

# ETAPES/ACTIONS

## D'UN PROCESSUS DE NUMERISATION

### Définition et périmètre :

A partir de [Étude sur la numérisation au service de la recherche](#), Cabinet Six&Dix, Septembre 2018, p.11

« Un programme de numérisation est compris comme une démarche scientifique coordonnée de reproduction numérique de documents sur support (par numérisation ou transcription), concernant un ou plusieurs corpus, dotée de moyens humains et financiers dédiés, et ayant produit des résultats. »

Ils peuvent avoir pour objectif la préservation et la valorisation du patrimoine documentaire ou avoir été réalisés dans un objectif de service à la recherche. Ils portent sur toutes les disciplines et sur tous les types de documents : imprimés, manuscrits, archives administratives, archives scientifiques, documents iconographiques, cartes, plans, objets, documents audios/vidéos, documents numériques issus de projets de numérisation précédant et inclus dans un corpus visé par un nouveau projet de numérisation et documents numériques natifs inclus dans un corpus visé par un projet de numérisation.

Les actions listées ci-dessous ne sont pas présentées dans un ordre chronologique par ailleurs propre à chaque projet de numérisation. L'aspect itératif entre les actions n'est pas représenté pour l'instant ici. Les actions décrites ne sont pas obligatoires.

### Sélection intellectuelle

#### Choix des modalités de numérisation en interne ou via un marché

Définition des spécifications de numérisation

#### Définition d'une politique de réutilisation des documents & Collecte des droits & Signature de conventions

#### Préparation des documents

Restauration/Dépoussiérage/Récolement/Complétion/Collecte/Acheminement/Masicotage

#### Catalogage

#### Constitution de trains de numérisation & Préparation de l'intégration de documents déjà numérisés ou numériques natifs

Mention de l'état de conservation

Indication des spécifications de numérisation

GIS CollEx-Persée

<https://www.collexpersee.eu/> @CollEx\_IR #CollExPersée

#GTNumérisation <https://www.collexpersee.eu/groupes/numerisation/>

## **Dématérialisation/numérisation/acquisition /intégration des documents déjà numérisés ou numériques natifs**

Nommage des fichiers

## **Post-traitements**

Redressement / nettoyage/ retouche / passage en binaire / OCR / OLR / HTR

## **Contrôle qualité des images numérisées ou intégrées**

Renumérisation

## **Contrôle qualité des métadonnées**

## **Contrôle qualité des post-traitements**

## **Contrôle qualité des enrichissements**

## **Stockage local ou en cloud**

## **Sauvegarde**

*Définition : redondance, contrôle d'accès, contrôle d'intégrité, traçabilité, système d'alerte bruts de scans et fichiers OCR (TIFF > JPEG) / masters/...*

## **Pérennisation (= versement dans un Système d'archivage électronique)**

## **Création de métadonnées descriptives (dont bibliographiques) liées aux documents numériques**

Import des métadonnées issues du catalogage et liaison entre images et notices /  
Intégration de métadonnées de bases de données existantes/génération  
automatique de métadonnées / Saisie manuelle de métadonnées / métadonnées  
enrichies

## **Création de métadonnées de structure (table des matières)**

Ajout de pagination, foliotation, création de la table des matières ou de la table  
d'index, création de l'arborescence des fichiers

## **Création de métadonnées administratives (juridiques ou techniques liées à la numérisation)**

## **Enrichissement post numérisation/post intégration**

Structuration documentaire

Indexation des entités nommées / liaison référentiels autorités / création d'index

Numérisation de manuscrits > Transcription, translittération

Annotations

Enrichissements sémantiques selon les principes du web sémantique via des  
référentiels de communautés scientifiques

Encodage TEI

## **Préparation à la mise en ligne & mise en ligne**

GIS CollEx-Persée

<https://www.collexpersee.eu/> @CollEx\_IR #CollExPersée

#GTNumérisation <https://www.collexpersee.eu/groupe/numerisation/>

## **Diffusion**

Exposition et diffusion technique des données : entrepôt OAI - PMH / API / IIIF/création d'un graphe spécifique / création d'un jeu de données RDF.  
Moissonnage de données extérieures sur le site  
Affichage des conditions de réutilisation des documents  
Médiation numérique / Editorialisation / Expositions virtuelles  
Développement de fonctionnalités  
Maintenance du site web / Gestion de la qualité et traitement des anomalies

## **Référencement web**

Référencement naturel SEO  
Intégration au linked open data

## **Évaluation de la diffusion et des usages**

Existence d'un outil  
Analyse des statistiques  
Études ponctuelles en ligne

## **Exploitation**

Fouille de texte et de données