

https://dspace.org/

Role of digitization in the conservation of heritage collections and the democratization of information

AAKKI, Hassan

2016-06-22
Chief digitization project to BNRM
http://hdl.handle.net/10539/20533

Downloaded from a DSpace repository. https://dspace.org/

Rôle de la numérisation dans la conservation des fonds patrimoniaux et la démocratisation de l'information

Préparé par : Hassan AAKKI

Chef de division " Numérisation et informatique" à la BNRM

Chef de projet de numérisation à la BNRM

Email : aakkihassan@gmail.com

Tél : 0661857417

Plan de la communication

- **Introduction**
- **Aperçu historique**
- **Axe 1: les enjeux de la numérisation**
  - **La numérisation des documents : Contexte général**
  - **Quelques indicateurs**
  - **La numérisation des documents : quel intérêt ?**
  - **Intérêt du fonds de la BNRM**
  - **Pourquoi numériser ?**
  - **Quelques pré requis**
  - **La BNRM : une structure dédiée à la conservation et la diffusion de l'information**
- **Axe 2 : Expérience de la BNRM**
  - **Les étapes de la numérisation**
  - **Schéma technique**
  - **Produit fini**
  - **Moyens logiciels**
  - **Moyens matériels : Scanners**
  - **Diffusion : Portail de la BNRM**
  - **Quelques partenaires de la BNRM**
  - **Documents numérisés**
- **Conclusion**
- **Bibliographie**

Introduction

La numérisation constitue l'un des chantiers prioritaire de la Bibliothèque Nationale du Royaume du Maroc. L'objectif étant de valoriser, préserver, promouvoir et faciliter l'accès, via Internet, à nos fonds documentaires patrimoniaux, notamment les documents précieux, fragiles, uniques, rares et les documents les plus sollicités par les usagers de la BNRM.

Historique BNRM

1924 : Création de la Bibliothèque Générale, et le fonds de l'Institut des Hautes Études Marocaine fut le 1er noyau de ses collections.

1926 : La Bibliothèque Générale, devint « Bibliothèque Générale et Archives

1927 : C. Funk Brentano succéda à Pierre de Cénival, qui prit en charge la direction de la Section historique du Maroc à Paris.

1932 : Promulguée la loi instituant le Dépôt Légal qui va renforcer les capacités d'acquisition de la Bibliothèque.

Après l'indépendance, les conservateurs de la Bibliothèque Générale et Archives furent successivement :

- Messieurs Abdallah Regragui,
- Mohamed Ben al Abbas al Kabbaj,
- Mohay Eddine al Machrafi,
- Aberrahmane el Fassi,
- Mohamed Bencherifa,
- Ahmed Toufiq,
- et actuellement Driss Khrouz depuis septembre 2003.

2003 :

- Décision : construction d'une bibliothèque moderne répondant aux exigences des normes internationales.
- Le dahir du 11 novembre 2003 dote la Bibliothèque Générale et Archives d'un statut d'établissement public et change sa nomination en « Bibliothèque Nationale du Royaume du Maroc ».

2008 Le 15 octobre 2008

- la nouvelle Bibliothèque Nationale du Royaume du Maroc, implantée au centre ville de Rabat entre Boulevard Ibn Hazm et boulevard Ibn Khaldoun, fut inaugurée par SM Mohamed VI.

Historique de la Bibliothèque numérique du Maroc (BNM)

2009 :

- Création d'un atelier de numérisation au sein de la BNRM
- Acquisition du matériel de numérisation
- Formation et encadrement des équipes de numérisation
- Démarrage de la production

2010 : Lancement d'une version BNM

2011 :

- Lancement du programme de numérisation des manuscrits
- Lancement du programme de numérisation des revues
- Numérisation des premiers numéros de certains titres des journaux
- OCRrisation des revues et des livres latins
- Lancement du programme de numérisation des collections spécialisées
- Lancement du programme de numérisation des documents audiovisuels

2012 :

- Projet de numérisation de masse des manuscrits (1 000 000 pages)
- Acquisition d'une solution de sauvegarde
- Conception et réalisation du portail de la BNRM

Axe 1:

Les enjeux de la numérisation

1. Quelques indicateurs

L'évolution rapide des NTIC bascule nos modes de vie :

- Maroc numérique 2013
- Internet : 4 millions abonnés à internet
- Tel fixe : 34 millions
- Mobile : 35 millions (115%)
- PC : 400000 unités/année (60% portables)
- Le nom de domaine : .ma : 44000 enregistrements
- Facebook : Plus de 5 millions de Facebookers au Maroc : 36 place mondiale
- Twitter : 700000 abonnés (5eme place)
- E. Commerce: 500 Mdh
- 4G en 2013
- Aujourd'hui on parle de la Bibliothèque 2.0 = Bibliothèque + Web 2.0

2. Intérêt du fonds de la BNRM

Les collections de la BNRM constituent le noyau dur de tout le patrimoine documentaire marocain ; Et dont la dimension est à la fois d'ordre nationale et internationale :

Types	Titres/volumes/unités
Manuscrits	13 486 volumes (34 000 titres)
Monographies	Plus de 600 000 volumes
Lithographies	615 titres
Périodiques	19 974 titres • Revues : 13 874 titres • Journaux : 6000 titres
Images animées	• 683 DVD • 119 Vidéo cassettes
Enregistrements sonores	• Cassettes audio : 869 • Disque vinyles : 442 • Bandes magnétique : 181
Ressources électroniques	• CD : 2099 • Disques vinyles : 442 • Autres support « jeux, disquettes…» : 155
Collections offertes	•Plus 165.812 volumes

3. La numérisation des documents : quel intérêt ?

La numérisation des documents s'inscrit dans les priorités des États qui ont relevé le défi de « la société de l'information ». En effet, elle contribue à la démocratisation de l'accès à l'information et au développement de l'économie de la connaissance.

4. Pourquoi numériser ?

La numérisation à pour objectifs de conserver, assurer l'accès à distance, diffuser et exploiter le contenu

a. Conserver

- Préserver les documents fragiles et/ou difficilement manipulables (manuscrits, lithographies, journaux…),
- Faciliter la consultation et la manipulation des documents via le support informatique,
- Permettre une consultation simultanée par plusieurs personnes.

b. Assurer un accès distant

Les ouvrages numérisés sont accessibles à distance via internet permettant aux usagers de les consulter sans avoir à se déplacer.

c. Exploiter le contenu

Avant la numérisation d'un document se fait en mode image. Autrement dit c'est une "photographie" du document, cependant il ne permet pas d'exploiter le contenu du document et en particulier d'y faire des recherches, juste une lecture linéaire. Mais aujourd'hui avec les NTIC on applique des opérations d'OCR ou d'ICR afin d'obtenir un Document en mode texte.

d. Diffusion de l'information

La mission de la BNRM s'articule autour des axes suivants : Collecter, Traiter, Diffuser et conserver

5. Quelques pré requis :

La numérisation nécessite les ressources suivantes :

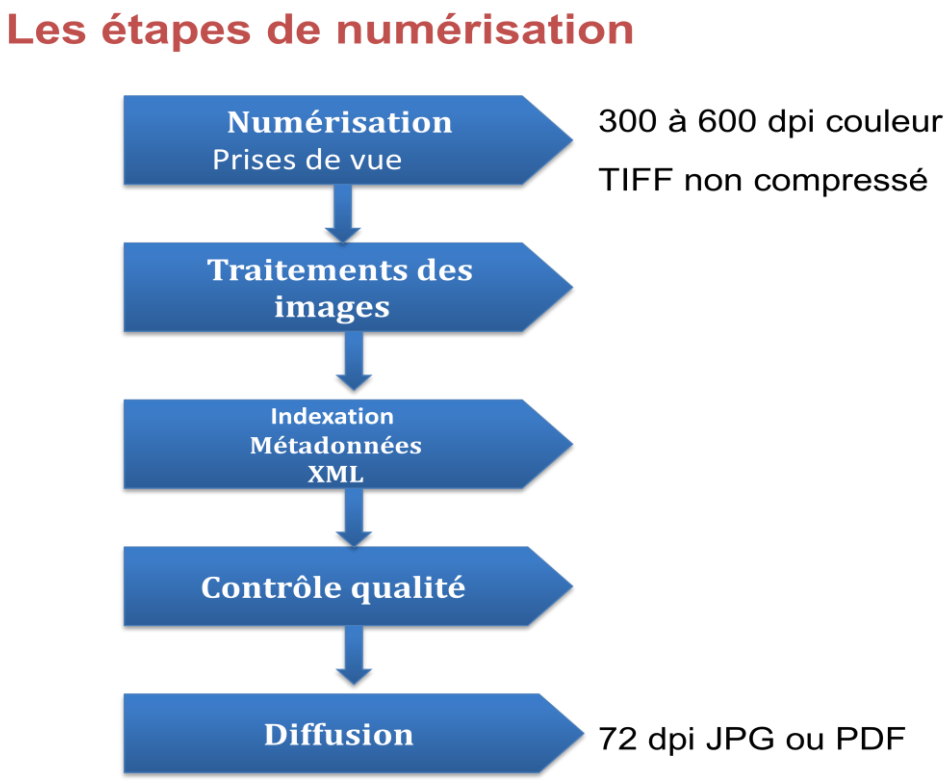
- Des moyens financiers conséquents ;
- Des moyens humains conséquents à la hauteur du projet envisagé;
- Des moyens technologiques importants et évolutifs;
- Une bonne maîtrise des évolutions technologiques pour assurer la pérennité des accès.

Axe 2 :

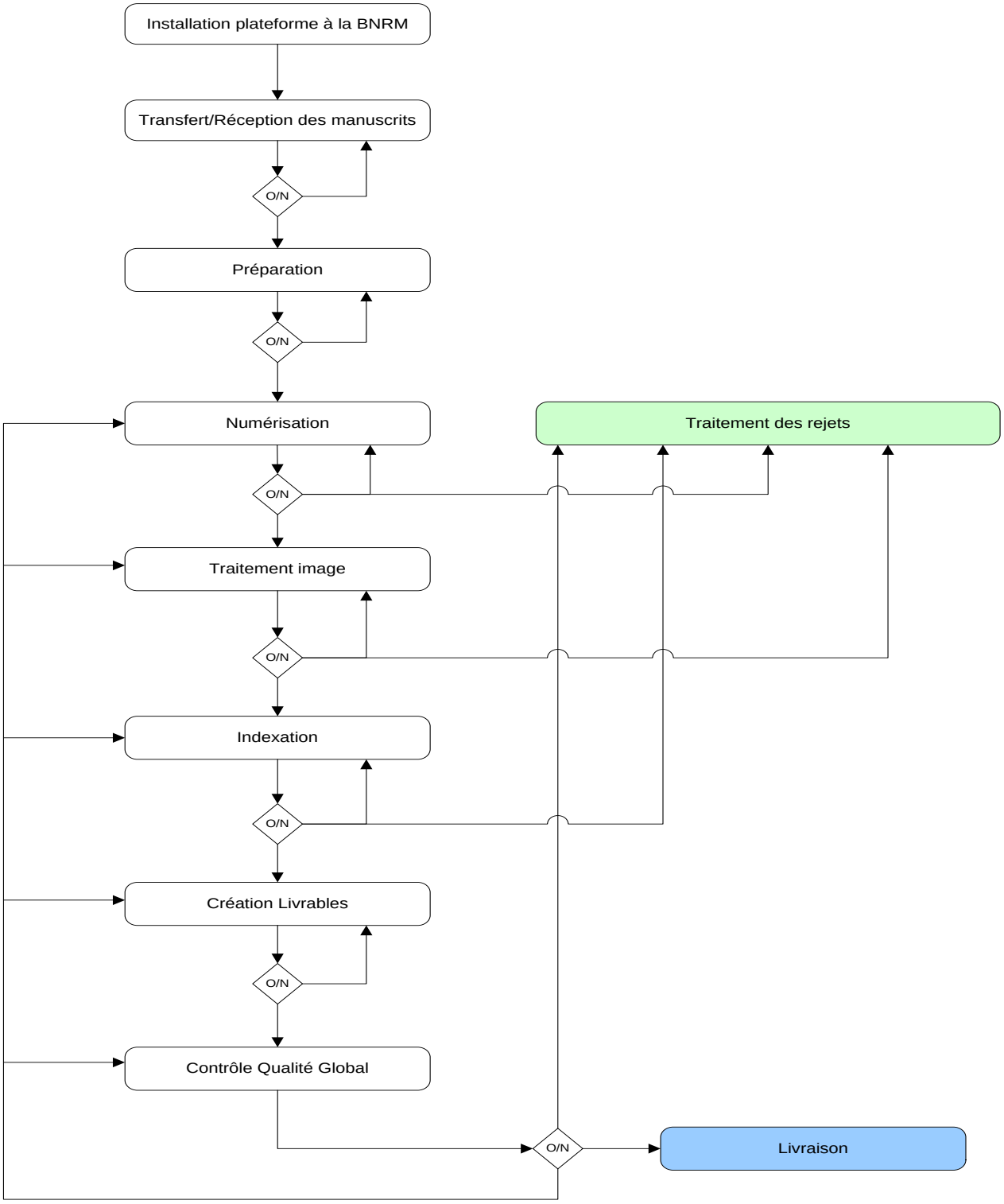
Expérience de la BNRM

6. Les étapes de la numérisation

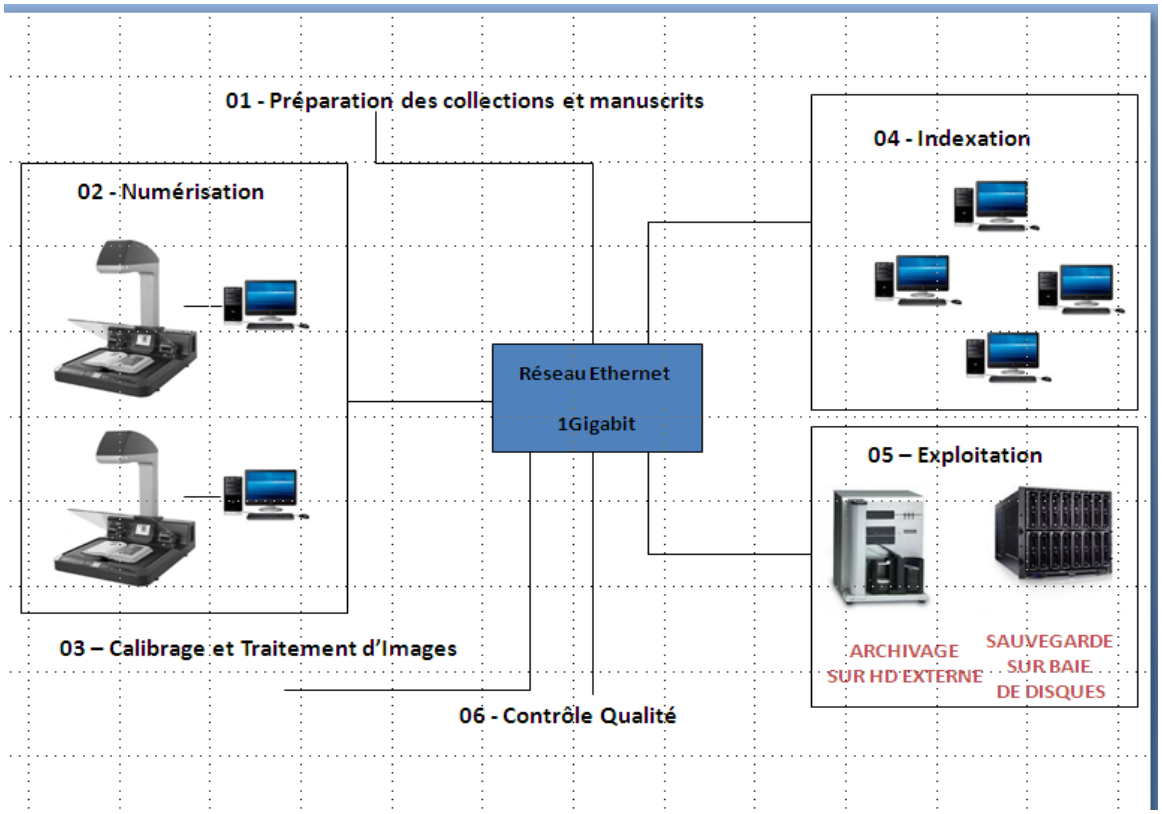
Chaque document subi un processus normal composé des étapes décrites ci-dessous :



7. Schéma générique du processus de numérisation



8. Schéma technique



9. Produit fini

Production des images en couleur des documents à numériser :

- Création des tables de correspondance pages physiques/pages électroniques ;
- Saisie des informations nécessaires à la création de l'exemplaire numérique ;
- La structuration suivant les schémas XML fournis ;
- Pour chaque image produite sur le scanner il faut lui attacher un profil ICC qui permettra de gérer au mieux la chaîne couleur.
- Le profil ICC Adobe RGB 98 doit être retenu. C'est une norme de référence dans ce domaine. Il faut fournir comme profil de sortie de ses images le Profil Adobe RGB 98.
- Préparation des données en vue de la fourniture d'un fichier pour lecture sur supports nomades ;
- Préparation des données images et des données associées ;
- Contrôle qualité et fourniture des rapports correspondants ;
- Stockage ;
- Reprise des documents rejetés.

A l'issue des divers traitements requis les éléments suivants :

- La production numérisée et les fichiers associés correspondants, les fichiers de métadonnées contenant les éléments de gestion de production livrés dans chaque manuscrit numérisé ;
- Document numérique : répertoire produit et correspondant à un exemplaire numérique. Le document numérique contient :
    - un fichier de métadonnées :
        - Meta données techniques ;
        - Meta données descriptives ;
        - Meta données informatives.
    - un répertoire d'images TIFF pour la sauvegarde [300 dpi - 600 dpi] ;
    - un répertoire d'images traité pour la diffusion ;
    - pour chaque image produite sur le scanner il faut lui attacher un profil ICC qui permettra de gérer au mieux la chaîne couleur.
    - Possibilité de conversion des fichiers en file system qui peut être supporté par le AIX.
- Répertoire document : dossier informatique contenant l'ensemble des fichiers correspondant aux traitements pour un document numérique.

11. Moyens à mettre en place

La BNRM dispose des matériels avec les caractéristiques suivantes :

a. Moyens matériels : Scanners

- Caméra CCD : de 200 à 600 dpi.
- Temps de numérisation sur format A2, A1 à 400 x 400 dpi
- optique : 3,4 s en couleur et 1,4 s en niveaux de gris.
- Caméra longue durée de vie, Prévisualisation en temps réel.
- Calibration et focus complètement automatique.
- Vitesse identique indépendamment du format du document
- Une station de numérisation autonome compacte.
- Une couleur neutre qui s'intègre parfaitement dans
- l'environnement de travail et ne perturbe pas l'analyse des couleurs.
- Lancement automatique ou intuitif de la numérisation
- (fermeture de la vitre, changement de page ou pédale).
- Porte-livre avec accès facile, Permet de numériser assis ou debout.
- Pas de lumière agressive pour l'opérateur

b. Moyens logiciels

i. Capture des images

Intègre les fonctionnalités de numérisation les plus courantes :

- Simplicité d'utilisation : en 2 clics, le document est numérisé et archivé dans l'application;
- Pas d'installation d'application spécifique pour numériser;
- Numérisation décentralisée disponible rapidement;
- Déploiement facile en mode client léger;
- Exploitation de documents optimisés par de nombreux traitements d'image disponibles

- Traitements d'image : redressement, détourage, ...
- Compression : la technologie d'hyper compression (ex. : 59 Ko au lieu de 1050 Ko pour un document A4 couleur 300 dpi)
- Pré-indexation avec saisie des données dans un formulaire
- Nommage des fichiers numérisés afin de faciliter la recherche du document numérisé sur le serveur et au format désiré
- Multi-format de sortie : TIFF noir & blanc G4, JPEG, PDF, PDF/A, PDF MRC (option)

ii. Logiciel de traitement des images

- la manipulation d'images (rotation, agrandissement, détourage, redressement) : recadrer le document afin d'obtenir une version semblable à l'original.

- la lecture et l'enregistrement d'images, conversion de formats (PDF, PDF /A, JPG, JPG2000, TIFF, etc.) : exploiter les fichiers grâce à des formats universels.

- l'amélioration de la qualité (binarisation, suppression de fond, réduction du bruit etc.) : optimiser la clarté et le rendu visuel du document numérisé.

- la gestion du multipages (création de signets dans un PDF à la volée, traitement d'arborescence complexe etc.) : rendre les pages des documents accessibles.

- l'amélioration des caractères (amincissement, épaississement, lissage, filtrage etc.) : optimiser la lisibilité du document.

- l'OCR (ABBYY FineReader) : permettre l'exploitation des contenus des documents.

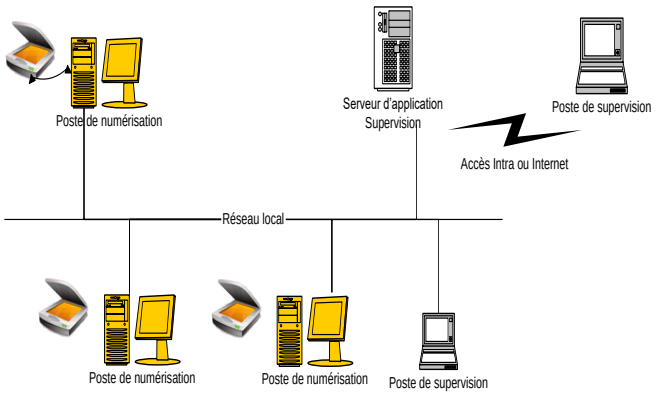
iii. Logiciel d'indexation

Caractéristiques et fonctionnalités

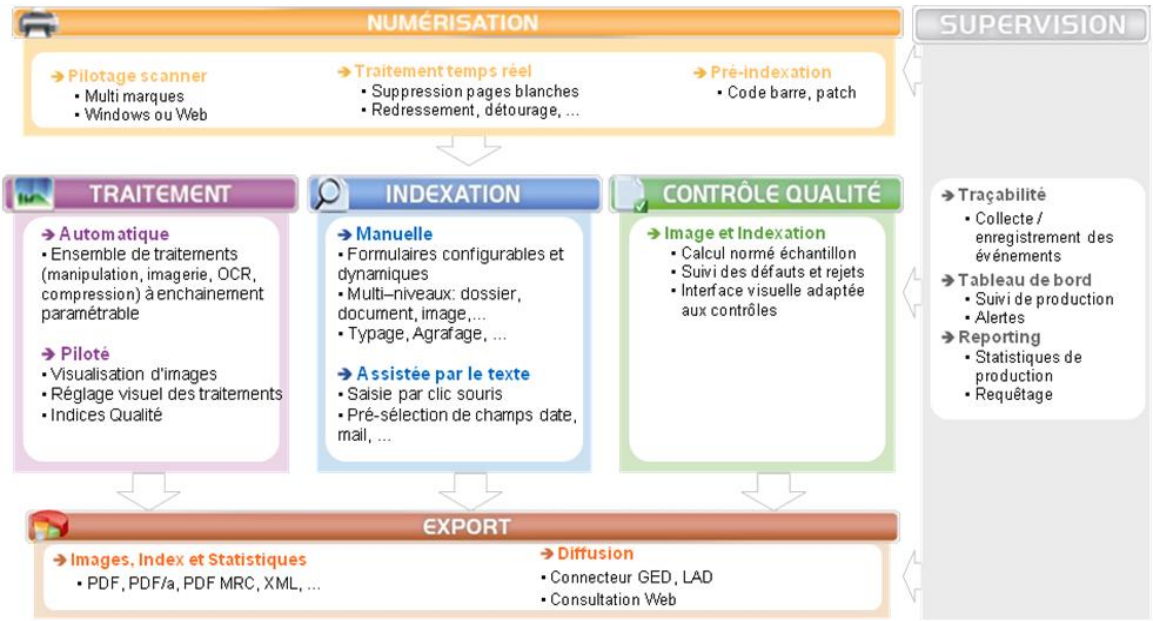
Différents niveaux d'indexation	Le formulaire d'indexation peut être divisé en plusieurs parties permettant l'indexation sur 3 niveaux : - indexation niveau lot : le ou les index associés au lot sont saisis une seule fois pour la totalité du lot, - indexation niveau dossier : le ou les index associés aux dossiers sont saisis pour chaque dossier - indexation niveau page : le ou les index sont saisis pour chaque page.
Facilité d'indexation	- Passage automatique au champ suivant dès la fin de saisie du champ en cours - Passage automatique d'un document à un autre lorsque tous les champs du document sont remplis - Champs chaînables (selon le choix effectué dans un premier champ, les champs suivants sont différents) - Différents types de champs : menus déroulants, boîtes à cocher, champs texte, bouton de choix - Raccourcis clavier paramétrables - Récupération d'informations dans une base de données à partir d'un index saisi
Sécurisation de l'indexation	Le contrôle du format des données utilise des expressions régulières. Il permet de détecter des erreurs de saisie dans des champs particuliers. Il est possible de définir des caractères obligatoires (@ et « . » pour les adresses mails), une longueur (5 caractères pour les codes postaux) ou bien un format (xx/xx/xxxx pour une date). L'interfaçage avec une base de données donne la possibilité : - de valider une information et mettre le champ en erreur si la valeur n'a pas été trouvée dans la base (par exemple, si le code client saisi n'existe pas cela signifie que la saisie n'est pas bonne et que l'utilisateur doit corriger le champ), - de récupérer des informations sans avoir à les saisir (par exemple, lorsqu'on saisie le code client, on peut récupérer de manière automatique les noms, prénoms, adresses, ou d'autres informations concernant le client déjà présentes dans la base).
Configuration matérielle et logicielle	- Windows 2000, XP ou Vista - Processeur Pentium IV ou équivalent (1,5GHz minimum) - 2 Go de RAM - Ecran conseillé : 1280 x 800 pixels
Interfaçable avec le module de scan	- Insertion de feuilles à partir d'un scanner dans un dossier existant - Interfaçage direct avec le scanner pour avoir une application de numérisation et d'indexation
Format d'entrée des images	- JPEG couleur ou niveaux de gris - TIF couleur, niveaux de gris ou noir et blanc - PDF couleur, niveaux de gris ou noir et blanc
Format de sortie	A la fin de l'indexation, les données sont exportées aux formats texte ou xml afin d'être intégrées dans une application tierce, GED, application métier, ERP. L'export peut également se faire dans un autre format grâce à un plugin Java. Lorsque l'application est interfacée à une base de données, on peut écrire les informations au fur et à mesure de la saisie.

iv. Logiciel de supervision

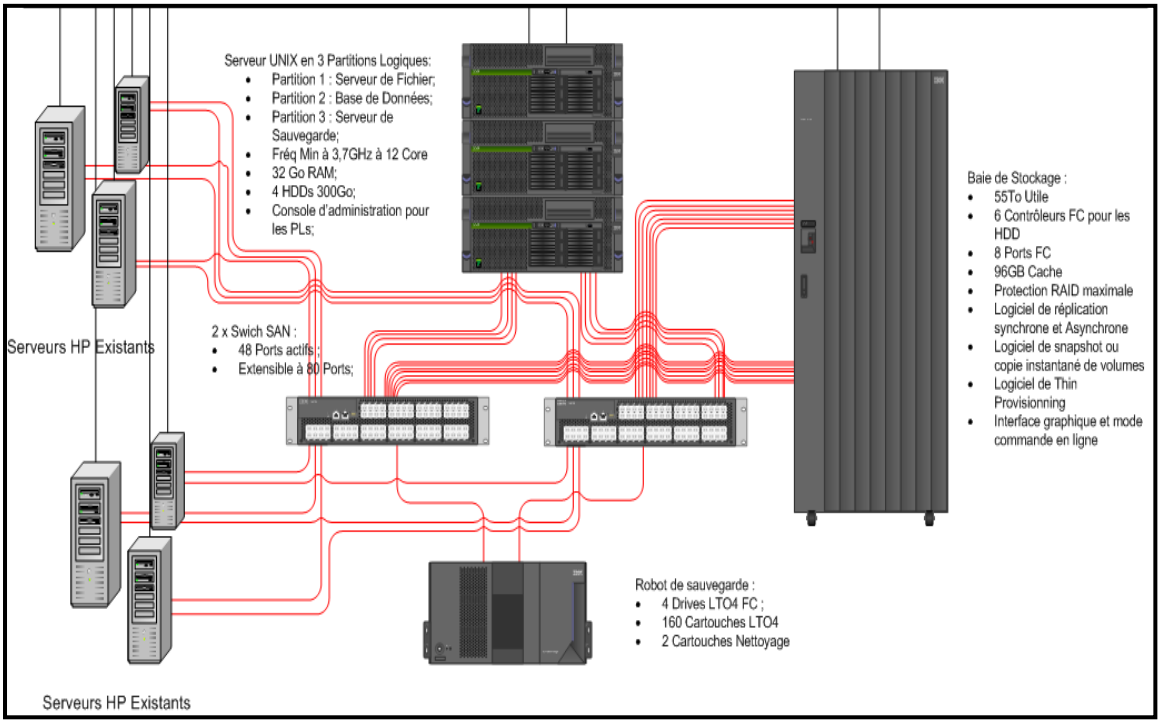
Architecture technique générique



i. Synthèse

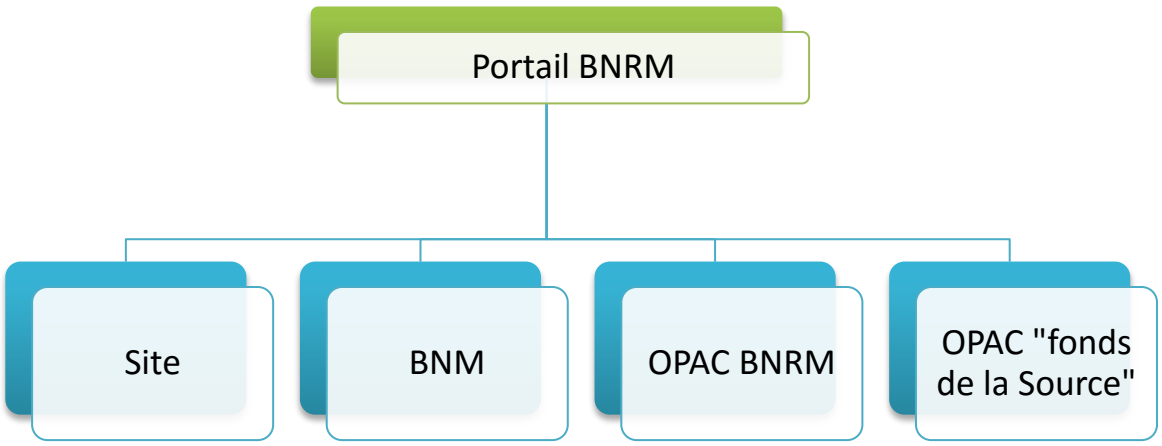


c. Solution de sauvegarde



12. Diffusion : Portail de la BNRM

a. Architecture du portail de la BNRM



b. Liste des logiciels Microsoft

Les logiciels nécessaires au développement et à la mise en route de la BNM sont :

- **Microsoft SharePoint Server ;**
- **Microsoft SQL Server ;**
- **Microsoft Windows Server ;**
- **Microsoft Visual Studio ;**
- **Microsoft Silverlight.**

c. Spécifications fonctionnelles du portail

La BNM offre les fonctions suivantes :

- Deux interfaces : arabe et française ;
- Charte graphique bien spécifique ;
- Back office pour la gestion de la bibliothèque, rajout de livres, gestion des livres ;
- Moteur de Recherche sur tout le contenu du site ;
- Mise à disposition des fonctionnalités de « turnning the pages » et de « Loupe » ;
- Utilisation des Technologies Microsoft : MOSS 2007 – SQL Server 2005/2008 – Silverlight - Open XML comme format de stockage ;

13. Documents numérisés

Liste des titres des Revues numérisées

Titre	Nombre de Numéros	Nombre de pages
Bulletin économique et social du Maroc	167	21808
Héspéris	43	17400
Archives Marocaines	35	14152
Lamalif	196	14000
L'avenir illustré	262	6000
France-Maroc	109	4360
Kalima	34	2632
Nord Sud	40	2340
Collection Hespéris	12	2090
L'indicateur marocain	46	1144
Souffles	22	1122
Maroc	22	1080
Les Archives Berbères	12	1056
Maghreb	22	946
Revue Marocaine	9	558
La Revue musicale du Maroc	12	296
Bâtir	7	378
HESPERIS Tamuda	46	13000
Les sources Inédites	8	4800
المغرب الجديد	17	780
المجلة المغربية	9	732
أنفاس	8	446
مجلة المغرب	45	1318
Total : 23 Titres	~ 1000	~ 100 000

Documents Patrimoniaux

Types de Document	Nombre de titre
Manuscrits	400 000 pages
Lithographies	67
Livres	43
Cartes & Plans	560
Photos	1018

14.Statistiques : Bibliothèque Numérique Marocaine : Période : 2012

Nombre de visites :

- En 2012 : 37**.246**
- Moyenne de consultation par mois : 3103 visiteurs/mois.

Nombre des pages consultées : 299.448

Moyenne de pages consultées par mois : **24.954** pages/mois

Documents les plus consultés :

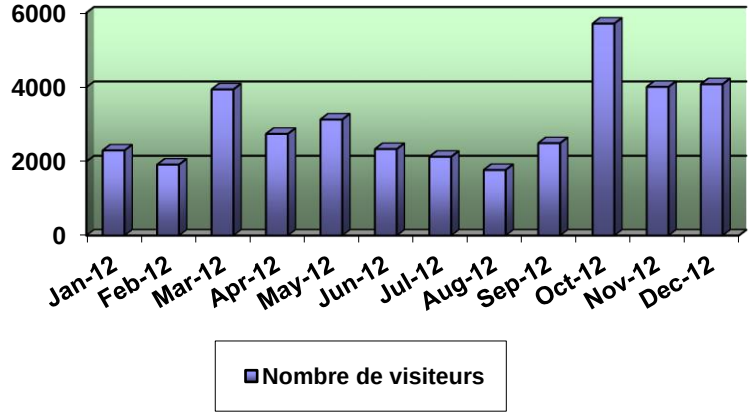
- Hespéris, Souffles, Bulletin économique et social du Maroc.

Synthèse géographique (2012)

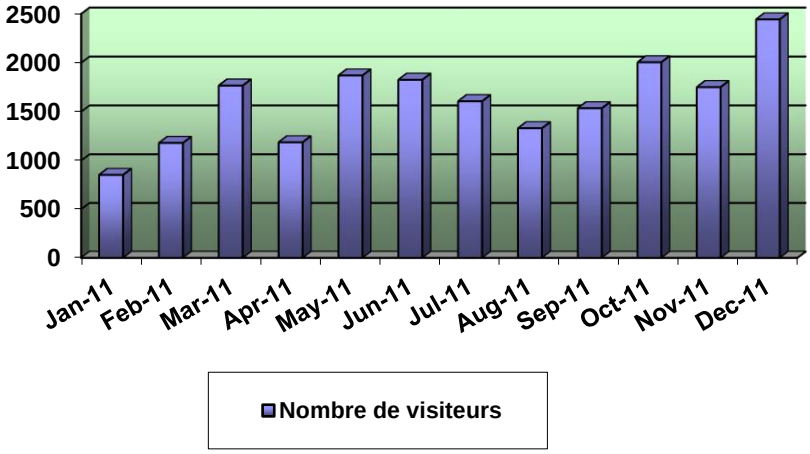
Pays	Nombre de visites	% Visites
Maroc	27 889	75 %
France	2 597	7 %
Algérie	1 567	4 %
États-Unis	740	2 %
Egypte	564	1,5 %
Tunisie	402	1 %
Espagne	391	1,05 %
Arabie saoudite	347	0,93 %
Autres	2749	7,5 %

Nombre de visites :

En 2012

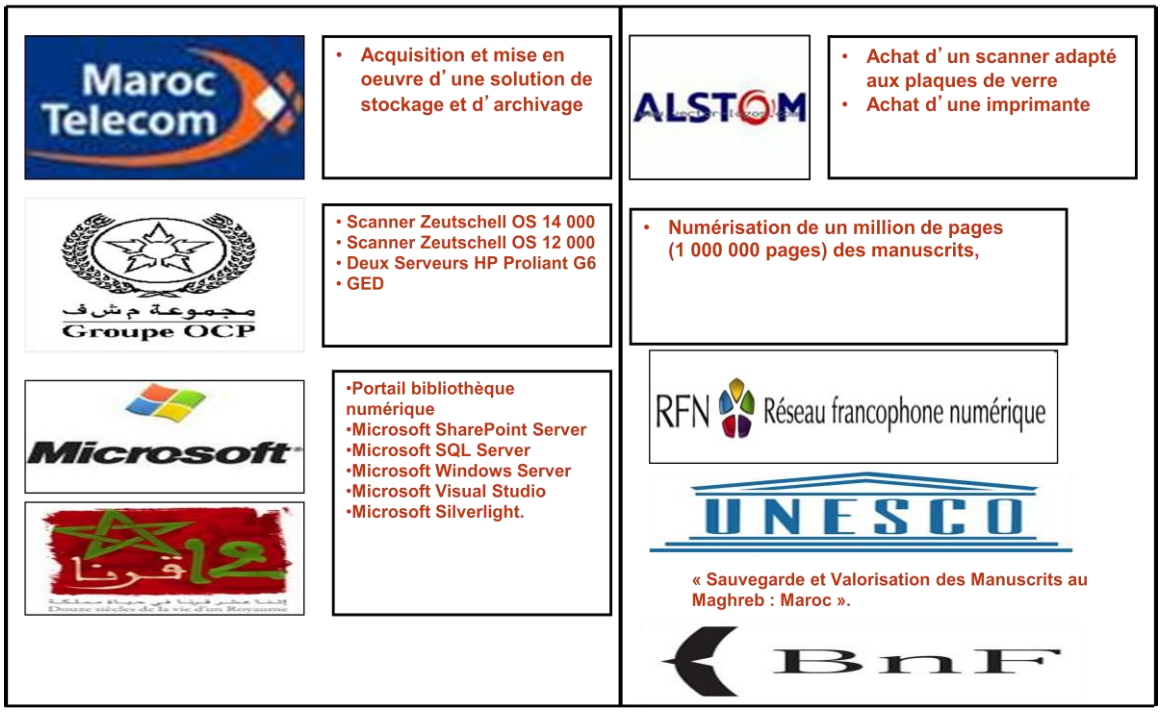


En 2011



15.Quelques partenaires de la BNRM

Partenaires de la BNRM



Conclusion

Les limites et les contraintes de la numérisation sont :

- Limites techniques ;
- Contraintes juridiques ;
- Limites humaines ;
- Limites financières ;
- Réticences socioculturelles.