII de l'Université, pour la fonction conseil relative aux outils de communication et la fonction d'installation et de maintenance à court terme.

Un magasinier, pour les opérations de manutention, de préparation à la numérisation et de dématérialisation (en interne s'il y a lieu).

1.1.2 Les partenaires extérieurs

Qu'il s'agisse de la phase de conception ou de la phase de réalisation, il apparaît indispensable de pouvoir associer des partenaires extérieurs au SCD dans la mesure où les compétences requises pour mener à bien le projet ne sont pas toutes réunies au sein des personnels du SCD. En premier lieu, il serait souhaitable de s'attacher les services d'un spécialiste de la systématique, enseignant, chercheur ou doctorant afin d'enrichir le contenu intellectuel du projet, notamment pour la collection des vélins dont le caractère spécialisé appelle une association de haut niveau. En second lieu, une collaboration au niveau technique avec un photographe et un restaurateur est nécessaire. Au cas où on opterait pour une intermédiation argentique, la présence au sein de la BIU d'un laboratoire photographique apparaît comme un atout en terme de proximité et de compétences. Pour la restauration, le cas est plus complexe puisque le laboratoire de la BIU dispose de personnels spécialisés dans la restauration des ouvrages imprimés et de l'iconographie sur supports anciens, mais que les plaques photographiques nécessitent des compétences propres, où l'expertise des personnels des grands établissements (BnF, Bibliothèque Centrale du Muséum, Institut National du Patrimoine) se révélerait d'une aide certaine. Les prestataires de services de dématérialisation et de logiciels de gestion électronique de document peuvent, à travers une fonction conseil, se poser comme des partenaires privilégiés.

1.2. La formation des personnels.

Les nouvelles tâches auxquelles vont être confrontés les personnels du SCD et auxquelles ils sont peu préparés appelleront des actions de formation. Après le recueil des besoins, il serait appréciable de monter un plan de formation afin d'accroître l'efficacité des procédures mises en oeuvre. Sans présumer des opérations qui seront menées en interne ou en externe, il est déjà possible d'établir certains besoins. Concernant la collection patrimoniale des plaques photographiques, une formation est d'ores et déjà programmée en matière de conservation et de restauration à l'Institut National du Patrimoine pour le bibliothécaire responsable du fonds. Pour ce même agent, qui devrait avoir en charge le catalogage des documents iconographiques originaux et de leurs reproductions numériques, une formation spécialisée sur ce point sera sans doute nécessaire. Pour les opérations de dégradation et de conversion des enregistrements de référence en enregistrements de communication, une mise à niveau des connaissances propres à l'utilisation du logiciel paraît elle aussi opportune. De même pour l'utilisation du logiciel de GED dont le fournisseur propose le plus souvent une action de formation que devrait suivre au moins deux personnels du service dans la mesure où il n'est pas souhaitable qu'un seul dispose des informations relatives au bon fonctionnement du système. Si on optait pour une structuration des données en XML, une formation spécifique à ce format est indispensable. A plus long terme, enfin, la maintenance des fonds numériques implique la création d'une nouvelle fonction au sein du service des nouvelles technologies de l'information et de la communication, fonction relative à la veille technologique en matière de conservation des documents numériques.

1.3. L'investissement minimal

L'équipement informatique minimal correspond aux opérations qui seront obligatoirement conduites en interne : opérations de stockage, de traitement des fichiers numériques et de contrôle qualité. Pour les mener à bien, le service devra disposer d'une station graphique de travail équipée d'un graveur de CD-R, d'une grande capacité de stockage ayant un taux de transfert élevé (200 Go type RAID), d'une mémoire vive d'1 Go minimum, 2 Go idéalement, d'un circuit vidéo performant de 256 Mo de VRAM, à laquelle s'ajoutera un écran de 21 pouces

accompagnée de sa sonde de calibrage colorimétrique. Pour un contrôle qualité visuel, le service s'équiperait d'une imprimante couleur de qualité photo et d'une imprimante adaptée aux sorties papier noir et blanc. Pour la partie logicielle, il faudra compter sur un logiciel de traitement des images, sur un logiciel de calibration colorimétrique et un logiciel de conversion des formats si le logiciel de traitement des images ne présente pas cette fonction. Pour la gestion et l'archivage des fichiers numériques image, on s'équiperait d'un logiciel de GED en version monoposte. Le jeu de copies des enregistrements de référence nécessitera 180 CD-R. Enfin, pour le travail d'inventaire du fonds des plaques photographiques, on adjoindra une table lumineuse équilibrée en lumière du jour à 5500°K et d'un compte fil équipé d'une jupe. Cet équipement minimum correspond à un budget de 4842€ hors taxes dont le détail figure à l'annexe n°13.

2. Internalisation versus externalisation.

Lors des considérations relatives aux spécifications techniques du processus de dématérialisation, il a été mis en évidence que les choix possibles impliquent des solutions techniques et la mise en œuvre de matériels différents. Les ressources techniques, humaines et financières du SCD sont limitées. En conséquence, certains matériels sont hors de portée, tant économiquement que du point de vue de la mise en œuvre d'une technique demandant un personnel qualifié. L'externalisation de certaines opérations permet d'y pallier. Mais les deux solutions présentent des avantages et des inconvénients selon le point de vue adopté.

2.1. Des délais différents.

On entend par là que certaines opérations selon qu'elles seront effectuées en interne ou sous-traitées n'exigeront ni le même rythme ni la même durée. Sans préjuger du délai réel et futur de restitution des documents originaux du SCD d'un prestataire de numérisation, il apparaît fort vraisemblable que l'opération de dématérialisation sera de beaucoup plus courte en externe qu'en interne. Pour une collection comme celle des vélins ou des plaques photographiques, elle se traitera en un petit nombre de lots. Le délai moyen de restitution maximum annoncé pour

l'année 2003 par les prestataires de services pour 1000 images est de 15 jours⁷¹. Ce chiffre est bien entendu relatif et dépend de la nature des documents mais il donne un ordre de grandeur cohérent avec le temps moyen estimé pour numériser un vélin avec une chambre numérique (4mn pour le scanning ; 3mn pour la manutention, les éventuels réglages, etc., soit 7mn). En comparaison, avec le concours du laboratoire photographique de la BIU, la campagne de dématérialisation des vélin prendrait entre 6 à 9 mois. Ce délai, estimé par le photographe de la BIU, tient compte de ses autres charges de travail. Les lots, compte tenu de la faible capacité de stockage du laboratoire et d'une plus longue durée d'immobilisation, peu souhaitable, seraient plus nombreux. En parallèle, suivant le rythme précédent, la dématérialisation des reproductions argentiques serait effectuée par le service des NTIC du SCD.

2.2. Les implications en terme de temps de travail des personnels.

Du point de vue de la disponibilité des personnels vis-à-vis du service, l'internalisation des opérations de dématérialisation représente un coût supplémentaire en terme de temps de travail. Dans le cas, par exemple, d'utilisation d'un scanner film pour dématérialiser des ektachromes au format 6x7, il faut compter sur un rythme minimum de 5 minutes par vue, soit environ 50 minutes par film de 10 vues, entre la manutention, le scanning et l'enregistrement du fichier. Sur la base d'un rythme hebdomadaire de 30 reproductions argentiques des vélin, soit 3 films, cela représente environ 3 heures par semaine de temps de travail pour l'opérateur qu'il faudra décharger d'autant par rapport à ses autres activités, et ce sur 6 à 9 mois. Au total l'opération durera 82 heures. Pour le fonds iconographique noir et blanc, dématérialisé avec un scanner à plat, une première estimation de 12 minutes par document prise pour référence, la durée de l'opération est de 200 heures pour 1000 documents, soit sur la base de 36 heures de travail hebdomadaire pour un opérateur à plein temps répartis sur 4,5 journées de travail, 25 jours ouvrables environ. A cela, il convient d'ajouter la durée des opérations d'installation, de calibrage et de caractérisation, d'adaptation de l'opérateur à l'outil de travail et de tests. Pour le fonds imprimé, si on opte pour

⁷¹Remize Michel. Tirer le gros lot. 2003. p.35. Une double page présente la synthèse détaillée d'une enquête auprès de

l'utilisation du scanner de livre de la BIU, il faut prendre en compte outre la durée des opérations de dématérialisation les temps de transports d'une section à l'autre et les répercutions possibles sur le service liées à l'absence des personnels.

2.3. La comparaison des coûts financiers.

Comparer les coûts d'une campagne interne à ceux d'une campagne externe en terme financier suppose que l'on adopte deux points de vue successifs : en premier lieu, pour un projet donné dont le nombre de documents est déjà fixé, on comparera le montant de l'investissement initial aux devis des prestataires de numérisation ; en second lieu, on envisagera le cas où une nouvelle collection serait candidate à une opération de dématérialisation. Pour le montant de l'investissement initial on ne prendra que le supplément de coût occasionné par le choix d'effectuer le travail en interne dans la mesure où l'investissement minimal est le même quelle que soit la solution retenue. Ce montant est alors égal au prix d'acquisition des matériels spécifiques augmenté du prix des consommables argentiques, soit environ 4550€ pour la dématérialisation des vélins et du fonds noir et blanc, soit, encore, un coût unitaire de 2,5€. A partir des informations en notre possession relatives à la tarification des prestataires de numérisation, le prix unitaire estimé varie entre 5 et 15€ hors taxes pour des documents comme les vélins ou les plaques photographiques. Le montant moyen estimé d'une campagne de numérisation pour ces deux collections s'élèverait à 18280€ environ. Si on envisage maintenant qu'une autre collection iconographique soit candidate dans le futur à une campagne de numérisation selon les mêmes procédés, on note alors qu'en interne le coût marginal (ie le supplément de coût) décroît proportionnellement au volume, le coût moyen diminuant suivant en cela le principe d'amortissement de l'investissement, tandis qu'en externe il restera, à tarification égale, proportionnel au volume pour des lots inférieurs à 1000 documents.

Pour les imprimés, il est, selon ce principe, plus rationnel d'utiliser les ressources existantes au service du patrimoine de la BIU. Les tarifs des prestataires de service varient selon la qualité et le format de l'original, les spécifications de la

dématérialisation et le post-traitement souhaité. A titre indicatif, le prix moyen estimé pour les documents imprimés patrimoniaux est d'environ 0,70 € par page.

Au total, on peut avancer que l'internalisation sera moins onéreuse suivant l'hypothèse de travail adoptée mais qu'elle représentera une durée de traitement beaucoup plus longue. De plus, elle immobilisera des personnels qu'il faudra alors seconder dans leurs autres activités, ce qui devrait entraîner une modification de l'organisation du service dont il faut évaluer l'impact.

3. Trois scénarios.

En conclusion, nous présenterons trois scénarios qui répondent, chacun, à une logique différente.

3.1. Scénario n°1

Il repose sur une logique de minimalisation des coûts financiers.

3.1.1 Dématérialisation

Le SCD prend alors techniquement en charge les opérations de dématérialisation avec le concours du laboratoire photographique et du service du patrimoine de la BIU. Un cahier des charges est établi par l'équipe projet ; il formalise les spécifications techniques et les méthodes de travail. Un planning des opérations est dressé, qui en décrit les flux. La collection des vélins, après inventaire et traitement documentaire, est reproduite sur support argentique que le service des NTIC est chargé de numériser. Cette phase s'inscrit nécessairement dans une durée de 6 à 9 mois. La collection des plaques de verre, après inventaire et traitement documentaire, est, elle aussi dématérialisée par le service interne. Les imprimés sont numérisés grâce au matériel du service du patrimoine de la BIU. Les fichiers numériques sont traités en parallèle selon le flux : archivage ; duplication ; retouche des duplicatas ; dégradation et conversion.

En terme d'investissement, ce scénario nécessite l'acquisition d'un scanner à plat de qualité et d'un scanner film moyen format. Il suppose aussi l'acquisition d'un graveur de CD-R dont les caractéristiques techniques répondent à la norme AFNOR de qualité de gravure.

En terme de consommables, trois postes sont à prévoir budgétairement : les supports de reproduction argentique ; les supports numériques d'archivage et de duplication ; les supports papier.

En terme de personnel, il implique la mise à disposition du photographe de la BIU et la dévolution des tâches techniques de dématérialisation au personnel de la BU sciences, soit 280 heures pour les deux collections iconographiques ; soit environ 8 semaines pour un opérateur à temps plein suivant le flux. On ajoutera une semaine pour l'installation, la calibration, la caractérisation et les tests d'ajustement, soit 36h. Le contrôle qualité des images numériques sera effectué au moment même de l'acquisition. Le temps de contrôle sur écran et sur papier est estimé à 5 minutes par image (chargement, contrôle visuel, temps d'impression), soit 152 heures. La durée des opérations de dégradation et de conversion de format est elle estimée à 3 minutes par image, soit 91 heures. Pour être complet, il faudrait encore ajouter le temps imparti à l'opération de duplication, qui, dépendant des performances du matériel, est très difficile à évaluer, et la durée du contrôle qualité des flux d'originaux. Au total, il faut donc compter sur environ 560 heures équivalent temps plein, réparties sur une durée de 6 à 9 mois pour les collections iconographiques, auxquelles on ajoutera la durée, difficilement évaluable à l'heure actuelle, de dématérialisation des imprimés.

3.1.2 Communication

Il prévoit la communication des documents électroniques grâce à deux applications différentes, l'une pour les collections iconographiques, l'autre pour les documents imprimés. Selon les possibilités d'intégration du système de gestion électronique de document, les premières seront consultables, soit sur l'ensemble du réseau intranet, soit sur des stations de consultation spécialisées. On notera la possibilité, moyennant l'acquiescement des droits de diffusion, de produire des cédéroms autonomes avec un moteur de recherche intégré selon les produits de GED. Pour les documents imprimés, la mise en œuvre d'une application HTML, type liste de liens vers des documents PDF, accessible par l'interface actuelle du système d'information, répondrait au mieux à la logique propre du scénario.

Financièrement, cela représente l'achat d'un logiciel de GED multiposte, du logiciel Acrobat d'Adobe et le développement d'une application HTML peu

coûteuse. En terme de personnel, il faudra compter sur le temps passé à la conception, à la recherche d'un produit satisfaisant, à la formation, à l'installation et aux tests, difficilement évaluable.

3.1.3 Traitement documentaire

Le signalement des collections est effectué dans le SUDOC à l'aide d'une notice mère qui indique les moyens de consultation en intranet de l'iconographie. Les images numériques sont toutes indexées dans le système de GED. Les facsimilés PDF sont catalogués à l'unité dans le SUDOC en tant que documents numériques en spécifiant les mêmes indications. En terme de personnel, l'opération la plus lourde est celle de l'indexation dans le système de GED. Sur la base des informations communiquées par la Bibliothèque Municipale de Toulouse⁷², elle s'élève, au rythme de 4 images par heure, à 457 heures pour un agent à temps plein pour les reproductions numériques des vélins et des plaques photographiques.

3.1.4 A quels coûts ?

Au total ce scénario nécessite au moins 1000 heures de travail équivalent temps plein, auxquelles on ajoutera 250 heures liées aux inventaires, 75 heures dues à la préparation des lots, et la durée du contrôle qualité des flux d'originaux, soit environ 1400 heures équivalent temps plein. Son budget s'établit entre 15180 et 16380 € HT auxquels on ajoutera les coûts de formation et ceux liés à la campagne de restauration. Une variante permettrait d'en diminuer significativement le coût : le produit Amardillo, en version monoposte, permet la création de cédéroms, avec un moteur de recherche intégré, qui prendraient place dans le système d'information actuel au prix d'un droit de diffusion à évaluer. L'annexe n°14 explicite ces résultats.

3.2. Scénario n°2

Il repose sur une logique de maximisation de la qualité

⁷² Deschaux, Jocelyne. Calculer les coûts induits de la numérisation. In Conduire un projet de numérisation. 2002. p. 127

3.2.1 La dématérialisation

Par rapport au scénario précédent, les opérations de dématérialisation sont externalisées. Le cahier des charges est établi en considération des spécifications techniques soutenables en terme de stockage. Le mode d'acquisition numérique est direct et sans contact. Un contrôle qualité est mis en place pour les flux entrants et sortants. Le prestataire peut être une entreprise de droit privé ou une institution de droit public. Le coût moyen estimé de la prestation pour les collections iconographiques s'élève à 15000 €. Pour la collection des imprimés, on numérise en mode image mais on pourrait étudier la possibilité de proposer une seconde version en mode texte après traitement OCR si le document le permet. On offrirait ainsi un choix supplémentaire à l'utilisateur qui élargirait l'éventail des usages.

3.2.2 La communication

Elle est réalisée en intranet grâce à une structuration XML des données, conforme à l'EAD pour les documents iconographiques et à BibliML pour les notices bibliographiques des imprimés. Un système de GED supportant XML doit être acquis ; il peut être du domaine public ou privé : SFX, SIM, Digitool... Les coûts d'intégration d'un tel logiciel auxquels on ajoutera les coûts de formation sont difficiles à quantifier mais on ne saurait les sous-évaluer. Les fonctionnalités de l'interface de consultation demeurent le point d'achoppement en matière iconographique. Concernant l'acquisition de Digitool, on soulignera la possibilité d'un financement plus large dans la mesure où d'autres services ou sections de la BIU pourraient être intéressés.

3.2.3 Le traitement documentaire

Le catalogage dans le SUDOC est réalisé au niveau de la notice de la collection. La saisie directe de l'inventaire est effectuée dans un document XML conforme à la description archivistique pour l'iconographie. Le catalogage des documents imprimés numérisés s'effectue dans le SUDOC puis les notices sont converties vers BibliML.

3.2.4 Pour quels avantages et inconvénients ?

La numérisation directe par un opérateur spécialisé disposant du matériel le plus adapté produira une meilleure qualité d'image que l'intermédiation argentique

ou la dématérialisation avec un scanner à plat. Une économie d'heures travaillées en interne est réalisée d'environ 280 heures équivalent temps plein en externalisant la dématérialisation. Mais le budget moyen estimé s'élève au moins à 22000€ HT sur la base d'un système de GED libre de droits comme l'illustre l'annexe n°15. En termes d'organisation et de planification des tâches, et compte tenu du volume des documents, la campagne complète de numérisation se révèle différente en sériant les opérations dans le temps, ce qui permettrait d'accroître l'efficacité des actions le plus souvent répétitives de traitement numérique et l'homogénéité du traitement documentaire. Jocelyne Deschaux souligne, en matière d'indexation, les avantages d'une telle organisation pour l'indexation, selon laquelle *c'est un travail qui doit être effectué de façon continue, à temps quasi complet, pendant le temps nécessaire ; faute de quoi l'indexation manquera de rigueur, d'homogénéité et finalement d'efficacité pour les recherches*⁷³. La phase de dématérialisation au sens strict de l'ensemble des collections peut être achevée en un laps de temps plus court, 4 semaines environ. On pourra dès lors engager les opérations de traitement numérique, puis de traitement documentaire, par blocs. Le choix d'une structuration XML assure une relative pérennité au travail accompli, alors qu'une structuration sur la base d'un logiciel de SGBDR implique le plus souvent le recours à un format « propriétaire » et un manque d'interopérabilité. Il permet, en outre, d'offrir a priori une qualité de service documentaire supérieure pour les imprimés.

3.3. Scénario n°3

Il se présente en définitive comme une variante des deux scénarios précédents dans la mesure où il envisage une mise en ligne sur Internet. Cette mise en ligne peut prendre trois formes : soit le développement d'un site Web propre au SCD dont il assure la conception, le développement, la publication et la maintenance ; soit par l'intermédiaire d'un fournisseur d'application hébergée ; soit, encore, par l'intégration des bases de données ou des documents numériques à un site partenaire comme pourrait l'être celui de la BIU ou de l'Université.

⁷³ Deschaux Joceline. Calculer les coûts induits de la numérisation. In Conduire un projet de numérisation. 2002. p.127.

La première solution sera coûteuse en temps de travail pour le personnel du SCD surtout si le développement est réalisé en interne, mais elle demeure envisageable puisque le SCD dispose déjà des ressources techniques et humaines avec la licence Dreamwaver et la compétence des bibliothécaires du service des NTIC. Elle nécessitera tout de même la fonction conseil d'un informaticien. Son coût financier est faible si l'on opte pour un système de publication libre de droits.

La seconde solution s'articule autour d'un service Internet de gestion et de diffusion de bases de données disponible sur le marché. Ce service permet de constituer, d'administrer et de diffuser des bases de données sans supporter les coûts de la publication et de la maintenance en terme de personnel. La tarification est souvent proportionnelle au nombre d'accès. A titre d'exemple, le service Phraseanet propose un accès pour la consultation à partir de 380 € HT et un accès pour l'administration à partir de 590 € HT.

La dernière solution demande la coopération de services extérieurs au SCD. Une première étude devrait démontrer la faisabilité d'une telle démarche car elle impose, dès la conception, la prise en considération des contraintes techniques, financières et humaines propres à ces derniers. Ce partenariat pourrait prendre la forme d'un binôme entre un membre du personnel du SCD et un membre du personnel du service informatique de la BIU ou du CR2I de l'UM2. Il pourrait aussi dans une perspective plus large de coopération s'articuler avec le projet d'une autre bibliothèque, dont les objectifs sont communs et les fonds patrimoniaux complémentaires, ou bien, encore, avec le projet MuseUM.

Bibliographie

Monographies et travaux universitaires :

Aspects patrimoniaux

Becdelièvre Danielle. *Pour une valorisation du patrimoine scientifique écrit de l'université.* Evaluation des collections du XIX siècle de la BU Sciences du SCD de Rennes 1. Villeurbanne: Enssib, 2002. 70 p.

Dulieu Louis. *La Faculté des Sciences de Montpellier de ses origines à nos jours.* Avignon : Les presses universelles, 1981. 196 p.

Collard Claude, Giamattasio Isabelle et Melot Michel. *Les images dans les bibliothèques.* Paris : éd. du cercle de la librairie, 1995. 390 p. (Coll. Bibliothèques).

France. Ministère de la culture et de la communication. Direction du livre et de la lecture. *Protection et mise en valeur du patrimoine des bibliothèques de France : recommandations techniques.* Paris : DLL, 1998. 174 p.

Néouze Valérie. *Quelle bibliothèque numérique pour une bibliothèque patrimoniale ? L'exemple de la bibliothèque centrale du Muséum d'histoire naturelle de Paris.* Villeurbanne: Enssib, 2000. 69 p.

Oddos Jean-Paul [et al.]. *Le Patrimoine histoire, pratiques et perspectives.* Paris : éd. du Cercle de la Librairie, 1997. 442 p. (Coll. Bibliothèques).

Rioux Jean-Antoine [et al.]. *Le jardin des plantes de Montpellier : quatre siècles d'histoire.* Graulhet (BP 171, 81304 Cedex) : éd. Odyssée, 1994. 230 p. [préf. de Théodore Monod].

Seta-Joannic Frédérique. *Le rôle de la numérisation dans la mise en valeur des fonds patrimoniaux : l'exemple de la bibliothèque inter-universitaire Cujas.* Villeurbanne: Enssib, 1999. 65 p. Disponible sur internet : <http://www.enssib.fr/bibliotheque/documents/dcb/seta.pdf>

Aspects techniques et normatifs

Images numériques : quels formats, quelles caractéristiques, pour quels usages ? Journée d'étude du 6 octobre 2003 organisée par le secteur audiovisuel de l'ADBS, [s.l.] : [s.n.], 2003. [Sans pagination ; document support de la journée].

La numérisation des textes et des images techniques et réalisations : actes des journées d'études organisées à la Maison de la recherche les 16 et 17 janvier 2003. Textes réunis et éd. par Isabelle Westeel et Martine Aubry ; éd. par le Conseil scientifique de l'Université Charles de Gaulle-Lille 3 Villeneuve d'Ascq : Université Charles de Gaulle-Lille 2003. 190 p.

Ballard Marie-Hélène, Delavanne Alexandre, Fortino David [et al.]. *Les bibliothèques virtuelles*. Villeurbanne : INSA, 1999. [Synthèse bibliographique]. 36 p.

Bélisle Claire [coordonnée par]. *La lecture numérique : réalités, enjeux et perspectives*. Villeurbanne : Presses de l'Enssib, 2004. 293 p. (Coll. Référence).

Buresi Charlette et Cédelle-Joubert Laure [sous la direction de]. *Conduire un projet de numérisation*. Villeurbanne : Enssib, éd. Tec&Doc, 2002. 326 p. (Coll. La boîte à outils).

Béquet Gaëlle et Cédelle- Joubert Laure. *La numérisation des documents patrimoniaux*. In France. Ministère de la culture et de la communication, Direction du Livre et de la Lecture. *Protection et mise en valeur du patrimoine des bibliothèques : recommandations techniques*. Paris : DLL, 1998. 174 p.

Bouanzi Latifa. *Les formats de numérisation des images fixes*. Villeurbanne ; Enssib, 1999. 59 p.

Bruckmann Denis [et al.]. *Le traitement documentaire de l'image fixe*. Paris: BPI, 1986. 92 p. (Coll. Dossiers Techniques – Bibliothèque Publique d'Information).

Cazabon Marie-Renée [et.al.]. *Le catalogage : méthode et pratiques*. Paris : éd. du Cercle de la librairie, 2003. 707 p. (Coll. Bibliothèques) .

Cazabon Marie-Renée et Witt Maria. *UNIMARC : manuel de catalogage*. Paris : éd. du Cercle de la Librairie, 1993. 205 p. (Coll. Bibliothèques).

Charton Ghislaine [sous la direction de]. *Les chercheurs et la documentation numérique. Nouveaux services et usages*. Paris : éd. du Cercle de la Librairie. 2002. 268 p. (Coll. Bibliothèques).

Ducharme Christian [sous la direction de]. *Du CD-ROM à la numérisation : développer des documents numériques en bibliothèque*. Institut de formation des bibliothécaires, Villeurbanne, 1997. 172 p. (Coll. La boîte à outils).

Etienney Monique. *Développement du fonds angevin numérisé de la bibliothèque universitaire d'Angers*. [s.l.] : [s.n.], 2000. 34 p.

Garnier François. *Thésaurus iconographique : système descriptif des représentations*. Paris, Le léopard d'or, 1984. 239 p.

Hernandez Jean-Alain et Joly René. *Les normes et les standards du multimédia : XML [...]*. Paris: Dunod, 2000. 213 p.

Jacquesson Alain et Rivier Alexis. *Bibliothèques et documents numériques. Concepts, composantes, techniques et enjeux*. Paris : éd. du Cercle de la Librairie, 1999. 369 p. (Coll. Bibliothèques).

Le Moal Claude et Hidoine Bernard [coordonné par]. *Bibliothèques numériques : cours INRIA, 9-13 octobre 2000*. Paris : ADBS, 2000. 246 p. (Coll. Sciences de l'information Série Etudes techniques).

Lee Stuart D. *Digital Imaging : a practical handbook*. New York. Neal-Schuman, Library Association Pub. 2001. 194 p.

Niemetzky Gérard. *Gestion de la couleur Calibration & profil ICC pour l'image numérique et la chaîne graphique*. Paris : éd. Eyrolles, 2002. 180 p.

Ponsele Véronique Van de. *Le projet de numérisation de la photothèque de la bibliothèque centrale du Muséum national d'histoire naturelle*. Villeurbanne: Institut de Formation des Bibliothécaires, 1993. 25 p.

Aspects juridiques

Alix Yves, Pierrat Emmanuel et Zylberstein Jean-Claude. *Le droit d'auteur et les bibliothèques*. Paris : éd. du Cercle de la Librairie, 2000. 237 p.

Bensoussan Alain. *Droits d'auteur et propriété intellectuelle. Numérisation et bibliothèques : document de synthèse*. Paris : Bensoussan Alain ; Ministère de l'éducation nationale, de la recherche et de la technologie, 1999. 14 p.

Bibent Michel. *Le droit du traitement de l'information*. Paris : ADBS, Nathan, 2000. 127 p. (Coll. 128).

France. Ministère de l'éducation nationale, de la recherche et de la technologie. *Programme numérisation pour l'enseignement et la recherche [...] aspects juridiques : synthèse des travaux, janvier 1999-sept 2002*. Paris : éd. Maison des Sciences de l'Homme, 2002. 561 p + cédérom.

Guernalec Carole. *Le multimédia dans l'enseignement supérieur: petit guide juridique à l'usage des responsables et des enseignants*. Orsay : Université Paris Sud, Faculté Jean Monnet, 1997. 108 p.

Nguyen Due Long Christine. *La numérisation des oeuvres aspects de droits d'auteur et de droits voisins*. Paris : Litec, 2001. 371 p. (Coll. Le Droit des affaires. Propriété industrielle).

Articles de périodiques :

Battisti Michèle. « Journée d'étude de l'ISDN: valorisation du patrimoine et numérisation des collections ». *Documentaliste*. 2002 39 (3), p.130-3.

Battisti Michèle. « Assises internationales du PNER: questions juridiques liées à la numérisation a des fins d'enseignement et de recherche ». *Documentaliste*; 2002, 39 (1-2), p.56-63.

Bequet Gaëlle et Cedille- Joubert Laure. « Numérisation et patrimoine documentaire ». *Bulletin des Bibliothèques de France*, 2000, t. 45, n°4, p.67-72. Disponible sur internet: http://bbf.enssib.fr/bbf/html/2000_45_4/2000-4-p67-bequet.xml.asp.

Collard Claude. « La numérisation des images fixes ». *Bulletin d'Informations Association des Bibliothécaires Français*; n°171, 1996, p.26-7.

Duchemin Pierre-Yves. « La numérisation des documents graphiques: enjeux techniques et financiers ». *Bulletin d'Informations Association des Bibliothécaires Français*; n°184-185,1999, p.136-42.

Duchemin Pierre-Yves. « La numérisation des documents graphiques ». *Bulletin d'Informations Association des Bibliothécaires Français*; n°174, 1997, p.131-4.

Le Saux Anne. « Numérisation et droits ». *Bulletin des Bibliothèques de France*, 2000, t.45, n°2, p.97-9.

Lupovici Catherine. « L'information bibliographique des documents électroniques ». *Bulletin des Bibliothèques de France*, 1998, t. 43, n°4, p.42-47. Disponible sur Internet: http://bbf.enssib.fr/bbf/html/1998_43_4/1998-4-p42-lupovici.xml.asp.

Lupovici Catherine. « Les stratégies de gestion et de la conservation préventive des documents électroniques ». *Bulletin des Bibliothèques de France*, 2000, t.45, n°4, p.43-54. Disponible sur Internet:http://bbf.enssib.fr/bbf/html/2000_45_4/2000-4-p43-lupovici.xml.asp.

Lupovici Catherine, Cloarec Thierry et de Charentenay France. « Les usages de Gallica ». *Bulletin des Bibliothèques de France*, 2003, t.48, n°4, p.40-44. Disponible sur Internet :http://bbf.enssib.fr/bbf/html/2003_48_4/2003-4-p40-lupovici.xml.asp.

Pellequer Bernard. « Projet MuseUM, la science en partage ». *Pôl'Actu*, n°11, juillet 2004. [Sans pagination].

Remize Michel. « Numérisation: tirer le gros lot ». *Archimag*, novembre 2003, n°169, p.35-37.

Savourat Lise. « Pour une appropriation du langage XML par les professionnels de l'information-documentation ». *Documentalistes-Sciences de l'information*, 2003, vol. 40, n°6 p362-369.

Schietse Bérengère et Vandooren Françoise. « L'Iconothèque numérique un nouveau service électronique pour l'université ». *Bulletin des Bibliothèques de France*, 2004, t.49, n°2, p.90-96.

Documents numériques :

Chung Nicolas et Cormenia Marie-Noelle. *Normes et standards de métadonnées*. Paris : Ministère de la Jeunesse, de l'Education Nationale et de la Recherche, 2003. [Consulté le 22/01/2004]. Disponible sur Internet : <http://www.educnet.education.fr/dossier/metadata/default.htm>.

Lecourt Dominique. *L'enseignement de la philosophie des sciences*. Rapport au ministre de l'Education nationale, de la Recherche et de la Technologie. Paris : Ministère de l'Education nationale, de la Recherche et de la Technologie, 1999. [Consulté le

12/02/2004]. Disponible sur Internet :

<http://www.education.gouv.fr/rapport/lecourt/default.htm>

France. Institut des Sciences du Document Numérique. *Le document numérique dans son cycle de vie. Etude pour la région Rhône-Alpes 2000-2003. Synthèse des travaux.*

[s.l.] : Institut des Sciences du Document Numérique, 30/09/2003. [Consulté le 20/01/2004]. Disponible sur Internet :

http://isdn.enssib.fr/archives/transversal/rapportfinal_isdn.pdf.

France. Ministère de la Culture et de la Communication, Mission de la recherche et de la technologie. UKOLN, Université de Bath. Ed scientifiques : Foulounneau Mureil et Bonnamy Alexandra. *Recommandations techniques pour les programmes de créations de contenus culturels numériques.* Paris : Ministère de la Culture et de la Communication, 11/01/2004. [Consulté le 20/01/2004]. Disponible sur Internet :

http://www.culture.gouv.fr/culture/mrt/numerisation/fr/eeurope/documents/guide_technique.pdf .

France. Ministère de la Jeunesse, de l'Éducation Nationale et de la Recherche. Direction de la technologie. *Normes et standards pour les activités numériques dans l'enseignement.* Séminaire de Lyon et Saint Clément les Places, 9 et 10 octobre 2003.[En ligne]. Paris : Ministère de la Jeunesse, de l'Éducation Nationale et de la Recherche.[Consulté le 20/01/2004]. Disponible sur Internet :

<http://www.educnet.education.fr/tech/normes/seminaire-lyon-cr.htm>.

France. Ministère de la Jeunesse, de l'Éducation Nationale et de la Recherche. *XML extensible markup language.* Paris : France. Ministère de la Jeunesse, de l'Éducation Nationale et de la Recherche, 12/11/2003. [Consulté le 20/01/2004]. Disponible sur Internet : <http://www.educnet.education.fr/dossier/xml/default.htm>.

France. Ministère de la Jeunesse, de l'Éducation Nationale et de la Recherche. Sous direction des bibliothèques et de la documentation. *A propos de la numérisation. Notions et conseils techniques élémentaires.* [Consulté le 15/01/2004]. Disponible sur Internet : <http://www.sup.adc.education.fr/bib/Acti/Num/notech.htm>

Roux-Fouillet Jean Paul. *Exemple de Cahier des charges d'une opération de numérisation : la BM de Dole (janvier 2000).* In Société Van Dijk Consultant, Journée d'étude ADDBN sur " Lancer un appel d'offres de numérisation - Méthodologie, cahier des charges, suivi de projet (juin 2002). ADDNB, 2002. [Consulté le 15/01/2004]. Disponible sur Internet : http://www.addnb.org/fr/docs/CR_Jprf.htm.

Table des annexes

TABLE DES ANNEXES	81
ANNEXE N°1: LES PRINCIPALES FILIÈRES D'ENSEIGNEMENTS À L'UM2.	83
ANNEXE N°2: DÉPARTEMENTS DE RECHERCHE ET FORMATIONS DOCTORALES	84
ANNEXE N°3: ÉTUDE PRÉLIMINAIRE DES COLLECTIONS ICONOGRAPHIQUES	85
ANNEXE N°4: LE RÉSEAU ET LE SYSTÈME D'INFORMATION DOCUMENTAIRE DU SCD MONTPELLIER 2	86
ANNEXE N°5: LES ÉTUDIANTS DE L'UNIVERSITÉ MONTPELLIER 2 EN 2004.....	89
ANNEXE N°6: LES CHAMPS INDEXÉS DU MODULE GESCO-MÉDIA À LA BCM	91
ANNEXE N°7: PRINCIPES DE NOMMAGE DES ENREGISTREMENTS DE RÉFÉRENCE (EXEMPLE).....	93
ANNEXE N°8: LES VOLUMES DE STOCKAGE	95
ANNEXE N°9: LES SUPPORTS DE STOCKAGE (UNE SÉLECTION).....	97
ANNEXE N°10: LES CONDITIONS DE CONSERVATION DES CÉDÉROMS	99
ANNEXE N°11: LES ÉLÉMENTS DU DUBLIN CORE.....	100

ANNEXE N°12: EXEMPLES DE LOGICIELS DE GESTION ÉLECTRONIQUE DE DOCUMENTS.....	101
ANNEXE N°13: BUDGET PRÉVISIONNEL DE L'INVESTISSEMENT MINIMAL	102
ANNEXE N°14: BUDGET ESTIMÉ DU SCÉNARIO N°1.....	103
ANNEXE N°15: BUDGET ESTIMÉ DU SCÉNARIO N°2.....	104
TABLE DES MATIÈRES.....	105

Annexe n°1: les principales filières d'enseignements à l'UM2.

Administration, Gestion, Commerce
Agronomie, Génie alimentaire
Biologie
Chimie
Environnement et Aménagement
Génie civil, Bâtiment et Travaux publics
Informatique
Electronique, Informatique industrielle, Télécommunications
Génie Industriel
Mathématiques
Métiers de l'enseignement
Physique
Santé
Sciences de l'ingénieur
Sciences de la Terre et de l'Univers

D'après <http://www.univ-montp2.fr/etudier/formations/filières.html>

Annexe n°2: départements de recherche et formations doctorales

Les départements de recherche :

Sciences mathématiques

Sciences de la terre, de l'eau et de l'espace

Biologie-Santé

Chimie

Physique des structures et de la matière

Sciences et technologies de l'information, de la communication et des systèmes

Biologie, Evolution et environnement

Les écoles doctorales :

Information, Structures et Systèmes

Sciences chimiques et physiques

Terre, Eau, Espace

Sciences chimiques et biologiques pour la santé

Biologie des systèmes intégrés

Science et procédé biologiques et industriels

Economie et Gestion

Informatique et information pour la société

D'après <http://www.univ-montp2.fr/recherche/>

Annexe n°3: étude préliminaire des collections iconographiques

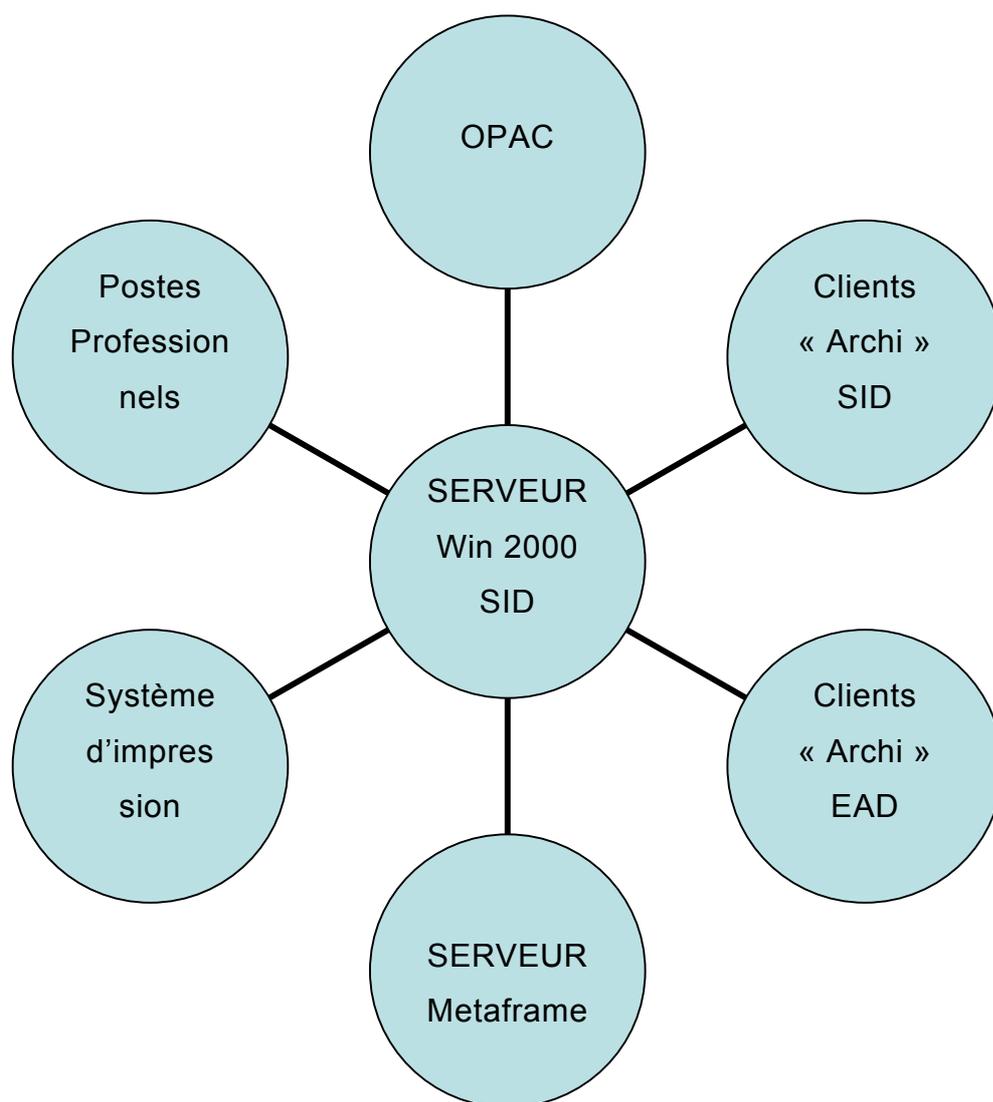
La collection des vélins Node-Véran

	Support Vélins	Support Papier
Nombre (collation)	158	87
En pourcentage	65%	35%
Format	32x47	32x47 à 42x55
Nombre (total estimé)	636	342
Brunissement à plus de 75% ou roussissure très importante	7 (4,4%)	2 (2,3%)
Jaunissement total	--	2 (2,3%)
Brunissement localisé ou roussissure localisée	39 (24,7%)	40 (46%)
Problème de planéité	77 (48,7%)	4 (4,6%)

La collection des plaques de verre

	Très bon état	Bon état	Etat moyen	Mauvais Etat	Total
Nombre	10	30	11	16	67
En pourcentage	15%	45%	16%	24%	100%

Annexe n°4: le réseau et le système d'information documentaire du SCD Montpellier 2



Base matérielle

Sous réseau du réseau de l'UM2

Câblage 10/100.

Architecture client serveur à l'intérieur du bâtiment

1 serveur Dell Pentium III 900 Mhz 768Mo de RAM 2 disques RAID 5

1^{er} disque C : (944Mo libres) D : (19,5Go libres)

2^{ème} disque E : (68Go libres)

Microsoft Windows 2000 Server

CD Line ; AIE

Serveur de fichiers pour le système d'information documentaire (cédéroms et ressources électroniques en ligne); serveur d'impression pour les imprimantes professionnelles

1 onduleur

1 unité de stockage PROCOM (6,5Go libres)

1 système de sauvegarde sur bande Hewlett Packard surestore Dat 40
Capacité 40Go

Clients AIE: Pentium III 700 Mhz, 128Mo de RAM

Navigateur sécurisé AIE

10 clients interface OPAC

28 clients interface de consultation de la documentation électronique (ouvrages de référence, bases de données, périodiques en ligne, sites internet)

6 clients interface Enseignements à distance (EAD)

2 serveurs d'impression Watchdoc ; 3 imprimantes réseau publiques

Nombre de postes professionnels : 30

3 imprimantes réseau professionnelles

1 scanner bureautique

1 imprimante couleur

1 lecteur enregistreur externe Zip 250Mo

1 graveur de CD-R et de CD-RW

2 PC PIII + écran 17" Iiyama

Architecture client léger pour les 6 postes distants

1 serveur IBM Pentium III 1,4 Ghz 1Go de RAM 1 disque RAID 5

C : (2Go libres) E : (10Go libres)

Windows 2000

Client AIE

Métaframe Server

Clients NEC: Pentium III 700 Mhz, 128Mo de RAM

Windows 2000

Métaframe Client

Annexe n°5: les étudiants de l'Université Montpellier 2 en 2004

Cycle	Inscrits UM2	Inscrits UFR	% UM2	Inscrits à la Bibliothèque Universitaire
1er Cycle	5944	2689	45,24%	1367
Deug MASS		129		
Deug MIAS		443		
Deug SM		458		
Deug STU		136		
Deug SV		1112		
Deug STPI		294		
Deust GTP		30		
Deust PAC		50		
DU		37		
Autres 1er Cycle	3255		54,76%	
ISIM	0			
IAE	0			
IUP BBI	28			
IUP GC	30			
IUP GEII	58			
IUP GMI	41			
IUP GMP	28			
IUT Montpellier	1671			
IUT Béziers	326			
IUT Sète	100			
IUT Nîmes	973			
2ème Cycle	3755	1983	52,81%	1417
Licence		1148	30,57%	
Biochimie		95		
Biologie-BGSTU		107		
Biologie-BO		130		
Biologie BCP-PA		74		
Biologie BCP-PVA		35		
Sc Terre et Univers		57		
Chimie		89		
Chimie-Physique		19		
Sciences Physiques		72		
Physique		58		
Physique-PA		21		
E,E,A		93		
Ingénierie Elect		52		
Informatique		128		
Mathématiques		118		
Maîtrise		835	22,24%	
Biochimie BMC		81		
Biologie BCP-BGSTU		92		

Biologie BCP-PA		54		
Biologie BCP-PVA		31		
Biologie Pop, Ecosys,		117		
Sc Terre et Univers		17		
Sc de l'environnement		45		
Chimie		57		
Sciences Physiques		43		
Physique-PA		29		
Physique et applicat.		18		
E,E,A,-A,I,I,		24		
E,E,A,-EMO		36		
Chimie-Physique		16		
Informatique		79		
Mathématiques		55		
Mathématiques-I,M,		16		
MASS		25		
Autres 2ème Cycle	1772		47,19%	
ISIM	892			
IAE	250			
IUP BBI	52			
IUP GC	78			
IUP GEII	91			
IUP GMI	126			
IUP GMP	85			
IUT Montpellier	122			
IUT Béziers	0			
IUT Sète	0			
IUT Nîmes	76			
3ème Cycle	2028	1767	87,13%	843
Agrégation		57		
DESS		355		
DEA		444		
DOCT		910		
DU		1		
Autres 3ème cycle	261		12,87%	
DRT ISIM	10			
Agregr IAE	8			
DESS IAE	195			
DEA IAE	11			
DOCT IAE	37			

D'après les statistiques de L'UFR Sciences et l'EGSBU en 2004.

Annexe n°6: les champs indexés du module Gesco-Média à la BCM

Champs indexés

Cote cliché ; n° de CD photo ; autres clichés ;
Support image ; format ; orientation ;
Sujet ; Nom du scientifique ;
Document original ; date ;
Signature ; matière et dimension ;
Source Auteur ; Source titre et tome ;
Nom de l'illustrateur ;
Source date édition ; source cote ; source éditeur ;
Référence planche ; légende ; n°d'enregistrement ;
Historique des vérifications ; date des vérifications ;

Interrogation des bases dans Gesco-Média

Recherche libre par équation :

La requête se construit à partir de :

- Rubriques (champs préalablement définis)

- Comparateurs = égal à

<= inférieur ou égal à

>= supérieur ou égal à

>< compris entre

- Opérateurs booléens ET, OU, SAUF, SOIT

Recherche globale sur rubriques :

- Liste verticale des rubriques que l'on coche avec la possibilité de les combiner avec les opérateurs ET et OU
- Zone de la valeur recherchée à saisir

Annexe n°7: principes de nommage des enregistrements de référence (exemple).

Documents imprimés

Soit un ouvrage illustré en 2 volumes, le Traité des arbres fruitiers de Duhamel de Monceau, dont la bibliothèque semble posséder une édition particulièrement soignée, numérisé en mode image et enregistré en TIFF. Le volume global des enregistrements de référence est inférieur à la capacité unitaire de stockage (Cédérom : 650 Mo); l'unité de base du document coïncide avec l'unité de stockage.

Chaque répertoire correspond à un volume. Le nom du premier répertoire sera « cote.1 ». Le fond numérisé s'organise sur la base de la cote du document original. Atteindre un répertoire est une opération facile à réaliser à partir d'un gestionnaire de fichier.

Chaque fichier correspond à l'image d'une page qui faut caractériser : numéro d'ordre du fichier (N), type de l'image (T), numéro de page ou de planche (X), extension (ext).

Le nom du fichier suit la syntaxe suivante : N.T.X.ext. Le principe de nommage inclus une table de correspondance.

Le numéro d'ordre du fichier est compris entre 0001 et 9999, la première image du volume est 0001 et la numérotation est réinitialisée à chaque volume. Le type de l'image décrit la nature de la page selon une nomenclature (V=page de titre, T=texte, P=planche hors texte, etc.). Pour les numéros de page ou de planche,

certaines conventions rendent le travail moins fastidieux (non numéroté=0, caractères romains, r1 pour I, r2 pour II, etc. L'extension renseigne sur le format d'enregistrement.

La page de titre sera contenue dans le fichier 0001.T.0.tif. La page n°4 qui se trouve être une planche hors texte sera contenue dans le fichier 0005.P.4.tif

Les fichiers des métadonnées produits (s'il y lieu) sont nommés selon la syntaxe suivante : cote du document original. Volume du document original. Indicateur du type de métadonnées. Ils sont ensuite enregistrés dans le répertoire correspondant. La notice bibliographique est ainsi nommée cote.1.desc, la table des matières cote.1.tmi, etc...

Documents iconographiques

Soit un vélin de la collection Node-Véran. Le volume global des enregistrements de référence est supérieur à la capacité unitaire de stockage (DON : 25 Go). L'unité de base (la collection) se trouve éclatée sur plusieurs unités de stockage. En réduisant l'unité de base au classement actuel, l'album, il en est de même pour la capacité d'un cédérom. L'unité de base devient donc la planche elle-même. Les numéros de page et le type de l'image deviennent inutiles. Il faut cependant ajouter un indicateur précisant la collection, la cote du document original par exemple. Le nom de fichier suit la syntaxe suivante : C.N.ext, où C représente la cote du document original. La notice bibliographique est nommée C.desc, la table des illustrations, C.tpi, etc... Les fichiers de métadonnées sont enregistrés sur chaque unité de stockage.

Le nom du fichier comportant l'image du vélin n°788 s'écrit « cote de la collection ».0788.tif

Annexe n°8: les volumes de stockage

Le volume de stockage peut être déterminé par la formule suivante : $v = RA \times RA \times L \times l \times p / 8$; où v représente le volume unitaire ; RA , la résolution d'acquisition ; $L \times l$, la surface du document à numériser en pouce ; p la profondeur de codage d'un pixel en bit, 24 bits minimum pour le mode RVB couleur, 8 bits pour le mode niveaux de gris, 1 bit pour le mode bitonal⁷⁴.

Volume de stockage pour la collection des vélins

VELINS					
Numérisation directe	Nb de bits couleurs				
Résolution d'acquisition		300	600	1200	2400
Volume unitaire en Mo	24	62	248	989	3953
	48	124	496	1978	7906
Nb de fichiers par CD	24	10	2		
	48	5	1		
Nb total de CD	24	100	489		
	48	196	978		
Volume total en Go	24	61	243		
	48	122	486		
Numérisation 6x7					
Résolution d'acquisition		900	1500	2400	
Volume unitaire en Mo	24	15	44	110	
	48	31	88	220	
Nb de fichiers par CD	24	43	14	5	
	48	20	7	2	
Nb total de CD	24	22	70	196	
	48	49	140	489	

⁷⁴ Stuart D Lee. Digital imaging a practical handbook, 2001. p.42.

Volume total en Go	24	15	44	108	
	48	31	87	216	

Volume de stockage pour les collections photographiques en RVB 24 bits

Plaques de verre	13x18				9x12		
Numérisation directe							
Résolution d'acquisition	500	700	900	1200	750	1050	1200
Volume unitaire en Mo	27	53	86	154	28	54	71
Nb de fichiers par CD	25	12	7	4	23	12	9
Nb total de CD	26	50	93	163	29	55	73
Volume total en Go	17,5	35	56	100	2,6	4,9	6,4

Volume de stockage pour les collections photographiques en niveau de gris

Plaques de verre	13x18				9x12			Tirages et gravures	
Numérisation directe									
Résolution d'acquisition	500	700	900	1200	750	1050	1200	400	600
Volume unitaire en Mo	9	17,5	29	51	9,5	18	24	11	24
Nb de fichiers par CD	72	37	22	12	69	36	27	59	27
Nb total de CD	10	18	30	55	2	3	4	4	8
Volume total en Go	6	11,5	19	33	0,8	1,6	2,2	2,2	4,8

Annexe n°9: les supports de stockage (une sélection).

Produits	Description	Utilisation	Type de lecteur
Bande magnétique classique	Support magnétique d'information à accès séquentiel ; sensible aux champs magnétiques	Sauvegarde ; Stockage à court terme : 2 ans	Spécifique
Digital Audio Tape (DAT)	Bande audio numérique ; Capacité variable	Sons, textes, images	Spécifique
Disque dur	Support magnétique ; Grande capacité ; Taux de transfert élevé ; Faible coût au Mo ; Nécessite parfois une sécurisation coûteuse ; Existe des tours	Sons, textes, images ; Diffusion ; Stockage à court terme : 2 ans	N.A
CD-ROM	Capacité : 650 Mo ; CD pressé, moulage créant des cuvettes, plus résistant ; Accessible en lecture seulement ; Normalisé ; Existe des tours ou juke-box ; Faible coût (croissant selon la qualité) ;	Sons, textes, images ; Conservation à moyen et long terme selon le matériau : 10 à 20 ans pour le polycarbonate, 100 ans pour le verre) ; Les données constructeur font état de durée plus importante en fonction du colorant, de 20 à 200 ans... Vérification tous les 5 ans	Lecteur de CD-ROM banalisé
CD-R	Capacité : 650 Mo ; Enregistrement créé par la déformation thermique du colorant ; Disque inscriptible en monosession (une seule fois) ou multisession ; Faible coût : en polycarbonate le coût de base est de 0,15 €	Conservation à court et moyen terme : 3-5 ans ; Dégradation progressive de l'information ; Vérification tous les 5 ans	Lecteur graveur de cédéroms ; Dépend du type d'enregistrement : un CD-R peut être enregistré pour devenir disque de données, CD-Audio ou CD photo
CD-RW	Idem CD-R mais réinscriptible	Sauvegarde temporaire et mise à jour	Lecteur graveur de cédéroms ;
DVD-ROM	Capacité : de 4 à 9,4 Go ; Accessible en lecture seulement ; Pas de standardisation	Diffusion d'un grand nombre de données Conservation à moyen terme ; Vérification tous les 5 ans	Spécifique mais qui se banalise rapidement
DVD-R	Capacité max : 7,9 Go ; Inscriptible mais non effaçable ;	Sauvegarde de données non compressées ; Vérification tous les 5 ans	Lecteur graveur spécifique
DVD-RAM	Capacité max : 5,2 Go ;	Sauvegarde de fichier	Lecteur graveur

	Réinscriptible	avec mise à jour	spécifique
DON WORN	Capacité de 10 à 25 Go ; Pas de standard et variété des procédés d'écriture ; Coût : 0,04 à 0,09 € le Mo	Conservation à long terme	Lecteur enregistreur spécifique
DON magnéto-optique	Capacité de 2,5 à 8 Go	Stockage temporaire et mise à jour	Lecteur enregistreur spécifique

Sources :

Buresi, Charlette et Cedille-Joubert, Laure. *Conduire un projet de numérisation*. Villeurbanne : Enssib, 2002. 326 p. (Coll. La boîte à outils).

Béquet, Gaëlle et Cedille- Joubert, Laure. *La numérisation des documents patrimoniaux*. In France. Ministère de la culture et de la communication, Direction du Livre et de la Lecture. *Protection et mise en valeur du patrimoine des bibliothèques : recommandations techniques*. Paris : DLL, 1998. 174 p.

Images numériques : quels formats, quelles caractéristiques, pour quels usages ? Journée d'étude du 6 octobre 2003 ; organisée par le secteur audiovisuel de l'ADBS, [s.l.] : [s.n.], 2003. [Sans pagination ; document support de la journée].

France. Ministère de la Jeunesse, de l'Education Nationale et de la Recherche. Sous direction des bibliothèques et de la documentation. *A propos de la numérisation. Notions et conseils techniques élémentaires*. [Consulté le 15/01/2004]. Disponible sur Internet : <http://www.sup.adc.education.fr/bib/Acti/Num/notech.htm>

Annexe n°10: les conditions de conservation des cédéroms

	Support en polymère organique	Substrat verre
Température	18 à 23°C	- 5 à 50°C
Humidité relative	30 à 50% HR	5 à 95% HR
Gradient de température	+/- 4° par heure max	+/- 15° par heure max
Gradient d'humidité relative	+/- 10% par heure max	+/- 10% par heure max

Source :

Jean-Gabriel Lopez In *Images numériques : quels formats, quelles caractéristiques, pour quels usages ?*

Annexe n°11: les éléments du Dublin Core

Contenu	Propriété intellectuelle	Instance particulière
Titre	Créateur	Date
Sujet	Editeur	Type
Description	Contributeur	Format
Source	Droits	Identifiant
Langue		
Relation		
Couverture		

Source :

Ministère de l'Education nationale. Métadonnées, normes et standards.
<http://educnet.education.fr/dossier/metadata/dc1.htm>

Annexe n°12: exemples de logiciels de gestion électronique de documents

Gesco Média

Module complémentaire à Taurus + pour l'archivage, la consultation et le traitement documentaire de l'image.

Références : Bibliothèque du Muséum d'histoire naturelle.

Armadillo Photo Média (Armadillo)

GED photo, son, vidéo ; fonctions d'archivage, d'indexation et de diffusion.

Références : Base images iconothèque de Vivendi Universal Education France, seulement le module de consultation Atlas relié via SQL Server à un module Access pour l'indexation

Phraséa Suite (Phraséa Software)

Système de gestion de base de données objet (SGBDO). Bases de données multimédia comprenant images fixes et animées, sons, fichiers textes, fichiers vectoriels.

Références : Libris (Lille3)

Orphéa studio (Algoba)

Références : Base Indigo par l'Institut de recherche pour le développement (IDR).

Micromusée (Mobydoc)

Références : Bibliothèque Municipale de Toulouse pour la gestion des fonds iconographiques

Système d'Information Multimédia (Archimed)

Environnement MS Windows 2000 Server ; modélisation de l'information en XML ; support de l'Unicode et support natif du Dublin Core ; prise en charge des annuaires LDAP ; compatibilité avec MS SQL Serveur 2000 et Oracle 9.2 ; module de moissonnage OAI-PMH ; indexation en texte intégral sur des documents XML répondant aux standards TEI, Docbook et EAD.

Configuration minimale : Pentium III 1 Ghz ; 512 Mo de RAM ; 4 Go de DD.

Un module complémentaire permet de gérer des fonds d'image et de photographies : SIM Photothèque

Digitool (Ex-Libris)

Interfacable avec Aleph 500

Outils nécessaires à la création, l'archivage, l'indexation, la recherche la gestion et l'affichage des collections numériques.

Environnement MS Windows 2000 Server ; prise en charge des annuaires LDAP ; supporte les normes et standards suivants : MARC 21, Dublin Core, EAD, Unicode, ODBC, HTML, XML, OpenURL, OAI-PMH.

Annexe n°13: budget prévisionnel de l'investissement minimal

Produit	Type	Fonction	Prix € HT
PC + écran 21"	N.A	Traitement des images ; contrôle qualité ; stockage temporaire ; gravure	1608
Epson Stylus Photo 1290S	Imprimante couleur	Contrôle qualité	400
Hewlett Packard Photosmart PS 7660	Imprimante noir et blanc	Contrôle qualité	160
Sonde Optix et logiciel Monaco Optix	Sonde colorimétrique + logiciel de calibration	Calibration et caractérisation couleur des périphériques	520
Adobe Photoshop CS	Logiciel de traitement des images	Contrôle qualité ; dégradation et conversion de format	1000
Armadillo Photo Média monoposte	Logiciel de GED image	Gestion ; archivage	750
Semelec LB 38	Table lumineuse	Inventaire du fonds photographique	120
Rodenstock x4	Compte fil	Inventaire du fonds photographique	104
CD-R	Support de stockage	Copie de sauvegarde des enregistrements de référence	180
TOTAL			4842

NB : Prix indicatifs 2004, établis à partir des catalogues et des sites Internet de fournisseurs.

Annexe n°14: budget estimé du scénario n°1

Produit	Type	Fonction	Prix € HT
Nikon Coolscan 9000	Scanner de film moyen format	Dématérialisation	2640
Epson P4870 Photo	Scanner à plat	Dématérialisation	400
Armadillo Photo Média version intranet / Phrasea Suite Multi	Logiciel de GED image	Archivage ; indexation ; recherche ; consultation...	6500 / 5300
Adobe Acrobat 6.0	Création de fichiers PDF	Consultation	480
100 Films 6x7 + Développement	Support argentique	Intermédiation argentique	1500
CD-R	Support numérique	Stockage et duplication	400
Encre et papier	Support papier	Contrôle Qualité	370
Report net de l'investissement minimal	N.A	N.A	4090
TOTAL			16380 / 15180

Annexe n°15: budget estimé du scénario n°2

Produit	Type	Fonction	Prix € HT
Prestation de dématérialisation de l'iconographie Prix moyen unitaire 7,5 € Quantité : 2000 documents	N.A	Dématérialisation	15000
Prestation de dématérialisation des imprimés Prix moyen unitaire 0,7 € Quantité : pour 1000 pages		Dématérialisation	700
SDX	Logiciel de GED XML	Archivage ; indexation ; recherche ; consultation...	0
Adobe Acrobat 6.0	Création de fichiers PDF	Consultation des imprimés	480
Encre et papier	Support papier	Contrôle Qualité	370
Report de l'investissement minimal	N.A	N.A	4840
TOTAL			21390
TOTAL iconographie			20210
TOTAL imprimés par tranche de 1000 page			1180

Table des matières

SOMMAIRE.....	2
INTRODUCTION.....	4
I° LE CONTEXTE DU PROJET DE NUMÉRISATION.	5
1. L’UNIVERSITÉ DE MONTPELLIER 2.	5
2. LA BIBLIOTHÈQUE INTERUNIVERSITAIRE DE MONTPELLIER.....	6
3. LE SCD DE MONTPELLIER 2.....	6
II° LA DÉFINITION DU PROJET.	8
1. LES OBJECTIFS DE RÉALISATION.	8
2. LES ENJEUX DU PROJET.	9
3. LES COLLECTIONS À NUMÉRISER.....	10
3.1. <i>La collection des vélins de Node-Véran</i>	10
3.2. <i>Le fonds patrimonial photographique</i>	12
3.3. <i>Les ouvrages anciens du SCD de Montpellier 2.</i>	14
3.4. <i>Les droits</i>	15
3.4.1 Dispositif général	15
3.4.2 Situation des documents soumis au projet.....	16
4. LES RESSOURCES EXISTANTES.	17
4.1. <i>Les ressources humaines</i>	17
4.1.1 Le service du patrimoine	17
4.1.2 Le service des nouvelles technologies de l’information et de la communication	17
4.2. <i>Les ressources matérielles</i>	17
4.2.1 Le système d’information.....	17
4.2.2 Les locaux.....	18
5. LES CONTRAINTES DU PROJET.....	19

5.1.	<i>Les contraintes liées aux documents.</i>	19
5.2.	<i>Les contraintes liées au service</i>	20
6.	LES PUBLICS VISÉS.	20
6.1.	<i>Un public des chercheurs :</i>	21
6.2.	<i>Un public élargi :</i>	22
III° LES RÉFÉRENCES EXTÉRIEURES.....		23
1.	LA BIBLIOTHÈQUE CENTRALE DU MUSÉUM D’HISTOIRE NATURELLE.....	23
1.1.	<i>La bibliothèque centrale</i>	23
1.2.	<i>La photothèque et les travaux de reproductions argentiques et numériques</i>	24
2.	CNUM, LA BIBLIOTHÈQUE NUMÉRIQUE DU CNAM.....	26
3.	NORDNUM ET LIBRIS : DEUX PROJETS NUMÉRIQUES AUTOUR DE L’UNIVERSITÉ DE LILLE 3	27
3.1.	<i>Libris</i>	28
3.2.	<i>NordNum</i>	29
4.	SYNTHÈSE	30
IV° LE CYCLE DE NUMÉRISATION : SPÉCIFICATIONS TECHNIQUES ET ORGANISATION GÉNÉRALE.....		31
1.	LA PRÉPARATION À LA NUMÉRISATION.	31
1.1.	<i>La collection des vélins</i>	31
1.2.	<i>Le fonds des documents photographiques.</i>	31
1.3.	<i>Le traitement documentaire des collections iconographiques</i>	32
1.4.	<i>La préparation des ouvrages imprimés.</i>	33
1.5.	<i>La constitution des lots de numérisation.</i>	34
2.	LA DÉMATÉRIALISATION.	34
2.1.	<i>La dématérialisation des vélins</i>	34
2.1.1	Analyse du document	34
2.1.2	Mode d’acquisition.....	35
2.1.3	Choix de la résolution	36
2.1.4	Le format des enregistrements de référence.....	38
2.1.5	Le profil ICC de l’outil d’acquisition numérique.....	40

2.2.	<i>La dématérialisation des plaques de verre et des documents iconographiques annexes</i>	41
2.2.1	Mode d'acquisition.....	41
2.2.2	Choix de la résolution	41
2.2.3	Le format des enregistrements de référence.....	41
2.2.4	Numérisation directe ou intermédiation argentique.....	41
2.3.	<i>La dématérialisation des ouvrages imprimés</i>	42
2.3.1	Mode d'acquisition.....	43
2.3.2	Choix de la résolution.....	44
2.3.3	Format des enregistrements de référence.....	45
2.3.4	Des outils de dématérialisation des documents imprimés.....	46
3.	LE STOCKAGE.....	46
3.1.	<i>Le volume de stockage des enregistrements de référence</i>	46
3.1.1	Pour la collection des vélin.....	46
3.1.2	Pour la collection des plaques photographiques.....	46
3.1.3	Pour les documents iconographiques noir et blanc annexes	47
3.1.4	Pour les ouvrages imprimés du fonds patrimonial.....	47
3.1.5	Le volume total	48
3.2.	<i>Les règles de nommage et d'archivage des fichiers</i>	48
3.3.	<i>Les supports de stockage</i>	49
4.	LE CONTRÔLE QUALITÉ.....	50
5.	L'EXPLOITATION DES DOCUMENTS.....	52
5.1.	<i>Une ou deux solutions ? Textuelle et/ou iconographique ?</i>	52
5.2.	<i>Les contraintes techniques de diffusion</i>	53
5.3.	<i>Les solutions techniques de communication</i>	55
5.3.1	Une diffusion en intranet.....	55
5.3.2	Une diffusion Internet	58
6.	LE TRAITEMENT DOCUMENTAIRE.....	59
6.1.	<i>Les principes généraux</i>	59
6.2.	<i>Les dispositifs particuliers</i>	60
7.	LA MAINTENANCE.....	61
7.1.	<i>La maintenance à court terme</i>	61

7.2. <i>La maintenance à moyen et long terme.</i>	62
V° QUELLES SOLUTIONS POUR LE SCD DE MONTPELLIER 2.	63
1. TROIS INVARIANTS : LA CONSTITUTION D'UNE ÉQUIPE, LA FORMATION DES PERSONNELS ET L'INVESTISSEMENT MINIMAL.	63
1.1. <i>La constitution d'une équipe projet</i>	63
1.1.1 L'équipe de base	64
1.1.2 Les partenaires extérieurs	64
1.2. <i>La formation des personnels.</i>	65
1.3. <i>L'investissement minimal</i>	66
2. INTERNALISATION VERSUS EXTERNALISATION.	67
2.1. <i>Des délais différents.</i>	67
2.2. <i>Les implications en terme de temps de travail des personnels.</i>	68
2.3. <i>La comparaison des coûts financiers.</i>	68
3. TROIS SCÉNARIOS.	69
3.1. <i>Scénario n°1</i>	70
3.1.1 Dématérialisation	70
3.1.2 Communication	71
3.1.3 Traitement documentaire	71
3.1.4 A quels coûts ?	72
3.2. <i>Scénario n°2</i>	72
3.2.1 La dématérialisation	72
3.2.2 La communication	73
3.2.3 Le traitement documentaire	73
3.2.4 Pour quels avantages et inconvénients ?	73
3.3. <i>Scénario n°3</i>	74
BIBLIOGRAPHIE	76
MONOGRAPHIES ET TRAVAUX UNIVERSITAIRES :	76
ASPECTS PATRIMONIAUX	76
ASPECTS TECHNIQUES ET NORMATIFS	76

<i>ASPECTS JURIDIQUES</i>	78
ARTICLES DE PÉRIODIQUES :	78
DOCUMENTS NUMÉRIQUES :	79
TABLE DES ANNEXES	81
TABLE DES MATIÈRES	105