

Numérisation des fonds patrimoniaux

1. Phases d'un projet de numérisation

La mise en place d'un projet de numérisation comporte trois phases principales :

La numérisation. Si certains établissements se dotent d'un laboratoire de numérisation pour traiter l'ensemble de la chaîne de travail, les projets de numérisation font souvent appel à des sociétés spécialisées qui sauront traiter dans les meilleurs délais des lots importants de documents.

L'indexation. Ce travail, proche des activités traditionnelles des bibliothèques et des archives est le plus souvent assuré par les établissements qui conservent les fonds. Cependant, les images numériques justifient souvent des descriptions fournies et un discours pédagogique étayé. Le recours à des spécialistes ou à d'autres établissements est souvent justifié (voir par exemple la base Medic@).

L'hébergement. Le recours à des bases de données spécifiques (GED) reste la règle. Mais l'hébergement, la mise en ligne et la sauvegarde des données numériques peuvent être confiées à des prestataires dans le cas où les établissements ne souhaitent pas se doter d'un équipement informatique lourd (serveurs, logiciels, maintenance). Les métadonnées s'appuient sur les éléments bibliographiques classiques et sur des informations techniques (noms d'images, formats, hiérarchisation des données). Le format XML favorise la diffusion sur le Web et le partage des données (notamment le moissonnage OAI).

2. Choix des fonds à numériser

Le choix des fonds à numériser doit tenir compte de critères matériels objectifs :

Redondance. Vérifier que les documents n'ont pas déjà été numérisés par d'autres établissements (voir les différents programmes existants à l'échelle nationale et européenne..

Droits d'auteur. Voir la fiche pratique de la BnF.

Partenariats. Etudier la possibilité de partenariats avec des institutions qui conservent des fonds complémentaires (autres bibliothèques, archives, musées, associations, particuliers...) ou qui constituent des références scientifiques pour l'expertise de ces fonds (universités, associations, chercheurs...).

Subventions. Répondre à des objectifs locaux ou nationaux permet de bénéficier de financements.

Manipulations. Vérifier si l'état des documents est compatible avec la numérisation et pour définir le mode de mise en place des ouvrages sous l'objectif. L'état des reliures conditionnera le mode d'ouverture des ouvrages et la prise de vue (1 ou 2 pages par image). Au cas par cas, on peut envisager de dé-relier un ouvrage.

Mode Texte ou Mode Image. Si l'on privilégie le texte, on choisira souvent un cadrage intramarginal pour éliminer les pixels inutiles. On anticipera également les traitements d'images ultérieurs destinés à renforcer les contrastes des caractères en vue du traitement par OCR. Si l'on privilégie le rendu des images, on choisira un cadrage qui permet de visualiser l'ensemble du document physique pour respecter un rendu réaliste. On sera également plus attentif à la fidélité des couleurs et à l'éclairage.

Numérisation sur site ou chez un prestataire. Si la numérisation est réalisée par un prestataire, on pourra exiger qu'elle soit effectuée sur le lieu de conservation des documents. Cette décision correspond surtout à la volonté ou la possibilité de laisser les documents quitter l'établissement.

Numérisation indirecte et traçage sur microfilm. La numérisation peut se faire à partir des supports de substitution déjà existants (microfilm, diapositive, ektachrome, tirage photo...). Les coûts seront sensiblement moins élevés (homogénéité des formats des supports, simplification des réglages et des cadences de numérisation), mais la qualité sera moindre en raison de la succession des opérations de reprographie (copie de copies). Ce choix reste lié à l'importance du rendu final des images et de la volumétrie du fonds. A l'inverse, la numérisation peut être associée à un traçage sur microfilm (création de microfilm à partir de fichiers numériques) pour produire un second support de substitution et de consultation.

Récolement et pré-indexation. Le processus de numérisation doit comporter la rédaction d'un fichier de récolement, qui récapitule les répertoires et les images produites, les problèmes rencontrés lors de la manipulation des ouvrages. Ce fichier doit s'appuyer sur les listes rédigées en amont du projet ou sur le cahier des charges en cas de prestation. Il constitue le document de suivi global des opérations : inventaire des documents à numériser, données techniques (noms des fichiers notamment), données de description et d'indexation.

Définition et volumétrie des images produites. On privilégiera une définition élevée pour les documents qui présentent un intérêt matériel (300 dpi en TIFF ou JPEG non compressé). Cette définition pourra être plus élevée encore pour des documents graphiques. En revanche, dans le cas de fonds purement documentaires (journaux, archives...), souvent très volumineux, on préférera des formats compressés (voir

les cahiers des charges de la BnF et des Archives de France). La définition doit être déterminée par l'utilisation future des images, sachant que le standard de référence est celui de l'imprimerie offset : 300dpi (ce standard nécessitera une définition de capture beaucoup plus grande si l'on prévoit des recadrage et des agrandissements). Attention toutefois, une définition trop importante entraînera des problèmes de netteté en réduisant la profondeur de champs lors de la prise de vue (problème important en cas de courbure des pages) et des fichiers très volumineux donc difficiles à stocker.

3. Prestations

Le cahier des charges doit être aussi précis que possible. En dehors des critères de prix et de délai, on peut baser la notation des propositions des différents prestataires sur les critères suivants :

Matériel utilisé : Dos numériques (capture directe en deux dimensions) ou caméras linéaires (balayage d'une ligne de capture le long du document). Le deuxième type permet des niveaux de définition beaucoup plus importants mais nécessite des temps de capture plus long donc plus coûteux. Il est important de vérifier que la définition du matériel est compatible avec le format des documents et la définition exigée dans le cahier des charges

Manipulation des ouvrages : Il sera difficile de vérifier qu'un prestataire respecte les prescriptions du cahier des charges. Seule une visite des ateliers permet de se rendre compte des méthodes de travail des prestataires. Elle permet également de vérifier que le matériel utilisé est celui annoncé.

Traitements numériques : correction des histogrammes colorimétriques, redressement des images pour corriger la courbure des pages, modification des contrastes voire ré-échantillonnage sont les principaux traitements numériques que le prestataire peut proposer. Si le dernier est à éviter (dégradation des images) les autres peuvent se justifier par l'utilisation prévue des images.

Calibrage du matériel. L'ensemble de la chaîne de production doit faire l'objet de calibrages à l'aide de sondes (scanners et écrans) utilisant des mires normalisées (Colorchecker et mire de résolution et quadrillage noir et blanc) : point blanc, point noir, gamma (luminosité), température de couleurs. L'utilisation de profils ICC (fichiers décrivant la manière dont un matériel informatique restitue les couleurs) permet d'assurer un suivi rigoureux du calibrage tout au long de la prestation.

Protocole et échantillonnage de contrôle. Certains critères comme le cadrage ou le nommage des fichiers doivent être vérifiés pour chaque image. D'autres, comme le contraste, la netteté et la colorimétrie peuvent être contrôlés en début et en fin d'ouvrage ou bien par séquence d'images. Chaque nouvelle séance de travail doit faire l'objet de réglages et de contrôles.

Tests préalables soumis au client : Des tests doivent être demandés en début de prestation pour valider les choix de numérisation.

Chaîne de travail : L'ensemble de la chaîne de travail doit être décrite dans la proposition. Elle est révélatrice de la rigueur et du professionnalisme du prestataire.

Conditions de sécurité : Le prestataire devra fournir des garanties en matière de transport, de stockage et d'assurance des documents (conditions de conservation, protection contre le vol et les catastrophes).

Liens	<p>Bibliographie et Webographie (liens consultés le 16 février 2009)</p> <p>Généralités :</p> <ul style="list-style-type: none"> - Groupe de travail Minerva n°6 : Identification des bonnes pratiques et des centres de compétences, <i>Guide des bonnes pratiques</i>, Rome, Minerva Project, 2004, 65 p. [en ligne] http://www.minervaeurope.org/structure/workinggroups/goodpract/document/bonnespratiques1_3.pdf - <i>Document numérique., Numérisation et patrimoine</i>. Paris, Hermès science publications, 2004. - <i>Conduire un projet de numérisation</i>, Villeurbanne, Tec et Doc Presses de l'Enssib (La boîte à outils), 2002. <p>Références citées dans le texte :</p> <ul style="list-style-type: none"> - Medic@ : http://www.bium.univ-paris5.fr/histmed/medica.htm - moissonnage OAI : http://www.culture.gouv.fr/culture/dll/OAI-PMH.htm#_Toc94010868 - Fonds numérisés à la BnF : http://gallica.bnf.fr/ - Fonds numérisés en France : http://www.numerique.culture.fr/mpf/pub-fr/index.html - Fonds numérisés en Europe : http://www.michael-culture.org/fr/home - Numérisation à la BnF : http://www.bnf.fr/pages/zNavigat/frame/infopro.htm?ancre=numerisation/num_pro.htm - Appel à projets du Ministère de la Culture : http://www.culture.gouv.fr/culture/mrt/numerisation/fr/f_01.htm - Informations techniques MCC : http://www.culture.gouv.fr/culture/mrt/numerisation/fr/f_04.htm - Cahier des charges Archives de France : http://www.archivesdefrance.culture.gouv.fr/static/1308
--------------	---