



Service d'Archives itinérant - CDG 90

FICHE ARCHIVES N°8

LA NUMERISATION

Qu'est-ce que la Numérisation ?

1. Les Objectifs
2. Usage et portabilité
3. Les Formats d'enregistrement

Attention : le stockage sur support numérique

Actualités...



Instruction
[DITN/RES/2008/001](#)
du 4 février 2008
Guide pour
l'élaboration d'un
cahier des charges de
numérisation.

Instruction
[DITN/RES/2005/004](#)
du 29 mars 2005
Recommandations
relatives à la gravure, à
la conservation et à
l'évaluation des
CD-R

Qu'est-ce que la Numérisation ?

La numérisation est un acte qui consiste à **créer une copie en mode image d'un document physique existant.**

Ces **documents numérisés** ainsi générés ne sont pas à confondre avec les documents dits « informatiques » créés dès l'origine sur ordinateur.

On utilise le terme « numérique » du fait que, par cette procédure, des données analogiques sont traduites en **langage binaire** (une suite de 1 et de 0) par un numériseur (scanner).

Cette « **traduction** » implique qu'il existe de bonnes et de mauvaises numérisations. La qualité de la numérisation se détermine alors en fonction de l'usage qui en est défini, et de la portabilité sur le réseau.

1. LES OBJECTIFS

Les objectifs sont les critères qui déterminent la **qualité de la numérisation**. Il peut s'agir :

- d'un moyen de sauvegarde de documents physiquement trop dégradés,
- de la rédaction d'un imprimé,
- de la mise en ligne d'un document,
- de la destruction de documents trop volumineux et sans valeur archivistique ou graphique (**Attention : numérisation ne veut pas dire destruction systématique du document physique ; même si depuis 2000 les documents électroniques peuvent avoir une valeur légale, l'authentification par signature, manuscrite ou électronique, reste la plus fiable**).

C'est en fonction de ces objectifs que sont déterminés le niveau de **résolution** (degré de détails) et de **fidélité** (la colorimétrie) de la numérisation. Ainsi des pertes de données peuvent être tolérées pour les trois premiers objectifs, tandis qu'elles sont inconcevables pour le quatrième puisque le document original sera détruit.

2. USAGE ET PORTABILITE

L'**usage** du fichier final est donc d'une extrême importance pour le choix de la numérisation. Cependant la **portabilité** est aussi déterminante. Il s'agit de choisir de bon format d'enregistrement dont la taille naturelle ne nécessitera pas de compression ultérieure.

Ainsi différentes catégories de documents à numériser peuvent être déterminées :

- **textes pouvant être lus sans exigence graphique** (documents administratifs) : le critère n'est pas d'avoir une qualité d'image optimale, mais plutôt un fichier léger sur le réseau. Le format d'enregistrement peut donc être de petite taille, et la numérisation uniquement bitonale (noir et blanc) ;
- **textes et images pouvant être consultés sans exigence graphique** : ces documents ont juste pour objectifs d'être consultés et non imprimés en qualité maximale. Ainsi une perte de données peut être acceptée en choisissant un format d'enregistrement léger;
- **ouvrages ou documents en téléchargement** : ici la qualité d'impression du document doit être d'assez bonne qualité. Une perte de données peut être admise mais de façon moindre que pour la catégorie précédente ;
- **documents numérisés pour la sauvegarde ou la reproduction** : ici l'usage implique une numérisation de haute résolution et un enregistrement sans perte de données ;
- **documents numérisés pour le web** : il s'agit de documents numérisés en très basse résolution, inexploitable à l'impression et tout juste tolérables à l'écran.

Si un fichier doit être utilisé pour différents usages, il est préférable de procéder à une numérisation de haute qualité, puis de générer d'autres fichiers de qualité moindre.

3. LES FORMATS D'ENREGISTREMENT

Le format d'enregistrement est déterminant pour le **poids** du fichier (sa portabilité) et pour sa **reconnaissance** par les applications informatiques. Existence donc des formats avec ou sans pertes de données.

- **les formats sans perte de données** : à utiliser pour créer une banque de fichiers sources. Il s'agit :
 - o du format **TIFF** (*Tagged Image File Format*) principalement pour les photographies ou les documents mono-pages,
 - o et du format **PDF** (*Portable Document Format*) utilisé pour les documents multipages ;

- **les formats avec pertes de données** : ces formats réduisent le nombre de couleurs et de pixels. Il s'agit des formats **JPEG ou GIF** pour les images (JPEG permettant de choisir sa qualité de résolution), **PNG** (propose une compression sans perte de données) ou **PDF** avec pertes de données ;

Tous ces formats sont **des formats dits « libres »**, à savoir qu'ils ne sont pas propres à une application, et ne nécessitent pas le paiement d'une licence.

Cependant certains systèmes de numérisation proposent des **formats d'enregistrement dits « propriétaires »** lisibles uniquement par leur application d'origine.

ATTENTION : LE STOCKAGE SUR SUPPORT NUMERIQUE



L'enregistrement des documents numérisés est réalisé sur un support numérique. Les plus utilisés sont les disques optiques (CD-Rom), les disques durs (serveurs), et bandes magnétiques.

Cependant comme tout support matériel, ils sont soumis aux **dégradations** et au vieillissement de leurs composants, ainsi qu'à **obsolescence** de leur technologie. Il s'agit de supports fragiles dont la durée de vie n'est pas garantie au delà de 10 ans. Des supports plus robustes existent tels que les **CD-R et DVD-R**, mais leurs coûts demeurent importants et ce n'est que l'objectif de la numérisation qui peut déterminer s'il convient de les utiliser.

De plus la disparition de certaines technologies (comme les lecteurs de disquettes sur les ordinateurs actuels) peut également engendrer la perte de données. Il convient donc de vérifier régulièrement que le support d'enregistrement est toujours lisible, et de procéder à la migration des données (fichiers) sur des supports plus récents.

ACTUALITE...

Guide méthodologique pour le choix de formats numériques pérennes dans un contexte de données orales et visuelles - Note d'information DGP/SIAF/2010/010 du 21 mai 2010

Circulaire DGP/SIAF/AACR/2010/010 du 29 juillet 2010. **Dérogations aux règles de communicabilité des archives publiques** : règles générales et procédure

Circulaire DGP/SIAF/AACR/2010/011 du 27 juillet 2010. Accès aux origines personnelles : **communicabilité des dossiers de pupille** pour lesquels le secret de l'identité du parent biologique a été explicitement opposé