



Di gi ti si ng
Content Together

Mi ni steri al Network
for Val ori si ng
Acti vi ti es i n Di gi ti sati on

0 1 > » [0 1 1 0
< 0
[0 1 1 0
œ ¥ 0 0 0 1
1 0 1 - -
- - - - 1 0 1 - -
0 1
0 1
0 1 1 0

œ ¥ 0 0 0 1
- - - - 1 0 1 - -
0 1
0 1 1 0
< < <

gui de des bonnes prati ques

version 1.3



édité par le Groupe de travail Minerva n° 6
Identification des bonnes pratiques
et des centres de compétence



guide des bonnes pratiques

guide des bonnes pratiques

version 1.3

édité par le Groupe de travail Minerva n° 6
Identification des bonnes pratiques
et des centres de compétence

3 Mars 2004

Minerva Working Group 6
Identification of good practices and competence centres

Coordinateur :

Borje Justrell (Riksarkivet, Suède)

guide des bonnes pratiques

version 1.3 (3 Mars 2004)

édité par le Groupe de travail Minerva n° 6
Identification des bonnes pratiques et des centres de compétence

Coordination générale

Rossella Caffo (Chef du projet Minerva)

Borje Justrell (Coordinateur du Groupe de travail
sur les bonnes pratiques)

Antonella Fresa (Coordinatrice technique Minerva)

Comité éditorial européen

David Dawson, Karl-Magnus Drake, Borje Justrell,
Muriel Foulonneau, Antonella Fresa, Anna Maria Tammaro

Secrétariat

Marzia Piccininno

(Ministero per i beni e le attività culturali, Italie)

Textes

Ciaran Clissmann (Pintail Ltd.),

Karl-Magnus Drake (Riksarkivet, Suède),

Borje Justrell (Riksarkivet, Suède),

Anna Maria Tammaro (Université de Parme, Italie)

Version Web

Maria Teresa Natale et Andrea Tempera

[http://www.minervaeurope.org/structure/workinggroups/goodpract/
document/goodpractices1_3.htm](http://www.minervaeurope.org/structure/workinggroups/goodpract/document/goodpractices1_3.htm)

Conception graphique

due_pavese

© 2004 Minerva Project

La version française a été réalisée pour le compte de la Mission de la recherche et de la technologie du Ministère de la culture et de la communication (partenaire français du projet Minerva) par Sarah Faraud (Relais Culture Europe) et Cécile Coquet-Mokoko (traductrice)

Remerciements:

Sans Ciaran Clissman, ce guide serait resté au stade des bonnes intentions. Outre son savoir concernant la structuration de l'information, l'art avec lequel il a su présenter les bonnes pratiques dans un anglais impeccable représente une contribution de taille à ce projet. S'il reste des imperfections dans le texte, la responsabilité en incombe aux seuls éditeurs, pour n'avoir pas toujours suivi les conseils de Ciaran Clissman.

1	Introduction	8
1.1	Présentation du document	
1.2	Structure du document	
2	Arrière - plan	11
2.1	Les Principes de l'und	
2.2	Le Projet Minerva	
3	Directives pratiques	12
3.1	Introduction	14
3.2	Planification du projet de numérisation	14
	<i>Origine du projet</i>	16
	<i>Ressources humaines</i>	17
	<i>Recherche</i>	18
	<i>Risques</i>	19
3.3	Sélection des sources pour la numérisation	20
	<i>Etablir des critères de sélection</i>	21
	<i>La sélection en fonction des critères</i>	22
3.4	Préparation à la numérisation	23
	<i>Équipement</i>	24
	<i>Logiciel</i>	27
	<i>Environnement</i>	29
3.5	Maniement des originaux	30
	<i>Déplacement et manipulation d'originaux</i>	31
3.6	Le processus de numérisation	32
	<i>Utilisation de scanners</i>	33
	<i>Utilisation d'appareils photo numériques</i>	35
	<i>Applications pour la Reconnaissance optique de caractères (OCR)</i>	36
3.7	La préservation d'originaux numérisés	37
	<i>Formats de fichiers</i>	38
	<i>Choix supports</i>	39
	<i>Stratégies de migration</i>	40
3.8	Les métadonnées	41
	<i>De l'intérêt d'utiliser les métadonnées en tant qu'objet</i>	42
	<i>Description</i>	43
	<i>Normes à appliquer aux métadonnées</i>	44
3.9	La publication	45
	<i>Traitement de l'image</i>	46
	<i>Questions liées à la 3D et à la réalité virtuelle</i>	47
	<i>Publication en ligne</i>	48

3.10	Propriété intellectuelle et droit de copie	49
	<i>Etablir le droit de copie</i>	50
	<i>Sauvegarder le droit de copie</i>	51
3.11	La direction de projets de numérisation	52
	<i>Diriger le processus de numérisation</i>	53
	<i>Monter une équipe</i>	55
	<i>Former le personnel</i>	56
	<i>Travailler avec des tiers pour l'assistance technique</i>	57
	<i>Travailler avec des tiers pour des projets de coopération et le partage des données</i>	58
	<i>Coûts</i>	59

introduction

1.1 Présentation du document

Ce document est le fruit du travail collectif du groupe de travail Minerva sur les bonnes pratiques. Il propose un guide pratique pour l'élaboration, l'exécution et la direction de projets de numérisation, en particulier dans le domaine culturel (bibliothèques, musées, archives). Il s'adresse en priorité aux équipes travaillant au sein d'institutions culturelles ou de concert avec elles, et qui prévoient ou ont déjà lancé des projets de numérisation. Ce document reflète les conclusions de la réflexion collective du projet Minerva, y compris l'important travail de recherche issu des questionnaires nationaux dépouillés en partenariat avec le Groupe de représentants nationaux (NRG) qui s'est réuni à Alicante en mai 2002.

1.2 Structure du document

Ce document se compose de deux chapitres:

- Arrière-plan
- Directives pratiques

Arrière-plan - Il s'agit de passer en revue les aspects principaux du projet Minerva et de décrire l'objectif de ce document dans l'économie générale du projet. On y présentera également le travail accompli jusqu'à présent, afin de fournir au lecteur une vision globale du contexte dans lequel ce document s'inscrit.

Directives pratiques - Les leçons pratiques les plus importantes qu'ait tirées l'équipe "bonnes pratiques" du projet Minerva ainsi que les informations qu'elle a collectées seront présentées ici. L'accent sera mis sur un certain nombre de points de repère utiles aux organisations qui prévoient, exécutent ou dirigent des projets de numérisation dans la sphère culturelle. Ces directives sont réparties en différents domaines, chacun correspondant à une étape précise dans le cours d'un projet de numérisation :

- Planification du projet de numérisation
- Sélection des sources pour la numérisation
- Préparation pour la numérisation
- Maniement des originaux
- Processus de numérisation
- Préservation d'originaux numérisés
- Métadonnées
- Publication
- Propriété intellectuelle et droit de copie
- Direction de projets de numérisation

Chaque directive est présentée de façon pragmatique, à l'intention d'équipes ayant à traiter des projets concrets.

On pourra trouver des informations complémentaires ainsi qu'un support technique sur le site internet de Minerva:

www.minervaeurope.org

Ce site propose des références utiles donnant des exemples de bonnes pratique (<http://www.minervaeurope.org/listgoodpract.htm>), des centres de compétences (<http://www.minervaeurope.org/competencecentre.htm>) des modèles actuellement employés dans le champ culturel européen ainsi que des liens utiles vers des ressources en ligne.

Il existe plusieurs autres sources proposant des directives sur la numérisation et la création de documents numérisés à contenu culturel. Les plus importantes sont énumérées dans la liste sélective de directives sur la numérisation (<http://www.minervaeurope.org/guidelines.htm>) incluse dans le présent guide et consultable sur le site web de Minerva.

Précisons enfin que certaines questions techniques pointues ont été délibérément laissées de côté dans ce guide, dans un souci de brièveté et de concision. Pour toute information complémentaire, s'agissant notamment des normes, on consultera les *Directives techniques Minerva* (<http://www.minervaeurope.org/structure/workinggroups/servprov/docindex.htm>), sur le site web Minerva.

2.1 Les principes de Lund

Le 4 avril 2001, des représentants et experts de la Commission européenne et des Etats-membres se sont réunis à Lund en Suède (durant la présidence suédoise) pour se concerter au sujet de la coordination et de la valorisation des programmes nationaux de numérisation au niveau européen. La réunion a débouché sur la publication d'une série de principes généraux devant orienter les initiatives publiques de numérisation et leur coordination. Ces principes, connus sous le nom de principes de Lund, sont devenus le plan d'action de Lund, qui établit une liste d'actions à réaliser par les Etats-membres, par la Commission, et par les deux conjointement, en vue d'améliorer l'approche de la numérisation à travers l'Europe.

2.2 Le projet Minerva

Ce document est issu du projet Minerva, lancé en 2002 sous l'égide du Ministère italien de la culture (contrat IST 2001-35461). Ce projet rassemble des représentants des ministères concernés ou d'agences publiques de nombreux Etats-membres, dans l'objectif commun de promouvoir une approche partagée et une méthodologie commune pour la numérisation de contenus culturels européens. Ce projet tient compte de la valeur propre au patrimoine culturel européen, ainsi que le rôle stratégique qui peut être le sien dans l'essor de l'industrie du numérique en Europe. Il tient aussi compte de l'importance de la coordination des efforts des gouvernements nationaux et des organisations culturelles, afin d'accroître la qualité de la synthèse et de la synergie entre les différentes initiatives de numérisation.

Le projet Minerva comprend plusieurs groupes de travail aux axes bien définis au sein d'un ensemble cohérent. Chacun de ces groupes de travail est constitué d'experts désignés par les partenaires du projet, qui travaillent ensemble sur un aspect précis des objectifs du projet. Les objectifs de chaque groupe de travail sont décrits sur le site web du projet à l'adresse suivante: <http://www.minervaeurope.org/structure/workinggroups.htm>.

L'organisation en groupes de travail a permis au projet d'explorer de front un certain nombre des domaines les plus importants de la numérisation.

Le projet comprend les groupes de travail suivants:

- Cadre d'évaluation;
- Interopérabilité et fourniture de services;
- Inventaires, découverte de contenus numérisés, questions liées au multilinguisme;
- Identification des besoins des utilisateurs, normes de contenus et de qualité pour les points d'accès communs;
- Identification des bonnes pratiques et des centres de compétences.

Les activités des groupes de travail sont les suivantes: réunions, ateliers publics, publications (dont le présent guide), coordination et coopération internationales, etc.

3.1 Introduction

Ce chapitre présente l'essentiel du contenu du guide. Il propose des directives pratiques à l'intention d'institutions ou d'organismes qui ont lancé des projets de numérisation ou envisagent d'en lancer. L'accent est mis sur la sphère culturelle; cependant, le contenu est largement applicable à d'autres domaines (tels que le tourisme ou la gestion de documents).

Le contenu de ce chapitre est détaillé en fonction des différentes étapes du processus de numérisation, de sorte que le lecteur puisse identifier aisément ce qui a trait à son propre travail, quel que soit le stade où il en est de son projet personnel. L'hypothèse de départ est que la plupart des utilisateurs de ce guide en seront à la première étape de leur projet (la planification); toutefois, une partie de l'information proposée ici devrait être utile à n'importe quel projet de numérisation.

Les étapes du processus de numérisation ont servi de trame à l'énumération des directives et sont détaillées dans le sommaire.

La description de chaque directive est structurée de la façon suivante:

- Titre de la directive
- Définition de la problématique amenant la directive et/ou présentant le problème que traite cette directive
- Liste de suggestions pratiques visant à faciliter un aspect précis concernant la mise en place ou l'exécution d'un projet de numérisation
- Notes ou commentaire fournissant des informations supplémentaires, le cas échéant. Leur présence ne sera donc pas systématique.

Des références complémentaires sont fournies sur le site internet :
www.minervaeurope.org

Ni les directives ni les références ne sont exhaustives; cependant, elles proposent les informations indispensables à un projet qui traite une (des) tâche(s) particulière(s) au sein d'un processus de numérisation.

3.2 Planification du projet de numérisation

La planification est la première étape de tout projet de numérisation. Le temps qui y sera consacré sera rentabilisé par une gestion et une exécution plus aisées du projet. En règle générale, il convient de répondre aux questions suivantes:

- Quoi (quel est le travail à accomplir) ?
- Qui (doit le faire) ?
- Où (doit-il se faire) ?
- Quand (se fera-t-il) ?
- Comment (se fera-t-il) ?

Un projet de numérisation doit avoir des buts et objectifs clairement spécifiés; ils détermineront précisément la sélection, le droit de copie et la publication. Le projet doit être servi par un personnel adéquat, possédant les compétences et le savoir-faire requis, ainsi qu'un dispositif de formation prêt à fournir toute expertise complémentaire qui pourrait être nécessaire à sa bonne exécution.

On ne devrait pas lancer un projet tant que d'autres projets participant au même domaine n'ont pas fait l'objet de recherches. Car ces recherches serviront à identifier les questions à traiter, à faire émerger de nouvelles idées et des domaines qui peuvent avoir été négligés, et ajouteront de la valeur et de la crédibilité au projet et à ses résultats.

Grâce aux recherches, on pourra également évaluer la somme de travail nécessaire à l'exécution du projet, par le biais de réunions ou d'échanges avec des organisations ayant déjà mis en place des projets similaires. De telles interactions permettront de savoir si votre organisation dispose du personnel, des compétences et de l'infrastructure technologique adéquats pour mener à bien votre projet, ou bien s'il vous faudra prévoir une formation et une préparation spécifiques.

On pourra consacrer avec profit un peu de temps à la vérification des droits des documents à numériser. Un projet de numérisation, même servi par une grande expertise technique et beaucoup d'expérience, peut échouer simplement parce qu'on aura négligé d'obtenir la permission de numériser et de publier des documents sur le Web.

On pourra enfin envisager de mettre en place un pilotage technique dès le lancement du projet, afin de s'assurer que d'éventuels anomalies ou problèmes liés au rythme de travail seront résolus avant le commencement du projet principal.

Origine du projet

Définition de la problématique

Chaque projet de numérisation a sa raison d'être. Souvent, il s'agit de fournir un accès Internet à des ressources culturelles qui, sans cela, seraient insuffisamment utilisées, ou bien de protéger des ressources fragiles de l'usure provoquée par un accès direct. Dans d'autres cas, les projets représentent des exercices de coopération entre différents organismes, et comprennent l'élaboration de portails, de réseaux, etc.

Les motivations à l'origine du projet seront déterminantes pour définir les critères de sélection des sources à numériser. Elles influenceront également sur la direction du projet, les méta-données, l'éventuelle publication en ligne des résultats du projet, le contrôle qualité, etc. Aussi la question "pourquoi" est-elle la plus importante à se poser avant de lancer un projet de numérisation.

Suggestions pratiques

- Le projet doit avoir des objectifs explicites et concrets, qui auront fait l'objet d'une description détaillée.
- Ces objectifs doivent être réalistes au regard des ressources dont on dispose.
- Chacune des étapes du projet doit être validée en fonction de ces objectifs, de façon à s'assurer que le travail accompli dans le cadre du projet est bien en conformité avec les directives à appliquer.
- Les objectifs du projet doivent présenter de façon détaillée l'apport qu'il fournira aux institutions qui s'y sont impliquées. Dès lors que l'on accepte d'investir du temps et de l'énergie dans un projet, la justification de ce dernier, d'un point de vue institutionnel, se doit d'être parfaitement claire.

Ressources humaines

Définition de la problématique

Avant de pouvoir monter un projet, il importe de disposer du personnel requis pour y travailler. De nombreux organismes culturels n'ont pas un grand nombre d'employés disposant d'assez de temps libre pour mettre en place des projets de numérisation en plus de leurs obligations ordinaires. En outre, le savoir-faire nécessaire à l'exécution d'un projet de numérisation peut différer des compétences correspondant au traitement des problèmes quotidiens de fonctionnement. Les solutions matérielles et logicielles que requiert un projet de numérisation doivent aussi être identifiées.

Suggestions pratiques

- S'assurer que l'on dispose d'assez de personnel pour mener à bien le projet.
- Assigner à chaque membre du personnel une tâche ou un ensemble de tâches identifiées au sein du projet.
- Identifier les besoins en formation, notamment en matière de formation à la technologie de l'information et d'apprentissage du maniement de documents et d'objets fragiles.
- Assurer la formation, si possible, à l'aide des solutions logicielles et matérielles qui seront utilisées pour exécuter le projet, avant le lancement de celui-ci (les vendeurs proposent parfois des solutions techniques à titre gratuit pour la formation; ou bien on peut louer les équipements nécessaires pour une courte durée).
- Mobiliser autour du projet un petit noyau de personnel compétent plutôt qu'un groupe important de personnel "occasionnel".

Notes/Commentaire

Même si le contenu présenté dans ce guide correspond à tous les scénarii de direction de projets, il n'est pas inutile d'y insister encore : en effet, on peut mettre en péril des documents et objets irremplaçables si ces ressources ne sont pas manipulées correctement.

Recherche

Définition de la problématique

Quelle que soit l'ampleur du projet, on peut partir du principe que des projets similaires ont été menés à bien par le passé. Il y a de fortes chances pour que l'on puisse trouver des informations sur de tels projets sur Internet, ou dans des publications spécialisées, etc.

Effectuer des recherches dans le domaine tandis que l'on en est encore au stade de la planification du projet peut permettre d'identifier les solutions logicielles et matérielles qu'il est préférable d'utiliser, de faire des projections sur le rythme de travail et le déroulement du processus, et enfin d'éviter les problèmes et écueils rencontrés lors de projets précédents.

Suggestions pratiques

- Lorsque vous planifiez votre projet, effectuez dès que possible des recherches sur d'autres projets traitant des mêmes questions que le vôtre. Le présent guide représente un point de départ, mais les contenus que l'on peut trouver sur Internet sont la source d'information la plus large et la plus complète.
- La recherche permet d'éviter de commettre des erreurs. Elle peut aussi mettre l'équipe chargée du projet en contact avec d'autres ayant mené des projets similaires, et lui donner l'occasion de bénéficier de leurs expériences.
- La recherche effectuée ajoute à la crédibilité et à la valeur du projet et de ses résultats. Le fait de savoir que vous n'avez pas exécuté votre projet en vase clos, mais en prenant en compte le travail fourni par d'autres, mettra mieux en valeur les résultats de votre projet.

Notes/Commentaire

De nombreux projets de numérisation sont financés par des fonds publics, et leurs organisateurs sont tenus de publier leurs résultats et leurs compte-rendus. Ce type de publications est presque toujours consultable sur Internet ou d'autres supports adaptés.

Les équipes chargées de projets sont généralement prêtes à partager leurs expériences et leurs résultats car cela ajoute de la valeur à leur travail.

Risques

Définition de la problématique

Lors du lancement d'un projet, il faut tout mettre en œuvre pour en garantir le succès. Cependant, l'objectif n'est pas d'éliminer tous les risques, mais de s'y préparer en créant un cadre qui soit en mesure de parer aux imprévus de manière efficace et imaginative. Le but est de monter un projet avec un personnel et des procédures aptes à suivre les changements. C'est pourquoi toute planification de projet se doit d'inclure une analyse des risques.

Suggestions pratiques

- La diffusion d'images numérisées sur Internet représente une forme de publication. Aussi est-elle couverte par les lois de droit de copie et le code de la propriété intellectuelle. Une analyse des risques comprend des questions telles que celles-ci :
 - Quelles pourraient être les conséquences de l'utilisation des contenus sans autorisation spécifique ?
 - A-t-on fait la démarche de trouver le détenteur des droits ?
 - S'est-on assuré que l'on détient les droits sur les sources ?Pour l'information du public, la valeur légale de l'information est une question importante. Quelles dispositions a-t-on prises pour garantir qu'une source numérisée n'a pas été altérée et a bien été produite par une institution autorisée ? L'authenticité doit aussi être garantie. Quelles actions a-t-on prises pour préserver les fichiers image, et quels outils a-t-on utilisés ?
- Le financement du projet peut être problématique et, partant, représenter un risque pour l'accomplissement des objectifs du projet.
- Le niveau de compétences requis par le projet est une question cruciale. Est-il possible d'engager du personnel hautement qualifié et expérimenté ? Sinon, cela aura-t-il des conséquences sur la planification du travail à accomplir sur le projet ?

3.3 Sélection des sources pour la numérisation

La sélection des sources à numériser représente une décision d'importance dans le cadre d'un projet de numérisation.

En règle générale, la meilleure option est de numériser l'intégralité d'une collection (de livres ou d'objets d'art); mais cela est rarement faisable, et il faut faire des choix.

Les critères de sélection différeront en fonction des objectifs du projet de numérisation : un site de ressources en ligne à destination des écoles optera pour la numérisation de contenus correspondant à un programme, tandis qu'un musée pourra numériser seulement ses collections les plus connues de façon à accroître le nombre de visiteurs, ou bien ses objets les plus fragiles de façon à réduire le plus possible les demandes de consultations directes. Ce ne sont là que quelques-unes des questions à se poser lorsqu'on définit les critères de sélection, car les motivations qui détermineront la décision de numériser (ou non) telle source varieront d'un projet à l'autre. Parmi ces motivations on peut encore citer les contraintes juridiques, les politiques institutionnelles, les problèmes techniques liés à la numérisation, l'existence d'autres copies numériques du même contenu, etc.

Etablir des critères de sélection

Définition de la problématique

Lorsqu'on planifie un projet de numérisation, le choix des sources à numériser est essentiel. Les critères de sélection dépendent non seulement des objectifs du projet, mais aussi des contraintes techniques et financières, des problèmes de droit de copie et de propriété intellectuelle, et enfin de l'existence d'autres projets dans le même domaine.

Suggestions pratiques

- Il est indispensable d'établir des critères pour sélectionner les sources à numériser.
- Les critères de sélection doivent refléter les objectifs de l'ensemble du projet. Les critères suivants doivent être pris en compte, mais la liste n'est pas exhaustive :
 - L'accès à des sources qui, laissées en l'état, seraient indisponibles ou d'une disponibilité limitée ;
 - Un accès plus large et plus aisé à des sources très demandées ;
 - L'état de conservation des originaux ;
 - La préservation d'originaux fragiles par la mise en circulation de versions numériques que l'on proposera à la consultation ;
 - Le thème du projet ;
 - Le droit de copie et la propriété intellectuelle ;
 - La disponibilité de versions numériques déjà existantes ;
 - Le coût de la numérisation ;
 - L'adéquation des sources à une présentation en ligne.
- Les critères de sélection doivent être présentés de façon explicite, faire l'objet de concertations et être entérinés par toutes les parties concernées, avant que ne s'opèrent la sélection ou la numérisation.
- Les critères de sélection doivent faire l'objet d'une présentation détaillée, de sorte que les motifs justifiant la décision de numériser ou de ne pas numériser soient clairs tout au long de la durée du projet.

Notes/Commentaire

Le plus souvent, les organismes culturels disposent de fonds de grande valeur et d'un grand intérêt pour les usagers, qui entrent par défaut dans tout projet de numérisation devant représenter l'institution.

Un nombre significatif de projets de numérisation visent la publication en ligne sur Internet. Ceci implique la nécessité, avant de procéder à toute sélection, de se poser la question des droits de reproduction et de la propriété intellectuelle pour chacune des sources que l'on envisage de numériser.

La sélection en fonction des critères

Définition de la problématique

Une fois posés les critères de sélection des sources à numériser, le processus de sélection peut être lancé. Ce guide vous propose quelques conseils sur la marche à suivre.

Suggestions pratiques

- Chacune des sources que l'on envisage de numériser doit être évaluée en fonction des critères de sélection. Si l'un d'eux n'est pas rempli, on doit le noter. Dans l'éventualité où cela aurait pour conséquence le rejet d'objets importants ou essentiels, on devrait alors réviser les critères de sélection. Au cas où cela se produirait, il faudrait alors noter les nouveaux critères.
- Dès lors qu'un objet a été sélectionné pour la numérisation, il faut entrer les détails le concernant dans la base de connaissances du projet de numérisation (voir le chapitre intitulé *La direction de projets de numérisation*).

Notes/Commentaire

A ce stade, le projet se dessine avec tous les objets qu'il prévoit de numériser. C'est là le moment opportun pour créer une base de connaissances rassemblant toutes les sources concernées par le projet.

Une telle base de connaissances permettra d'appuyer la direction du projet et de s'assurer que, par exemple, on a acquis les compétences nécessaires à la manipulation d'objets rares, ou, plus banalement, que l'on sait où se trouvent les originaux.

3.4 Préparation à la numérisation

Un environnement approprié ainsi qu'un système matériel/logiciel doivent être mis en place avant de procéder à la numérisation. Les composants d'un tel environnement comprennent le matériel pour le processus de numérisation proprement dit (tels que des scanners, des appareils photo numériques, des supports de copie, etc.), une infrastructure informatique à laquelle ce matériel sera connecté, du logiciel pour la capture de l'image et son traitement, ainsi que des logiciels pour les métadonnées et le contrôle qualité. L'environnement de travail doit être approprié aux sources à numériser : on veillera tout particulièrement à la qualité de la lumière, au taux d'humidité, aux vibrations, au rangement et au déplacement des originaux, etc.

Équipement

Définition de la problématique

Pour que l'on puisse procéder à la numérisation, il faut avoir mis en place un équipement technique adéquat. Il s'agit généralement d'appareils de capture d'images numériques (appareils photo numériques, scanners pour les livres, documents et microfilms, matériel audio et vidéo, si nécessaire), connectés à une plate-forme informatique adaptée (ordinateur, système d'exploitation, réseau, etc.) On peut distinguer deux méthodes de numérisation, reposant sur du matériel différent : l'une a recours au scanner et l'autre à l'appareil photo numérique.

Suggestions pratiques

- Avant de procéder à la numérisation, le matériel adapté doit être installé et sa qualité et son bon état de fonctionnement avoir été vérifiés.
- On doit procéder à des tests afin d'évaluer les outils de capture d'images numériques.
- Tant que l'environnement matériel n'a pas été totalement installé, testé et validé à l'aide d'objets non sensibles, ne faut pas traiter les sources concernées par le projet. Pour la plupart des projets de numérisation, on utilisera un scanner à plat pour toutes les sources qui ne risquent pas d'être endommagées par ce dispositif (telles que des documents imprimés non reliés ou des feuillets manuscrits). Le projet doit être équipé du scanner le plus grand qui existe sur le marché: il faut éviter de plier les sources ou de les scanner "en mosaïque". Il ne faut pas toutefois perdre de vue le fait que le transport de grands scanners (de type A0) n'est pas une mince affaire.
- En général, il ne faut avoir recours au scanner à plat que lorsque les documents eux-mêmes sont déjà plats et ne risquent pas d'être endommagés par une pression exercée sur une surface plane. Un scanner pourvu d'un repose-livres peut être adapté à de nombreux documents reliés, tant que les limites de format sont respectées. La plupart des projets de numérisation nécessiteront un appareil photo numérique, pour la capture de sources qui ne peuvent pas être aplaties ou maintenues sur le repose-livres d'un scanner à plat.
- Si l'on a recours à un scanner, l'idéal est qu'il soit aussi large que le document à numériser.

- On doit procéder à la capture d'images (que ce soit par scanner ou appareil photo numérique) avec la plus haute résolution possible, dans les limites du raisonnable. En conséquence, on aura des fichiers originaux très lourds; mais on pourra en extraire des fichiers plus petits pour des utilisations telle que la diffusion sur Internet. En revanche, on ne pourra jamais tirer une image de qualité supérieure à partir d'une image de qualité inférieure.
- On définira ce qu'est une résolution "dans les limites du raisonnable" en fonction de la nature des sources à numériser, et des usages que l'on prévoit de faire de l'image numérisée. Si, par exemple, les images numérisées ne doivent pas servir à autre chose qu'à être des vignettes sur l'écran, on peut se permettre une basse résolution. De même, la résolution doit capturer les détails les plus spécifiques de l'objet; donc si la numérisation à haute résolution ne donne pas plus d'informations qu'une numérisation à basse résolution, il sera difficile de justifier le choix de la haute résolution.
- La capture de l'image doit être faite selon un format de fichier qui évite les pertes de données, c'est-à-dire qu'il ne doit pas être compressé. C'est pourquoi on emploie le plus souvent le format TIFF (Tagged Image File Format). Il est préconisé de conserver une copie originale de l'image au format RAW utilisé par l'appareil photo ou par le scanner. Cela sera très utile lors des futurs traitements des images dans le cadre d'éditions en haute définition, ou de zooms.
- Si l'on a recours à un appareil photo numérique, il faut choisir l'appareil le plus puissant et sensible que le budget permette d'acheter.
- Les limites du matériel de numérisation ne peuvent être palliées par un quelconque traitement ultérieur de l'image. Précisons que le "zoom numérique" ne donne pas une image de meilleure qualité mais affiche simplement moins de pixels par unité de vue. Pour capturer les détails, trois paramètres principaux entrent en ligne de compte: le nombre de pixels de l'image, le nombre de bits et la qualité de la lentille optique utilisée.
- Il importe de disposer de supports adéquats pour y déposer les sources pendant leur numérisation.
- Il faut utiliser un appareil numérique doté d'un support de copie adapté. L'appareil photo doit être monté sur un trépied ou sur une colonne élévatrice et disposer d'autant de sources de lumière, de filtres etc. que nécessaire. Il est recommandé de consulter un expert en photographie numérique possédant une expérience dans des projets similaires, dans la mesure du possible, avant de mettre en place l'environnement matériel.

- Le plan photographique et le plan de la source à numériser doivent être exactement parallèles, afin de ne pas provoquer de distorsion de l'image.
- Il faut prévoir une luminosité adéquate lorsqu'on s'apprête à photographier les sources avec un appareil numérique : il est très rare que la lumière ambiante suffise.
- On doit utiliser des filtres adaptés afin de réduire les distorsions de couleurs.
- Les appareils doivent être connectés à un ordinateur aux capacités de stockage suffisantes. Cet ordinateur doit faire l'objet de sauvegardes régulières : cet impératif témoigne du coût élevé en temps et en technologie, ainsi que de l'usure possible des originaux, qu'implique le processus de numérisation.
- Si un objet doit être scanné en plusieurs fois, il faut prévoir un chevauchement de plusieurs centimètres afin de s'assurer qu'il n'y aura pas de discontinuité entre les parties. Il faut utiliser le même environnement, les mêmes lumières, etc., pour toutes les parties afin d'éviter l'effet "patchwork".

Notes/Commentaire

Le matériel employé représente une contrainte majeure pour la qualité du résultat final d'un projet de numérisation. A moins que le projet ne consiste à ne numériser que des documents plats que l'on peut scanner sans endommager les reliures, les cadres ou la source elle-même, l'utilisation d'un appareil photo numérique sera d'une grande importance. Même si l'on peut avoir recours à un appareil analogique, puis scanner les diapositives ou les tirages, un appareil numérique de haute qualité permet d'épargner beaucoup de temps et d'efforts en offrant de nombreux avantages en termes de qualité.

Si le projet a une durée de vie limitée, il peut être indiqué de louer du matériel informatique. Une autre solution est de faire appel à des agences extérieures pour exécuter la numérisation pour le compte des organismes culturels impliqués dans le projet.

Logiciels

Définition de la problématique

Une fois que l'on a créé une version numérique de l'objet, il est probable que le fichier correspondant devra être traité pour pouvoir être utilisé. Il se peut qu'il faille ajuster les couleurs, supprimer des détails superflus sur les bords de l'image, etc. De plus, les fichiers maîtres sont souvent très lourds, donc on aura probablement besoin d'un fichier plus petit, de format compressé (par exemple, un format vignette, pour diffusion sur Internet).

Suggestions pratiques

- Lorsqu'on met en marche le scanner ou l'appareil photo numérique, un processus de mise au point doit démarrer automatiquement.
- On aura besoin de logiciels adaptés au traitement de l'image pour retraiter les fichiers maîtres selon les objectifs poursuivis par le projet. Or, bien que le matériel de numérisation soit presque toujours vendu avec quelques logiciels, ceux-ci ne sont généralement pas assez puissants ni souples pour la plupart des projets.
- Les caractéristiques que doivent présenter les logiciels ne sont pas les mêmes d'un projet à l'autre. Il est utile de préciser que, tandis que les fichiers maîtres ne pourront être modifiés en aucune façon, on pourra recourir à plusieurs types de logiciels différents pour les traiter. Toutefois, cela peut impliquer un grand investissement en temps et en énergie, qui sera plus coûteux en définitive qu'un investissement dans un logiciel plus puissant.
- Le projet devra donc être pourvu des logiciels les plus puissants et adaptés que son budget lui permettra.
- Le logiciel doit, au minimum, être en mesure d'exécuter les tâches suivantes :
 - ouvrir des fichiers images très lourds ;
 - modifier la résolution et la profondeur des couleurs ;
 - sauvegarder plusieurs versions différentes, dans différentes tailles de fichiers ;
 - sélectionner et copier une partie d'une image pour la sauvegarder en un fichier indépendant ;
 - exporter des images dans différents formats de fichiers, notamment dans les normes communes du Web que sont les formats JPEG et GIF.

Il existe un certain nombre de logiciels gratuits qui offrent ce niveau de fonctionnalité; cependant l'investissement dans un produit commercial sera probablement plus rentable en termes d'énergie, de temps, de documentation et de support technique.

- Si le projet de numérisation a une composante OCR (reconnaissance optique de caractères), le choix d'un logiciel adapté est tout aussi essentiel. Toute procédure d'OCR implique, jusqu'à un certain point, des corrections et une édition manuelles; et la manière dont le logiciel utilisé répondra à cette contrainte aura des conséquences directes sur le temps et l'énergie consacrés à l'exécution du projet. Les meilleurs logiciels OCR permettent de d'afficher et de modifier le texte sur un seul écran, suggèrent des corrections pour des mots que le logiciel n'a pas pu lire correctement, acceptent les présentations de textes en plusieurs colonnes (comme dans les maquettes de journaux), etc.

L'évaluation de plusieurs logiciels d'OCR peut s'avérer extrêmement utile si l'ampleur du projet réclame, par exemple, la mobilisation de plus d'une personne sur une année.

Notes/Commentaire

Choisir le bon logiciel, c'est faire gagner au projet beaucoup de temps et d'énergie. S'il est de longue durée (s'il mobilise, par exemple, plus de deux personnes pendant plus de six mois), cela vaudra la peine d'évaluer plusieurs logiciels afin de déterminer quel est celui qui correspond le mieux aux besoins du projet.

Environnement

Définition de la problématique

De nombreux objets rares ou fragiles exigent un environnement spécifique. Il est donc essentiel de veiller à ce que le processus de numérisation cause le moins de dommages possible aux sources. C'est pourquoi de nombreux projets de numérisation requièrent un environnement adéquat.

Suggestions pratiques

- L'environnement du processus de numérisation est extrêmement important.
- On doit demander l'avis d'experts afin de s'assurer que le maniement d'originaux soit envisagé de la meilleure façon possible. Cela inclut la mise en place d'un environnement convenable pour leur numérisation.
- Les locaux prévus pour la numérisation doivent être exclusivement consacrés à l'exécution du projet pendant toute la durée de celui-ci. Trop de mouvements, de réorganisations de l'espace de travail peuvent conduire à l'endommagement, à la perte de sources, sans compter la perte de temps pour l'exécution du projet.
- Si les sources exigent des conditions particulières de luminosité, d'humidité, etc., il faut reproduire ces conditions aussi scrupuleusement que possible dans l'environnement du processus de numérisation. Pour certains objets, tels que des documents en cuir, le fait d'augmenter brièvement le taux d'humidité peut faciliter l'assouplissement du matériau avant qu'on ne l'aplatisse pour le photographier ou le scanner.
- Dans pratiquement tous les cas, il est déconseillé d'exposer les sources à une lumière vive (telle que la lumière du soleil) pendant de longues périodes. Bien entendu, il faut interdire que l'on fume, boive ou mange à proximité des objets.

Notes/Commentaire

L'ampleur et le budget du projet peuvent rendre impossible la mise à disposition d'un environnement exclusivement consacré à son exécution. Néanmoins, les points soulignés ici, tels que de réduire au maximum le mouvement, la manipulation et le changement brusque des conditions de conservation des sources, ne doivent pas être perdus de vue.

3.5 Le maniement des originaux

Cette section décrit la façon dont un projet de numérisation doit anticiper le traitement des sources à numériser. Dans de nombreux cas, ces sources sont rares ou de grande valeur; aussi doit-on réduire au maximum les effets néfastes que peut avoir sur elles la numérisation.

En tout état de cause, il faut souligner que les compétences des spécialistes chargés de la conservation des sources tout au long de l'année seront d'un secours précieux à l'équipe chargée du projet.

Déplacement et manipulation d'originaux

Définition de la problématique

Dans un grand nombre de cas, les objets à numériser sont particulièrement sensibles ou fragiles. Leur numérisation est d'ailleurs souvent demandée pour leur éviter une trop grande manipulation par les usagers. Il est donc extrêmement important qu'un projet de numérisation inclue toutes les mesures nécessaires pour éviter que les sources soient endommagées durant leur numérisation. Ces mesures vont du choix d'un matériel adapté à la mise en place d'un micro-climat propice à leur conservation, en passant par le déplacement du centre de numérisation jusqu'au site où les sources sont conservées, plutôt que l'inverse.

Pratiques Suggestions

- Consulter la personne responsable de la conservation des sources avant de les déplacer ou de les manipuler. Entrer les informations sur leur maniement dans la base de connaissances du projet de numérisation (voir le chapitre sur la Direction du processus de numérisation).
- Rester souple : on peut surmonter un contretemps dans un projet de numérisation, tandis qu'un dommage causé sur un objet unique peut être irrémédiable.
- Si nécessaire, apporter le matériel de numérisation (appareil photo numérique) sur le site où est conservé l'objet, plutôt que de le transporter hors du site.
- Éviter de défaire la reliure de livres ou de dossiers reliés. Préférer à l'utilisation d'un scanner à plat celle d'un scanner équipé d'un repose-livres, ou d'un appareil photo numérique.
- Toujours retirer les trombones, agrafes et autres attaches : ils peuvent endommager aussi bien les appareils de numérisation que les objets à numériser.
- Demander l'avis d'un expert (en particulier la personne chargée de la conservation de la source à numériser) avant toute manipulation de l'original.
- Demander cet avis avant de lancer la numérisation, de préférence au moment où l'objet est sélectionné pour être numérisé. Enregistrer ces conseils dans la base de connaissances du projet, puis les consulter avant de déplacer ou de numériser l'objet. Si besoin, présenter à l'expert les avantages et inconvénients de chaque appareil de numérisation que l'on envisage d'utiliser.

Notes/Commentaire

Même si la plupart de ces conseils coulent de source, il est très important de poser et de maintenir une discipline tant que l'on manipule les originaux.

3.6 Le processus de numérisation

Dans ce chapitre, on trouvera des directives pratiques pour mener à bien le processus de numérisation. Il existe différentes solutions techniques de capture numérique. On traitera ici plus particulièrement des scanners, des appareils photo numériques ou d'applications et logiciels de reconnaissance optique de caractères (OCR), car la plupart des projets de numérisation y ont recours. On ne traitera pas de la numérisation d'originaux transparents tels que les microfilms.

Utilisation de scanners

Définition de la problématique

Les scanners à plat sont très communément utilisés pour numériser. Les modèles A4 et A3 les plus courants sont relativement bon marché, n'exigent pas de grandes compétences techniques, et ont une bonne capacité de traitement, dès lors qu'on a établi un rythme de travail. Les grands modèles (jusqu'à A0) de scanners à plat et de scanners équipés d'un repose-livres sont très onéreux et correspondent donc à des projets à long terme, devant numériser un grand nombre de sources, ou des sources de dimensions exceptionnelles.

Suggestions pratiques

- Ne scanner sur un scanner à plat que les documents qui ne seront pas endommagés par la pression sur une surface dure. Consulter les experts en cas de doute.
- S'assurer que la vitre du scanner est toujours parfaitement nette. Cela garantit une meilleure qualité de l'image et évite que le document ne soit sali.
- Dans la mesure du possible, ne scanner que les objets aux dimensions du scanner, qu'il s'agisse d'un scanner à plat ou d'un appareil équipé d'un repose-livres en une pièce.
- Si l'on ne peut faire autrement que de scanner un document en plusieurs fois, s'assurer que le chevauchement est suffisant pour permettre un réassemblage de l'image lors du traitement postérieur (où l'on aura recours à des logiciels de mosaïque).
- Tester le scanner et évaluer ses capacités sur des documents non sensibles avant de commencer à numériser des originaux. Pour former les utilisateurs, on emploiera les mêmes documents non sensibles.
- Etablir une convention de nommage des fichiers pour tous les fichiers que produira le scanner: utiliser par exemple le système de catalogage existant ou leur donner des noms qui fassent sens. Le nom du fichier doit permettre de faire le lien entre le fichier et la source numérisée.
- Afin de faciliter au maximum l'exportation des fichiers d'une plate-forme informatique à une autre, il est conseillé de s'en tenir à des noms de fichiers d'au maximum huit caractères, suivis d'une extension d'au maximum trois caractères.
- Avant de définir le rythme de travail ou sa répartition, exécuter quelques tâches de numérisation et de traitement de l'image en entier, afin de s'assurer que le résultat final du plan de travail correspondra bien à ce qu'on a prévu.

- Scanner avec la plus haute résolution possible et la plus grande profondeur de bits, en tenant compte des motivations à l'origine du projet, des limites du scanner, des conditions de stockage des données et des particularités de la source.
- Scanner avec la plus grande profondeur de couleurs possible, en tenant compte des limitations énoncées ci-dessus.
- Sauvegarder tous les jours le contenu du disque dur où l'on aura stocké les données. Le contrôle qualité de l'image numérique et des méta-données est très important; il vaut mieux se soucier des questions de qualité au moment de scanner. On gardera à l'esprit les points suivants :
 - Etablir une résolution minimale et des paramètres couleur (en particulier la résolution spatiale et la profondeur de bits) pour les groupes d'objets à scanner.
 - Examiner le résultat scanné sur écran, sur papier et dans tout autre format prévu pour son utilisation (comme par exemple sur un support mobile).
 - S'assurer que les écrans utilisés sont réglés de façon fiable.
 - Eviter de laisser d'autres objets sur ou autour de l'écran, car ils pourraient altérer la perception de l'objet à numériser.
 - Les images maîtresses doivent être créées avec des règles graduées bien visibles, et les images en couleurs ou en différents niveaux de gris doivent également inclure un témoin de couleurs ou de nuances de gris de référence.

Notes/Commentaire

Scanner est, en soi, une opération relativement simple à mener. Toutefois, afin d'améliorer l'efficacité et réduire le risque d'erreurs, il sera utile d'avoir mis en place un rythme de travail systématisé.

Il faut beaucoup de temps et d'énergie pour scanner un objet aux dimensions exceptionnelles, comme pour faire des opérations de très haute qualité. On pourra en économiser en utilisant du matériel informatique adapté à chaque objet (par exemple, un scanner plus grand, ou équipé d'un repose-livres); au cas où l'on ne pourrait pas avoir accès à du matériel informatique plus grand, il faut alors prévoir un laps de temps plus important. Il ne faut pas non plus oublier l'intérêt de la formation au maniement d'objets de dimensions exceptionnelles ou irrégulières.

Au moyen d'appareils photo numériques

Définition de la problématique

L'utilisation d'appareils photo numériques est de plus en plus courante dans les projets de numérisation. Cela est dû à leur large capacité à numériser des objets non-plats, tels que des livres reliés, des manuscrits pliés ou froissés, et des objets en trois dimensions. On préférera cependant un scanner équipé d'un repose-livres pour numériser des livres reliés et des documents aux dimensions exceptionnelles tels que des cartes ou des dessins.

Suggestions pratiques

- Envisager la location d'un appareil de haute qualité, si le projet est d'ampleur limitée.
- Monter l'appareil sur un pied motorisé et placer les objets à numériser sur un support fixe entouré de lumières spécialement étudiées. Organiser la formation à l'aide d'un spécialiste de la photo numérique.
- S'assurer que l'arrière-plan présentera clairement l'objet.
- Eviter de changer les conditions de luminosité entre les prises et entre les images des différentes parties d'un objet, car cela peut entraîner des perceptions déformées des variations de couleur.
- Utiliser des lentilles apochromatiques et des filtres photographiques adaptés pour réduire la distorsion des images et un mauvais enregistrement des couleurs.

Notes/Commentaire

L'emploi de plus en plus fréquent d'appareils photo numériques dans les projets de numérisation résulte de leur statut d'objet de consommation courante, largement accessible au grand public en raison de leur prix de plus en plus abordable. Cependant, il demeure une différence importante entre des appareils numériques pour spécialistes et les appareils d'entrée de gamme, produits en masse.

Applications logicielles pour la reconnaissance optique de caractères (OCR)

Définition de la problématique

De nombreux projets de numérisation s'appliquent à des documents imprimés, tels que des livres et des journaux. Le plus souvent, on utilise pour cela des scanners, même si cela n'est pas obligatoire. Le recours à des logiciels OCR est fréquent pour extraire des informations de documents scannés et permettre de traiter ces informations. Les logiciels OCR reconnaissent les lettres et les nombres qui composent l'image scannée (c'est-à-dire le fichier image bitmap) et les exportent ensuite sous le format de fichiers textes ASCII, et non sous le format de fichiers image. Ceci permet de procéder à toute une série d'opérations de traitement des données, telles que la recherche, l'indexation ou la conversion de format.

Suggestions pratiques

- Evaluer plusieurs logiciels OCR avant de se fixer sur un produit en particulier. Si des logiciels OCR sont souvent fournis avec les scanners, les logiciels les plus puissants sont généralement vendus séparément.
- Un enjeu crucial dans tout projet de reconnaissance optique des caractères est l'identification et l'édition manuelle des erreurs, ambiguïtés et endroits où le texte n'a pas pu être traité. Un logiciel OCR pourvu d'une interface conviviale pour mener à bien ces tâches peut représenter un gain significatif de temps et d'énergie.
- Le système OCR est d'autant plus efficace que les documents sont en bon état. Les erreurs et défaillances seront plus nombreuses avec des sources pliées, froissées ou décolorées. Afin d'éviter cela, il convient, si c'est possible, de les traiter au préalable.
- Lorsque les sources ne sont pas en bon état, l'utilisation de logiciels de traitement de l'image est à envisager pour éliminer les décolorations et améliorer les contrastes avant de lancer le logiciel OCR.
- On doit vérifier la présence, dans le logiciel OCR, de dictionnaires dans la langue de la source à numériser.

3.7 La préservation d'originaux numérisés

L'objectif à long terme d'un projet de numérisation est de protéger les données qu'il aura créées et de les garder accessibles le plus longtemps possible. Ceci implique de prendre en compte le fait inévitable que les formats de fichiers numériques et les différents supports informatiques de stockage des données deviendront tôt ou tard obsolètes. La préservation d'originaux numérisés et des méta-données correspondantes permet d'éviter qu'on ait à re-numériser des objets, et protège donc les originaux fragiles en leur épargnant la répétition du long processus de numérisation et de création des méta-données.

Formats de fichiers

Définition de la problématique

Le résultat du processus de numérisation est généralement la création d'un fichier maître dans un format TIFF non-compressé, avec incrustation de quelques métadonnées (voir le chapitre intitulé *De l'intérêt d'utiliser des métadonnées en tant qu'objet*). Le format des fichiers ainsi que leur compression détermineront directement les possibilités d'utilisation du résultat de la numérisation. C'est donc à ce stade du projet qu'il faut s'intéresser à des questions telles que le format des fichiers, les normes de taille des fichiers, le temps de transmission du réseau et les différents types de résultats (sur moniteur ou sur imprimante).

Suggestions pratiques

- Avant de se fixer sur un format de fichier, il faut prendre en compte les normes à appliquer dans ce cas précis, les spécificités de la base consultée par les utilisateurs et les capacités du logiciel employé par votre organisation et votre public cible à manipuler les différents formats de fichiers.
- La taille de la base à laquelle les utilisateurs ont accès est un bon indicateur du support à prévoir pour un format de fichier donné. Elle indique également les possibilités de mettre en place des protocoles de migration viables pour les fichiers lorsque les formats auront changé.
- Le format fichier par défaut pour le résultat d'une numérisation est le format TIFF (Tagged Image File Format). A moins que le projet en cours n'ait une bonne raison d'utiliser un autre format de fichier, le résultat de la numérisation, et donc les fichiers maîtres, seront créés dans ce format.
- Le fichier résultant de la numérisation sera forcément assez lourd. Il est courant d'avoir un fichier maître assez important, que l'on stocke sur un disque local et qu'on ne transmet pas sur Internet. A partir de ce fichier maître, on pourra créer des versions plus petites en recourant à des logiciels de traitement de l'image, soit en format TIFF, soit, plus souvent, dans des formats tels que JPEG 2000, PNG ou GIF
- Même si un format national ou breveté peut paraître tentant d'un point de vue technique, il est important de garder à l'esprit que si l'on ne respecte pas les formats standards, cela posera un obstacle majeur à l'échange international de fichiers image maîtres et des métadonnées incrustées correspondantes, ainsi que la création d'un réseau de ressources.

Notes/Commentaire

Le choix d'un format de fichier doit être soumis à l'impératif de produire un résultat de numérisation de la plus haute qualité possible, ainsi que par la disponibilité de protocoles de migration pour assurer la préservation des originaux numérisés. Le rôle des normes dans ce domaine est d'une extrême importance.

Choix des supports

Définition de la problématique

La question du choix des supports est d'importance pour les projets qui prévoient de maintenir leurs collections numériques pendant plusieurs années. Des projets de grande ampleur, tels que l'initiative anglaise du Domes day Book (recueil cadastral établi à la fin du XIème siècle sur l'ordre de Guillaume le Conquérant afin d'évaluer les droits fiscaux sur les terres d'Angleterre, N.D.T) , ont été perdus en raison de l'obsolescence de leurs supports.

Suggestions pratiques

- Le résultat du projet de numérisation sera conservé sur des serveurs, y compris Internet. Ces machines nécessitent des sauvegardes régulières. Par ailleurs, si un serveur n'est pas spécifiquement affecté à un projet de numérisation, il faudra stocker les contenus numérisés sur des supports amovibles, séparément des autres données gardées sur le serveur.
- Tous les fichiers maîtres (y compris les métadonnées) doivent normalement être sauvegardés sur deux types de supports, stockés dans des lieux distincts.
- A ce jour (début 2004), le recours au CD-ROM pour la sauvegarde des données est en voie d'être supplanté par l'utilisation du DVD. Le DVD a des capacités de stockage bien plus grandes. Les graveurs de DVD restent d'un coût plus élevé, mais sont déjà abordables pour la plupart des projets.
- Toutefois, il est peu probable que dans un futur proche le DVD remplace les supports cassette magnétiques tels que la cassette linéaire numérique (DLT) comme support de stockage de préférence pour la sauvegarde des données informatiques. On peut envisager l'utilisation de ces deux technologies pour stocker des contenus numériques.
- Quel que soit le choix du support, on doit garder à l'esprit que ces supports deviendront obsolètes à plus ou moins long terme. En l'espace de cinq ans, la migration vers de nouveaux supports de stockage de données sera très probablement indispensable.

Notes/Commentaire

Le changement rapide des supports, dû principalement aux exigences de l'industrie des consommables électroniques, a eu des conséquences importantes sur les projets de numérisation par le passé. Cependant, la tendance à stocker des données sur de gros serveurs, ou bien sur des disques durs amovibles, facilite la migration de données d'un site à un autre et d'un support à un autre. Dès lors qu'on aura sauvegardé le contenu des serveurs et qu'on l'aura transféré vers de nouveaux serveurs, la dépendance envers les supports amovibles comme seules solutions d'enregistrement d'un processus de numérisation tendra à décroître.

Mais en attendant, la question de la sélection des supports demeure importante. Rien n'indique qu'on ait déjà atteint les limites des solutions de stockage numérique compressé et à encombrement réduit.

Stratégies de migration

Définition de la problématique

Ainsi qu'on l'a précisé plus haut, le choix du format de fichier et du support de stockage doit tenir compte de la possibilité de transférer les données vers un nouveau format de fichier et/ou un autre support de stockage.

Suggestions pratiques

- Examiner les normes en vigueur pour les formats de fichier et le support de stockage, selon les directives suggérées dans les chapitres précédents. La conformité aux normes permet d'être relativement certain qu'un format ou un support continueront à disposer d'une assistance technique à l'avenir.
- Les formats de fichiers brevetés et le stockage sur des supports vierges doivent être l'exception et non la règle.
- En transférant les données d'un format vers un autre, il faut éviter de faire migrer les fichiers numériques maîtres d'un format sans perte de données (tel que TIFF dans le domaine de l'image) vers un format qui en fait perdre (tel que le format JPEG). Dès lors que les données ont été perdues, elles ne peuvent être remplacées.
- Il faut garder à l'esprit que, quels que soient le format de fichier et/ou le support de stockage choisis, ils deviendront obsolètes dans un avenir proche (peut-être moins de cinq ans, sûrement moins de dix ans).
- L'offre du marché des supports de stockage des données permet de prévoir la possibilité de migrations futures d'un support vers un autre au moment où le support initial sera devenu obsolète.
- Une fois créé le contenu numérique, les supports de stockage (CD-ROM, DVD) devront être actualisés périodiquement (une fois tous les deux ou trois ans) pour éviter la perte de données. Cela implique de migrer toutes les données vers de nouveaux supports.
- Le statut des sources numérisées doit être enregistré sur un fichier compte-rendu accompagnant le projet.

3.8 Métadonnées

Le domaine des métadonnées est l'un des plus dynamiques et l'un de ceux qui font l'objet du plus grand nombre de recherches dans toute la sphère du numérique, ainsi que dans des domaines tels que la recherche sur Internet, l'échange de données, l'intégration des applications, etc.

Le modèle de métadonnées sélectionné est particulièrement important, car il influe sur le choix des attributs servant à la description d'objets. Le choix d'un modèle normatif y est aussi relié, ainsi qu'on le verra dans le chapitre suivant, qui porte sur les normes.

De l'intérêt d'utiliser les métadonnées pour la description d'objets

Définition de la problématique

Avant d'opter pour un modèle de métadonnées pour le projet de numérisation, il faut passer en revue les sources que l'on va décrire avec ces métadonnées. Cela permettra d'identifier les modèles existants de métadonnées et de repérer des omissions ou des écarts entre ce que peut couvrir un modèle existant et le type de métadonnées dont vous avez besoin pour votre projet.

Suggestions pratiques

- Il est extrêmement important d'avoir les bonnes métadonnées pour faciliter la recherche et la récupération de sources dans des collections numériques. C'est encore plus vrai lorsqu'on veut mener une recherche dans plusieurs collections, conservées dans différents sites (telles que des catalogues communs à plusieurs bibliothèques, ou musées, etc.)
- Il existe déjà de nombreux modèles de métadonnées. Aussi chaque projet doit-il opter pour un modèle correspondant à ses propres objectifs. Il est recommandé d'éviter d'en créer un nouveau, à moins que les normes existantes ne soient très loin de correspondre aux besoins de votre projet.
- On fera bien d'investir du temps à définir les caractéristiques majeures des sources à numériser, et à identifier leurs principaux attributs et descripteurs. On pourra ensuite comparer ce modèle aux capacités et caractéristiques des modèles de métadonnées existants.
- Il faut identifier les éventuels listes d'autorités (par exemple, pour décrire un site ou un artiste). Il existe déjà plusieurs lexiques de cet ordre; les recherches ont grâce à eux davantage de chances de succès. Pour plus de détails, voir le chapitre sur les normes de métadonnées et les listes d'autorités.

Notes/Commentaire

Le projet historique de la Bibliothèque du Congrès, "The Making of America II" utilisait trois catégories de métadonnées:

- Descriptive - pour décrire et identifier les informations;
- Structurale - pour la navigation et la présentation;
- Administrative - pour la gestion et le traitement.

Les modèles de métadonnées qui seront sélectionnés pour un projet de numérisation sont particulièrement importants, notamment le choix des attributs que l'on utilisera pour caractériser les ouvrages et objets à numériser, les images qui en résulteront, la description des procédés, techniques et technologies employés, la gestion des droits, etc.

On peut trouver un modèle très efficace auprès de la Bibliothèque nationale australienne.

La pléthore de modèles existants et de normes concurrentes pour les métadonnées a conduit de nombreux projets à ne se focaliser que sur le passage d'un modèle normatif à un autre.

Les normes de métadonnées en vigueur

Définition de la problématique

Il existe d'ores et déjà certaines normes importantes de métadonnées. Dans le domaine bibliographique, et de plus en plus dans d'autres domaines culturels, la norme "Dublin Core" est une référence.

Suggestions pratiques

- Passez en revue les modèles existants et les normes de métadonnées avant de créer les vôtres.
- Mieux vaut éviter de créer des modèles de métadonnées totalement inédits pour des collections culturelles.
- Le travail accompli sur les métadonnées lors de projets antérieurs et similaires au vôtre peut vous être utile ; les modèles de métadonnées s'exportent fort bien d'un projet à un autre.
- A moins que votre projet ait une bonne raison de ne pas l'utiliser, les champs du modèle "Dublin Core" devraient être intégrés dans votre modèle de métadonnées. Même si les musées peuvent lui préférer le modèle CIMI (<http://www.cimi.org/>), qui convient mieux à leurs collections, on doit tendre vers une liste d'attributs commune pour permettre la recherche entre les collections.
- Si l'on veut malgré tout recourir à un modèle de métadonnées inédit, on doit aussi mettre en place des passerelles entre celui-ci et le modèle "Dublin Core".
- Un système de dénomination ou une convention nationale de nommage peuvent être d'une grande utilité; néanmoins, il est préférable de disposer d'un modèle complet de métadonnées, non seulement en raison du volume de données qui peuvent être entrées sur chaque objet mais aussi pour permettre des recherches plus poussées et une coopération avec d'autres projets et d'autres pays.

3.9 La publication

A ce stade du projet, les fichiers maîtres numériques ont déjà été créés et stockés ou sauvegardés. On a également identifié un modèle adapté de métadonnées et créé les métadonnées associées à chaque article.

Pour se préparer à la publication, on doit au préalable traiter les sources nouvellement créées. La publication se fait le plus souvent sur Internet et le traitement consiste à réduire les fichiers au format et à la qualité image, audio ou vidéo et à les télécharger pour qu'ils correspondent aux contraintes d'exploitation d'Internet.

La traitement de l'image

Définition de la problématique

Les fichiers TIFF créés lors du processus de numérisation sont généralement très lourds (ils peuvent atteindre un certain nombre de mégabits). De tels fichiers ne conviennent pas à la publication sur Internet, en raison du temps qui serait nécessaire pour les télécharger jusqu'à l'utilisateur final.

Suggestions pratiques

- Créer des versions de diffusion des fichiers maîtres.
Il doit y avoir au moins une version de diffusion.
Une deuxième version, une "vignette", peut aussi être utile, selon l'agencement du site web sur lequel on entend publier les sources.
- On crée des versions de diffusion en ouvrant le fichier maître TIFF avec un logiciel de traitement de l'image, puis en l'exportant sous un format JPEG ou PNG .
- La résolution des couleurs peut généralement être réduite à 256 couleurs. Si cela implique une perte de qualité importante, on peut recourir à une plus grande résolution. Le choix de la bonne résolution de couleurs nécessite une décision plus ou moins subjective dans la plupart des cas.
- Une image créée à 72 ppp apparaîtra à peu près dans sa taille originale sur la plupart des écrans d'ordinateurs. Aussi le choix de 72 ppp est-il conseillé pour de nombreuses images qui seront visionnées sur écran. Pour des résolutions plus faibles, il faudra une décision subjective pour définir ce que sera une qualité "acceptable".
- Pour choisir le format de fichier, la résolution de couleurs et la résolution de pixels, il faut avoir décidé ce qu'est une qualité "acceptable". Il faut donc trouver un moyen terme entre la qualité et la taille du fichier.
- En général, la totalité des fichiers image sur une page web ne doit pas excéder 100 kilo-octets. On peut bien sûr publier de plus grandes images, mais alors il faudra prévoir un accès par un lien placé sur la page web, ainsi qu'un texte avertissant l'utilisateur que le téléchargement sera prolongé.
- Sauf si les fichiers sont répartis, les sources audio et vidéo impliqueront le plus souvent des fichiers de grande taille, qu'il faudra télécharger avant de les visionner hors connexion. Cependant, le temps de téléchargement peut être modulé en jouant sur la résolution de la vidéo, le temps d'échantillonnage du fichiers audio, etc.

Notes/Commentaire

Les décisions concernant le traitement de l'image dépendent en grande partie de jugements personnels. Les directives proposées ici peuvent paraître trop strictes ou trop lâches, selon la nature du projet et du public visé.

Des logiciels de traitement de l'image sont téléchargeables gratuitement en ligne. Les logiciels d'édition de sources audio et vidéo peuvent aussi être téléchargés gratuitement en ligne. Mais des logiciels plus puissants peuvent faire économiser assez de temps et d'énergie pour que leur coût soit justifié.

Questions liées au 3D et à la réalité virtuelle

Définition de la problématique

Les directives énoncées ci-dessus concernant la publication d'images ne sont pas immédiatement applicables aux rendus numériques en 3D ou en réalité virtuelle. Cependant, la recherche de l'équilibre entre la qualité et la taille du fichier est un lieu commun sur Internet.

Suggestions pratiques

- Les visualiseurs d'images en 3D et de réalité virtuelle ne sont pas encore largement distribués dans les configurations de base des ordinateurs personnels, à la différence des visionneuses d'images et des lecteurs audio et vidéo.
- S'assurer qu'on dispose déjà de visualiseurs pour tout fichier en 3D ou en réalité virtuelle. Faire en sorte que les logiciels associés à ces visualiseurs puissent être téléchargés à partir du site où est présenté le fichier. Cela permettra de surmonter les problèmes qui peuvent se poser lorsque d'autres sites de téléchargement de ces logiciels deviendront indisponibles.
- Evaluer plusieurs visualiseurs avant de porter son choix sur l'un d'eux. La compatibilité entre les formats de fichiers et les visualiseurs n'est pas aussi standardisée que dans le domaine de l'image.
- Les ordinateurs personnels modernes, très axés sur les jeux, possèdent souvent des cartes accélératrices et une mémoire augmentée pour les graphiques. Ceci peut avoir un impact considérable sur la façon dont on visualisera la réalité virtuelle.

Notes/Commentaire

Un visualiseur utilisant le langage VRML (langage de modélisation de la réalité virtuelle) a déjà été utilisé avec succès dans l'un des projets de référence (le projet irlandais ACTIVATE); il s'agit du visualiseur Blaxxun Contact.

La publication en ligne

Définition de la problématique

De nombreux projets de numérisation relevant du domaine culturel mènent à la création de ressources culturelles en ligne: le plus souvent un site web contenant des images, des méta-données, des objets en 3D, etc. Ils vont du site aux contenus les plus simples jusqu'aux portails nécessitant des visualiseurs spécifiques. On dispose d'une vaste somme de connaissances sur la création de sites web; aussi ne trouvera-t-on ici que quelques directives, ainsi que des liens vers des exemples de sites web choisis par les partenaires de Minerva comme des exemples des meilleures pratiques.

Suggestions pratiques

- Il doit être facile de naviguer sur un site web ; on doit pouvoir trouver, tout au long de la navigation, des liens vers la page d'accueil ou vers un sommaire.
- Il convient de veiller à ce que les sites web puissent offrir un accès universel et être utilisés par les mal-voyants et les autres personnes handicapées.
- Les pages web doivent être assez courtes pour que l'utilisateur n'ait pas à les dérouler sur plusieurs pages-écrans.
- Les images doivent avoir une taille réduite pour ne pas interrompre la navigation.
- Les images de grande taille doivent être accessibles via des liens apparaissant sur les pages web et précisant dans une note que l'image est grande et que son téléchargement peut prendre du temps.
- Il ne faut pas abuser des animations, des menus locaux apparaissant au-dessus ou en dessous de la page, des technologies Flash et autres. On doit pouvoir accéder au menu principal sans être contraint de passer par de longues séquences animées d'introduction au site.
- Dans l'idéal, un site web doit être multilingue et comprendre au moins la langue du pays d'accueil ainsi qu'une ou deux autres langues.
- On doit périodiquement mettre à jour les liens vers des ressources extérieures, afin de réduire au maximum les liens périmés et la perte de temps qu'ils induisent.

Notes/Commentaire

On maîtrise bien la marche à suivre pour mettre des sources en ligne et on dispose d'une documentation riche sur la question. Le propos de ce guide n'est pas d'expliquer comment on crée un site web et des programmes en HTML ni comment on construit des bases de données en ligne, ni comment on exécute d'autres tâches nécessaires à la création et à l'entretien d'une présence sur Internet. Il est prévisible que la plupart des institutions culturelles qui auront recours à ces directives disposeront déjà d'un serveur web qu'elles exploiteront pour leur projet de numérisation.

3.10 Propriété intellectuelle et droits de reproduction

Toute publication en ligne doit s'accompagner d'une prise en compte du code de la propriété intellectuelle et de la façon dont il s'applique aux sources concernées. Pour celles qui appartiennent au domaine public (telles que des ouvrages ou des journaux très anciens, ou des documents ou objets explicitement placés dans le domaine public), on rencontre assez peu de difficultés. Cependant, de nombreuses institutions culturelles tirent un revenu de l'utilisation d'images d'objets ou de tableaux appartenant à leurs collections; elles veillent donc aux droits de reproduction. La publication de sources appartenant à des tiers ne peut avoir lieu sans le consentement desdits tiers. Heureusement, il existe toute une gamme de solutions techniques pour protéger les droits de reproduction des sources mises sur Internet. On en trouvera le détail ci-dessous.

Etablir les droits de reproduction

Définition de la problématique

La première démarche à suivre lorsqu'on étudie la question des droits de reproduction pour un objet culturel est d'établir la propriété de ces droits.

Suggestions pratiques

- Etablir le régime juridique en matière de droits de reproduction et de publication qui prévaut dans le pays où l'on conduit le projet. Chaque pays a ses propres lois concernant les droits de reproduction, dont la plupart remontent au moins au XIX^{ème} siècle. De telles lois s'appliquent généralement à toutes les formes de publication, y compris sur Internet. Elles peuvent ou non couvrir l'acte de numérisation, que l'on peut extrapoler comme étant un procédé d'archivage, ou qui peut être tenue pour de la copie.
- En aucun cas il ne faut procéder à une publication en ligne sans avoir demandé au préalable les droits de reproduction.
- Certaines sources, comme par exemple de vieux journaux, sont gouvernées par des lois très claires en matière de droits de reproduction. Celles-ci autorisent généralement la copie à titre gratuit dès lors que les journaux ont dépassé un certain âge. Les sources qui entrent dans cette catégorie peuvent alors être numérisées et publiées sans qu'on ait à s'acquitter d'aucun droit.
- Dans le cas de sources dont les droits de reproduction sont détenus par l'institution qui a commandité le projet, on demandera une autorisation interne pour procéder à la numérisation et à la publication en ligne.
- Dans le cas de sources dont les droits de reproduction sont détenus par un tiers, tel que le prestataire ou le donateur d'une collection d'objets historiques, il faut demander par écrit l'autorisation de ce tiers. On ne procédera à la publication qu'une fois cette autorisation accordée.
- Pour obtenir l'autorisation de numériser et de publier, on peut donc avoir à payer. Les dépenses à prévoir devront être justifiées au regard de l'intérêt qu'on attachera à la présence de telle ou telle source dans la ressource en ligne.

Notes/Commentaire

Les lois régissant le droit de reproduction varient d'un pays à un autre.

Sauvegarder les droits de reproduction

Définition de la problématique

Publier en ligne sur Internet, c'est ouvrir la porte à la copie des sources ainsi mises à la disposition du public. Il est impossible d'empêcher toute copie de sources mises en ligne; toutefois, il existe un certain nombre de procédures auxquelles on peut avoir recours, et dont chacune contribue à protéger le droit de reproduction.

Suggestions pratiques

- Déterminer si le droit de reproduction doit ou non être protégé.
- Se mettre d'accord avec les détenteurs des droits de reproduction sur les procédures à utiliser pour protéger ces droits.
- On peut envisager, entre autres, les procédures suivantes :
 - L'apposition d'un filigrane ou d'un logo de droit de copie sur chaque image.
 - L'apposition d'un filigrane numérique invisible à l'œil nu sur chaque image. De tels signes distinctifs peuvent être utilisés pour prouver la propriété d'une image "volée", et également pour assurer la traçabilité de l'image sur Internet.
 - Le cryptage d'images, qui ne pourront être consultées que par des utilisateurs enregistrés, auxquels on aura donné la clef ad hoc. Mais ce type de solution réduit forcément l'intérêt de l'image en ligne pour le reste du public.
 - La restriction de la publication à des images à faible résolution, comme 72 ppp pour une présentation sur écran. Ceci restreint les possibilités d'utilisation de ces images dans d'autres domaines, tels que l'impression, les vêtements, etc.
 - La restriction de la publication à de petites parties de l'image.
 - La présentation d'images aux seuls membres enregistrés d'une communauté définie.
 - Tester les résultats du dispositif de protection des droits de reproduction avec les premières sources, pour vérifier qu'il n'induit pas d'effets inattendus ou indésirables.

Notes/Commentaire

L'approche la plus adaptée à chaque projet dépend en grande partie de ses objectifs et de ceux de l'institution culturelle concernée, ainsi que de la nature des sources traitées. En général, la publication d'une sélection restreinte d'images, à faible résolution, est une solution retenue par beaucoup de galeries et musées en ligne. Le caractère plus ou moins unique de nombreuses collections culturelles est en soi une preuve de propriété des droits de reproduction dans de nombreuses situations.

3.11 La direction de projets de numérisation

La réussite d'un projet, et notamment d'un projet de numérisation, dépend dans une grande mesure de la façon dont il a été dirigé. Dans ce chapitre, on trouvera quelques directives traitant spécifiquement de la direction de projets de numérisation

Diriger le processus de numérisation

Définition de la problématique

Un projet de numérisation comprend des dizaines, des centaines, voire des milliers d'objets à numériser. Pour mener à bien efficacement un projet, il importe de déterminer un rythme de travail qui rentabilise au maximum les efforts de l'équipe. En outre, les sources d'information telles que la base de connaissances du projet jouent un grand rôle.

Suggestions pratiques

- Définir de manière détaillée chacune des étapes du processus de numérisation.
Par exemple :
 - récupération de l'objet sur le site où il est conservé
 - nettoyage ou préparation
 - numérisation par scanner ou prise de vue directe
 - retour au site de conservation
 - nommage du fichier
 - stockage du fichier
 - création de versions diffusables en ligne des fichiers maîtres
 - sauvegarde des serveurs ou supports de stockage des données.
- Développer une base de connaissances qui assure la traçabilité de l'objet tout au long du processus de numérisation et permette de vérifier à tout moment où en est le projet. Cette dernière peut prendre la forme d'une base de données (par exemple en MS Access, Oracle, MySQL, etc), ou d'un simple tableur, ou même, d'une collection de documents. Ce qui importe n'est pas son format, mais le procédé par lequel on s'assurera de l'enregistrement des tâches au fur et à mesure de leur exécution.
- Le nom de chaque objet à numériser, son identifiant et toute autre information importante devront être entrés dans la base de connaissances du projet dès lors qu'on a sélectionné l'objet en question. On devra aussi y enregistrer régulièrement son statut, c'est-à-dire la dernière étape qu'il aura franchie).
- On devra faire des choix quant à la procédure à suivre : par exemple, déterminer si les objets doivent être rassemblés sur le site de numérisation au commencement de chaque journée, de chaque semaine, ou bien un par un.
- Les objets qui réclament des tâches similaires ou un même équipement informatique devront être numérisés ensemble, afin de rentabiliser le temps passé à monter les appareils photos numériques, à configurer les scanners,

etc. On conservera le détail des configurations informatiques de façon à pouvoir répéter la procédure de numérisation en cas de perte d'un fichier ou autre incident de parcours.

- On fera bien de noter et de garder sous la main tout au long du projet les coordonnées complètes de tous les membres du personnel chargé du soutien.

Notes/Commentaire

Plus un projet a d'ampleur, plus on aura intérêt à définir une procédure et un rythme de travail. L'efficacité qu'on y gagnera compensera largement le temps passé à les mettre en place.

Monter une équipe

Définition de la problématique

Les projets de numérisation mettent souvent le personnel des informations culturelles aux prises avec les nouvelles technologies pour la première fois, qu'il s'agisse d'équipements de numérisation, de publication sur Internet, de traitement de l'image, de balisage des métadonnées, de développement et de gestion de bases de données, etc.

Suggestions pratiques

- Veiller à ce que l'équipe chargée du projet comprenne au moins une personne dotée des compétences nécessaires en matière de technologie de l'information.
- Très en amont du projet, évaluer le degré de connaissances du personnel chargé de travailler sur le projet ainsi que les connaissances dont il aura besoin en matière de technologie de l'information. Identifier les besoins de formation et y répondre avant de lancer le projet.
- Les connaissances en matière de technologie de l'information ne sont pas les seules requises. On pourra avoir besoin de compétences spéciales, comme on l'a déjà évoqué plus haut, pour la manipulation de documents et objets fragiles, etc. Une formation ad hoc pourra être dispensée par les personnes responsables de la conservation des sources.

Notes/Commentaire

Il est préférable de disposer d'un petit groupe de personnel très qualifié pour travailler sur un projet plutôt qu'un groupe plus large de participants occasionnels. Cependant, bien qu'il soit plus efficace pour la bonne marche du projet de développer une compétence en particulier, il se peut que le personnel préfère maîtriser l'intégralité du processus de numérisation. L'exécution des tâches de numérisation et de balisage des métadonnées n'est pas en soi un travail très épanouissant, tandis que la familiarisation aux autres étapes du projet accroîtra la motivation du personnel.

Former le personnel

Définition de la problématique

On aura besoin de former le personnel, à moins que celui-ci ne possède déjà une expérience significative liée à des projets antérieurs. La formation portera sur deux domaines très différents : la technologie utilisée et le maniement des sources.

Suggestions pratiques

- Ne pas partir du principe qu'une formation du personnel n'est pas nécessaire, ni que le personnel des archives, de la bibliothèque ou du musée possède déjà toute l'expertise nécessaire.
- S'assurer qu'on a bien identifié les besoins de formation du personnel chargé du projet, et ce, dès le début du projet, c'est-à-dire lors de la phase de planification. Ces besoins devront être entrés dans la base de connaissances du projet, et devront être traités avant que les difficultés ne se posent durant l'exécution du projet.
- Il se peut que certaines connaissances, telles que l'utilisation de la technologie de la numérisation, puissent s'apprendre "sur le tas"; mais d'autres, telles que la manipulation de sources, nécessitent une formation préalable.
- Un personnel restreint, formé, qui se construit une expérience tout au long du projet, vaut mieux qu'un groupe plus large et informel dont les membres changent fréquemment.
- On peut faire assurer la formation technologique à partir d'un autre projet financé par la même institution ; ou bien une agence extérieure, spécialisée dans la numérisation, peut assurer la formation.
- La formation à la conservation sera mieux assurée par les personnes responsables de la conservation des sources originales.

Notes/Commentaire

Faute d'une formation adéquate, le projet peut pâtir très rapidement d'accidents malheureux et irréversibles. Il peut se passer la même chose lorsqu'on retire du personnel de l'équipe chargée du projet pour le remplacer par du nouveau personnel. C'est pourquoi une équipe réduite et bien formée est ce qui convient le mieux à de tels projets.

Le temps investi dans la formation au début du projet sera compensé par une meilleure productivité et moins de problèmes tout au long du projet.

Travailler avec des tiers pour l'assistance technique

Définition de la problématique

Il est fréquent d'engager les services d'un ou de plusieurs tiers durant l'exécution d'un projet de numérisation. Les services les plus communément fournis comprennent la procédure de numérisation proprement dite, la direction du projet, l'intégration aux systèmes tiers, le développement de logiciels, etc. Cela permet à l'organisme culturel de se concentrer sur ses propres domaines d'expertise, sans être contraint de former ou d'engager du personnel.

Suggestions pratiques

- Comme pour tout autre projet, la relation entre les partenaires techniques et les autres membres du projet doit être régie par des contrats stricts et clairs. Avant d'entamer quelque tâche que ce soit, il faut conclure un accord spécifiant de façon très détaillée, dans un document signé par toutes les parties, quels sont les produits ou services à fournir.
- Le travail doit être vérifié très régulièrement, afin de s'assurer que le résultat correspond effectivement aux objectifs et besoins du projet.
- Même s'il peut être utile de recourir à des tiers, on doit bien garder à l'esprit que toute expérience ou expertise qui pourra être gagnée durant l'exécution d'un projet externalisé sera perdue pour l'institution culturelle commanditaire à la fin du projet. Ceci vaut aussi pour le personnel temporaire employé pour la durée du projet. Il peut être préférable d'affecter au projet un membre du personnel déjà ancien et de le remplacer momentanément par un intérimaire pour l'exécution de ses tâches habituelles.

Notes/Commentaire

Certain projets de grande ampleur, tel que le programme national français de numérisation, ont identifié un fournisseur de services privilégié, avec qui la relation pourra s'étendre sur plusieurs années et couvrir plusieurs projets. Une fois qu'on a établi une relation de travail avec un fournisseur, il ne faudra pas en changer à la légère entre deux projets.

Travailler avec des tiers pour des projets de coopération et partager des données

Définition de la problématique

De nombreux projets de numérisation sont soit des efforts conjoints impliquant au moins deux organismes culturels, soit des projets cadres financés par l'UE, avec des partenaires multiples dans plusieurs pays. Il existe déjà beaucoup de directives pour ce type de projets, qui dépassent les ambitions du présent document. On a cependant inclus les indications suivantes:

Suggestions pratiques

- S'assurer que tous les partenaires sont bien conscients de leur rôle et de leurs responsabilités dans le projet, et qu'ils les acceptent. Les leur rappeler périodiquement.
- Etablir un mode de communication commun aux différents partenaires, et s'assurer que tous les partenaires ont bien reçu l'information qui leur est destinée. Le courrier électronique est un moyen idéal pour ce faire, pourvu que les partenaires lisent ces courriers et y répondent.
- Les sous-traitants doivent être encadrés par des contrats commerciaux stricts, spécifiant clairement et sans ambiguïté leurs obligations.
- Les droits de propriété intellectuelle de chaque partenaire doivent être clairement détaillés et faire l'objet d'une signature formelle de leur part. Il faut établir en amont du lancement du projet un accord de partenariat qui énonce clairement les droits de propriété intellectuelle s'appliquant aux sources que doit traiter le projet, ainsi que celles qui seront créées par le projet.
- Chaque partenaire doit avoir un rôle précis dans le projet. Si le rôle de l'un d'entre eux n'est pas clair, il faut remettre en question l'opportunité de sa participation au projet.

Notes/Commentaire

Les points qui précèdent ne sont qu'un petit échantillon de ce qui peut être fait pour la définition et la direction de projets à plusieurs partenaires. Les partenaires et les prestataires de services sont souvent à l'origine des délais et des problèmes de communication dont souffrent les projets. Des accords clairs et une acceptation officielle des rôles et responsabilités de chaque partenaire à tous les stades du projet peuvent permettre d'éviter ces problèmes.

Coûts

Définition de la problématique

Il est très onéreux de monter une collection numérique. Les projets doivent donc tenir compte de tous les coûts de démarrage et d'infrastructure ainsi que des coûts de gestion du projet. Ceux-ci englobent les coûts de la planification initiale, des spécifications de données, des systèmes de traçabilité et de documentation, de formation du personnel, etc., ainsi que le coût croissant de la numérisation des sources.

Suggestions pratiques

- Les projets de numérisation doivent prendre en compte les coûts suivants, induits par la planification, la mise en oeuvre et la maintenance d'une collection numérique :
 - Gestion du personnel
 - Moyens requis
 - Coûts opérationnels
 - Coûts des systèmes de stockage et de diffusion des données
- La gestion du personnel recouvre notamment les salaires pour la gestion du projet, ceux du programmeur internet, des chargés de formation, de l'équipe d'assistance technique, etc., mais aussi les frais de déplacements et de formation.
- Les dépenses engagées pour les moyens sont liées aux équipements nécessaires, qui correspondent à des solutions de compromis. Les directeurs de projets doivent ainsi souvent décider s'il vaut mieux, pour le succès du projet, maintenir des coûts aussi bas que possible ou bien atteindre la meilleure qualité possible pour la capture d'images.
- Les coûts opérationnels à prendre en compte sont les suivants:
 - Le temps nécessaire au maniement des sources (depuis le rayonnage jusqu'au système de capture d'images, et vice-versa), à intégrer comme une part du salaire journalier total.
 - La préparation des sources (conservation, nettoyage, etc).
 - Le temps de capture (depuis le montage des appareils jusqu' au nommage et à la sauvegarde des fichiers), à intégrer comme une part des salaires journaliers totaux des opérateurs.
 - Le catalogage et la gestion des métadonnées, à intégrer comme une part des salaires totaux à régler.
 - Le coût des équipements et logiciels informatiques pour chaque source numérisée (à calculer de préférence à partir des coûts d'amortissement ou de remplacement plutôt que sur les prix d'achat).

- Le temps nécessaire pour assurer la garantie de qualité, à intégrer comme une part des coûts salariaux.
- La maintenance du matériel et des logiciels informatiques.
- Le temps d'assistance technique nécessaire à la capture d'images.
- Le temps de gestion du projet lié à la capture d'images.
- La formation liée aux tâches de capture d'images.
- Etre bien conscient du fait que la capture d'images est souvent la partie la moins coûteuse d'un projet de numérisation. En moyenne, un tiers des coûts totaux sont liés à la conversion en numérique, un peu moins d'un tiers à la création des métadonnées, et un peu plus d'un tiers à des tâches administratives et à la garantie de qualité. Le reste correspond aux coûts de maintenance à long terme.
- Les coûts de stockage à prendre en compte doivent être les coûts totaux de maintenance par gigaoctet.

Notes/Commentaire

La viabilité à long terme du projet est souvent reléguée au bas de la liste des priorités parce que d'autres préoccupations sont plus immédiates et pressantes. Pourtant, quelles que soient la qualité et la robustesse des sources numériques créées par un projet de numérisation, celles-ci ne dureront pas longtemps si le projet en question ne trouve pas le financement nécessaire à leur maintenance.

