



Aspects informatiques et contrôles qualité

Colloque BSN5-2013 « Numérisation des PV du BDL 1795-1932 »

Alain COURGEY, DIO, Observatoire de Paris

Organisation du projet

- Réunion de lancement
- Test portant sur une centaine de vues
- 3 livraisons successives d'environ 7500 vues (soit 15 000 fichiers) puis 1 livraison finale pour les reprises
- support de la livraison : disque dur externe

Volumétrie

- 22 179 vues
- 2 jeux de fichiers: TIFF + JPEG
- 1 To de données

Contrôles qualité

Plusieurs éléments sont vérifiés (arborescence, nommage des fichiers, conformité des formats, métadonnées, colorimétrie...)

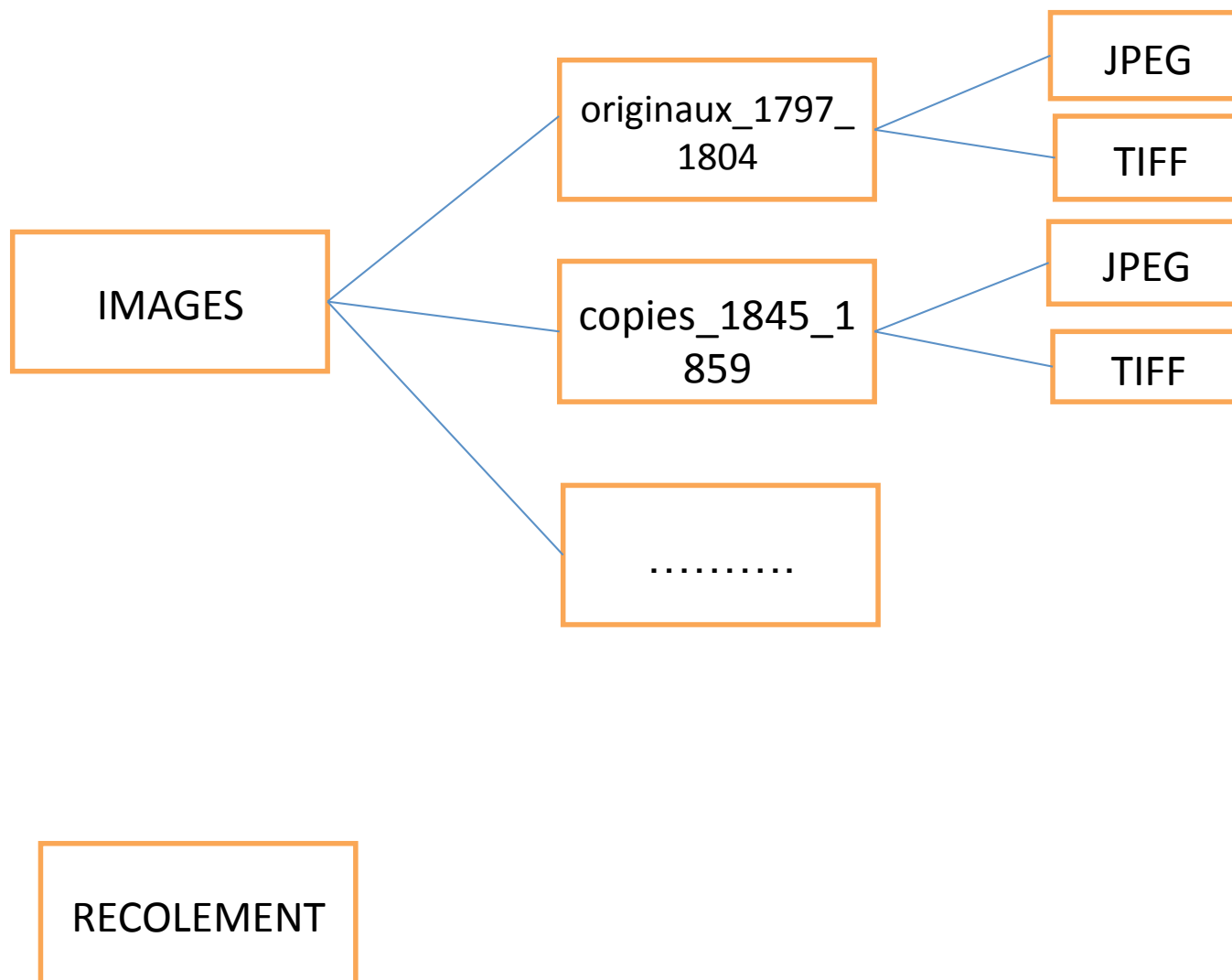
selon plusieurs types de contrôles :

- exhaustif/par échantillon
- Automatiques (script) / manuels (visuels)

Contrôle qualité

- **Respect de l'arborescence de livraison**

Architecture de la livraison



Contrôle qualité

- Respect de l'arborescence de livraison
- **Respect du plan de nommage des fichiers**

Plan de nommage

FR751142302_006_XXXXXX_A.EXT

- FR751142302 : Identifiant de la Bibliothèque (n°ISIL)
- 006 : Numéro d'ordre du projet
- XXXXXX : Numéro d'ordre séquentiel de l'image.
- A : Caractère alphabétique
- EXT est l'extension du format physique du fichier

Plan de nommage

Exemples:

- FR751142302_006_000246_A.TIF
- FR751142302_006_000389_A.JPG

Contrôle qualité

- Respect de l'arborescence de livraison
- Respect du plan de nommage des fichiers
- **Contrôle du fichier de récolement**

Fichier de récolement

Nombre de vues numériques réalisées	Nom du premier fichier numérique	Nom du dernier fichier numérique
138	FR751142302_006_000001_A	FR751142302_006_000138_A
297	FR751142302_006_000139_A	FR751142302_006_000435_A
665	FR751142302_006_000436_A	FR751142302_006_001100_A
621	FR751142302_006_001102_A	FR751142302_006_001722_A
861	FR751142302_006_001722_A	FR751142302_006_002582_A
...
...
...
...

Fichier de récolement

Nombre de vues numériques réalisées	Nom du premier fichier numérique	Nom du dernier fichier numérique
138	FR751142302_006_000001_A	FR751142302_006_000138_A
297	FR751142302_006_000139_A	FR751142302_006_000435_A
665	FR751142302_006_000436_A	FR751142302_006_001100_A
621	FR751142302_006_001102_A	FR751142302_006_001722_A
861	FR751142302_006_001722_A	FR751142302_006_002582_A
...
...
...
...

Contrôle qualité

- Respect de l'arborescence de livraison
- Respect du plan de nommage des fichiers
- Contrôle du fichier de récolement
- **Contrôle de la présence des fichiers dans les répertoires**

Présence des fichiers dans les répertoires

- Nommage
- Fichiers manquants
- Fichiers en trop

Contrôle qualité

- Respect de l'arborescence de livraison
- Respect du plan de nommage des fichiers
- Contrôle du fichier de récolement
- Contrôle de la présence des fichiers dans les répertoires
- **Conformité des fichiers**

Conformité des fichiers

- erreurs dans le format de fichier
- fichier altéré
- compression ou résolution non conforme

→ *réalisés à l'aide de scripts*

Conformité des fichiers

Jhove:

- Identification des formats et validation de leur conformité
- Formats pris en charge: AIFF, GIF, HTML, JPEG, JPEG 2000, PDF, TIFF, Wave, XML
- Interface graphique ou ligne de commande

Conformité des fichiers

```
jhove -c conf/jhove.conf -m tiff-hul -t temp -h xml C:\IMAGES\originaux_1797_1804\TIFF  
\FR751142302_006_000139_A.tif >C:\controles\format_fichiers\originaux_1797_1804\TIFF  
\FR751142302_006_000139_A.tif.txt
```

```
jhove -c conf/jhove.conf -m tiff-hul -t temp -h text C:\IMAGES\originaux_1797_1804\TIFF  
\FR751142302_006_000140_A.tif >C:\controles\format_fichiers\originaux_1797_1804\TIFF  
\FR751142302_006_000140_A.tif.txt
```

```
jhove -c conf/jhove.conf -m tiff-hul -t temp -h text C:\IMAGES\originaux_1797_1804\TIFF  
\FR751142302_006_000141_A.tif >C:\controles\format_fichiers\originaux_1797_1804\TIFF  
\FR751142302_006_000141_A.tif.txt
```




















```
jhove -c conf/jhove.conf -m tiff-hul -t temp -h text C:\IMAGES\originaux_1797_1804\TIFF  
\FR751142302_006_000142_A.tif >C:\controles\format_fichiers\originaux_1797_1804\TIFF  
\FR751142302_006_000142_A.tif.txt
```

.....

.....

.....

Conformité des fichiers

Organiser		Nouveau		Ouvrir		Sélectionner	
> bibli > projets2016 > projetBSN5 > controles > resultats_controles_jhove							
Nom	Modifié le	Type	Taille				
 FR751142302_006_000139_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000140_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000141_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000142_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000143_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000144_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000145_A.tif.txt	21/10/2016 09:51	Document texte	1 Ko				
 FR751142302_006_000146_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000147_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000148_A.tif.txt	20/10/2016 15:28	Document texte	1 Ko				
 FR751142302_006_000149_A.tif.txt	20/10/2016 15:28	Document texte	1 Ko				
 FR751142302_006_000150_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000151_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000152_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000153_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000154_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000155_A.tif.txt	21/10/2016 09:51	Document texte	1 Ko				
 FR751142302_006_000156_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				
 FR751142302_006_000157_A.tif.txt	20/10/2016 15:28	Document texte	3 Ko				

Conformité des fichiers

Jhove (Rel. 1.11, 2013-09-29)

Date: 2016-10-20 15:28:54 CEST

RepresentationInformation: C:\IMAGES\tiff\FR751142302_006_000139_A.tif

ReportingModule: TIFF-hul, Rel. 1.7 (2012-08-12)

LastModified: 2015-03-16 14:05:44 CET

Size: 6841930

Format: TIFF

Version: 6.0

Status: Well-Formed and valid

SignatureMatches:

TIFF-hul

MIMEtype: image/tiff

Profile: DLF Benchmark for Faithful Digital Reproductions of Monographs and Serials: color

TIFFMetadata:

ByteOrder: little-endian

IFDs:

Number: 1

Conformité des fichiers

fichier	Format	Version	Statut	Compr.	Orient	Unit	XSamplingf	YSamplingf
C:\IMAGES\tiff\FR751142302_006_000139_A.tif	TIFF	6.0	Well-Formed and valid	LZW	normal	inch	400	400
C:\IMAGES\tiff\FR751142302_006_000140_A.tif	TIFF	6.0	Well-Formed and valid	LZW	normal	inch	400	400
C:\IMAGES\tiff\FR751142302_006_000141_A.tif	TIFF	6.0	Well-Formed and valid	LZW	normal	inch	400	400
C:\IMAGES\tiff\FR751142302_006_000142_A.tif	TIFF	6.0	Well-Formed and valid	LZW	normal	inch	400	400
C:\IMAGES\tiff\FR751142302_006_000143_A.tif	TIFF	6.0	Well-Formed and valid	LZW	normal	inch	400	400
C:\IMAGES\tiff\FR751142302_006_000144_A.tif	TIFF	6.0	Well-Formed and valid	LZW	normal	inch	400	400
C:\IMAGES\tiff\FR751142302_006_000145_A.tif	TIFF		Not well-formed					
C:\IMAGES\tiff\FR751142302_006_000146_A.tif	TIFF	6.0	Well-Formed and valid	LZW	normal	inch	400	400
C:\IMAGES\tiff\FR751142302_006_000147_A.tif	TIFF	6.0	Well-Formed and valid	LZW	normal	inch	400	400

Contrôle qualité

- Respect de l'arborescence de livraison
- Respect du plan de nommage des fichiers
- Contrôle du fichier de récolement
- Contrôle de la présence des fichiers dans les répertoires
- Conformité des fichiers
- **Contrôle des métadonnées**

Contrôle des métadonnées

Exiftool:

- Extraire, éditer, ajouter des métadonnées dans une multitude de variétés de fichiers dont les fichiers images de format tif ou jpg
- Différents formats de sortie possibles (dont csv)
- Ligne de commande

Contrôle des métadonnées

- `exiftool -g1 -ee -csv C:\bibli\projet_bdl\images\originaux_1804_1810\tif*.tif > C:\bibli\projet_bdl\metadonnees\metadonnes_originaux_1804_1810_tiff_exiftool.txt`
- `exiftool -g1 -ee -csv C:\bibli\projet_bdl\images\copies_1845_1859\tif*.tif > C:\bibli\projet_bdl\metadonnees\metadonnes_copies_1845_1859_tiff_exiftool.txt`
- `exiftool -g1 -ee -csv C:\bibli\projet_bdl\images\originaux_1795_1797\tif*.tif > C:\bibli\projet_bdl\metadonnees\metadonnes_originaux_1795_1797_tiff_exiftool.txt`
- `exiftool -g1 -ee -csv C:\bibli\projet_bdl\images\originaux_1891_1895\tif*.tif > C:\bibli\projet_bdl\metadonnees\metadonnes_originaux_1891_1895_tiff_exiftool.txt`
- `exiftool -g1 -ee -csv C:\bibli\projet_bdl\images\copies_1868_1874\tif*.tif > C:\bibli\projet_bdl\metadonnees\metadonnes_copies_1868_1874_tiff_exiftool.txt`
- `exiftool -g1 -ee -csv C:\bibli\projet_bdl\images\originaux_1914_1918\tif*.tif > C:\bibli\projet_bdl\metadonnees\metadonnes_originaux_1914_1918_tiff_exiftool.txt`
- `exiftool -g1 -ee -csv C:\bibli\projet_bdl\images\copies_1874_1881\tif*.tif > C:\bibli\projet_bdl\metadonnees\metadonnes_copies_1874_1881_tiff_exiftool.txt`

Contrôle des métadonnées

SourceFile	...	XMP- dc:Coverage	XMP-dc:Creator	XMP-dc:Rights	XMP-dc:Title
FR751142302_006_000001_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97
FR751142302_006_000002_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97
FR751142302_006_000003_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97
FR751142302_006_000004_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97
FR751142302_006_000005_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97
FR751142302_006_000006_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97
FR751142302_006_000007_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97
FR751142302_006_000008_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97
FR751142302_006_000009_A.tif	1795/1797	Bureau des longitudes (France)	Laboratoires d'histoires des sciences et de philosophie - Archives Henri Poincaré / Bibliothèque de l'Observatoire de Paris / Bureau des longitudes	Bureau des Longitudes - Séances 1e à 160e - 1795-97

Contrôle qualité

- Respect de l'arborescence de livraison
- Respect du plan de nommage des fichiers
- Contrôle du fichier de récolement
- Contrôle de la présence des fichiers dans les répertoires
- Conformité des fichiers
- Contrôle des métadonnées
- **Contrôles visuels**

Contrôles visuels

Contrôles exhaustifs

- sens d'orientation
- cadrage
- présence du filigrane
- consignes particulières

→ réalisés grâce à un affichage en mode vignette des jpeg

Contrôles visuels

Contrôles par échantillonnage

- Netteté et contraste
- Colorimétrie : double contrôle (visuel et valeurs des noirs et des blancs dans les mires)



Contrôles visuels

Contrôles par échantillonnage

- Netteté et contraste
- Colorimétrie : double contrôle (visuel et valeurs des noirs et des blancs dans les mires)

Trois problèmes :

- l'environnement de contrôle (écrans calibrés, lumière naturelle)
- concilier reprises et sécurité et conservation des originaux
- normalement, nécessité d'effectuer les contrôles visuels avec les originaux en regard

Résultats des contrôles

Très peu de problèmes détectés

- Quelques vues manquantes et une vue en double induisant un décalage dans la numérotation des fichiers
- Un fichier jpeg corrompu et illisible
- Une erreur de majuscule dans le filigrane
- Un problème de fidélité des couleurs constaté sur les reliures en peau verte qui n'a pas pu être réglé

+ Un problème détecté après validation : deux vues manquantes. Nécessité de revoir le plan de nommage pour ne pas avoir à renommer l'ensemble des fichiers

Politique d'archivage à l'OP

Ce qui est conservé:

- les fichiers images
- fichiers de récolement
- les mires
- les fichiers d'extraction des métadonnées
- la documentation liée au projet (CCTP, rapport de production)

Architecture des dossiers

- RECOLEMENT
- MIRES
- IMAGES/cote du document/TIFF ou JPEG
- DOCUMENTATION
- METADONNEES

Conservation

- données répliquées sur deux sites (Paris et Meudon)
- stockées sur disque à Paris
- archivées sur deux jeux de bandes régulièrement testées
- un jeu stocké dans un coffre ignifugé
- l'autre stocké sur le deuxième site géographique

Éléments de bilan

- Nécessité de disposer de moyens et compétences informatiques pour bien mener un contrôle qualité exhaustif
- Choix de privilégier la préservation des originaux (contrôle qualité effectué sans les documents pour éviter un nouveau transport)
- Un prestataire efficace et compétent pour un projet complexe et d'envergure
- Une politique de conservation du numérique qui ne répond pas encore aux critères d'un archivage pérenne