

Stage Dakar
24 janvier – 1 février 2011

**La numérisation au service de la
conservation et de la diffusion
du patrimoine documentaire**

LA CHAÎNE DE NUMÉRISATION

Dominique Maillet
Bibliothèque nationale de France

PLAN

I Qu'est ce qu'un projet de numérisation : Pourquoi Numériser ?

- Les critères de gestion.
- Les critères scientifiques.

II La description de la chaîne de numérisation : les différentes étapes, les différents métiers intervenants, la répartition des rôles :

II.1 La chaîne de numérisation

- 0 et 1) Sélection de la collection numérisée et définition des axes stratégiques
- 2) La gestion des droits d'auteur
- 3) Catalogage des documents numérisés.
- 4) La préparation des documents à numériser.
- 5) La numérisation des documents.
- 6) La post-production (création des méta données)
- 7) Le passage en OCR pour les documents imprimés
- 8) Le contrôle qualité des documents numérisés.
- 9) Le versement des documents numérisés validés sur les serveurs informatiques
- 10) La mise en consultation des documents numérisés.

II.2 Les Intervenants

III Les éléments concernant la numérisation :

III.1 Définitions de base (les pixels, la résolution, bitmap, niveaux de gris, couleurs, les formats)

III.2 Les différents matériels : numériseurs, appareils photos numériques

III.3 Les règles de prises de vues

Filière Imprimés

Filière Images

A - Choix de traitement couleurs niveaux de gris, noir et blanc.

B - Choix des résolutions et des formats de fichiers

C - Les Mires et les profils ICC

D - Définition de l'organisation des prises de vues.

III.4 Les règles de manipulation des documents

III.5 Les règles de contrôles.

IV Exemples de documents numérisés

V Exemple fonds d'écran de l'outil de consultation Gallica

I Qu'est ce qu'un projet de numérisation : Pourquoi numériser ?

Deux objectifs essentiels président à la constitution d'un fonds numérique

• Les critères de gestion.

- **Améliorer** la préservation, en offrant aux lecteurs une copie électronique de substitution, évitant ainsi de fournir le document original.
- **Rationaliser** la communication en fournissant aux lecteurs sur place une version électronique consultable sur écran. Cela évite par exemple, d'acheter en trop grand nombre d'exemplaires un ouvrage qui peut être souvent demandé, ou de laisser un lecteur en attente si le document demandé est indisponible ou déjà communiqué.
- **Permettre la communication** de documents incommunicables par leur taille, leur fragilité.
- **Transmettre à distance**, donc ouvrir la communication hors les murs.

• Les critères scientifiques.

- Constituer des bibliothèques électroniques composées de fonds organisés de façon scientifiquement cohérente et les offrir au public sur place et à distance
- Valoriser les collections par la production et la diffusion de documents éditoriaux électroniques issus des fonds numérisés et portant sur un thème, une discipline, une période, un événement particulier pour lesquels une demande forte est suggérée.

Après le « pourquoi numériser » on passe au « quoi numériser » et au « comment numériser »

Il est important de préciser que :

« **Comment** numériser », n'est pas la question prioritaire à se poser et elle ne peut-être correctement traitée que lorsque les documents à numériser ont été sélectionnés. Le comment s'appuie sur l'état de l'art du marché de la numérisation (en évolution constante rapide) des possibilités techniques qu'il offre par rapport aux documents à traiter et au budget que l'on souhaite allouer à ces projets. Des expertises et des tests sont indispensables à chaque réalisation de projet.

« **Quoi** numériser » n'est qu'une conclusion de la politique de communication que l'on veut mener sur un fonds de documents originaux identifiés et dont l'état physique est connu.

II La description de la chaîne de numérisation : les différentes étapes, les différents métiers intervenants, la répartition des rôles :

II.1 La chaîne de numérisation

La constitution d'un fonds numérique s'insère dans une chaîne composée de différentes étapes que l'on peut lister de la façon suivante:

- 0) La sélection des fonds : Identification de la collection originale et de son état physique. Choix des documents qui feront l'objet du transfert en vue de constituer le fonds électronique
- 1) Définition des axes stratégiques de mise à disposition de ces fonds, sous leur forme originale ou sous leur forme électronique de substitution : permet de définir les règles de numérisation. .
- 2) La gestion des droits d'auteur
- 3) Catalogage des documents à numériser. Préalable indispensable à la numérisation. S'il est plus ou moins facile de retrouver un livre physique dans un magasin, tout document numérisé sans avoir été identifié est un document perdu.
- 4) La préparation des documents à numériser (extraction, dépoussiérage, restauration éventuelle, etc.)
- 5) La numérisation des documents
- 6) La post-production (création des métadonnées)
- 7) Le passage en OCR pour les documents imprimés
- 8) Le contrôle qualité des documents numérisés.
- 9) Le versement des documents numérisés validés sur les serveurs informatiques*
- 10) La mise en consultation des documents numérisés.

0 et 1) Sélection de la collection numérisée et définition des axes stratégiques

Ce sont les conservateurs détenteurs des collections à numériser qui ont en charge le choix scientifique des documents qui feront partie du fonds. Ce sont eux qui connaissent les collections leurs caractéristiques et leur état physique, les demandes du public, les usages potentiels du fonds numérisé, et la qualité souhaitée.

2) La gestion des droits d'auteur

Étape indispensable pour permettre la mise à disposition des fonds numériques. Cette étape se déroule en fait en parallèle avec les autres et démarre dès que sont identifiés les documents à numériser. De la réussite de la négociation dépendra bien évidemment le droit de diffuser la numérisation sur Internet.

Il y a plusieurs étapes successives: droits de numériser, droits de consulte en Intranet ou en Internet, droits de dupliquer, droit d'imprimer, et ce pour chaque collection.

Rappelons pour la BnF que la loi permet que les documents antérieurs à 50 ans deviennent libres de droit ce qui nous permet d'offrir un très grand nombre de documents libres de droits.

3) Catalogage des documents numérisés.

Ce sont les catalogueurs qui vont identifier chaque document numérisé. Tout document numérisé sans avoir été identifié au préalable par une notice est un document inaccessible donc inexploitable donc perdu. En effet s'il est possible de se promener dans des magasins de documents et de les extraire et les manipuler pour les identifier, les supports informatiques empêchent ce type d'identification.

Cette phase d'identification peut être simple ou très lourde selon la complexité des documents à traiter et leurs volumes.

Une notice simplifiée est suffisante pour la numérisation : côte du document, titre, auteur
Mais ces notices simplifiées ne sont pas toujours suffisantes pour permettre au public internaute de mieux sélectionner ce qu'il souhaite consulter. S'assurer que les fonds numérisés soient facilement identifiables donc accessibles permet d'éviter des transferts, longs donc coûteux sur les réseaux et sur des stations de lecture, de documents numérisés qui s'avèreraient après consultation par le lecteur non pertinents pour sa consultation.

Une approche peut consister à créer des notices simplifiées et compléter en post-production (étape 6) des éléments d'identification complémentaire (saisie de légende d'images, saisie de tables des matières et d'éléments d'index par exemple).

4) La préparation des documents à numériser.

Ce sont tous les travaux éventuels nécessaires à la bonne numérisation d'un document : dépoussiérage, restauration éventuelle, défroissage, dépliage, etc.

Il peut être intéressant de concilier des traitements de restauration préalables à la numérisation ou au contraire profiter d'un document à la reliure en mauvais état pour le numériser avant restauration.

5) La numérisation des documents

Elle se fait soit dans des ateliers en interne soit elle est sous traitée à des prestataires de services qui assurent la numérisation soit chez eux soit en installant un atelier dans les locaux de la bibliothèque si les documents sont trop précieux pour sortir.

Choisir un prestataire permet d'avoir une souplesse plus grande et de bénéficier d'un choix de matériel de numérisation plus important.

A la BnF 80% de la numérisation est effectuée par des prestataires.

Cette étape fait l'objet d'un paragraphe particulier (III) qui décrit la numérisation et ses règles.

6) La post-production (création des métadonnées)

Etape indissociable de la numérisation, la post-production est effectuée dans l'atelier qui a produit la numérisation. Dans les cas le plus simple elle consiste à effectuer un lien entre l'image numérisée et sa notice ; Pour les documents reliés elle consiste à faire la pagination ou la foliotation, identifier la table des matières si elle existe, ou encore la table d'index de façon à permettre des parcours de consultation multiples.

7) Le passage en OCR pour les documents imprimés

Norme XML Alto utilisée par la BnF pour océriser ses documents.

Selon la qualité des documents, on demande au prestataire un Dirty OCR ou OCR de très haute qualité (99,99 %, soit un caractère erroné tous les 1000 caractères)

Le Dirty OCR n'est pas affiché il permet de faire des recherches plein texte et seule la page en mode image est affichée.

L'OCR à 99,99 % est par contre affiché.

8) Le contrôle qualité des documents numérisés.

Phase indispensable et consécutive à toute production pour en vérifier la qualité, la BnF dispose de ses propres stations de contrôle et de personnels affectés à vérifier la qualité des documents produits.

Ce contrôle qualité, effectué systématiquement sur tous les documents mais plus souvent par échantillonnage, vu les volumes importants à contrôler, pose le problème de l'élaboration des

critères de qualité qui sont demandés aux prestataires et qui sont ensuite repris par la BnF pour juger de la qualité du travail fourni.

Car, selon la nature des documents à numériser, et selon les usages qui en seront faits il peut être extrêmement difficile de juger si un document a fait l'objet du traitement que l'on demandait.

Ainsi, s'il est facile de vérifier qu'un livre a été bien numérisé - pas de pages manquantes, pas de pages inclinées, une lisibilité facile sur un écran de résolution 150 dpi - il est beaucoup plus difficile pour un document iconographique, ancien ou abimé de donner des règles aussi précises.

Trois éléments essentiels rendent complexes ces contrôles:

- L'élément subjectif de la personne qui effectue le contrôle qui devient très fort pour les images et les sons, et pour lesquels les seuls critères de lisibilité et d'audibilité sont loin d'être satisfaisants et discriminants.

- Les attendus sur le document numérisé par rapport à son original: doit on respecter l'original au plus près ou doit on le rendre le plus communicable possible?

Par exemple, pour la numérisation d'un autochrome. Doit-on essayer de retrouver le rendu original obtenu par les techniques photographiques de l'époque, ou transférer le document en gardant les altérations dues à l'usure du temps?

- En l'absence du document original comment peut on savoir que des pages sont réellement manquantes ou que l'illisibilité du document numérique est due à une mauvaise numérisation et non à une défaillance du document original.

On donnera comme exemple le contrôle des documents iconographiques :

Comment savoir si les couleurs du document original ont été fidèlement traduites à la numérisation ? Pour s'affranchir de tout critère de subjectivité au contrôle et pour effectuer cette comparaison au mieux, il faut avoir :

Un écran calibré sur lequel on affiche l'image numérisée à contrôler

Une station d'observation du document en lumière naturelle sur laquelle on ouvre le document original sur la page à contrôler.

Ce qui implique la participation des conservateurs qui connaissent le fonds à traiter.

9) Le versement des documents numérisés validés sur les serveurs informatiques

Ce versement des documents sur les serveurs sert d'archivage de conservation. Est ainsi sauvegardé le fichier original numérique en TIFF et une copie de consultation est effectuée en JPEG au moment de la migration. C'est ce fichier JPEG qui est mis à disposition du public en Internet ou Intranet. Le fichier original TIFF sert de sauvegarde et peut être communiqué (en général vendu) à des éditeurs ou des imprimeurs pour faire des tirages particuliers.

10) La mise en consultation des documents numérisés.

Cette application de mise en consultation est développée selon la politique de communication de l'établissement ou selon les objectifs d'un projet en partenariat international en lien avec les conservateurs et selon les objectifs que l'on se donne.

A la BnF, il y a 3 grandes applications de consultation des documents numérisés :

La plus connue est Gallica qui permet de fournir sur Internet la totalité des documents libres de droits (ou pour laquelle les droits ont été obtenus) de la BnF.

L'application de consultation du catalogue générale de la BnF qui permet d'afficher lorsque la notice a été trouvée le document numérique associé

L'application de consultation du catalogue des Manuscrits de la BnF qui permet d'afficher lorsque la notice a été trouvée le manuscrit numérique associé

II.2 Les Intervenants

4 types d'intervenants :

- Les conservateurs qui font les sélections des fonds
- Les catalogueurs qui font le catalogage
- Les opérateurs qui font la numérisation
- Les informaticiens qui développent les logiciels d'archivage et de consultation des documents numérisés.

Les rôles sont répartis entre les différents intervenants, et les tâches sont organisées en fonction des personnels affectés, des rythmes de production, et des contraintes d'enchaînement de ces tâches.

Les détenteurs des collections assurent les phases 0,1,2,3,4 participent à la phase 8, et expriment leurs besoins en terme de spécifications fonctionnelles sur la phase 10.

Les phases 5,6,7,8,9,10 sont des phases techniques réservées à des opérateurs photographes et à des informaticiens.

Si les phases 0 à 4 sont en général toujours effectuées par des personnes de la bibliothèque les phases 5 à 10 peuvent et sont souvent sous-traitées.

III Les éléments concernant la numérisation :

III -1 Définitions de base (les pixels, la résolution, bitmap, niveaux de gris, couleurs, les formats)

III - 2 Les différents matériels : numériseurs, appareils photos numériques

Les 2 grands principes d'acquisition :

Les capteurs CCD en ligne. Principe d'acquisition par balayage utilisés par les scanners.

Les capteurs CCD en matrice : sur les dos numériques. Prise de vue en one shot.

Les grands constructeurs de numériseurs : Metis (Italie), Zeutschell (Allemagne), I2S (France)

Les dos numériques : Phase One, Sinar

Les dispositifs : porte modèles et balance de livre qui permettent de numériser des documents sans les mettre à plat.

Photos montrées pendant la formation

III - 3 Les règles de prises de vues

On organise les collections en 2 filières qui permettent ainsi de créer des grandes règles par filière :

- Filière Imprimés : Imprimés courants, périodiques et presse

- Filière Images : Iconographie, manuscrits, imprimés exceptionnels

4 points sont à prendre en compte :

A- Le choix de traitement couleurs, niveaux de gris ou noir et blanc

B - Le choix des résolutions

C - L'utilisation des mires et des profils ICC

D - La définition de l'organisation des prises de vues.

A - Choix de traitement couleurs niveaux de gris, noir et blanc.

Filière Imprimés : La lisibilité du document est le critère le plus important. Ils sont candidats à l'océrisation et à l'identification des éléments peritextuels (type tables des matières.)

Les Imprimés sont donc numérisés en noir et blanc et s'ils comportent des illustrations(cas des documents mixtes), elles sont numérisées en niveaux de gris. Quelques pages particulières pourront être numérisées en couleur (pages de couverture) .

Une exception cependant pour des imprimés entièrement en couleurs dont la dégradation systématique en noir et blanc n'aurait aucun sens (catalogues d'art).

La presse est numérisée en niveaux de gris pour être plus lisible.

Filière Images : La demande des usagers évolue pour toute l'iconographie mais également pour l'ensemble des documents manuscrits vers une numérisation couleurs même si les documents d'origine sont en noir et blanc (gravures, estampes au trait, cartes monochromes)

La lisibilité et le rendu des documents sont en général meilleurs et la couleur donne une apparence flatteuse.

Cependant deux contraintes importantes ne permettent pas d'appliquer systématiquement le choix de la couleur :

- le poids des fichiers couleurs est 3 fois plus important que le même fichier numérisé en niveaux de gris.
- les matériels de numérisation couleurs sont plus chers même si les progrès technologiques tendent à faire baisser les coûts.

Par conséquent, les images (iconographie) sont numérisées en couleur que le document original soit couleur ou monochrome sauf si le résultat obtenu n'est plus conforme au document original. (Exemple des positifs de projection noir et blanc qui, numérisés en couleur, prennent une teinte sépia que l'on ne retrouve pas sur l'original) et les Manuscrits et imprimés exceptionnels sont numérisés sauf exception en couleurs. Il peut y avoir une exception de numérisation en niveaux de gris si les fonds ne justifient pas de la couleur (brouillons d'écrivain, cahiers de compte, etc)

B - Choix des résolutions et des formats de fichiers

- **Filière Imprimés** : Le format utilisé est le TIFF V6 monopage UIT GIV en 1bit pour le trait , TIFF V6 monopage non compressé 24 bits pour la couleur ou 8 bits pour le niveau de gris. La résolution de 300 dpi est utilisée, sauf exception nécessitant du 400 dpi

- **Filière Images** : Le format utilisé est le TIFF V6 monopage non compressé 24 bits pour la couleur ou 8 bits pour le niveau de gris.

Le JPEG n'est pas un format d'archivage n'est pas utilisé à l'acquisition.

La totalité des documents images traités dans le cadre des filières images font l'objet d'un tuilage au moment de l'archivage sur les serveurs de façon à permettre le zoom.

Les résolutions sont uniques et sont exprimées en dpi. Ce qui a l'avantage que tous les documents traités, et ce quelque soit leur taille, aient la même finesse d'analyse.

Les résolutions de 300 et 400 dpi sont utilisées à l'exception des documents d'un format inférieur à A6 ou comportant des détails inférieurs à A6 pour lequel la résolution de 600 dpi est appliquée.

De plus une petite proportion de documents présente des caractéristiques qui nécessitent d'appliquer une résolution pouvant varier de 50 à 2800 dpi. Ce sont les documents de petite taille ou de très grande taille (Timbres postes, Monnaie et Médaille et Affiches de Cinéma)

Cas particulier des documents transparents :

S'il s'agit de reproductions sur support argentique de documents originaux pour lesquels il y a une échelle d'agrandissement : les documents seront traités selon les règles ci-dessus en appliquant 300 ou 400 dpi sur le document original ayant fait l'objet de la reproduction.

Pour les documents originaux transparents (plaques de verre, autochromes, etc.) et les documents transparents de substitution pour lesquels l'échelle n'est pas connue : les documents feront l'objet d'un test préalable à la numérisation pour définir quelle résolution est la mieux adaptée au document à traiter.

On utilise alors la règle du 1000 points au cm (2500 dpi) qui donne en général de bons résultats puisqu'elle s'appuie sur un agrandissement de 8 d'un négatif photo. Par exemple une diapositive 24X36 devrait être numérisée à 2400 X 3600.

C - Les mires et les profils ICC.

Uniquement concernés : les fichiers numériques produits dans la filière Image pour tous les documents couleurs.

Le profil ICC Adobe 98 est inscrit dans l'entête TIFF du fichier. De plus, pour permettre aux imprimeurs de caler leurs impressions, une mire de couleurs est associée à la fin du document couleurs (manuscrits ou imprimés exceptionnels) ou derrière chaque image.

C'est La Mini charte Color Checker qui est retenue à la BnF et chaque numériseur a sa propre charte. Elle est numérisée dans les mêmes conditions que les documents reproduits et est réduite en 600X800 pixels par une procédure automatisée pour prendre moins d'espace de stockage.

D - Définition de l'organisation des prises de vues.

Filière Imprimés :

Règle n°1: Selon les cas, un plateau aspirant ou une glace pourront être utilisés pour maintenir les documents à traiter et assurer une numérisation optimum. Evidemment leur utilisation sera faite en adoptant toutes les précautions d'usage, en particulier pour la glace en limitant la pression. D'autres dispositifs adaptés, pourront être utilisés le cas échéant dans le même respect des documents à traiter.

Règle n°2 : On optimisera la numérisation du document en fonction des contraintes du matériel de numérisation dans le but d'alléger et de sécuriser les manipulations sur les documents. Il se peut par conséquent que l'acquisition ne soit pas systématiquement effectuée dans le sens du document. C'est la post-production qui remettra dans le sens du document tous les éléments composant le document. **De plus, lorsque toutes les pages de l'original ne sont pas dans le même sens, l'opérateur effectuera les rotations nécessaires pour les remettre dans le sens de la lecture.**

Règle n° 3 :La numérisation s'attache à reproduire au plus près le document original sans chercher à l'améliorer. Les réglages sont optimisés au moment de la numérisation. Pas de retouche en post-production.

Règle n°4 : On numérise en une page image sauf exception. (*)

Le corps de l'ouvrage est numérisé dans son intégralité. On numérise l'intégralité du document du plat supérieur au plat inférieur en incluant les pages vierges foliotées ou non entrant dans la pagination ou non. Si le département de collection souhaite des prises de vues complémentaires il les indique sur le bordereau de consignes de prises de vues (tranche de queue, de tête, de gouttière, reliure)

.Pour éviter aux ateliers des traitements différents selon l'imprimé, son ancienneté et son origine, il est souhaitable que les éléments de reliure soient traités avec le même systématisme.

(*)La prise de vues portera sur la double page si l'image occupe les 2 pages. La pagination devra être conforme à la prise de vues.

Règle n° 5 : Le cadrage est effectué en recadrant chaque page le long du bord extérieur des marges de manière à montrer l'intégralité de la page. On évitera de garder un liséré noir autour de la page. Lorsque la marge côté reliure est inexistante ou que des lignes de texte paraissent coupées le cadrage se fera au niveau du fond de cahier, de façon à montrer que la numérisation n'a pas tronqué d'informations. Sinon le cadrage ne prendra pas le fond de cahier. *On notera que cette règle diffère des règles appliquées aux manuscrits où la prise de vues mord systématiquement sur la page opposée de façon à montrer que rien n'a été perdu.*

Règle n° 6 : Les pages sont numérisées majoritairement en noir et blanc. Le traitement en niveaux de gris ne sera utilisé que lorsqu'il améliore la lisibilité de la page. Les pages seront numérisées en couleurs si la couleur est indispensable à la compréhension du document (planches de dessins de couleurs, typographie de plusieurs couleurs, etc.) Un référentiel sur l'utilisation du niveau de gris et du noir et blanc rédigé par le responsable du pôle contrôle numérisation permet de fixer quelques éléments de choix. Ce document sera complété en intégrant la partie couleur. Il servira de base de travail dans les ateliers et sera joint en annexe au document synthétisant les règles mises en œuvre et nos recommandations.

Règle n°7: Pas de mire couleur insérée à la fin des documents mixtes présentant de la couleur.

Règle n° 8 : Il n'y a pas de remise de la page au format ISO immédiatement supérieur.

Filière Images :

Règle n°1: Selon les cas, un plateau aspirant ou une glace seront utilisés pour maintenir les documents à traiter et assurer une numérisation optimum. Evidemment leur utilisation sera faite en adoptant toutes les précautions d'usage, en particulier pour la glace en limitant la pression. D'autres dispositifs adaptés, pourront être utilisés le cas échéant dans le même respect des documents à traiter.

Règle n°2 : On optimisera la numérisation du document en fonction des contraintes du matériel de numérisation dans le but d'alléger et de sécuriser les manipulations sur les documents. Il se peut par conséquent que l'acquisition ne soit pas systématiquement effectuée dans le sens du document. C'est la post-production qui remettra dans le sens du document tous les éléments composant le document. **De plus, lorsque toutes les pages de l'original ne sont pas dans le même sens, leurs présentations seront respectées à la numérisation et c'est l'application de consultation qui permettra d'effectuer les rotations nécessaires à l'affichage du document dans le sens de la lecture.**

Règle n° 3 : La numérisation s'attache à reproduire au plus près le document original sans chercher à l'améliorer. Les réglages sont optimisés au moment de la numérisation. Pas de retouche en post-production.

Règle n°4 : Documents **type manuscrits ou imprimés de réserve numérisés en couverture complète** : On numérise en une page image sauf exception.(*)

On numérise l'intégralité du document du plat supérieur au plat inférieur en incluant les pages vierges foliotées ou non entrant dans la pagination ou non. Une exception cependant : la DCO pourra demander de limiter la prise de vues à 2 pages blanches non paginées ou n'entrant pas dans le cadre de la pagination, dès lors que la séquence de pages est élevée (cas des cahiers de manuscrits par exemple où seules les premières pages sont écrites) Si le département de collection souhaite des prises de vues complémentaires il les indique sur le bordereau de consignes de prises de vues (tranche de queue, de tête, de gouttière, reliure)

On pourra ainsi recréer un fac simile sans avoir à ressortir le document. Un logo particulier sera créé pour identifier ces pages.

(*)La prise de vues portera sur la double page si l'image occupe les 2 pages. La foliotation devra être conforme à la prise de vues.

Règle n°5 : Documents reliés : On numérise l'intégralité de la page (cadrer au bord extérieur de la page et mordre dans le fond de cahier) pour permettre de montrer que rien n'a été coupé à la prise de vues.

Documents pièces : Même règle que pour les documents reliés. De plus, pour les documents montés, on cadre sur le document original en incluant la légende ou la côte inscrite éventuellement sous l'image.

Règle n°6 : Les documents seront numérisés pleine page sans vue de détail sauf exception dûment justifiée. Mais la qualité de la résolution permettra au moment de la consultation de zoomer dans le document et d'agrandir les détails souhaités.

Règle n° 7 : Toutes les images d'un même document auront la même résolution sauf exception.

En particulier :

- les cartes de grands formats qui font l'objet d'une vue globale à une résolution inférieure puis de vues de détails à des résolutions de 300 dpi. Dans ce cas, le principe des cartes nécessitant des vues de détail sera décidé lors de la phase d'expertise et les modalités de prises de vues avec éventuellement des légendes particulières seront fournies par la DCO à l'atelier ou au prestataire.
- Les ensembles iconographiques pour lesquels on numérise des images et des objets (types daguerréotype dans son coffret)
- Les ouvrages dont certaines pages nécessitent une résolution supérieure (miniature de petite dimension) ou avec prise de vues objet.

Règle n° 8 : Il n'y a pas de remise de l'image au format ISO immédiatement supérieur.

Règle n°9 : La prise de vue devra permettre de cadrer de façon à isoler la page de son contexte mais sans couper d'information.

Un cache devra systématiquement être utilisé pour toutes les pages présentant des lacunes des trous ou déchirures ou pour masquer les écritures parasites dues à la transparence éventuelle du papier.

L'utilisation du cache peut être lourde en manipulation et coûteuse. Par conséquent son utilisation sera laissée à l'entière appréciation du photographe dès lors que la règle énoncée ci-dessus est respectée.

Le cache devra être le même pour tous les ateliers et il sera fourni aux prestataires. Il sera de couleur coquille d'œuf. Dans certains cas particuliers, un cache de couleur différente pourra cependant être exceptionnellement utilisé (si la couleur coquille d'œuf transparait en changeant la couleur du document, par exemple).

On trouvera sur la base de production dans le dossier des exemples de prises de vues auxquelles il conviendra de se référer.

III.4 - Les règles de manipulation des documents

Recommandations pour le transport :

Par sécurité pour les documents lors des transports (aller et retour) : Chaque bac devra porter clairement une mention d'appartenance sur 2 côtés ainsi que sur le couvercle gravée en caractères majuscules de 30 millimètres de hauteur minimum.

Les parois et le fond seront pleins et lisses en intérieur. Il est demandé la présence d'une zone granitée en extérieure pour la pose d'étiquette autoadhésive.

Recommandations élémentaires de manipulation des ouvrages :

- Un plan de travail propre et dégagé exclusivement réservé à la manipulation des ouvrages (pas de nourriture, de liquide, de végétaux en pot ou en vase, d'objet contenant, d'outil d'écriture autre qu'une pointe graphite, ...)
- Des mains propres pour l'opérateur (cela signifie qu'un opérateur doit se laver fréquemment les mains au cours de sa journée car les ouvrages sont sources de salissures des mains)
- Privilégier la manipulation unitaire des ouvrages (surtout pour les ouvrages de moyen et grands formats)
- Un livre se prend par le centre du dos, en étau entre le pouce et les 4 doigts (cela signifie que l'on ne tire pas un livre pour le saisir par sa coiffe)
- Sur un plan de travail, un livre se pose à plat (sur un des plats de couverture, première ou quatrième de couverture). En cas de constitution de pile, en privilégiant des formats proches les uns des autres et toujours les formats en ordre décroissant (le plus grand à la base de la pile). Les ouvrages tête-bêche, pour une meilleure stabilité de la pile et d'une hauteur raisonnable, le tout en dehors du périmètre extérieur du plan de travail afin éviter tout risque de percussion de la pile d'ouvrage par des mouvements de personnes ou d'objet alentour.
- Sur un chariot, un livre se pose verticalement sur sa tranche de queue ou sa tranche de gouttière, en appui sur une joue de bord du chariot. Les livres serrés les uns contre les autres et en utilisant un serre-livre ou un bloc de calage pour le dernier livre.
- Un livre s'ouvre toujours poser à plat sur un plan de travail, en l'ouvrant par son milieu pour revenir au début. Cette ouverture doit être faite dans la limite mécanique d'ouverture de l'ouvrage sans contrainte de l'opérateur pour obtenir une ouverture à plat (180°).
- On tourne les pages des livres toujours par le haut de la page et par glissement de la main à plat au verso de la page. (On ne mouille pas son doigt, ni ne corne une page).
- On referme toujours un livre en dehors de son temps d'utilisation (même pour une pose de l'opérateur)
- En cas de nécessité, on peut utiliser un signet en papier blanc comme marque-page (pas de Post-it, trombone, ou tout autre objet)
- Manipulations des boîtes contenant un document ; La boîte doit rester associée à son document. La cote est inscrite sur le rondage du document et sur le portage (inscrit sur la page de titre ou la garde volante).
- Les boîtes de grand format doivent être stockées couchées.
- Les documents reliés ne doivent pas être stockés debout sans être maintenu par d'autres ouvrages ou des serres livres.

- Ne pas entasser les livres en grandes piles. Les piles doivent être en forme de pyramide. Jamais de petit format sur lequel repose un plus grand format. De préférence en tête-bêche les uns par rapport aux autres

- Utilisation d'un papier neutre ou d'une feutrine en dessous du document à numériser

- Eviter le frottement du dos sur le bord du porte livre en utilisant des mousses et de façon plus générale utiliser les mousses pour caler au mieux le document sous les 2 vitres en évitant les pressions.

- Une fois le document positionné, éviter les mouvements de la reliure.

- Remonter entièrement la vitre qui tient la page opposée à celle que l'on numérise de façon à ce que l'opérateur ait les 2 mains libres pour manipuler le document.

- La vitre qui est sur la page à numériser doit aller au plus près de la page opposée sans la racler.

- Lorsqu'un ouvrage est fourni dans une boîte, pour sortir le document de la boîte on le fait glisser pour pouvoir le saisir et c'est la manipulation inverse qui est faite pour entrer le document dans sa boîte après numérisation.

- Manipulation des documents pliés dans un corps d'ouvrage : Ils sont à traiter avec précaution ; l'opérateur procédera au dépliage en respectant l'ordre des plis. Lors de la numérisation, il ne faut jamais les mettre en porte à faux et s'assurer que tout ce qui est hors prise de vue soit maintenu sur le même plan. Après la numérisation, l'opérateur repliera les documents en respectant les plis d'origines.

- **Documents en feuille conditionnés en boîte** : ne jamais toucher le document original, on utilisera le support sur lequel le document est fixé pour toutes les manipulations. Ne jamais faire tourner une feuille d'un document grand format. La prendre, la numériser puis la remettre dans le couvercle de la boîte face vers le haut, les autres images viendront s'empiler au dessus. Lorsque la totalité des documents de la boîte sont traités, on rempile ces documents un à un dans l'ordre initial de rangement dans le fond de la boîte. Ne jamais laisser un document original exposé à la lumière, le couvrir d'un papier permanent.

- **Documents photographiques** : certains fonds photographiques seront manipulés à la demande des conservateurs avec des gants de coton.

III.5 Les règles de contrôles.

On trouvera ci-dessous la liste codifiée des types d'erreurs qui sont gérés par les applications de contrôle à la BnF et qui permettent d'utiliser un vocabulaire commun entre celui qui a fait la numérisation et celui qui la contrôle de façon à ce qu'en cas d'erreur la compréhension de ce qui est erroné soit immédiate. Les codes erreurs s'appliquent aussi bien à la filière imprimés qu'à la filière images.

Information tronquée
Pagination incohérente

Incohérence de référence
Page inutile
Image mal positionnée ou non centrée
Travers hors tolérance ($>1,3^\circ$)
Changement d'échelle en cours d'ouvrage
TdM mal renseignée
Index mal renseigné
Page TdM non renseignée
Page INDEX non renseignée
Erreur de liens TdM/Index
Erreur de balisage TdM/Index
Polarité inversée
Erreur OCR HQ
Autre erreur majeure - commentaire libre
Visibilité insuffisante - trop clair/trop sombre
Image mal détournée
Trait(s) de saturation parasite(s)
Échelle de restitution incorrecte
Élément(s) étranger(s) à l'image
Mode traitement des images non homogène
Effet mosaïque
Marges noires ou blanches
Absence de commentaires
Erreur de saisie TdM/Index
Erreur de segmentation OCR
Erreur OCR
Autre erreur mineure - commentaire libre
Distorsions géométriques
Halos sur le contour des images
Ombres portées
Cache non utilisé / éléments hors contexte page visibles

Effet bombé
Défauts colorimétriques

IV Exemples de documents numérisés

V Exemple fonds d'écran de l'outil de consultation Gallica