

# Université Kasdi Merbah de Ouargla

Faculté des Sciences, de la Technologie et des Sciences de la Matière

Département des Mathématiques et d'Informatique



Mémoire

## MASTER ACADEMIQUE

Domaine : Mathématiques et Informatique

Filière : Informatique

Spécialité : Informatique Industrielle

Présenté par : *Nasria Bouhyaoui et Hadjer Zaghabi*

Thème

*Etude comparative des solutions de stockage et d'archivage*

*électronique en local et sur le « Cloud »*

*Optimisation et Intégration d'une solution OPEN SOURCE*

Soutenu publiquement

Le : / 06 / 2013

Devant le jury :

<b>M. Mohamed Salim MEFLAH (MAB)</b>	Président	UKM Ouargla
<b>Mme. Chafika BENKHEROUROU (MAB)</b>	Examinatrice	UKM Ouargla
<b>Dr. Dris KORICHI (MCA)</b>	Encadreur/ Rapporteur	UKM Ouargla

Année Universitaire : 2012 / 2013

# *Dédicaces*

*Je dédie ce modeste travail à :*

*A mes chers Parents,*

*Pour tous vos sacrifices pour moi, nul mot ne saura exprimer mon amour envers vous. Que Dieu vous protège et vous accorde une longue vie, car je ne pourrais jamais oublier la tendresse et l'amour dévoué par lesquels ils m'ont toujours entouré depuis mon enfance.*

*A mes frères et à mes sœurs,*

*Que Dieu vous garde, Je vous aime et je vous souhaite une vie pleine de succès et de réussite et surtout la femme de mon frère.*

*A mes oncles et à mes tantes*

*A tous les familles Zaghabi, Djaghab, chaabi*

*À tous mes cher(e)s ami(e)s*

*Surtout En particulier Sid Ahmed pour son aide permanant et son implication, et encouragement et pour avoir été là pour moi chaque fois que j'en avais besoin.*

*Merci pour m'avoir donné du goût et du sens à notre amitié. Je tiens aussi à remercier vivement ET votre père Ismail*

*A mon cher binôme Nasria, pour son aide, son soutien et son encouragement durant mon cycle universitaire.*

*ET... A tous ceux que j'amie et à toutes les personnes qui m'ont prodigué des encouragements.*

*A tous les membres de la promotion de master LMD en informatique 2012-2013 de l'université Kasdi Merbah d' Ouargla.  
...Et à tous ceux qui me connaissent...*

**HADJER**

# *Dédicace*

*Je dédie ce modeste travail à celle qui m'a donné la vie, qui ma bercée et a veillé mes nuits, celle qui ma vécue que pour me voir un jour réussir, à ma très chère mère.*

*À celui qui a puisé sa vie et sa jeunesse, et qui n'a jamais su dire non pour subvenir à mes besoins, à mon très cher père.*

*Nul mot ne saura exprimer mon amour envers vous. Que Dieu vous protège et vous accorde une longue vie.*

*À mes frères et mes sœurs.*

*À tous les membres de la famille Bouhyaoui & Kama.*

*Je vous remercie pour votre amour inconditionnel.*

*Que dieu vous garde, Je vous aime et je vous souhaite une vie pleine de succès et de réussite.*

*À mes chères amies et sur tout à mon âme sœur Zineb et ma binôme Hadjer.*

*À tous les membres de la promotion de master LMD en informatique 2012-2013.*

*À tous ceux que m'aide et à tous ceux que j'aime que cette modeste dédicace puisse vous témoigner ma profonde gratitude pour votre immense amour, votre confiance et vos paroles apaisantes qui m'ont toujours fait garder la sourire.*

*Nasria*

## *Pour notre encadreur*

*« Nous tenons à remercier sincèrement **Dr. Dris KORICHI**, d'avoir proposé ce sujet et qui, en tant que Directeur de mémoire, s'est toujours montré à l'écoute et très impliqué tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'il a bien voulu nous consacrer et sans qui ce mémoire n'aurait jamais vu le jour »*

# Remerciements

*Nous souhaitons adresser nos vifs remerciements à :*

*Notre créateur (Allah) pour nous avoir donné de la force à accomplir ce travail.*

*Notre gratitude pour notre encadreur, monsieur le Docteur **Dris Korichi**  
Maître de Conférences Habilité -A- Au département Mathématiques et  
Informatique à UKMO, pour ces précieux conseils, sa disponibilité, la confiance  
qu'il nous a toujours témoigné et la sollicitude dont il nous a entouré, et ce tout  
au long de l'élaboration du présent travail.*

*Nous remercions vivement les membres du jury d'avoir accepté de siéger et de  
juger notre travail. Nous les remercions aussi pour l'honneur qu'ils nous font en  
acceptant d'examiner ce mémoire.*

*Nous remercierons vivement monsieur le directeur de la CNR d'avoir accepté de  
nous recevoir au sein de son établissement. Nos chaleureuses remerciements à  
Mohamed Korichi, le responsable informatique, ainsi que au directeur de  
l'administration au CNR pour leurs aides tout au long de notre séjour au CNR.*

*Enfin, nous adressons une pensée particulièrement affective à nos amis qui ont  
rendu agréable nos longues années d'études. Nous tenons à remercier tous ceux  
qui ont collaborés de près ou de loin à l'aboutissement de ce travail ; Qu'ils  
acceptent nos humbles remerciements.*

### *Liste des abréviations*

<i>Abréviation</i>	<i>Désignation</i>
<b>GED</b>	Gestion électronique des documents
<b>LGPL</b>	Lesser General Public License
<b>SSO</b>	Single Sign-On
<b>XML</b>	Extensible Markup Language
<b>LDAP</b>	Lightweight Directory Access Protocol
<b>ECM</b>	Entreprise Content Management
<b>CIFS</b>	Common Internet File System
<b>ODF</b>	Open Document Format
<b>TIFF</b>	Tagged Image File Format
<b>SAE</b>	Système d'Archivage Electronique
<b>RAD</b>	Reconnaissance Automatique de Documents
<b>PNG</b>	Portable Network Graphics
<b>OCR</b>	Optical Character Recognition
<b>LAD</b>	Lecture Automatique de Documents
<b>JPEG</b>	Joint Photographic Experts Group
<b>J2EE</b>	Java 2 Enterprise Edition
<b>ISO</b>	International Organization for Standardization
<b>ICR</b>	Intelligent Character Recognition
<b>GIF</b>	Graphics Interchange Format
<b>DPI</b>	Dot Per Inch

## *Liste des abréviations*

---

<b>RM</b>	Record management
<b>DAM</b>	Digital Asset Management
<b>WCM</b>	Web content management
<b>DM</b>	Document management

# Sommaire

Table des Figures .....	VI
Table des tableaux.....	VII
Introduction Générale.....	IX

## *Chapitre 1 Etat de l'art d'Archivage et Stockage Électronique*

Introduction .....	1
1. Archivage et stockage électronique .....	2
1.1. Qu'est ce qu'un document numérique? .....	2
1.2. Qu'est ce qu'un archive ? .....	2
1.3. C'est quoi l'Archivage électronique ? .....	2
1.4. Pourquoi archiver ? .....	3
1.5. Cycle de vie des documents .....	3
1.5.1. Les archives courantes .....	3
1.5.2. Les archives intermédiaires .....	4
1.5.3. Les archives définitives .....	4
1.6. Dématérialisation .....	4
1.6.1. Deux types de dématérialisation .....	4
1.6.2. Le scanner «numériseur» .....	5
1.6.3. Formatage .....	5
1.6.4. Les technologies de traitement automatisé .....	7
1.7. Les enjeux de l'archivage électronique .....	7
1.8. Les exigences d'archivage électronique .....	8
2. Les supports de stockage .....	9
2.1. Les supports de stockage Traditionnel : .....	9
2.2. Stockage en local .....	11

2.2.1. NAS (Network Attached Storage) .....	11
2.2.2. SAN (Storage Area Network) .....	11
2.2.3. DAS (Direct Attached Storage) .....	11
2.3. Le stockage dans le nuage (Cloud).....	12
2.3.1. Le cloud.....	12
2.4. Tableau comparatif .....	15
Conclusion.....	16

## *Chapitre 2: GED et Etude comparative de quelques solutions Open Sources*

Introduction .....	17
1.La Gestion Electronique de Documents .....	18
2.Les étapes de GED .....	19
2.1. L'acquisition des documents.....	19
2.2. Classement.....	19
2.3. Stockage.....	19
2.4. Diffusion de documents.....	19
2.5. Archivage.....	19
3.les fonctionnalités de GED .....	19
3.1. L'indexation/Métadonnées .....	20
3.1.1. Indexation manuelle .....	20
3.1.2. Indexation automatique .....	20
3.2. Moteur de recherche .....	21
3.3. La gestion des versions.....	22
3.3.1. Le système de réservation (check-in/check-out).....	22
3.3.2. Incrémentation des versions .....	22
3.4. Import/export.....	23
3.4.1. Import / export en masse .....	23
3.4.2. Export sur support physique.....	23
3.5. Intégration de la GED.....	23

3.5.1. Comme un espace de stockage.....	23
3.5.2. Intégration avec les outils de bureautique.....	23
3.6. Gestion des droits .....	23
3.6.1. Niveaux de gestion.....	24
3.7. Workflow (processus) .....	24
3.8. La gestion des dossiers .....	24
3.9. Numérisation .....	25
3.10. Gestion de cycle de vie.....	25
4.Les avantages de la GED .....	26
5.La GED en mode SaaS.....	26
6.Les solutions de GED .....	26
6.1. Logiciel libre ou logiciel propriétaire.....	26
6.2. Les open sources étudiées .....	27
6.2.1. Alfresco .....	27
6.2.2. Nuxeo .....	28
6.2.3. Knowledge Tree .....	29
6.2.4. Jahia.....	29
6.2.5. Maarch.....	29
6.2.6. Freedom.....	30
6.3. Tableau comparatif.....	31
Conclusion.....	33

### *Chapitre3: Etude et comparaison des plateformes Alfresco et Nuxeo*

Introduction .....	34
1.La solution Nuxeo .....	35
1.1. Technologies utilisées .....	35
1.2. Des pré-requis.....	35
1.3. La prise en main de l'application .....	36
1.4. Les espaces disponibles .....	38

1.5. Les concepts .....	39
1.6. Déplacement .....	41
1.7. Droits d'accès.....	41
1.8. La Recherche .....	42
2.La solution Alfresco.....	42
2.1. Technologies utilisées .....	42
2.2. Des pré-requis d'Alfresco.....	43
2.3. Prise en main de l'application .....	43
2.4. Vocabulaire contient dans Alfresco .....	44
2.5. Les espaces disponibles .....	45
2.6. Les droits d'accès .....	47
2.7. La Recherche .....	48
3.Comparaison et Synthèse (Alfresco et Nuxeo).....	48
Conclusion.....	50

## *Chapitre 4 : Mise en œuvre d'une solution open source « Alfresco » au sein du CNR*

Introduction .....	51
1.Présentation général de la CNR .....	52
2.Présentation du projet.....	53
2.1. Les Besoins de la société CNR : .....	53
2.2. Etude de la mise en place de la solution GED Alfresco.....	54
2.2.1. L'amélioration de fonctionnement de CNR :.....	54
2.2.2. L'architecteur réseau d'application.....	55
2.3. Solution et mise en route de l'installation :.....	57
2.3.1. Outils utilisés pour l'installation .....	57
2.3.2. Les étapes de l'assistant d'installation .....	57
2.3.3. Configuration et développement de l'environnement.....	58
2.3.4. Interface graphique de la solution .....	60
Conclusion.....	66

Conclusion Général et Perspectives..... 67  
Bibliographie..... 68

## *Table des Figures*

<b>Figure 1</b> : Cycle de vie des documents.....	4
<b>Figure 2</b> : Dématérialisation de documents papiers .....	5
<b>Figure 3</b> : Dématérialisation de processus .....	5
<b>Figure 4</b> : Les réseaux de stockage en local .....	12
<b>Figure 5</b> : les 3 couches du cloud .....	13
<b>Figure 6</b> : Dorpbox .....	14
<b>Figure 7</b> : Google drive .....	14
<b>Figure 8</b> : Icloud .....	14
<b>Figure 9</b> : Skydrive .....	15
<b>Figure 10</b> :Alfresco.....	28
<b>Figure 11</b> : Nuxeo.....	28
<b>Figure 12</b> : Knowledge Tree.....	29
<b>Figure 13</b> : Maarch .....	30
<b>Figure 14</b> : Freedom .....	30
<b>Figure 15</b> : l'architecture de Nuxeo .....	35
<b>Figure 16</b> : installer des programmes supplémentaires .....	36
<b>Figure 17</b> : Nuxeo démarré.....	36
<b>Figure 18</b> : configuration de l'application.....	37
<b>Figure 19</b> : Etape la configuration de l'application.....	37
<b>Figure 20</b> : sélectionner les modules fonctionnels .....	37
<b>Figure 21</b> : télécharger les modules sélectionnés .....	38
<b>Figure 22</b> : configuration définie.....	38
<b>Figure 23</b> : Connexion à Nuxeo EP.....	38
<b>Figure 24</b> : tableau de bord.....	39
<b>Figure 25</b> : domaine du serveur default.....	39

<b>Figure 26:</b> les espaces dans Nuxeo.....	40
<b>Figure 27 :</b> Les espaces de travail .....	40
<b>Figure 28:</b> l'architecture d'Alfresco .....	42
<b>Figure 29 :</b> Page d'identification .....	44
<b>Figure 30 :</b> Tableau de bord Alfresco Share .....	46
<b>Figure 31 :</b> Personnalisation du tableau de bord .....	46
<b>Figure 32:</b> Tableau d'administration.....	47
<b>Figure 33:</b> Espace documentaire .....	47
<b>Figure 34 :</b> Organigramme de la direction de CNR .....	52
<b>Figure 35 :</b> L'architecteur globale de réseau d'application.....	56
<b>Figure 36 :</b> L'interface bureautique FTP.....	59
<b>Figure 37:</b> L'interface WebDav .....	59
<b>Figure 38:</b> Page d'identification de l'application .....	60
<b>Figure 39:</b> page de tableau de bord .....	61
<b>Figure 40:</b> profil d'administrateur (Directeur général de CNR) .....	61
<b>Figure 41 :</b> utilisateurs de CNR.....	62
<b>Figure 42:</b> Prévisualisation des documents .....	62
<b>Figure 43:</b> la modification des métadonnées.....	63
<b>Figure 44:</b> File Plan de record management .....	64
<b>Figure 45:</b> L'importation des fichiers électroniques .....	64
<b>Figure 46:</b> La Déclaration des fichiers comme records .....	65
<b>Figure 47:</b> Disposition schedule.....	65

## *Table des tableaux*

<b>Tableau 1</b> : les caractéristiques des supports de stockage .....	10
<b>Tableau 2</b> : tableau comparatif des services Cloud .....	15
<b>Tableau 3</b> : tableau comparatif entre les solutions open source de GED ; .....	32
<b>Tableau 4</b> : Comparaison et Synthèse (Aflresco - Nuxeo) .....	48

## Introduction Générale

De nos jours, les entreprises, et même les particuliers, devront faire face à un fond de plus en plus volumineux de documents à archiver et à stocker. En effet, elles émettent, reçoivent et accumulent une masse importante de documents.

Le développement des nouvelles technologies de l'informatique ont données naissance à une large panoplie de solutions pour traiter la problématique d'archivage et stockage d'informations. Il s'agit des solutions dites « solutions GED » : Gestion Electronique des Documents.

La Gestion Electronique des Documents (GED) fournit les outils nécessaires afin de gérer électroniquement les documents qui circulent dans l'entreprise.

L'objectif de notre travail de recherche consiste en l'étude de ce type de solutions GED.

Après avoir réalisé une étude comparative des différentes solutions GED, nous avons retenu la solution la plus riche en matière de fonctionnalités que nous avons intégré dans une PME<sup>1</sup> Algérienne.

Le premier chapitre sera complètement consacré à l'état de l'art des outils d'archivages et de stockage, que ce soit en local ou sur le « Cloud ».

Une étude plus détaillée des solutions GED sera présentée au deuxième chapitre. Nous allons également présenter une large étude comparative des solutions Open Source les plus utilisées dans le monde.

Au troisième chapitre, nous avons jugé plus efficace d'étudier en profondeur les deux meilleurs outils : « Alfresco » et « Nuxeo ».

Dans le quatrième et dernier chapitre, nous avons présenté une mise en œuvre de la solution open source choisie, « Alfresco », au sein d'une PME Algérienne. Il s'agit du Caisse Nationale de la Retraite (CNR) de Ouargla. En plus de notre volonté de se rapprocher du monde professionnel, nous avons voulu faire profiter une administration locale de nos travaux.

La conclusion générale et les perspectives à donner à notre travail seront présentées à la fin de ce mémoire.

Un petit lexique utile relatif au métier d'archivage sera présenté dans l'annexe.

---

<sup>1</sup>PME : Petites et Moyennes Entreprises

# *Chapitre 1*

## *Archivage et Stockage Electronique*

*Etat de l'art*



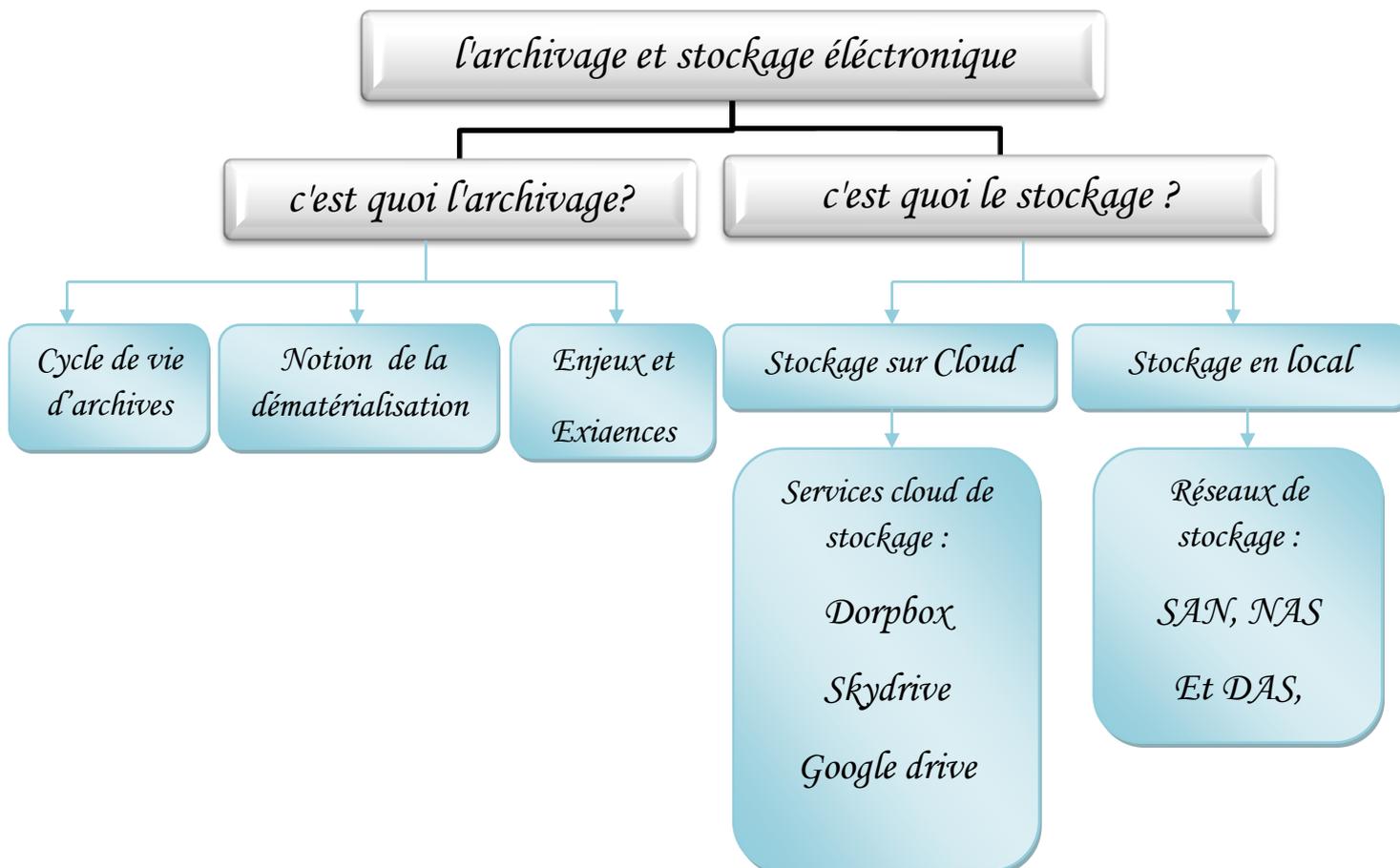
## Introduction

L'entreprise moderne d'aujourd'hui devra faire face à un fond de plus en plus volumineux de documents. Ces documents devraient être archivés et stockés soigneusement pour pouvoir les trouver facilement en cas de besoin. C'est pourquoi, l'archivage et le stockage électronique sont devenus de plus en plus indispensables pour les sociétés ou même pour les particuliers. Ils ont besoins d'un système d'archivage fiable et efficace dans le temps.

L'archivage électronique permet la manipulation des documents de différentes natures; le papier qui doit être dématérialisé en fichier informatique, souvent en format PDF ou JPEG pour les images, ou des documents produit d'une manière électronique nativement.

Ce premier chapitre portera principalement sur l'aspect théorique des questions d'archivage et de stockage. Nous énoncerons la définition de l'archivage électronique et le stockage, la dématérialisation des papiers et des processus, le cycle de vie d'archivage, les enjeux et les exigences et les différents espaces de stockage en local et sur le Cloud.

### Plan du chapitre



## 1. Archivage et stockage électronique

A l'effet d'assurer la viabilité et augmenter la productivité, les entreprises et les particuliers ont toujours besoin de trouver solution idéale pour optimiser le stockage et l'archivage de leurs documents. Dans cette partie nous allons identifier les contextes de base et définir les concepts généraux de notre étude.

### 1.1. Qu'est ce qu'un document numérique?

- C'est un objet informatique immatériel et manipulable avec un ordinateur. Il peut être une image, un fichier son, un ensemble de données organisées en fichier, un écrit électronique (la création est réalisée sur ordinateur), etc.
- Un document numérique permet de séparer les caractéristiques d'un document classique, à savoir sa présentation (métadonnées), son contenu (informations), son architecture, offrant alors la possibilité d'une exploitation séparée.
- Définition de l'ISO  
*« Ensemble formé par un support et une information, généralement enregistrée de façon permanente, et tel qu'il puisse être lu par l'homme ou la machine » [1]*
- **Information** : est un ensemble de données liées entre elles et porteuses de sens. [2]

### 1.2. Qu'est ce qu'un archive ?

Avant d'aborder la notion d'archivage électronique nous devons rappeler qu'est ce qu'un archive. « Les archives sont l'ensemble des documents, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité.»[3]

### 1.3. C'est quoi l'Archivage électronique ?

**L'archivage électronique n'est ni du stockage ni de la sauvegarde**, Il convient donc de bien dissocier l'archivage à celle de stockage et la sauvegarde, alors :

- **Le stockage**, s'apparente aux actions, outils et méthodes permettant d'entreposer des contenus électroniques et servant de base au traitement ultérieur des contenus ;
- **La sauvegarde «backup»**, qui peut être définie comme étant l'ensemble des actions, outils et méthodes destinés à dupliquer des contenus électroniques

d'origine dans un but sécuritaire pour éviter leur perte. La durée de la sauvegarde est fonction de sa périodicité et porte le plus souvent sur du court terme. [29]

L'archivage est un ensemble de processus tandis que le **stockage** est un outil qui permet d'entreposer des contenus électroniques et qui sert de base au traitement ultérieur de ces derniers.

- La **sauvegarde** est une copie de secours destinée à éviter l'effacement des données en cas de dysfonctionnement du dispositif sur lequel elles sont enregistrées. En d'autres termes, la sauvegarde permet la restauration des données perdues en cas d'incident. Quant à l'archivage, il permet la conservation des données afin de respecter les obligations légales et règlementaires. [29]

Donc l'**Archivage électronique** c'est l'ensemble des méthodes et des outils qui permettent la conservation, pour exploiter et rendre accessible dans le temps tout document électronique à titre de preuve ou à titre informatif. Après l'archivage définitif, le contenu est figé il ne peut pas être modifié.

**Les méthodes peuvent être** : identifiés, recueillis, classés et conservés des informations numériques.

#### 1.4. Pourquoi archiver ?

L'archivage électronique permet une meilleure structuration et organisation des données, il réduit le volume des données non structurées, et évite l'archivage aléatoire.

En effet les besoin d'archivage sont nombreux :

- pour laisser des preuves de toutes les réalisations des entreprises.
- Pour la réutilisation des données.
- Pour faciliter la recherche.
- Pour « communiquer envers ses clients, ses partenaires, ses salariés, voire la société. »[5] et la conservation des données patrimoniales sur le long terme.

#### 1.5. Cycle de vie des documents

Le cycle de vie des documents doit gérer les documents depuis leurs créations jusqu'à leurs utilités opérationnelles.

##### 1.5.1. Les archives courantes

(Archives actifs) : sont les documents qui sont d'utilisation courante, généralement consultés dans le cadre de la production.

### 1.5.2. Les archives intermédiaires

Les documents qui ne sont plus d'usage courant, conservés dans des délais limités, pour des raisons juridiques et/ou gestion.

### 1.5.3. Les archives définitives

Sont les dossiers destinés à être conserver dans une langue durée, pour des raisons historiques ou patrimoniales, leurs consultation est exceptionnel.

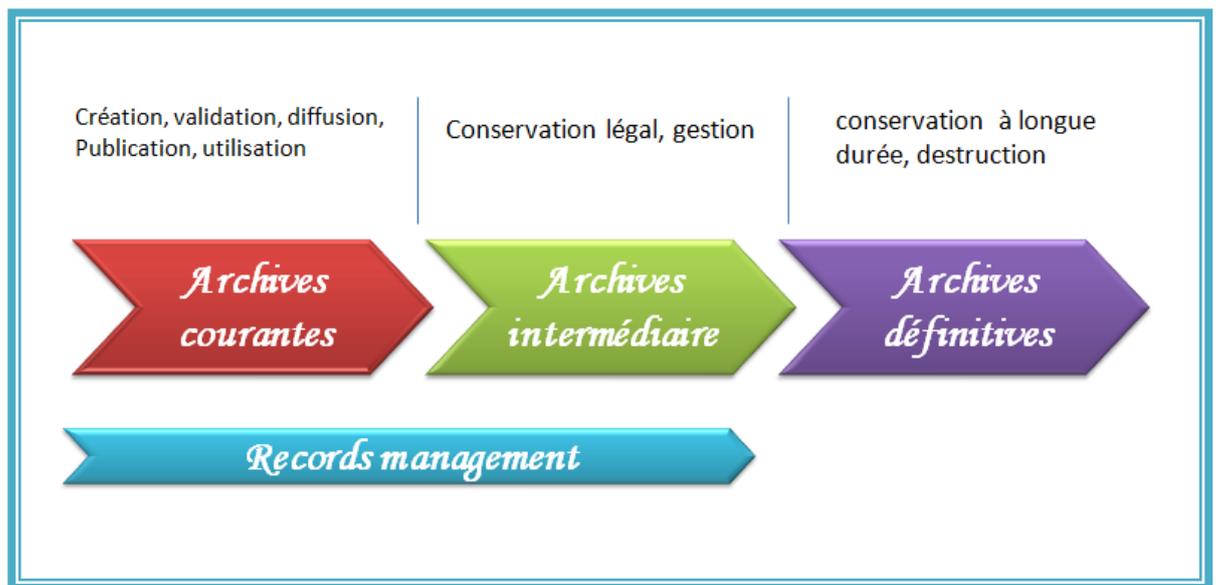


Figure 1 : Cycle de vie des documents

## 1.6. Dématérialisation

Avec l'évolution technologique, la dématérialisation est devenue une véritable révolution pour l'entreprise. La dématérialisation est la transformation des documents ou des processus sur support numérique, pour la gestion numérique des flux de l'ensemble des données et qui permettent l'échange, le stockage et l'archivage de l'information en toute sécurité, assurent des gains d'espace de stockage et de temps et la conservation de la traçabilité des données, avec l'efficacité et la rapidité.

### 1.6.1. Deux types de dématérialisation

#### 1.6.1.1. Dématérialisation de documents papier

La dématérialisation repose sur le processus de numérisation de documents originaux papier. [30]



**Figure 2:** Dématérialisation de documents papiers

### 1.6.1.2. Dématérialisation de processus

La dématérialisation des processus désigne la production de documents nativement électroniques. Les documents sont créés, validés, diffusés, et conservés sous forme numérique, sans avoir recours au papier



**Figure 3:** Dématérialisation de processus

### 1.6.2. Le scanner «numériseur»

Les Scanners sont des périphériques d'acquisition des documents qui permettent la transformation des documents sous forme d'image numérique.

### 1.6.3. Formatage

L'action qui suit l'opération de numérisation est le formatage, qui consiste à l'enregistrement sous un format de fichier. Le choix de format doit achever la pérennité souhaité.

Dans la plus part de temps le choix de format est liée à l'importance de fichier, pour les documents le format le plus adopté est le PDF.

Voici les principaux critères qui doivent dicter le choix du format de fichier :

- la garantie de l'intégrité des données ;

- la qualité de compression, qui conditionne la future exploitation de l'image (OCR, impression en grand format, etc.) ;
- la rapidité de numérisation ;
- le poids du fichier, qui conditionne le volume de stockage et la vitesse d'affichage du document. [31]

#### 1.6.3.1. Les formats

- **Formats usuels:**

- ✓ Formats textuels (TXT, WORD, RTF, HTML, XML)
- ✓ Formats image (BMP (bitmap), TIFF (Tagged Image File Format), PNG (Portable Network Graphics), JPEG (Joint Photographic Experts Group) )
- ✓ Formats audio, vidéos
- ✓ Formats mixtes PDF (Portable Document Format) normalisé dans sa forme PDF/A. [1]

- **Les formats plus utilisés**

- **PDF (Portable Document Format)**

Les PDF peuvent être simples ou complexes, structurés ou non structurés. Ils peuvent contenir du texte, des images, des graphiques et même des contenus multimédia. Les annotations, les métadonnées, les liens hypertextes, les signets ainsi que la signature électronique sont également supportés par ce format de fichier.

- **Le JPEG (Joint Photographic Experts Group)**

*Les images JPEG* ont l'extension « .jpg », « .jpeg », « .jpe » ou « .jfif », le **JPEG** est très adapté aux photographies ou images volumineuses qui contiennent des millions de couleurs telles qu'un fond d'écran, *capture d'écran vidéo etc.*[32]

- **TIFF (Tagged Image File Format)**

Le format TIFF sauvegarde des images raster (graphisme pixel ou bitmap), une succession de pixels sous forme de tableau (éléments d'image), auxquelles sont attribuées des couleurs. [6]

#### 1.6.4. Les technologies de traitement automatisé

- **La technologie OCR « Reconnaissance Optique de Caractères »** L'OCR est une technologie qui permet la conversion des documents de type image obtenu par la numérisation des documents papiers ou fichiers PDF vers des formats texte pour les traiter et exploités, le principe est de reconnaître la suite des caractères, par la comparaison de la forme lue avec une base de données.
- **La RAD « Reconnaissance Automatique de Document »** est une technique très récente, qui fait l'identification automatique des document numérisé, elle permet de reconnaître le type de document traité, à travers la comparaison de forme reconnu à celle qui existe dans la base de connaissance des formes de document.
- **LAD « Lecture Automatique de Document »** est une technologie qui permet d'extraire des informations contenu dans un document numérisé de type formulaires structurés ou semi-structurés, et donne la possibilité d'extraire facilement les métadonnées, cette technologie dépend de trois techniques indispensables OCR, IRC et RAD.
- **L'ICR « Intelligent Character Recognition »**, ou reconnaissance intelligente de caractères), est un système d'OCR avancé intégrant des technologies d'intelligence artificielle. Un logiciel d'ICR peut compléter sa base de connaissance au fur et à mesure de la reconnaissance et donc étendre sa capacité de reconnaissance. Ces techniques sont utilisées notamment pour la reconnaissance de caractères manuscrits. [30]

#### 1.7. Les enjeux de l'archivage électronique

Les enjeux de l'archivage sont au nombre cinq :

**Sécuritaire** : garder la confidentialité des données et les protégés.

**Juridiques** : La validité des données destinées a des preuves pour un audit ou un juge, les données doivent respectes l'authenticité, l'intégrité, accessibilité, pérennité.

**Logistique** : l'existence des matériels ou moyen nécessaire pour faciliter la recherche et l'exploitation des données.

**Technique** : l'enjeu technique est double dans l'espace avec les problèmes d'interopérabilité entre systèmes, et dans le temps avec le défi de pérennité des

données sur le long terme, face à l'obsolescence récurrente des formats, supports et outils de restitution. [6]

**Financier** : l'inexactitude ou l'absence de données peut générer des coûts inutiles pour l'entreprise (amende, redressement fiscal, condamnation judiciaire). [33]

### 1.8. Les exigences d'archivage électronique

Depuis quelques années, aux questions d'archivage viennent s'ajouter/juxtaposer les questions « d'archivage légal » ou « à valeur probante », c'est-à-dire la conservation sous forme électronique des documents pouvant servir de preuve [31]. Cet aspect n'est pas couvert directement par les solutions qu'il existe, mais par le respect des exigences d'archivage légale.

Pour l'implémentation d'un archivage légale, il est nécessaire de garantir:

- **L'intégrité des documents**

Elle assurer que le document électronique n'a subi aucune modification ni altération depuis qu'il a été créé. La mise en place de la signature numérique peut prévenir ces risques et garantir l'intégrité des documents.

- **Pérennité**

Pour assurer la pérennité, il faut garantir l'intégrité des données archivées dans le temps, donc le document doit être lisible dans n'importe quel moment.

- **Traçabilité**

La traçabilité est l'historique de l'ensemble des traitements sur les objets numériques. la traçabilité assure l'enregistrement de la trace de tous les traitements opérés sur le document électronique ou données. Il permet également de garantir l'intégrité.

- **Authenticité**

Qualité d'un document ou d'une donnée dont l'origine et l'auteur sont certifiés et incontestables. Dans le monde numérique, la signature électronique est le procédé permettant de garantir cette qualité. [32]

- **Accessibilité**

L'accessibilité est la capacité d'accéder facilement au document archivé.

« La Norme ISO 15489 définit l'accès comme "les droits et les moyens de rechercher de l'information".

- *Droit d'accès : seules les personnes autorisées aient accès aux archives et au système ;*
- *Modalité et moyens de rechercher l'information : les documents doivent être classés, identifiés et indexés selon un processus assurant la recherche d'un ou plusieurs documents, les informations d'indexation devant être conservées dans le système. »*

## 2. Les supports de stockage

### Stockage de documents

La problématique du stockage est incontournable. Le non prise en compte de cet aspect peut entraîner des situations critiques. Les problématiques sont les suivantes :

- **Le support de stockage** doit être adapté le mieux possible avec le volume des documents. Il doit aussi, en fonction de la fréquence de consultation et de l'importance des données, offrir un faible temps d'accès.
- **L'organisation du stockage** peut être hiérarchisée en fonction du contenu des documents (texte, vidéo, image, etc.) de leurs provenances, états, types, etc.
- **La durée de conservation** doit aussi être considérée afin de permettre une épuration périodique du système, en vue de faciliter le stockage et d'alimenter les archives. À cet effet, le système doit prendre en compte une sortie vers un archivage définitif.
- **Le lieu de stockage** doit être différent pour plus de sécurité de données en cas d'incendie.
- **Les sauvegardes** doivent être mise en œuvre avec une périodicité en adéquation avec les impératifs de l'entreprise. [45]

#### 2.1. Les supports de stockage Traditionnel :

- **Les supports magnétiques:** qui sont (disques durs, bandes magnétiques...etc.) les supports magnétiques offrent des caractéristiques importantes « la rapidité d'accès et de transfert des données, la fiabilité du support, la migration technologique et la possibilité de supprimer un fichier. »[10] et la ré-écriture peut être protégée par des logiciels.

- **Les supports optiques :**

**Les disques optiques de technologie WORM** (Write Once Read Many) qui ont la propriété d'interdire toute modification des données enregistrées. C'est actuellement le support le plus utilisé pour archiver des volumes relativement réduits.

**Les disques optiques réinscriptibles.** Plusieurs technologies sont disponibles. À des fins d'archivage, on utilise principalement des supports dont la réécriture est protégée par un procédé logiciel. [11]

**Les supports optiques:** WORM (*Write Once Read Many*), CD-ROM, DVD et Blu-Ray).

- **Les supports magnéto-optiques :** ce support de stockage allie la technologie magnétique pour la lecture et l'écriture, à la technologie optique pour le positionnement des têtes de lecture/écriture. Cette caractéristique permet d'obtenir une densité de stockage très importante. Ces disques constituent une nouvelle catégorie de supports de stockage permanents, appelée mémoire de masse amovible. [33]

Les supports	Durée de vie	Capacité
CD	100 ans	700 MO
DVD	100 ans	4,7 GO
Blu-RAY	50 ans et plus	25 GO en simple face 50 GO en double face
Disque dur	2 à 5 ans	Peut arriver à 500 GO
Disquettes	5 à 15 ans	1,4 MO
Bands magnétiques	30 à 100 ans	Peut arriver à 1,6 Téra octets
Magnéto-optique	Estimé 50 ans	60 GO une face 120 GO deux face

**Tableau 1 :** les caractéristiques des supports de stockage

Les supports de stockage leur durée de vie est limité, et pour assurer la pérennité d'accès aux données voulue il faut faire des sauvegardes et/ou des migrations, les disques

magnétique peuvent être sécurisés à l'aide de la technologie de RAID (Redundant Arrays of Inexpensive Disks) qui permet de dupliquer le disque et ainsi de le remplacer à chaud.

## 2.2. Stockage en local

### 2.2.1. NAS (Network Attached Storage)

Est un moyen de stockage (serveur de stockage) directement attaché aux réseaux locaux (Ethernet, TCP/IP), l'accès fournit est au niveau de fichier.

- Parmi ces avantages :
  - ✓ facile à mettre en place
  - ✓ plus économique au SAN
  - ✓ adapté au partage de fichier
  - ✓ l'ajout des capacités de stockage sans avoir à immobiliser le réseau.
- l'inconvénient majeur qu'il est plus lent que le SAN ;

### 2.2.2. SAN (Storage Area Network)

Est une technologie de stockage en réseau en fibre optique basé sur le protocole Fibre Channel, « Le SAN offre une infrastructure de stockage autonome (composé de plusieurs baies de disques formant un réseau spécifique) qui se connecte directement aux serveurs. ». [13]

Le traitement fournit est au niveau blocs (comme dans le cas d'un disque directement connecté au serveur)

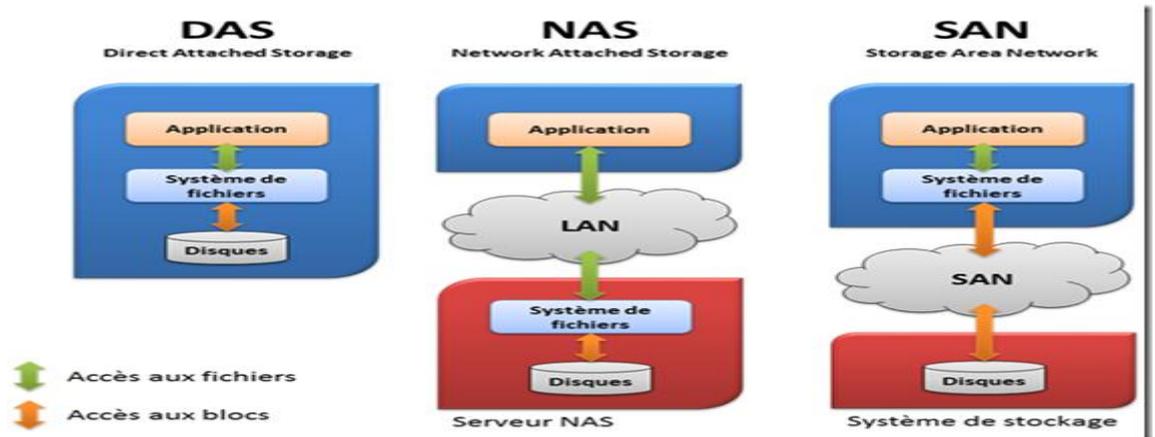
- Parmi ces avantages :
  - ✓ Fourni une performance d'accès très haut.
  - ✓ Centralisation des données.
  - ✓ La prise en charge d'un nombre quasi illimité de matériels. [20]
  - ✓ La prise en charge des techniques de type : reprise à chaud, mise en miroir et réplication(RAID). [20]
- Les inconvénients : complexité de réseau. Prix plus élevé (onéreux).

« **Fibre Channel** = 1 seul protocole pour gérer les connexions (entre ordinateurs, entre périphériques et ordinateurs). » [20]

### 2.2.3. DAS (Direct Attached Storage)

Est une unité de stockage externe connectée directement à un ordinateur en USB. Le DAS est plus imposant que le disque dur externe, car il peut supporter plusieurs

disques durs. Vous pouvez ainsi cumuler un espace de stockage important pour centraliser vos données multimédias et les partager via le réseau par l'intermédiaire d'un ordinateur. [46]



**Figure 4 :** Les réseaux de stockage en local

### 2.3. Le stockage dans le nuage (Cloud)

Le Cloud est un service de stockage des données en ligne, le stockage des informations sur des serveurs distants accessibles de façon facile et transparents pour les utilisateurs via l'internet de n'importe quel lieu. Ses serveurs ont des capacités de stockage énorme (avec un abonnement).

Cet aspect recouvre **l'ensemble des solutions de stockage distant.**

#### 2.3.1. Le cloud

Le nuage ou « cloud » désigne un modèle technologique qui héberge applications et données sur des centres de données distribués, hors site, qui sont accessibles via l'Internet par les ordinateurs et appareils mobiles équipés d'un navigateur Web. [45]

##### 2.3.1.1. Les différents services de cloud

Le cloud computing peut être décomposé en trois couches :

- Applicative (SaaS, Software as a Service);
- Plateforme (PaaS, Platform as a Service);
- Infrastructure (IaaS, Infrastructure as a Service) ;

La **Figure 5 ci-dessous** représente les différentes couches du cloud computing de la couche la moins visible pour les utilisateurs finaux à la plus visible. L'infrastructure as

a Service (IaaS) est plutôt gérée par les architectes réseaux, la couche PaaS est destinée au développeur d'applications et finalement le logiciel comme un service (SaaS) est le « produit final » pour les utilisateurs.



**Figure 5** : les 3 couches du cloud

### 2.3.1.2. Types de Cloud Computing :

Le concept de Cloud Computing est encore en évolution. On peut toutefois dénombrer trois types de cloud computing :

- **le cloud privé (ou interne)**: réseau informatique propriétaire ou un centre de données qui fournit des services hébergés pour un nombre limité d'utilisateurs.
- **le cloud public (ou externe)** : prestataire de services qui propose des services de stockage et d'applications Web pour le grand public. Ces services peuvent être gratuits ou payants.
- **le cloud hybride (interne et externe)** : un environnement composé de multiples prestataires internes et externes. [16]

### 2.3.1.3. Les services de Cloud

Avec cette nouvelle technologie, les internautes ont vu apparaître de nombreuses applications permettant de synchroniser leurs données entre divers appareils (ordinateurs, smartphones, tablettes). Ces données étant stockées en ligne, via l'application cliente. Parmi les plus connues, dropbox, skydrive, Google Drive et icloud.

- **Dropbox**

Il s'agit d'un service qui permet de stocker tous types de données en ligne (documents, photos, vidéos, musique, etc.) et d'y accéder depuis n'importe quels systèmes. Dropbox permet de sauvegarder, partager et synchroniser les

données. Il se compose d'une interface en ligne accessible depuis n'importe quel navigateur. [36]



**Figure 6** : Dorpbox

- **Google Drive**

Google Drive est un service de synchronisation de fichiers; il permet aux utilisateurs de synchroniser leurs documents à partir d'ordinateurs, de tablettes et de terminaux. [17]



**Figure 7** : Google drive

- **iCloud**

iCloud est un nouveau service qui stocke les contenus de façon automatique et sécurisée, afin qu'ils soient accessibles par divers appareils (ordinateurs, smartphones, tablettes). [41]



**Figure 8** : Icloud

- **SkyDrive**

SkyDrive est un disque dur virtuel sur Internet. Permet facilement le téléchargement des fichiers, des photos, des vidéos. Il permet d'accéder à tous les fichiers en toute sécurité avec n'importe quel appareil connecté à Internet.

Et donne la possibilité de partager des fichiers, des photos ou des vidéos avec les personnes. [42]



Figure 9 : Skydrive

#### 2.4. Tableau comparatif

Caractéristiques	Dropbox	skydrive	Google Drive	icloud
Espace de stockage gratuit	2 Go	7 Go	5 Go	5 Go
Accès mobile	Oui	Oui	oui	oui
Applications mobiles	iOS, Android, Blackberry	Windows Phone, iOS, Android (prochainement)	Android et iOS (prochainement)	iOS
contrôle de version	30 jours	25 jours	30 jours	/
Partage de fichiers / dossiers	oui	oui	oui	Oui
Synchronisation sélective	Possibilité de sélectionner les répertoires à synchroniser	non	Possibilité de sélectionner le répertoire à synchroniser	Non
Taille maximal pour l'upload d'un fichier	300 Mo/illimité	2 Go	10 GO	25 Mo/250 Mo pour version payante

Tableau 2 : tableau comparatif des services Cloud

**Conclusion**

Pour les organismes modernes d'aujourd'hui, le document numérique est devenue une réalité quotidienne. Pour rechercher ou réagir rapidement dans le temps avec tout document archivé et stocké, les organismes ont besoin d'outils fiables.

Ce premier chapitre qui a traité l'état de l'art de ce type de problématique. Nous à montré qu'après l'étape d'acquisition (dématérialisation) des documents, il convient de passer aux étapes d'indexation, de classement et de stockage. Ces étapes seront abordées dans le deuxième chapitre qui sera consacré aux systèmes de gestion électronique des documents(**GED**). On présentera en détail toutes les étapes d'archivage, de stockage et beaucoup plus de fonctionnalités qui apportent des gains incontournables aux entreprises.

# Chapitre 2

GED

## Etude comparative de quelques solutions Open Source

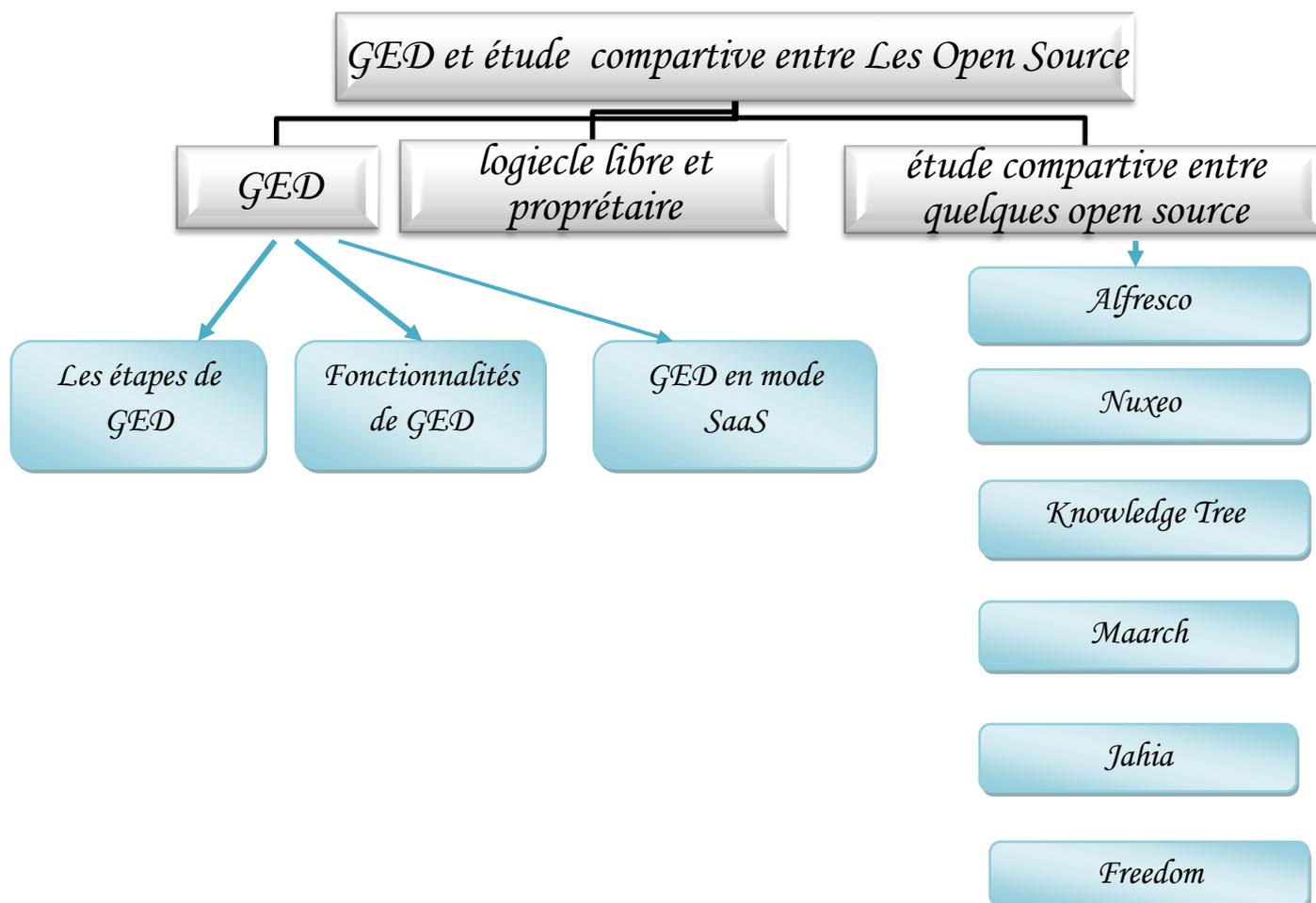


## Introduction

Au regard de l'évolution des nouvelles technologies, et avec l'accroissement des volumes de documents numériques, l'apport des solutions GED est devenu indéniable pour les entreprises dans lesquelles elles introduisent des fonctions de gestion, de partage, de circulation et d'archivage des documents. Historiquement, la GED a été utilisée pour l'archivage électronique, avec l'ajout des fonctionnalités diverses. Elle est devenue aujourd'hui un véritable outil de gestion de l'information.

Dans la première partie de ce chapitre, nous essayons de donner une explication de la GED, de ces principales fonctionnalités et de ces quatre étapes importantes à savoir : l'acquisition, le classement, le stockage et la diffusion. La deuxième partie du chapitre sera consacrée à l'étude et la comparaison de quelques solutions de gestion électronique des documents : **Alfresco**, **Nuxeo**, **Knowledge Tree**, **Maarch**, **Freedom**, et **Jahia**. Enfin, nous présentons une synthèse sous forme de tableaux comparatifs des différentes caractéristiques des solutions proposées.

## Plan du chapitre



Les organisations qui se lancent dans des projets de dématérialisation de documents, sont directement touchées par les enjeux d'une gestion électronique de documents.

Il existe deux époques marquantes de l'histoire de la GED :

- **Une première période**, allant de l'apparition de la GED au milieu des années 80 jusqu'aux environs de 1995 où les premières solutions développées correspondaient à des besoins d'archivage électronique. La justification d'une GED était alors d'ordre légal et administratif. Les solutions techniques de GED étaient alors très lourdes, très coûteuses, et proposaient une alternative à la micrographie.
- **Une deuxième période**, survenant vers 1995, au moment de l'expansion des réseaux et du multimédia. La justification d'une GED est alors devenue économique et stratégique, avec une recherche de gisements de productivité, de facteurs d'innovation et de qualité, à travers la mise en place de processus de travail collaboratif, de capitalisation et d'échanges informationnels. [30]

### 1. La Gestion Electronique de Documents

La Gestion Electronique des Documents abrégée en GED en anglais, Electronic Content Management (ECM), c'est l'ensemble des logiciels et /ou matériels pour la gestion des documents électroniques (textuel, graphique, sonore, etc.), généralement définie par « système informatisé d'acquisition, classement, stockage, archivage des documents » [18]

**D'autre part** : « La gestion électronique des documents (ou gestion électronique de l'information ou de documents existants – GEIDE) recouvre l'ensemble des techniques qui permettent de gérer les flux de documents qui pénètrent, sortent ou circulent à l'intérieur de l'entreprise. Ces techniques ont pour fonction de capturer ou dématérialiser des documents, afin de gérer, indexer, stocker, rechercher, consulter, traiter et transmettre des fichiers numériques de toutes origines ». [19]

La GED est donc une application qui vis a gérée et organiser les documents au cours de leur cycle de vie, allant de sa création, sa modification, sa publication, sa diffusion, etc., jusqu'à sa destruction ou archivage, et de les structurés pour faciliter leur circulations et leur utilisation.

## 2. Les étapes de GED

### 2.1. L'acquisition des documents

L'acquisition des documents se réalise avec différentes façons :

- La numérisation de document papier existant s'effectue par l'intermédiaire de scanners, afin de rendre l'information gérable par l'ordinateur. Les documents numérisés peuvent alors être triés via une technologie de RAD. Il est possible d'extraire des informations de l'image numérisée par l'utilisation de la technologie de LAD. Cette dernière peut utiliser des techniques d'OCR, de reconnaissance de codes à barres et d'ICR.
- L'intégration s'effectue au document qu'il existe déjà sous forme électronique.

### 2.2. Classement

Pour fournir aux utilisateurs une recherche instantanée des documents, les documents classés au sein de GED s'indexent de différentes manières :

- **L'indexation par type** : L'utilisation des métadonnées qui sont les informations liées aux documents. (Incluent par exemple le titre, l'auteur, la description, la langue, la date de publication, date de création, dates de modification successives, date de validation, format, type de documents etc.).
- **L'indexation par concepts ou mots-clés** : cette indexation est réalisée par la désignation des mots clés les plus expressifs au contenu (langages libres). Ou à travers des langages contrôlés par la sélection à partir de thésaurus.

### 2.3. Stockage

Le stockage des documents se réalise afin de « garantir l'intégrité, la sécurité et la sauvegarde des documents du système de GED. »[20]

### 2.4. Diffusion de documents

La diffusion des documents électronique s'effectue par le biais d'internet ou intranet.

### 2.5. Archivage

Certains documents doivent légalement être conservés mais n'ont plus de pertinence pour les opérations courantes. Ils deviennent alors des candidats à l'archivage, ce qui peut avoir des conséquences sur leur accès et leur indexation. [44]

## 3. les fonctionnalités de GED

### 3.1. L'indexation/Métadonnées

Est le fait de décrire ou de représenter un document pour permettre son exploitation. Elle doit être «la représentation la plus fidèle et la plus exhaustive du contenu, du sens du document afin de faciliter les recherches et d'en garantir la pertinence ». [22]

Cette indexation regroupe les métadonnées du document et le contenu du document, Les métadonnées sont les informations liées aux documents. « Informations structurées ou semi-structurées qui permettent la création, la gestion et l'utilisation des documents archivés au cours du temps, au sein du domaine d'activité qui les a créés. ». [23]

#### 3.1.1. Indexation manuelle

Cette méthode manuelle consiste à créer une notice descriptive ou une fiche descriptive, l'identification et la sélection des concepts se réalise au moyen d'un formulaire. L'opération de remplissage de ce formulaire se fait ou bien par :

- **les langages libres** (mots-clés et de descripteurs libres) choisie librement par la personne qui fait l'indexation ou l'auteur;
- **les langages contrôlés** (des listes d'autorité et de thésaurus) pour mieux utiliser les métadonnées;

L'indexation manuelle réalisée par les documentalistes conduit à la subjectivité d'une part, et d'autre part elle est couteuse puisqu'elle nécessite plus de documentalistes.

#### 3.1.2. Indexation automatique

Cette méthode d'indexation automatique est totalement prise en charge par l'ordinateur.

- **Extraction de métadonnées**

Il existe plusieurs modes d'extraction de métadonnées.

- **Récupération d'informations**

Chaque fichier informatique est automatiquement doté d'informations issues du système de gestion de fichiers, qui peuvent être pertinentes comme métadonnées : le nom du fichier, les dates de création et de modification, sa taille, son emplacement...etc.

- **Extraction de données structurées**

Certains formats de fichier disposent d'une structure lisible. C'est notamment le cas pour nombre de formats ouverts, ODF par exemple, mais également des formats MS Office, par le biais de leurs propriétés.

- **Induction de métadonnées**

Certaines solutions, les plus élaborées, permettent de déterminer automatiquement les métadonnées les plus pertinentes pour indexer un document. Ce type d'induction de métadonnées est souvent réalisé par :

- ✓ des outils statistiques, qui analysent les chaînes de caractères les plus présentes et donc potentiellement les plus représentatives ;
- ✓ des outils sémantiques, capables d'extraire automatiquement les mots et expressions les plus pertinents, voire de reconnaître s'il s'agit de mot-clé, de date, du titre. [45]

### 3.2. Moteur de recherche

Cette fonctionnalité est considérée parmi les plus importantes dans le système de gestion d'information, elle vient après le stockage des documents dans le système pour faire des recherche, donc le moteur de recherche est la principale étape qui donne la possibilité de l'exploitation d'outil GED.

#### Fonctions de base

- Indexation (des métadonnées).
- L'indexation de tout le contenu de document en texte intégrale ou full text en anglais.

La recherche avec l'utilisation : des opérateurs booléens (ET, OU, SAUF), désignation de format de document, l'utilisation des caractères génériques par exemple «\*» qui remplace une chaîne de caractère...etc.

« Le moteur de recherche doit filtrer ses résultats de recherche selon les habilitations en lecture de l'utilisateur, c'est-à-dire ne lui proposer que des documents qu'il a le droit de lire. C'est souvent ce croisement, entre résultats de recherche et habilitations, qui est le plus difficile à mettre en œuvre avec des moteurs autres que celui nativement intégré à l'application. » [31]

## Fonctions avancées

Parmi les fonctionnalités qui existent dans les moteurs de recherche avancés :

- ✓ **Gestion des paniers de recherche**: permet d'enregistrer des recherches, ou plutôt les critères de ces recherches, pour les lancer à la demande. Cela évite aux utilisateurs de perdre du temps à saisir plusieurs critères parfois de nombreuses fois dans la même journée.
- ✓ **Combinaison de plusieurs techniques de recherche** : à titre d'exemple, utilisation de l'ontologie, accompagnée de l'apprentissage réalisé sur les recherches antérieures, etc. [44]

### 3.3. La gestion des versions

La gestion des versions a pour but d'organiser le travail de plusieurs intervenants sur le même document, pour différencier la dernière version valide du document. Avec l'utilisation des techniques intéressantes qui empêchent des interventions simultanées sur le même document.

#### 3.3.1. Le système de réservation (check-in/check-out)

C'est un système qui permet à un utilisateur quand il réserve un document pour faire des modifications, de le bloquer afin d'interdire d'autre utilisateur de le modifier jusqu'au déblocage par le premier utilisateur. Donc il existe deux conditions. Le blocage du document c'est le check-out, et le déblocage du document c'est le check-in.

#### 3.3.2. Incrémentation des versions

GED permet de donner un numéro séquentiel aux différentes versions dans l'ordre des mises à jour effectuées au document. Il peut utiliser les notions des versions mineures/majeures « les versions (déclinées sous la forme x.y où x est la version majeure et y la version mineure. Par exemple si un document est en version 1.2 (c'est à dire version majeure 1 et version mineure 2), la version 1.3 sera donc une évolution mineure du document et la version 2.0 une évolution majeure du document. ». [24]

Pour l'incrémentation, et le passage d'une version à une autre peut être effectué selon des règles qui sont définies par l'utilisateur ou l'administrateur.

Cette fonctionnalité permet de faire la traçabilité des évolutions des documents.

### **3.4. Import/export**

#### **3.4.1. Import / export en masse**

Cette fonctionnalité pour but d'entrer/sortir d'une solution une grand quantité des documents rapidement pour gagner de temps. « Ces fonctions permettent de prendre en compte les métadonnées et des fichiers et si possible la gestion des autorisations. »[41]

#### **3.4.2. Export sur support physique**

L'export pour but d'archivage consiste à exporté des documents avec leurs métadonnées sur des supports physique qui garantissent leurs pérennité.

L'outil GED garde les métadonnées pour permettre de faire des recherches sur les documents déjà exportés, ou pour faire des consultations hors ligne.

### **3.5. Intégration de la GED**

#### **3.5.1. Comme un espace de stockage**

Dans ce cas la GED est utilisée comme un espace de stockage au même titre que l'utilisation d'un disque réseau, elle permet d'utiliser des fonctionnalités similaires aux fonctionnalités des outils bureautique par exemple « glisser-déposer, copier/coller... ». L'interface utilisée ressemblant à celle utilisée par les systèmes d'exploitation. Cette intégration peut être réalisée par l'utilisation d'une interface Web ou bureautique dédiée à la GED.

Ces interfaces sont *WebDav*, *CIFS* et *FTP*.

#### **3.5.2. Intégration avec les outils de bureautique**

Les utilisateurs travaillent directement sur leur outil bureautiques et de façon transparente sous la couche de GED, donc ils travaillent sur les deux environnements. Cette intégration fourni en plus de stockage le travail collaboratif. Cette approche n'est pas fournit directement par les solutions open sources donc elle nécessite une intégration (couteuse).

### **3.6. Gestion des droits**

C'est un point délicat pour les entreprises, elle influence le choix de la solution GED.

Un système de gestion des documents doit prendre en considération les droits d'accès des utilisateurs aux différents documents.

### 3.6.1. Niveaux de gestion

#### Gestion des profils

La notion de profil sert à donner aux utilisateurs des accès à des fonctions (traitement, manipulation, usage) bien définies sur les documents. [42]

#### Annuaire et SSO

Vu la haute confidentialité de certains documents et vu le risque d'erreur sur la gestion des profils des utilisateurs, il est fortement conseillé de la réaliser sur un annuaire d'entreprise LDAP (Lightweight Directory Access Protocol). Ce dernier permettra de décider et de paramétrer le ou les profils à donner à chaque utilisateur. Le système GED ne fera que prendre le profil total correspondant à chaque utilisateur depuis l'annuaire directement. [45]

Le SSO est le « *Single Sign On* ». pour permettre une authentification unique (un seul nom d'utilisateur et un seul mot de passe) des utilisateurs. [44]

### 3.7. Workflow (processus)

Le Workflow est considéré parmi les fonctionnalités les plus importante dans l'outil GED, il permet de circuler les documents entre plusieurs utilisateurs, avec la prise en compte des droits d'accès, donc c'est l'ensemble des opérations effectuées au document. Allant de ça création, rédaction, stockage, validation,...jusqu'à ça publication, archivage ou suppression.

### 3.8. La gestion des dossiers

La gestion des dossiers de GED est semblable à la gestion des dossiers papier et à la gestion de bureau de poste de travaille. Un dossier dans GED peut inclure plusieurs documents ou informations, ces documents ont des caractéristiques quelconques communes par exemple les métadonnées. Ces documents apparaissent dans plusieurs dossiers sans duplication ce qui apporte des gains d'espace de stockage.

La gestion de dossiers de GED permet:

- ✓ le partage des informations en définition pour chaque dossier les métadonnées associé.
- ✓ sa structure de groupement de document.
- ✓ les liens avec les autres dossiers et les droits des utilisateurs.
- ✓ la gestion des dossiers se fait par des déclencheurs automatiques pour l'archivage

- ✓ Assurer la traçabilité de toute action sur chaque dossier pour permettre des audits sur le traitement des dossiers. [25]

### 3.9. Numérisation

En général la numérisation ne se fait pas directement par les solutions GED, elle se fait par des solutions dédiées, depuis la conversion du support, jusqu'à l'injection du document dans la GED. Les outils (logiciels et matériels) de scannage sont installés au poste de travail dans lequel est réalisé la saisie des métadonnées initiaux (c'est la préindexation). Il existe des solutions de GED qui proposent des moteurs dédiés et commerciaux pour la réalisation de ces fonctionnalités.

### 3.10. Gestion de cycle de vie

#### Politique de conservation

C'est l'archivistique, discipline recouvrant toutes les fonctions liées à l'archivage, qui permet de répondre aux questions liées à la politique de conservation.

On trouve également le vocable de « records management » dont l'optique est plus adaptée au domaine numérique dans la mesure où il prend en compte une sous-sélection de l'ensemble des documents à archiver, sous sélection qui correspond aux « documents vitaux », c'est-à-dire nécessaires à l'activité de l'organisation.

#### Archivage

L'archivage peut être analysé de différentes façons en fonction du contexte de l'application. Du simple tag « archive » à la mise hors ligne des sélections du contenu sur des supports optiques, par exemple.

Les applications intégreront donc des fonctions qui permettent de:

- ✓ gérer les informations de conservation dans les métadonnées des documents.
- ✓ déclencher automatiquement un traitement d'archivage.
- ✓ choisir de rechercher ou non dans les éléments archivés.

En fonction des besoins réels, on envisagera différents types de traitements pour répondre à des problématiques de pérennité, de fiabilité et de traçabilité des archives.

Parmi les fonctions envisageables, on citera :

- ✓ figer les éléments archivés, en calculant des empreintes cryptographiques, éventuellement signées, de façon à garantir l'inaltérabilité des archives.
- ✓ faire valider, via un workflow, l'archivage «le versement en archive » d'un ensemble de documents par leur auteur ou le responsable du service concerné.

- ✓ extraire les documents archivés et leurs métadonnées pour les stocker sur des supports à longue durée de vie. [31]

#### 4. Les avantages de la GED

La GED, rassemble en un point central l'ensemble de l'information produite ou reçue par une entreprise

- Recherche plus simple et plus rapide des documents
- Offre plus de liberté au niveau du classement
- Permet à l'ensemble de collaborateurs, sur place ou sur un site distant, de retrouver instantanément un document papier ou électronique
- Supprime les tâches de photocopies et de reclassement
- Permet d'indexer, classer et diffuser automatiquement toutes les pièces d'un dossier client (courrier, dossiers lients,...)
- Permet de gérer de façon électronique les circuits de validation (bon de commande, factures, ...). [26]

#### 5. La GED en mode SaaS

**GED peut être en mode SaaS** (Software as a service). Cette technologie permet à une entreprise qui ne peut pas se permettre de lourds investissements matériels de profiter des technologies offertes par la GED sans en subir les coûts. Les documents sont consultables via internet et bénéficient d'un accès sécurisé. [45]

**Il est à noter que les solutions de GED en mode SaaS sont proposées avec un abonnement elles ne sont pas gratuites.**

#### 6. Les solutions de GED

##### 6.1. Logiciel libre ou logiciel propriétaire

Pour pouvoir faire le choix entre un logiciel libre et un logiciel propriétaire, il convient de faire un rappel sur la philosophie du logiciel libre.

Le logiciel libre ou « open source », sont l'ensemble des logiciels qui respectent certains critères à savoir:

- la libre redistribution.
- la possibilité d'avoir accès au code source, d'ajouter ou de modifier des fonctionnalités.

Le logiciel propriétaire «logiciel non libre » est généralement distribué au public sous sa forme binaire uniquement. Cette forme n'étant pas compréhensible par l'humain, le détail

de son fonctionnement est ainsi conservé secret. la duplication, la modification et l'usage reste limité, Ceci se matérialise par des limitations légales matérielles ou logicielles.

Les différentes open sources de GED qui sont offertes pour la communauté adoptent des licences GNU (GNU Not UNIX) :

- La GNU General Public License (GNU GPL) : son principe essentiel est que si un logiciel qui a été modifié est distribué à un tiers. Le code source doit être livré avec. sous la même License GNU GPL.
- La GNU Lesser Public License (GNU LGPL) : est moins contraignante en tolérant l'emploi d'un logiciel libre comme module d'un logiciel propriétaire. dans ce cas. le module reste libre. Mais pas le logiciel qui en résulte.
- La GNU Affero General Public License (GNU AGPL) : est la licence GNU GPL à laquelle est ajoutée une clause permettant d'étendre l'obligation de délivrance des sources lors de l'usage du logiciel sur internet ou par réseau. [25]

## 6.2. Les open sources étudiées

Six solutions open sources ont été étudiées : Alfresco, Nuxeo, Jahia, Knowledge Tree, Maarch et Freedom. L'objectif étant la sélection des plus performantes, en matière du nombre de fonctionnalités.

### 6.2.1. Alfresco

Alfresco est solution d'ECM (Enterprise Content Management), La société Alfresco a été fondée en juin 2005 en Grande-Bretagne. Il est le produit d'une équipe composée du co-fondateur de Documentum et un ancien cadre de Business Objects qui se sont appuyés sur 15 ans d'expérience respective afin de concevoir une solution innovante en matière de fonctionnalités et de bénéfiques aux utilisateurs. Basé sur la technologie J2EE (Java Enterprise Edition).

Alfresco est un outil de gestion des documents électronique il offre beaucoup de fonctionnalité.

Alfresco est distribué en deux versions :

- ✓ Une version « Community », gratuite sous licence GPL.
- ✓ Une version « Entreprise », offre le support, les garanties, la fiabilité, ainsi que l'accès au support d'Alfresco et de ses partenaires. [46]



Figure 10 : Alfresco

### 6.2.2. Nuxeo

Nuxeo est une société française qui édite depuis 2001 une solution de GED open source. en environnement Java J2EE la première solution était développée en Zope Python, ( Nuxeo a migré en java sa solution de GED au début de 2007), et elle est aujourd'hui présente en France et aux Etats Unis. Nuxeo est une **solution complète de gestion de contenu d'entreprise**, Nuxeo est distribué sous Licence LGPL.

La solution de Nuxeo est très modulaire :

- Nuxeo EP (Enterprise Platform) est le socle commun à toute l'offre,
- Nuxeo DM (Document Management) est le module dédié à la GED collaborative,
- Nuxeo DAM (Digital Asset Management) permet de gérer les contenus multimédia,



Figure 11 : Nuxeo

### 6.2.3. Knowledge Tree

Knowledge Tree est une solution de GED développée par la société Sud-africaine JamWarehouse, Knowledge Tree est développé en PHP. Il est plus léger que les solutions précédentes et plus simple à installer.

Knowledge Tree présente un ensemble de fonctionnalités et modules qui permettent une bonne intégration dans l'environnement bureautique.

La version Community Edition, gratuite et libre, permet d'avoir une solution GED simple qui ne dispose pas de fonctionnalités telles que l'intégration à Microsoft Office ou Navigation sont des modules commerciale [47]

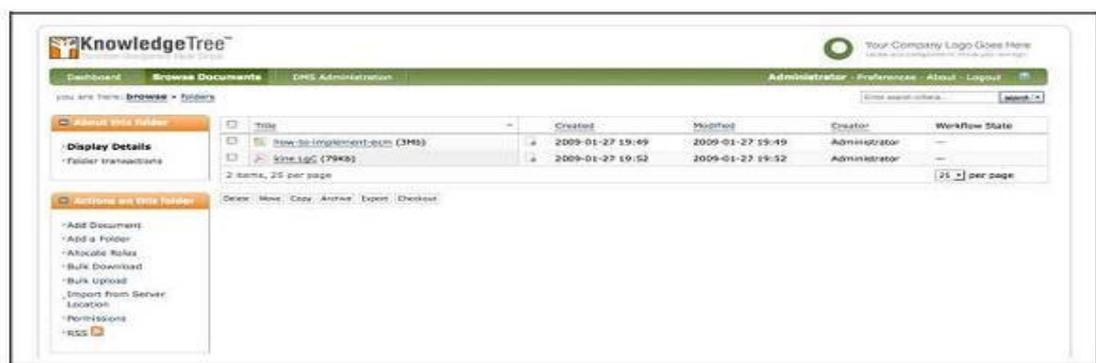


Figure 12 : Knowledge Tree

### 6.2.4. Jahia

Jahia est un serveur applicatif remplissant les fonctions de portail d'entreprise, de serveur de gestion de documents, de serveur de gestion d'activité, de moteur de recherche et de suite collaborative. Il est fondé sur le langage Java et distribué sous licence collaborative (concept de « contribuer ou payer »). Jahia porte l'accent sur la création et la gestion de contenu Web. Plus qu'une GED au sens strict du terme, il s'agit d'un module très complet de diffusion de contenu sur le Web. Jahia est disponible en une version Community (communautaire), libre et gratuite (distribuée sous une licence GPL), et une version Enterprise. La version Enterprise fournit une assistance professionnelle et des modules supplémentaires tels que la prise en charge de l'annuaire LDAP (Lightweight Directory Access Protocol). [48]

### 6.2.5. Maarch

Maarch système d'archivage électronique et gestion documentaire open source, c'est une solution développée par la société Maerlys. L'offre comporte deux applications : Maarch LetterBox, une application dédiée à la gestion de courrier, Et Maarch

Enterprise est une solution de GED complète et modulaire dédiée à la conception d'applications ; elle inclut les fonctions nécessaires à une GED d'entreprise. Maarch permet de stocker différentes catégories de documents et inclut des fonctions importantes. Les applications Maarch sont écrites en PHP, et La solution est distribuée sous licence GPL.

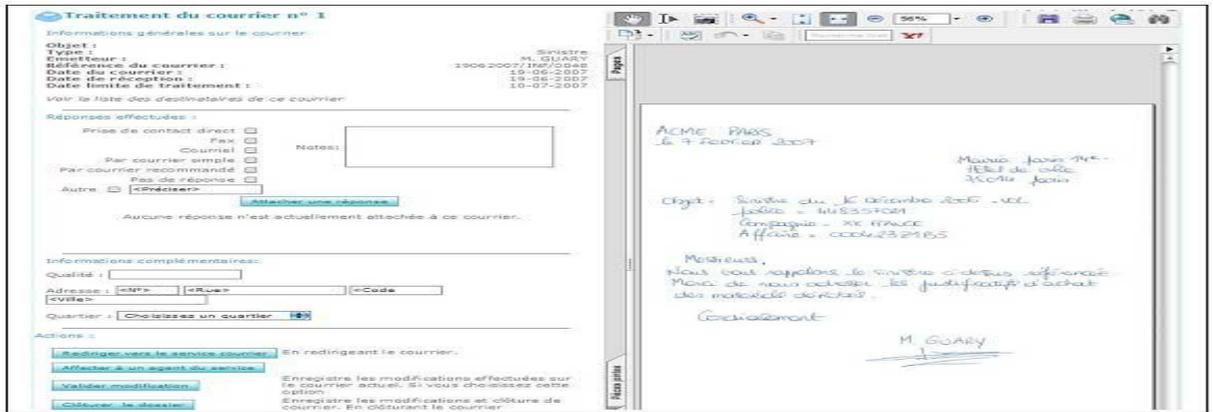


Figure 13 : Maarch

### 6.2.6. Freedom

Freedom est une solution **open source** de la gestion électronique des documents sous la licence GPL, développé en PHP par la société française Anakeen en 1998.

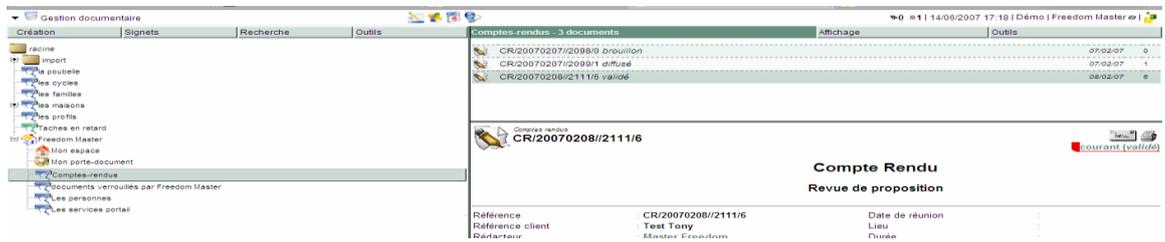


Figure 14 : Freedom

6.3. Tableau comparatif

Caractéristiques Système	Alfresco	Nuxeo	Knowledge tree	Maarch	Jahia	Freedom
Année de création	2005	2001	2006	2008	2002	1998
Langage	J2EE	J2EE	PHP, MySQL	PHP	J2EE	PHP, SQL
Dernière version	4.2	5.6	3.7	1	6 (6.5 beta)	3
Licence	GNU GPL	GNU LGPL	GNU LGPL	GNU LGPL	GNU LGPL	GNU AGPL
Architecture	<b>Plate-forme</b> Linux, Windows, Mac <b>Serveur d'application</b> JBoss Apache Tomcat IIS	<b>Plate-forme</b> Linux, Windows, Mac <b>Serveur d'application</b> JBoss Apache Tomcat, IIS	<b>Plate-forme</b> Linux Windows, Mac <b>Serveur d'application</b> Apache MySQL.	<b>Plate-forme</b> Linux, Windows Mac. <b>Serveur d'application</b> Apache MySQL	<b>Plate-forme :</b> Linux, Windows, Mac. <b>Server d'application:</b> Tomcat, Apache, IIS	<b>Plateforme</b> Linux Windows, Mac <b>Serveur de base de données :</b> PostgreSQL
Fonctionnalités						
Tableau de bord	✓	✓		X	✓	X
Intégration Ms Office	✓	✓	X	X	X	X
Intégration application web service	✓	✓	✓	X	✓	X
Import/export	✓	✓	✓	Export commercial	✓	✓
Versionning	✓	✓	✓	X	X pas de versions mineurs	✓ sauf les page web
Traçabilité	✓	✓	X	✓	✓	X
Historique de connexion	✓	✓	X	-	✓	X
Archivage Email	✓	✓	X	X	X	X

Métadonnée	<b>(indexation): Automatique</b>	✓	✓	–	–	Automatique Sauf (PDF,MP3)	–
	<b>(indexation Manuel)</b>	✓	✓	✓	✓	✓	✓
Referential de classement	Plan de classement	Plan classement	Lexique	Type de document	Plan de classement	Plan de classement	
Texte integral	Majorité des documents bureautique	Majorité des documents bureautique	Majorité des documents bureautique	X	Les fichiers plus simples	Tous les documents bureautiques	
Integration de la GED	Webdav FTP CIFS	webdav	webdav(version commercial)	–	webdav	Seulement l'ajout et la suppression	
Gestion de droit	LDAP, SSO	LDAP, SSO	LDAP	Peu de choix dans les autorisations LDAP	LDAP	LDAP, SSO	
Workflow	✓	✓	✓	X	✓	✓	
Gestion des dossiers	✓	✓	X	X	X	✓	
Interface avec scanner	Commercial	Commercial	Commercial	✓	X	X	
Replication de la base de données	✓	✓	X	X	✓	X	
Integration d'un moteur de recherché	✓	✓	✓	✓	✓	X	

**Tableau 3:** tableau comparatif entre les solutions open source de GED ;

**Conclusion**

Les solutions GED proposent les outils adéquats pour la gestion, la sécurité, le stockage, l'archivage et beaucoup plus des fonctionnalités nécessaires au traitement et à la gestion des nombreux documents au sein des entreprises. C'est souvent un autre moyen pour augmenter le niveau ou la valeur de ces entreprises. Alors qu'à ses débuts, ses principes se résumaient à numériser un document et à le caractériser par des champs descriptifs (métadonnées) afin de faciliter son accès et sa recherche ultérieurement. Aujourd'hui les solutions GED sont de plus en plus riches en nouvelles fonctionnalités importantes en intégrant par exemple la gestion de processus « Workflow », check-in /check-out, l'import et l'export, la gestion des droits...etc.

Dans ce chapitre, l'étude comparative que nous avons réalisée sur les six solutions open sources nous a permis de retenir les deux solutions intéressantes : il s'agit des plateformes « Alfresco » et « Nuxeo » qui incluent la majorité des fonctionnalités, elles sont plus riches par rapport aux autres open sources. En effet, elles fournissent un environnement collaboratif complet pour les différents utilisateurs.

Dans le chapitre suivant, nous allons présentés une étude détailler de ces deux plateformes retenus.

# Chapitre 3

## Alfresco et Nuxeo

### Etude et comparaison des plateformes



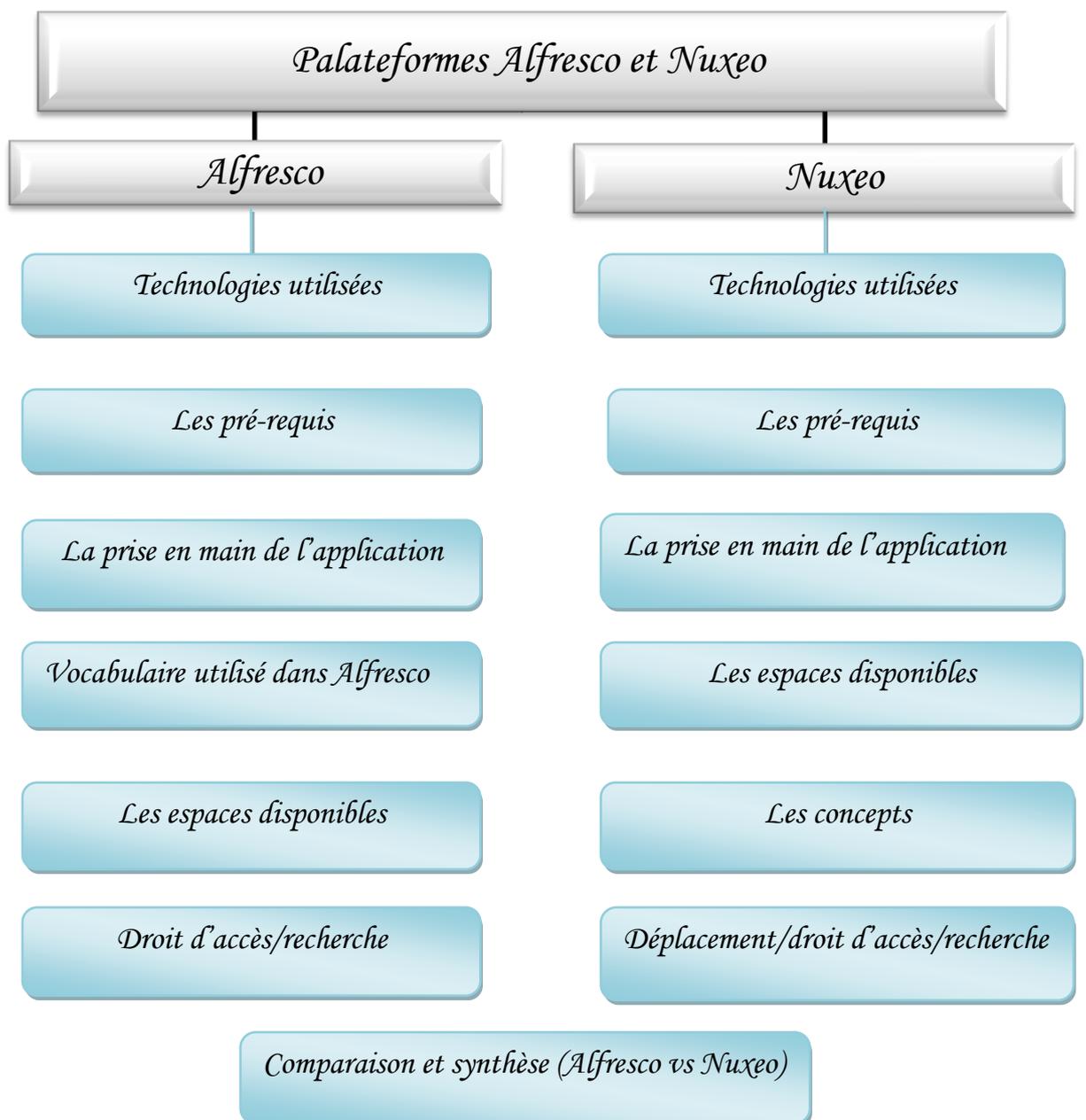
**Introduction**

Comme nous l’avons montré au précédent chapitre, les plateformes GED « Alfresco » et « Nuxeo » incluent la majorité des fonctionnalités nécessaires.

C’est pour cette raison que nous avons jugé intéressant de consacrer ce chapitre complètement à l’étude détaillée de ces deux solutions.

Dans ce chapitre, nous présentons les solutions « Alfresco community » et « Nuxeo Entreprise Edition (EP) », leur installation ainsi la réalisation d’une étude détaillée du fonctionnement des ces deux plateformes.

***Plan du chapitre***



## 1. La solution Nuxeo

### 1.1. Technologies utilisées

Nuxeo est basé sur la technologie Java avec la spécification Java EE (Java Enterprise Edition). Il n'est lié à aucun serveur particulier ni aucune base de données.

Le schéma ci-dessous présente l'architecture de Nuxeo:

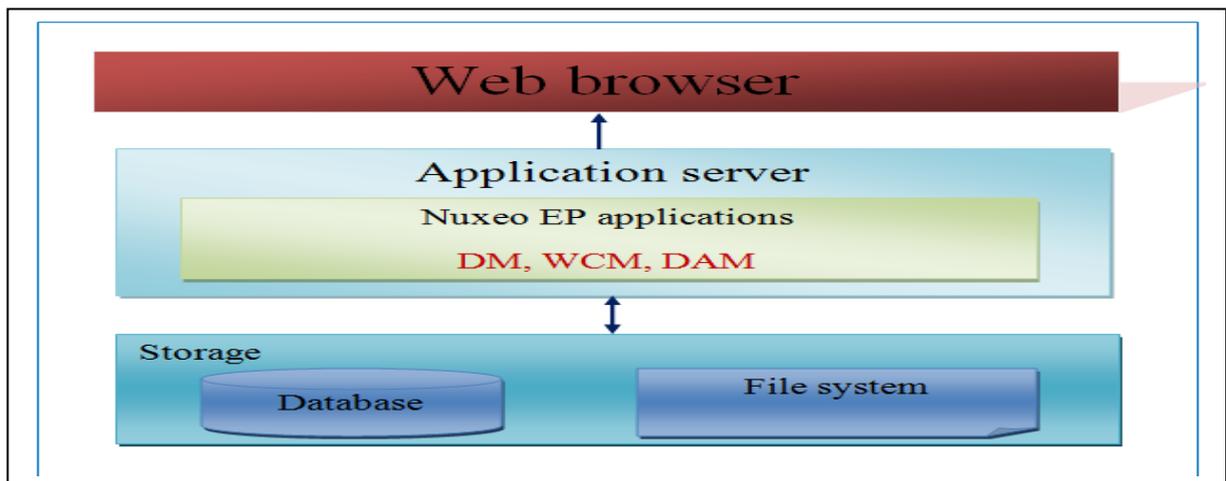


Figure 15 : l'architecture de Nuxeo

### 1.2. Des pré-requis

#### Serveur:

**OS:** Windows 2003 +, Mac OS X, Unix, Linux.... tout système sur lequel on peut installer une JVM Sun 1.5 ou 1.6.

**Base de données :** Oracle, SQL Server, Postgres, MySQL Server.

**Serveur d'application :** Tomcat 6.0, JBoss 4.2.

**Logiciels optionnels annexes à installer :** Open Office, ImageMagick (permet la manipulation des images).

#### Client:

**OS :** pas de contraintes.

**Navigateur web :** Internet Explorer 7+, Firefox 3+, Chrome 1+, Safari 3+...

**Éléments annexes à installer sur le poste client:** Pour bénéficier du drag and drop depuis le poste, un plugin disponible pour Firefox et Internet Explorer.

### 1.3.La prise en main de l'application

#### Processus d'installation

Téléchargement du package Nuxeo pour windows depuis le lien

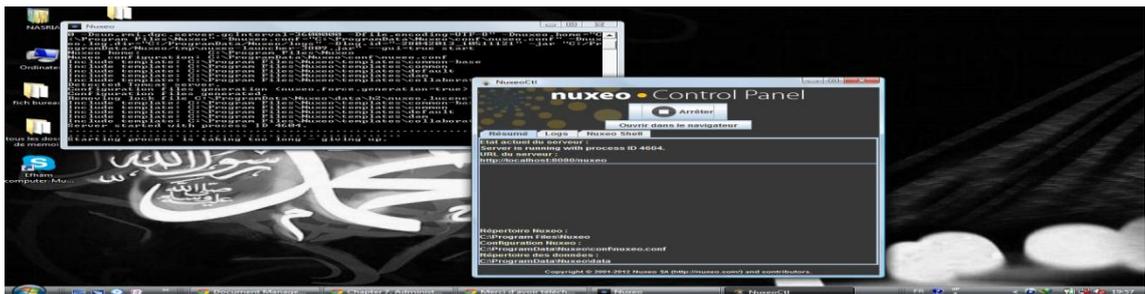
<http://www.nuxeo.com/fr/telechargez-nuxeo>

L'exécution de fichier .exe et le suivie des étapes de l'installateur. L'installateur propose de télécharger et d'installer des programmes supplémentaires qui peuvent être recommandés pour une utilisation optimale de la plateforme.



**Figure 16 :** installer des programmes supplémentaires

- Cliquer sur le raccourci "Nuxeo" créé lors de l'installation ou double-cliquer sur le fichier "Start Nuxeo.bat" dans le dossier de Nuxeo. Ceci lance une fenêtre de terminal et le Control Panel de Nuxeo, qui permettent le contrôle de serveur Nuxeo.
- Une fois Nuxeo démarré, il est possible d'ouvrir le navigateur et aller à l'adresse suivante : <http://localhost:8080/nuxeo> Ou de choisir à partir du control panel «Ouvrir dans le navigateur».



**Figure 17 :** Nuxeo démarré

- Un assistant de configuration s'affiche pour aider à la configuration de l'application.



Figure 18 : configuration de l’application

- Cette étape de l'assistant permet d’inscrire à un essai gratuit de 30 jours de Nuxeo Connect, ce qui donne accès à Nuxeo Studio et à la Marketplace Nuxeo pour adapter facilement la plateforme Nuxeo.
- il est possible de passer l’étape Nuxeo Connect pour l’instant, et inscrire à cette offre à tout moment depuis [nuxeo.com](http://nuxeo.com) ou directement depuis l’application Nuxeo.



Figure 19: Etape la configuration de l’application

- Cette étape permet de sélectionner les modules fonctionnels à installer sur la plateforme Nuxeo. Ces modules apportent des fonctionnalités de gestion de documents, de gestion des actifs numériques, de case management et de collaboration Sociale prêtes-à-l'emploi.



Figure 20 : sélectionner les modules fonctionnels

- Cette étape permet de télécharger les modules sélectionnés.

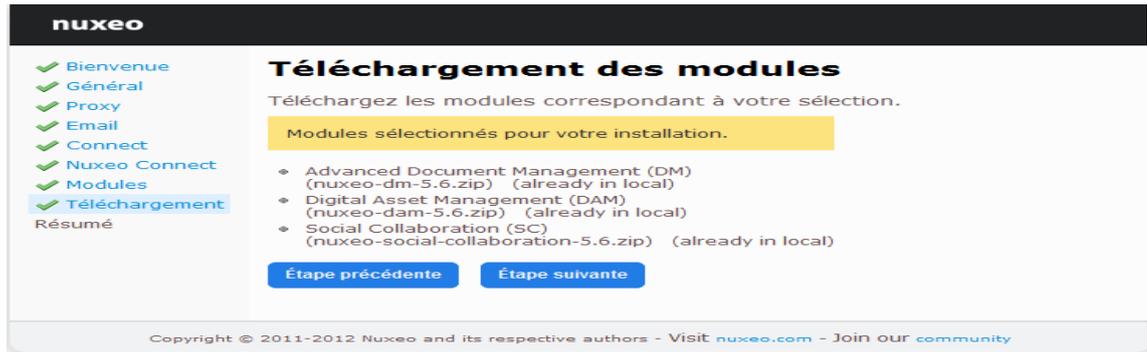


Figure 21 : télécharger les modules sélectionnés

- Cette étape affiche les paramètres de configuration définie.

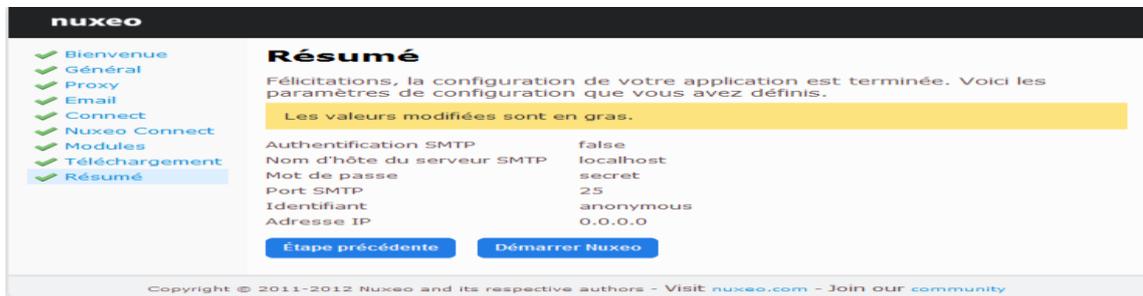


Figure 22 : configuration définie

- L'application Nuxeo redémarre.

### Connexion à Nuxeo EP

Il faut s'authentifier pour accéder à l'application (un nom d'utilisateur et un mot de passe)

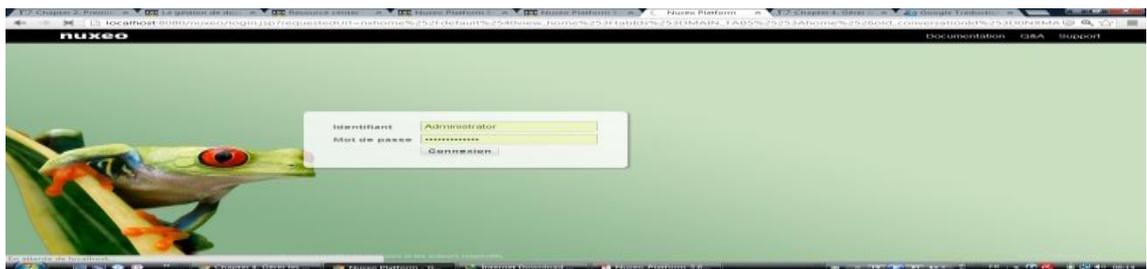


Figure 23 : Connexion à Nuxeo EP

### 1.4. Les espaces disponibles

- **L'onglet Home:** est l'endroit où l'utilisateur peut trouver toutes les informations le concernant (son tableau de bord)

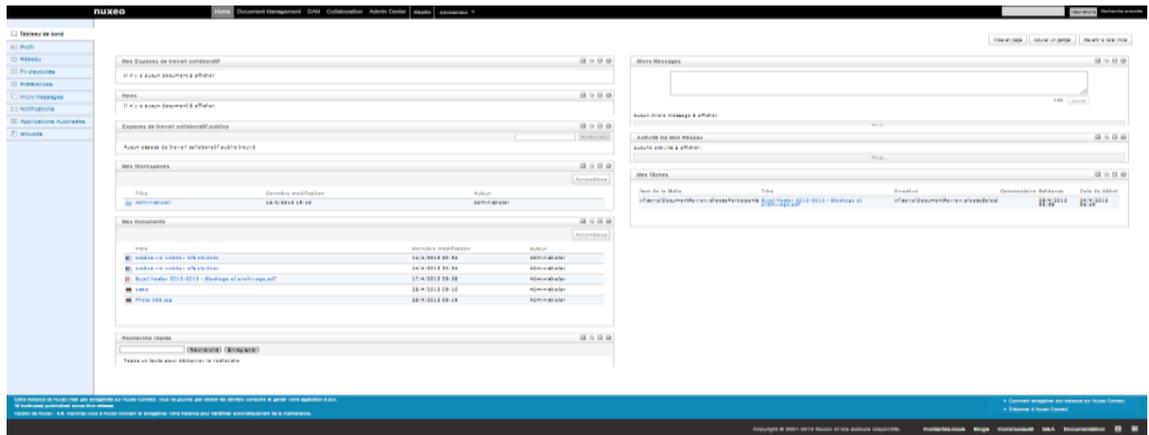


Figure 24 : tableau de bord

- **L'onglet Document Management:** est l'onglet principal. C'est la partie où l'utilisateur active au sein de Nuxeo.
- **L'onglet Digital Asset Management (DAM):** gérer Les contenus numériques et multimédias, comme les images, les fichiers audio et les vidéos, demandent une gestion particulière. [49]
- **L'onglet Collaboration:** l'espace de travail collaboratif permet d'avoir une vue orientée sur le projet des contenus tout en y intégrant les personnes, communications et interactions afin d'obtenir un espace de travail projet facile à utiliser. Grâce à une vue sous forme de tableau de bord. [50]

Les administrateurs disposent de deux onglets supplémentaires:

- **l'onglet Admin Center:** depuis lequel ils peuvent gérer l'application,
- **l'onglet Studio :** offrant un accès à **Nuxeo Studio**, l'environnement de personnalisation et de configuration en ligne.

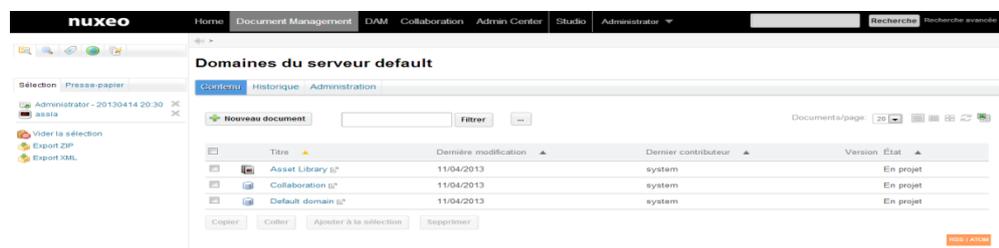


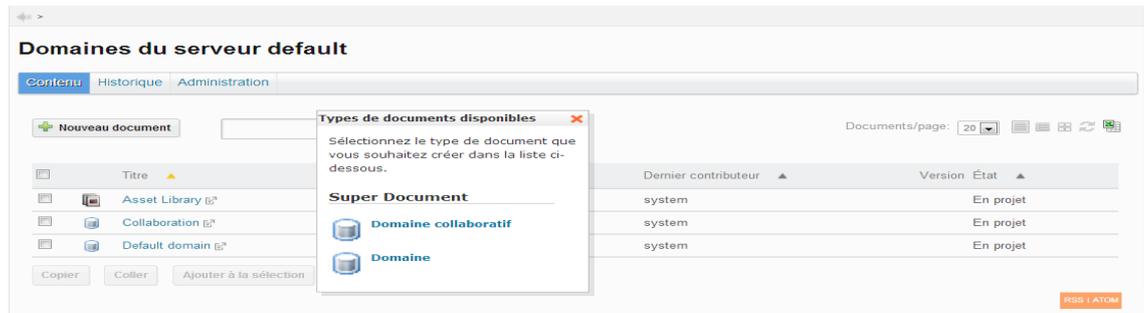
Figure 25 : domaine du serveur default

### 1.5. Les concepts

Nuxeo EP est organisé autour de domaines, qui contiennent des espaces de travail, des modèles et des espaces de publication. Les documents sont créés et diffusés dans ces

espaces de travail et de publication. Les actions disponibles pour les utilisateurs dans les espaces sont déterminées par les droits d'accès.

**Domaines:** est un espace créé seulement par l'administrateur à la racine de l'application, l'utilisation des domaines pour organiser la structure de l'application de façon plus précise et adaptée.

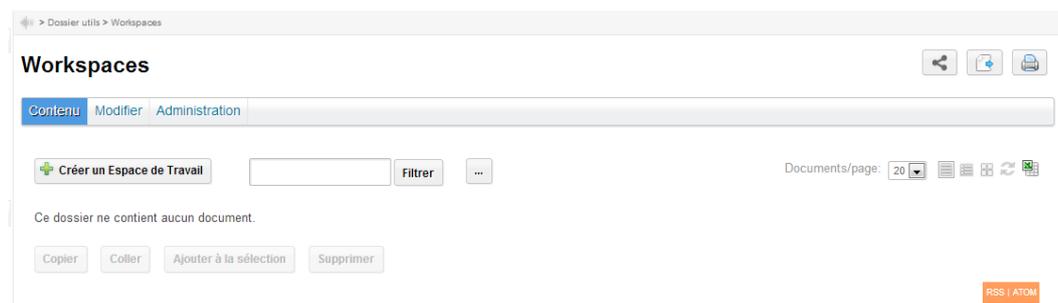


**Figure 26:** les espaces dans Nuxeo

Trois espaces sont automatiquement créés dans le domaine:

- Workspaces, racine des espaces de travail,
- Sections, racine des espaces de publication,
- Templates, racine des modèles.

**Espaces de travail:** Les espaces de travail pour but de la création et de la modification des documents, pour être partagé et modifié par plusieurs utilisateurs.



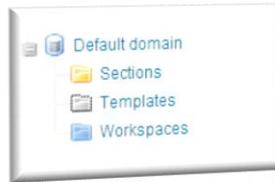
**Figure 27 :** Les espaces de travail

**Espaces de publication:** sont des espaces de publication des documents. Destiné à être uniquement consultés.

**Modèles:** il est possible de créer des documents et gérer les droits d'accès des espaces de travail modèles. Dans le modèle le nouvel espace de travail est créé avec la structure, les droits d'accès et le contenu du modèle utilisé.

## 1.6.Déplacement

Pour passer d'un espace de travail à l'autre:



L'arbre de navigation : Il affiche les espaces de travail et de publication d'un domaine, ainsi que les dossiers et les forums. Les documents n'y sont pas affichés.



Le chemin de navigation : Le chemin de navigation est la liste de liens située au-

Et aussi le tableau de bord: il est possible d'utiliser le tableau de bord pour accéder aux documents. Il s'affiche par défaut :

- les derniers documents modifiés,
  - les espaces de travail auxquels il est possible accéder,
  - les documents créés par l'utilisateur,
  - les derniers documents publiés,
  - les workflows sur lesquels l'utilisateur doit intervenir,
  - les workflows dont l'utilisateur est l'initiateur.

Le tableau de bord est accessible depuis toutes les pages de l'application.

- ✓ Sur un document, accéder au moyen des onglets aux différentes actions disponibles.

## 1.7.Droits d'accès

L'accès aux documents est défini par les droits d'accès. Nuxeo permet de constituer des groupes d'utilisateurs pour la gestion des droits d'accès. Les droits d'accès sont accordés aux: les espaces de travail, les sections, les modèles, les domaines, mais aussi les dossiers, les blogs, les sites web, etc. et les sous-espaces héritent ces droits.

Il existe quatre droits d'accès :

**Lecture:** l'utilisateur peut lire, consulter les méta-données et l'historique, et ajouter des commentaires sur ces documents.

**Ecriture:** l'utilisateur peut consulter, créer et modifier du contenu dans un espace de travail.

**Suppression:** l'utilisateur peut supprimer du contenu d'un espace de travail ou de publication

**Gestion:** l'utilisateur peut consulter, créer et modifier des documents, et il peut modifier l'espace de travail ou de publication et gérer les droits d'accès.

### 1.8.La Recherche

Nuxeo EP intègre son propre moteur de recherche, conçu pour prendre en compte tous les paramètres particuliers de l'application.

**Recherche rapide:** La recherche rapide est accessible depuis toutes les pages de l'application.il est possible de rechercher un document ou un espace.

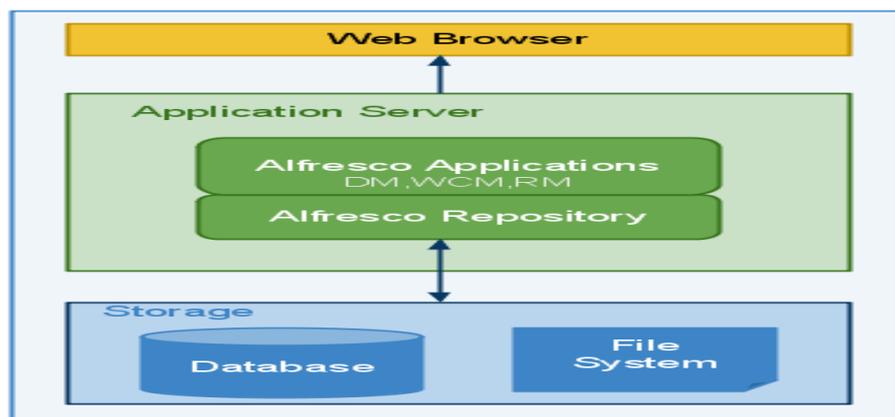
**Recherche avancée :** La recherche avancée permet de rechercher un document avec l'utilisation de plusieurs critères.

## 2. La solution Alfresco

### 2.1.Technologies utilisées

Alfresco est basé sur la technologie Java avec la spécification Java EE (Java Enterprise Edition). Il n'est lié à aucun serveur particulier ni aucune base de données, bien qu'il soit livré par défaut avec le serveur Apache Tomcat et la base de données MySQL.

Le schéma ci-dessous présente l'architecture d'Alfresco :



**Figure 28:** l'architecture d'Alfresco

## 2.2. Des pré-requis d'Alfresco

### Serveur

**OS:** Windows 2003 +, Mac OS X, Unix, Linux.... tout système sur lequel on peut installer une JVM Sun 1.5 ou 1.6.

**Base de données :** Oracle, SQL Server, Postgres, MySQL Server.

**Serveur d'application :** Tomcat, JBoss.

**Logiciels optionnels annexes à installer :** Open Office, ImageMagick, Flash Player (pour upload multiples), Share Point (permet d'exploiter l'environnement share point de Microsoft avec Alfresco), Web Quick Start (permet la création facile des modules).

### Client

**OS :** pas de contraintes.

**Navigateur web :** Internet Explorer 7+, Firefox 3+, Chrome 1+, Safari 3+,...

## 2.3. Prise en main de l'application

Téléchargement du package **Alfresco** pour windows depuis le lien :

[http://wiki.alfresco.com/wiki/Download\\_Community\\_Edition](http://wiki.alfresco.com/wiki/Download_Community_Edition)

### Processus d'installation

- L'exécution de fichier .exe et le suivi des étapes de l'installateur.
- Il existe deux types d'installation dans l'assistant d'installation:
  - Facile:** Type standard installe Alfresco en utilisant les options par défaut et la configuration.
  - Avancé:** Type avancée installe Alfresco, permet de configurer les ports du serveur et les propriétés de service. Et le choix des composants à installer.
- Une fois l'installation terminée, Le serveur Alfresco démarre et Alfresco Share lance dans le navigateur par défaut. A l'adresse <http://127.0.0.1:8080/share>.

### Connexion à Alfresco

Les deux clients d'Alfresco ( Share et Explorer) :

Alfresco Share est la solution Open Source de gestion collaborative d'Alfresco qui étend le produit de GED. Alfresco Document Management avec des fonctionnalités collaboratives avancées autour des documents avec une interface Web 2.0 plus intuitive et pensée pour les utilisateurs. Maintenant quasiment il est possible de tout administrer depuis l'interface

Share. L'ancienne interface web Alfresco Explorer n'est que pour de très rares tâches d'administration comme l'export d'espace, la récupération de données effacées, la gestion des catégories que l'interface Alfresco Explorer est nécessaire. Sachez aussi que tous les contenus créés et gérés par Share sont en fait tout simplement stockés dans des dossiers Alfresco. [27]

La première page envoyée lors de la connexion au site d'Alfresco Share est une page de connexion. (Nom d'utilisateur et mot de passe).

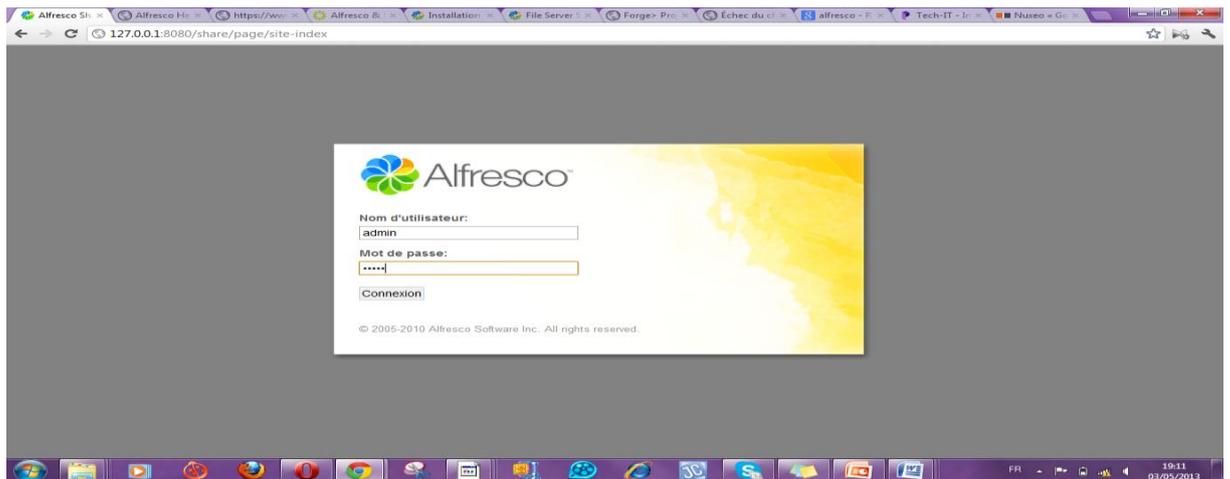


Figure 29 : Page d'identification

#### 2.4. Vocabulaire contenu dans Alfresco

- **Espace** : Un espace correspond à un dossier ou un répertoire. Il peut contenir des contenus (fichiers) ou d'autres sous-espaces.
- **Contenu** : Un contenu correspond à l'information finale désirée. C'est en général un fichier à télécharger (PDF, Word, Excel, etc.), avec en plus des informations spécifiques, des métadonnées.
- **Aspect** : un aspect est un ensemble de métadonnées spécifiques à Alfresco, associées à un comportement particulier. Par exemple, il existe dans Alfresco un aspect s'appelant «Validité ». Il permet d'ajouter une date de début et de fin de validité à un contenu.
- **Action** : Les actions dans Alfresco sont des actions basiques réalisables par l'utilisateur, telle que copier, coller un document, envoyer un courriel, gérer les droits sur un espace, ajouter un aspect à un document, etc.

## 2.5. Les espaces disponibles

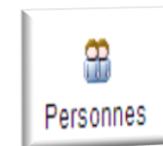


**L'onglet Mon tableau de bord:** est l'endroit où l'utilisateur peut trouver toutes les informations le concernant



**L'onglet Sites :** permet l'accès à

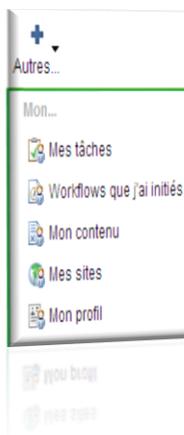
- Recherche des sites ;
- Créer un site.



**L'onglet Personne :** permet de trouver des personnes.



**L'onglet Entrepôt :** permet l'accès au plan de classement, c'est l'espace où sont stockés tout les documents. L'accès est défini par les droits d'accès.



**L'onglet Autres :** permet l'accès à :

**Mes tâches :** les taches de l'utilisateur.

**Workflows que j'ai initiés :** les taches et les documents qui l'utilisateur a demandé de les réviser.

**Mon contenu :** permet d'afficher les documents récemment ajouté et les documents récemment modifié.

**Mes sites :** permet d'afficher les sites où l'utilisateur est membre.

**Mon profil :** permet l'accès au profil de l'utilisateur.

**Tableau de bord:** Le tableau de bord de l'utilisateur se compose de modules qui lui permettent de suivre toutes les informations qui lui sont pertinentes. Chaque module est un élément de synthèse relatif à une fonctionnalité du site.

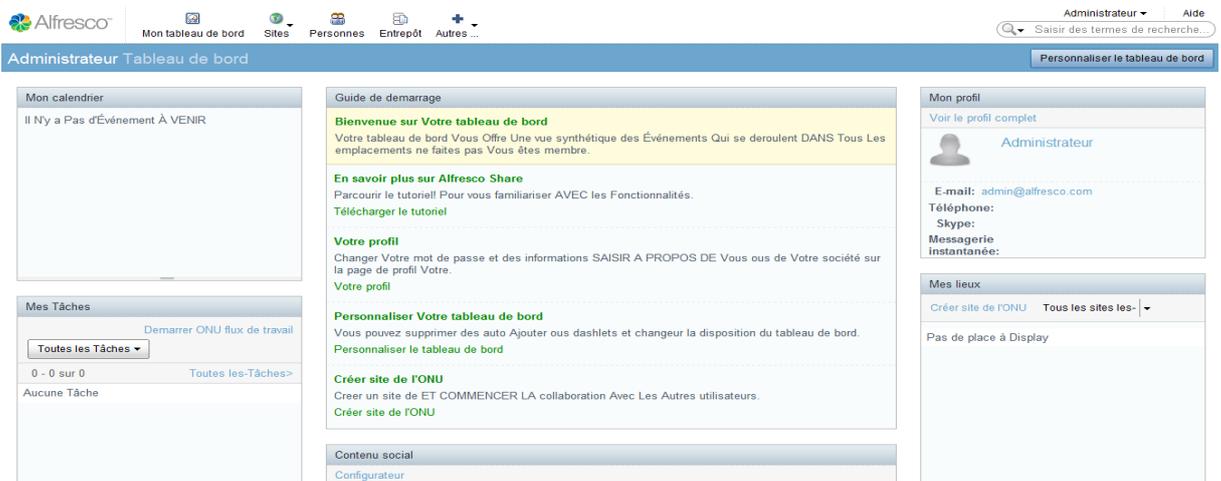


Figure 30 : Tableau de bord Alfresco Share

Le tableau de bord se personnalise en ajoutant ou supprimant des modules selon les besoins.



Figure 31 : Personnalisation du tableau de bord

**Les sites:** un site est un espace virtuel de collaboration réservé à une unité de travail ou à un projet. Les utilisateurs peuvent créer des sites publics et privés.



Tous les utilisateurs peuvent voir les sites publics, mais seuls les membres de ces sites peuvent y exécuter des tâches. Ils peuvent exécuter les tâches suivantes :

- ✓ travailler avec le contenu d'Espace documentaire;
- ✓ créer des événements dans le calendrier du site;
- ✓ créer, commenter et publier des messages sur un blogue;

- ✓ créer et commenter les liens des sites;
- ✓ traiter des thèmes de discussion;
- ✓ travailler avec des pages de wiki.

**Gestion de l'administration** l'administrateur est capable de gérer le thème du site, les groupes, les tâches de réplication, l'entrepôt et les utilisateurs grâce à une partie «administration» présente sur le site Alfresco Share.

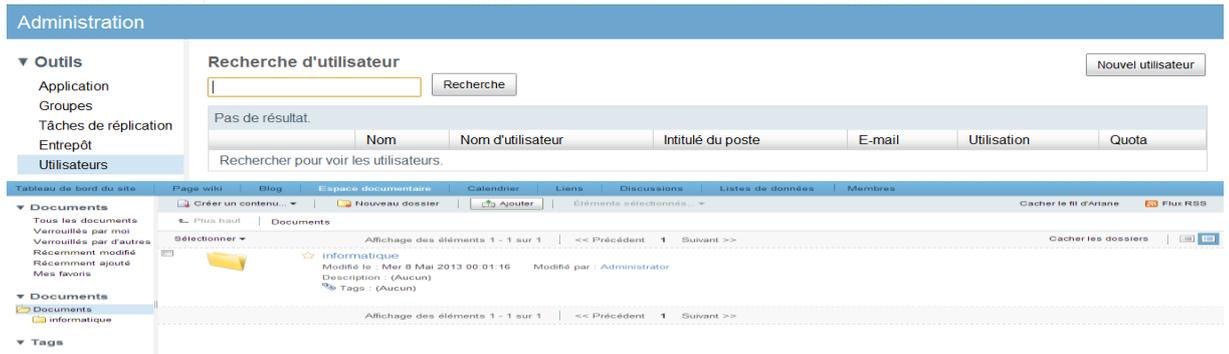


Figure 32: Tableau d'administration

**Espace documentaire:**

L'application permet aux utilisateurs de stocker et gérer tout contenu relatif à un site.

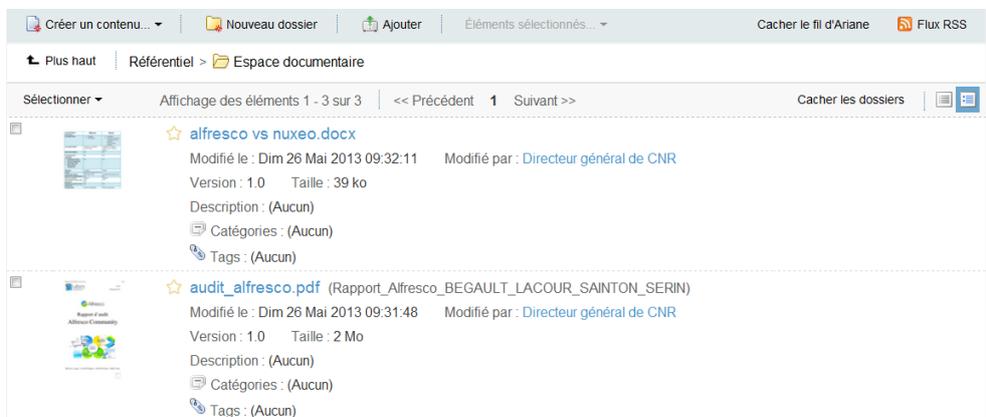


Figure 33: Espace documentaire

**2.6. Les droits d'accès**

**Lecteur :** Lire.

**Contributeur:** Lire et ajouter.

**Collaborateur :** Lire, ajouter, copier, supprimer, archiver, extraire et télécharger.

**Gestionnaire** : tous les droits d'accès, créer des sites, inviter d'autres utilisateurs, lire, copier, supprimer, archiver, extraire, télécharger et mettre à jour/modifier le contenu créé par d'autres utilisateurs.

### 2.7. La Recherche

**Recherche d'un document par mots clés**: Pour chercher un document, il est possible d'utiliser la zone de la recherche située dans le coin supérieur droit de l'interface.

**Recherche avancée** : La recherche avancée permet de rechercher un document avec l'utilisation de plusieurs critères.

### 3. Comparaison et Synthèse (Alfresco et Nuxeo)

caractéristiques fonctionnel	Alfresco	Nuxeo
Les points forts	<ul style="list-style-type: none"> <li>• Rapidité</li> <li>• Robustesse</li> </ul>	<ul style="list-style-type: none"> <li>• Flexibilité</li> </ul>
Les points faibles	Complexe à configurer	Complexe à configurer
Interface	Simple	Simple
Recherche en texte intégrale	✓	X
Sauvegarder de recherche	✓	✓
Module de Record mangement	✓	X
Gestion des droits	✓	✓
Archivage l'égal	✓	X
Cloud	✓ (la version 4.0 et 4.2)	✓
Extraire les métadonnées	✓	✓
Accès au contenu directement à travers des protocole	WebDAV ; CIFS ; FTP.	WebDav
Gestion de cycle de vie de documents	✓	✓
Workflow par défaut	✓	X
Communauté	Grands Communauté	Petite Communauté

**Tableau 4:** Comparaison et Synthèse (Alfresco - Nuxeo)

**La synthèse :**

Les deux solutions Alfresco et Nuxeo se sont distingués, car ils répondent aux besoins des entreprises. Pour la concrétisation de nos objectifs, nous avons retenu la solution Alfresco comme solution la plus adéquate et ce à cause de plusieurs critères.

- ✓ Alfresco est plus grand avec plus de sujet couvert,
- ✓ Il se caractérise par une communauté internationale assez importante,
- ✓ Alfresco a un module de Record Management dédié à l'archivage qu'il n'existe pas à Nuxeo,
- ✓ l'un des critères qui nous a poussé à choisir Alfresco est que Nuxeo n'intègre pas par défaut certaines fonctionnalités importantes, par exemple le moteur de recherche, pour la recherche en texte intégrale et que Alfresco donne la possibilité d'accéder aux documents de différentes façons, à savoir l'interface CIFS, le webdav (avec les fonctionnalités de versionning et le check-in/check-out) et le FTP. Par contre Nuxeo ne supporte que WebDav.
- ✓ Alfresco est un véritable outil de travail collaboratif qui offre non seulement la possibilité de partager des documents, mais aussi de gérer les versions de documents et la mise en place de circuits de validation grâce au système de workflow intégré.

Pour tous ces avantages Alfresco constitue le meilleur choix pour la gestion électronique de documents.

**Conclusion**

Les deux solutions « Alfresco » et « Nuxeo » servent de systèmes de gestion de documents. Chaque solution a ses propres avantages et inconvénients. Mais d'après notre étude comparative, nous avons conclu que la solution « Alfresco » est le meilleur système de gestion de documents électroniques avec ses fonctionnalités puissantes de gestion de documents. Aussi, « Alfresco » possède une interface utilisateur conviviale et une communauté d'utilisateurs dynamiques.

Dans le chapitre suivant, et pour mettre en pratique une partie de nos connaissances, nous avons voulu profiter une entreprise Algérienne de nos résultats. Nous avons alors intégré et adoptée la plateforme « Alfresco » au sein du Caisse Nationale de la Retraite (CNR) de Ouargla.

# Chapitre 4

*Mise en œuvre d'une solution open source « Alfresco » au sein du CNR<sup>2</sup>*



---

<sup>2</sup>Caisse Nationale de la Retraite

## Introduction

Aujourd'hui, les sociétés gèrent un nombre considérable de dossiers et de documents de différentes natures. L'utilisation de solutions modernes de GED facilite énormément le travail et donne plus de valeur à l'entreprise.

Nous avons choisi de mettre en place l'open source « Alfresco » au sein du CNR (Caisse Nationale de la Retraite). Nous pensons que le CNR représente un exemple idéal qui illustre les difficultés liées à la gestion des documents papiers et électroniques au sien d'une entreprise. Nous avons traité les documents dans un processus de bout en bout : dès sa création ou numérisation jusqu'à l'archivage ou la destruction.

Dans ce chapitre, nous allons d'abord présenter la société CNR et ces besoins. Puis nous exposons l'objectif de la solution, les grandes lignes de la réalisation de la solution ainsi que toutes les étapes de la mise en place de l'open source « Alfresco ».

### *Plan du chapitre*

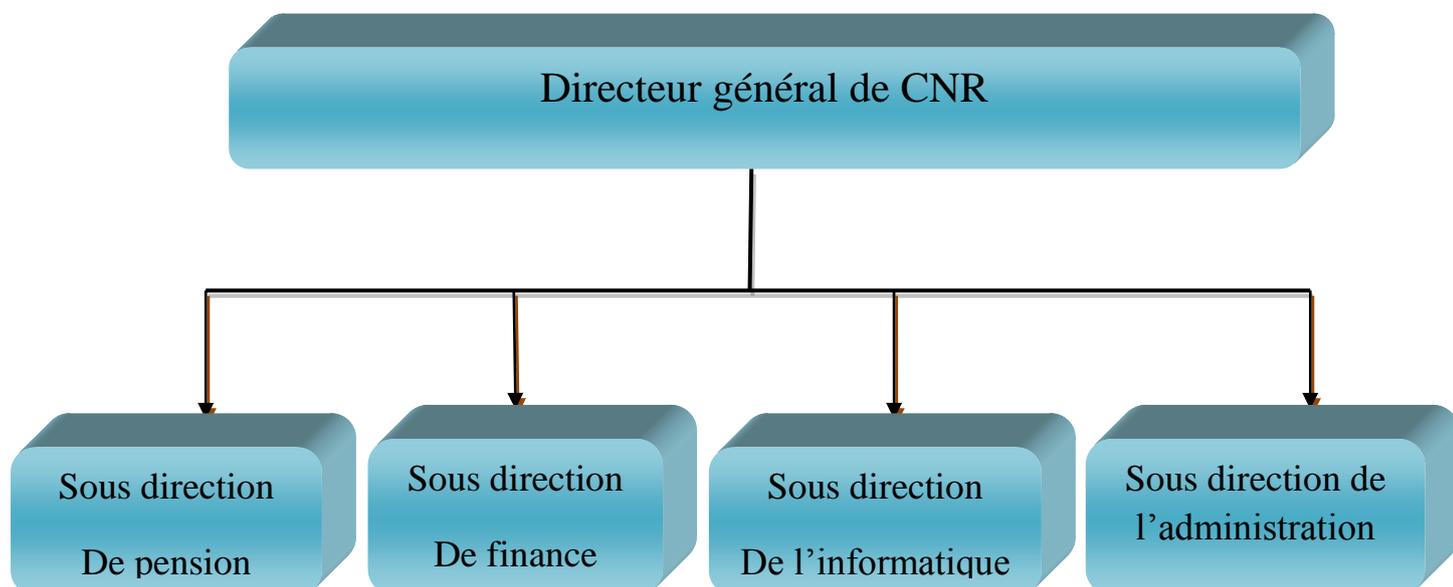


## 1. Présentation général de la CNR

La Caisse Nationale des Retraites (CNR) est une société qui a été créée par décret n°:85-223 du 20 août 1985 abrogé et remplacé par le décret N°: 92-07 du 04 janvier 1992 portant statut juridique des Caisses de Sécurité Sociale et organisation administrative et financière de la Sécurité Sociale. Chargée de la gestion des différents régimes de retraite existant avant l'institution en 1983, d'un régime national unique de retraite, offrant les mêmes avantages à tous les travailleurs quel que soit leur secteur d'activité. [51]

Son siège est situé : Route de Ghardaïa, de la wilaya d'Ouargla.

Chaque sous-direction est divisée en plusieurs unités. Ci-dessous l'organigramme général:



**Figure 34** : Organigramme de la direction de CNR

### Fonctionnement Administratif

Les organes essentiels chargés d'assurer le fonctionnement de la caisse sont :

- Le conseil d'administration.
- Le directeur Général.

### Le Conseil d'Administration

Il administre, contrôle et anime la Caisse. Il est composé de 29 membres répartis comme suit :

- **18** représentants des travailleurs par les organisations syndicales les plus représentatives,

- 9 représentants des employeurs dont 2 représentants de la fonction publique,
- 2 représentants du personnel de la Caisse.

### **Le Directeur Général**

Le directeur Général dirige la Caisse et assure son fonctionnement sous le contrôle du conseil d'administration.

#### **1.1.Missions de la CNR**

Elles sont les suivantes :

- Gérer des salaires et de retraite des subventions, en plus des salaires et d'accorder ces droits
- Gérer La conduite des salaires et des subventions conformément à la législation nationale pour Janvier 01, 1984 jusqu'à la fin de leurs droits;
- Assurer la fourniture d'informations pour les bénéficiaires et des travailleurs;
- Gérer jusqu'à extinction des droits, les pensions et allocations servies au titre de la législation antérieure au 1er Janvier 1984,
- Assurer le recouvrement, le contrôle et le contentieux du recouvrement des cotisations destinées au financement des prestations de retraite,
- Mettre en application les dispositions relatives prévues par les conventions et accords internationaux de sécurité sociale,
- Gérer le fonds d'aide et de secours en application de l'article 52 de la Loi N° : 83-12 du 02 Juillet 1983 relative à la retraite

**NOTA** : Les tâches Liées à l'affiliation et au recouvrement sont assurées pour l'essentiel par la **CNAS** en coordination avec les services de la **CNR**.

## **2. Présentation du projet**

L'objectif essentiel est la mise en place d'une solution de gestion électronique de documents dans la société de CAISSE NATIONNALE DE RETRAITE(CNR) afin de gérer, organiser et archiver les documents accumulés suite à la réalisation de la dématérialisation.

### **2.1. Les Besoins de la société CNR :**

La société CNR gère un nombre intensif des documents chaque jour, ce qui mène aux différents problèmes, parmi ces problèmes, la difficulté liée à la gestion de flux des documents papiers et électroniques qui circulent quotidiennement au sein de cette organisation, Le

manque de la centralisation et le partage des documents, la difficulté d'accès dans le temps aux documents archivés, la perte des dossiers liée aux problèmes de circulation des dossiers papiers manuellement ou à leurs recherches longues, et ainsi le grand espace physique d'archivage papier (traditionnel).

Ces difficultés diminuent le rendement en performance et productivité. La bonne réponse à ces préoccupations citées au par avant, *sera la mise en place de la plateforme de GED Alfresco*, qui permet de structurer toutes les informations, alors elle permet d'organiser, d'accéder instantanément à n'importe quel contenu et de concevoir des flux de travail automatisés au sein de la société.

## **2.2. Etude de la mise en place de la solution GED Alfresco**

Cette étude contient deux étapes principales (Amélioration de fonctionnement de CNR, l'architecteur réseau d'application).

L'objectif de notre projet est de tester les fonctionnalités et d'intégrer une open source Alfresco dans l'architecture actuelle du réseau de CNR. Et d'assurer un ensemble de fonctionnalité fourni par Alfresco.

### **2.2.1. L'amélioration de fonctionnement de CNR :**

#### **Les démarches suivies pour la mise en œuvre :**

- 1- **Premier cas** : L'archivage directe des dossiers qui existent dans l'archive (les retraités et les documents de l'administration) de l'entreprise CNR en utilisant une chaîne de dématérialisation. le processus passe en suite par une indexation, une identification, une sécurité des dossiers, une définition des droit d'accès aux archives et une définition des procédures requises pour la conservation des archives dans le système, jusqu'à leurs destruction ou leurs transfert vers un autre emplacement.
- 2- **Deuxième cas** : La gestion des dossiers avec la plateforme Alfresco allons de ça création, et le passage éventuellement par toutes les fonctionnalités ou les étapes nécessaire fournit par la solution d'Alfresco jusqu'à leur Archivage ou leur destruction.

#### **Processus de la réalisation :**

La numérisation de document papier existant par l'intermédiaire d'un scanners, ou l'intégration de document électronique existant.

L'enregistrement des documents par l'attribution d'une référence (nom de fichier, numéro,...) conforme aux règles de gestion de CNR. En suite le déclenchement d'un

circuit de revue et de validation (workflow) selon les règles organisationnelle. Sachant que le workflow lié à la validation est paramétrable, il prend en compte les droits d'accès et les profils d'utilisateurs système (annuaire LDAP). Chaque service à un responsable, chaque document est initié par un responsable (le document rédigé par un premier rédacteur, qui passe en suite en revue pour une validation par d'autres acteurs à chaque étape du workflow et une validation finale qui s'effectue par le responsable assurant que ce dernier est conforme aux procédures et accessible aux personnes concernés).

- Le classement et l'indexation des documents sous forme de :
  - Type (auteurs, dates, sources, ...) ou;
  - Concepts ou mots-clés contenu dans les documents.
  - Cette indexation vise principalement des opérations de recherche, pour fournir aux utilisateurs une recherche instantanée.
  - La gestion des documents recouvrent tout ce qui se passe sur le document après sa création : la sécurité et les droits d'accès, L'administration (modifications des métadonnées liées à l'attribution et la localisation du document) et la variation du document qui concerne la modification (la traçabilité).
  - La conservation du document vise à maintenir dans le temps sa disponibilité, cela induit à une notion d'une durée de fin de conservation de ce dernier (archivage pérenne, révision ou destruction). Il devra prendre en compte les besoins d'intangibilité mais de disponibilité permanente de certains documents à portée légale.

### **2.2.2. L'architecteur réseau d'application**

La mise en place de l'outil Alfresco pour notre exemple d'application nous a suffi selon l'existant de : trois machines

- Une machine serveur pour l'installation de l'open source ;
- Deux machines clients, un Switch et un Scanner.

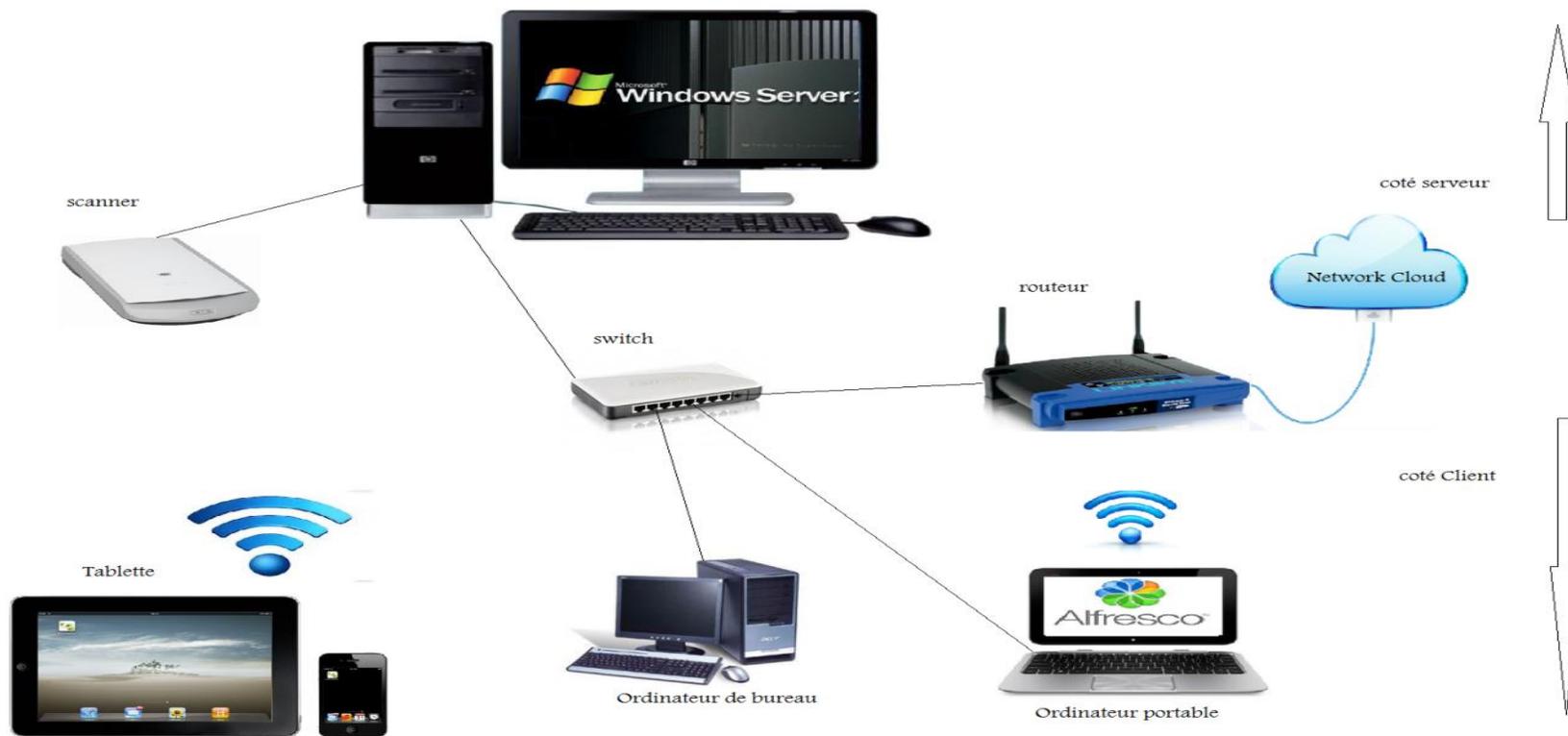


Figure 35 :L'architecteur globale de réseau d'application

## 2.3. Solution et mise en route de l'installation :

### 2.3.1. Outils utilisés pour l'installation

Les outils utilisés pour l'installation du présent projet sont les suivants :

- **Le JDK de Sun :** Le Java développement Kit(JDK) est l'environnement dans lequel le code Java est compilé pour être transformé en byte code afin que la machine virtuelle Java (JVM) puisse l'interpréter.
- **Java et Java Enterprise Edition (Java EE) :** Java EE est un standard de la technologie Java particulièrement adapté aux applications d'entreprise. En particulier pour la gestion d'application multi-niveaux, dont l'exemple le plus connu est l'architecture client-serveur. L'utilisation de Java EE et de nombreux modules permet à Alfresco d'être une application robuste, rapide et portable, c'est-à-dire qu'elle peut être utilisée sur plusieurs systèmes d'exploitation (en particuliers Windows, Linux et Mac OS).
- **Apache Tomcat :** Apache Tomcat est un conteneur de servlets Java EE. C'est un logiciel libre produit par la fondation Apache. Il est écrit en java comme l'application Alfresco, ce qui le rend utilisable partout où l'est Alfresco. Il permet le déploiement des servlets et des JSP (Java server pages) et intègre également un serveur HTTP. [50]
- **PostgreSQL :** PostgreSQL est un moteur de bases de données relationnelles.C'est un moteur adapté à des bases métier, donc riche en fonctionnalités et puissant. Son installation est cependant plutôt simple. [28]
- **Eclipse (Eclipse Java EE for Web Développeur, version) :** Eclipse est un environnement de développement intégré libre extensible, universel et polyvalent, permettant de créer des projets de développement mettant en œuvre n'importe quel langage de programmation. [50]

### 2.3.2. Les étapes de l'assistant d'installation

L'installation d'Alfresco via l'assistant d'installation pour Microsoft Windows installe tous les logiciels et les composants nécessaires pour faire fonctionner Alfresco.

- Sélectionner la langue d'installation, qui définira la langue à utiliser pour le reste de l'assistant d'installation.
- La sélection des composants à installer :
  - Share Point ;
  - Gestion des enregistrements
  - Web Quick Start
  - Web Project Management (AVM)

- Prise en charge du connecteur Quick
- Open Office ;
- Les composants par défaut sont :
  - PostgreSQL
  - Java
  - Alfresco
- Le choix de type d'installation **Avancé** pour la configuration des ports de serveur d'application Tomcat , et les propriétés de service :
  - **Web server Domain** : 192.168.1.3
  - **Port de serveur Tomcat** : 8080
  - **Port d'arrêtTomcat** : 8005
  - **Tomact SSL port** : 8443
  - **Port AJP Tomcat** : 8009
- La saisie de mot de passe pour l'utilisateur *admin*
- Une fois l'installation terminée, Le serveur Alfresco démarre et AlfrescoShare lance dans le navigateur par défaut. A l'adresse <http://192.168.1.3:8080/share> .

### 2.3.3. Configuration et développement de l'environnement

Pour la configuration et le développement d'Alfresco, plusieurs cas sont possibles. L'ajout de classes ou de codes source.

#### **Intégration du module record management (RM) :**

Dédié à l'archivage électronique,

- Téléchargement d'Alfresco communauté dod5015-3.4.d.zip.
- Déplacement d'alfresco-community-dod5015-3.4.d.amp et dod5015-part-3.4.d.amp dans les **répertoires amps et amps\_share**.
- Arrêter le serveur Alfresco
- Exécution de la **commande** apply\_amps, dans Alfresco root (sudobash apply\_amps.sh)
- Lancement de serveur Alfresco ;

#### **Configuration de la prévisualisation des documents :**

Pour activer la prévisualisation quelques configurations ont été apportées à la classe d'Alfresco **alfresco-global.properties** :

```

### External executable locations ###
ooo.exe=C:/Alfresco/openoffice/App/openoffice/program/soffice.exe
ooo.enabled=true
ooo.port=8100
img.root=C:/Alfresco/imagemagick
img.dyn=${img.root}/lib
img.exe=${img.root}/convert
swf.exe=C:/Alfresco/swftools/pdf2swf.exe
jodconverter.enabled=false
jodconverter.officeHome=C:/Alfresco/openoffice/App/openoffice
jodconverter.portNumbers=8100

```

### Configuration de serveur FTP :

La modification de la propriété ftp.enabled.

```

### FTP Server Configuration ###
ftp.enabled=true
ftp.port=21
ftp.ipv6.enabled=false

```

### Les interfaces utilisées pour l'accès à la plateforme

- ✓ L'interface bureautique FTP

[FTP://\[serveur\]](#)

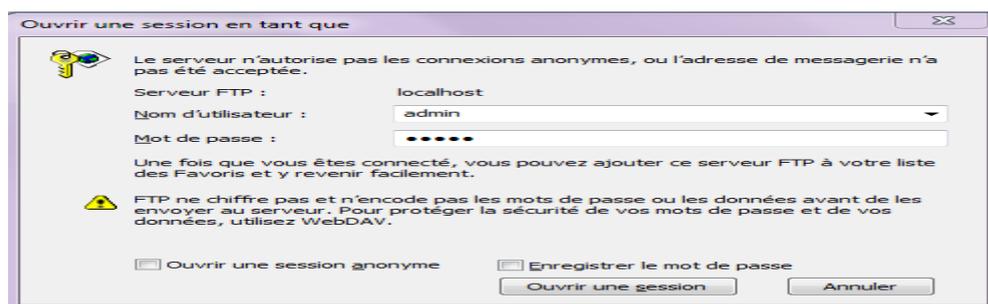


Figure 36 : L'interface bureautique FTP

- ✓ L'interface WebDav

[http://\[serveur\]:8080/alfresco/webdav/](#)

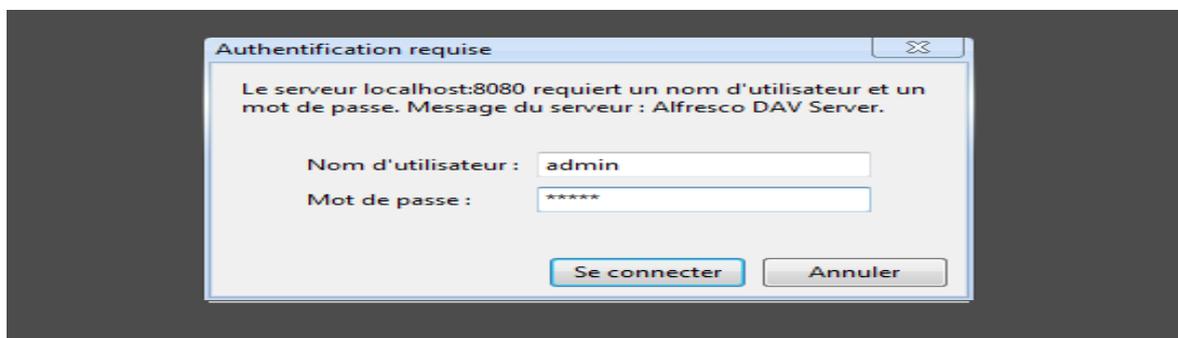


Figure 37:L'interface WebDav

Qui s'affiche :

Contenu du répertoire /

Nom	Taille	Type	Date de Modification
<a href="#">Dictionnaire de données</a>			Wed, 22 May 2013 20:32:15 GMT
<a href="#">Espace invité</a>			Wed, 22 May 2013 20:32:16 GMT
<a href="#">Espaces Utilisateurs</a>			Wed, 22 May 2013 20:32:16 GMT
<a href="#">Sites</a>			Wed, 22 May 2013 20:32:35 GMT
<a href="#">Projets Web</a>			Wed, 22 May 2013 20:33:07 GMT
<a href="#">Web Deployed</a>			Wed, 22 May 2013 20:33:07 GMT

✓ **URL clients Web**

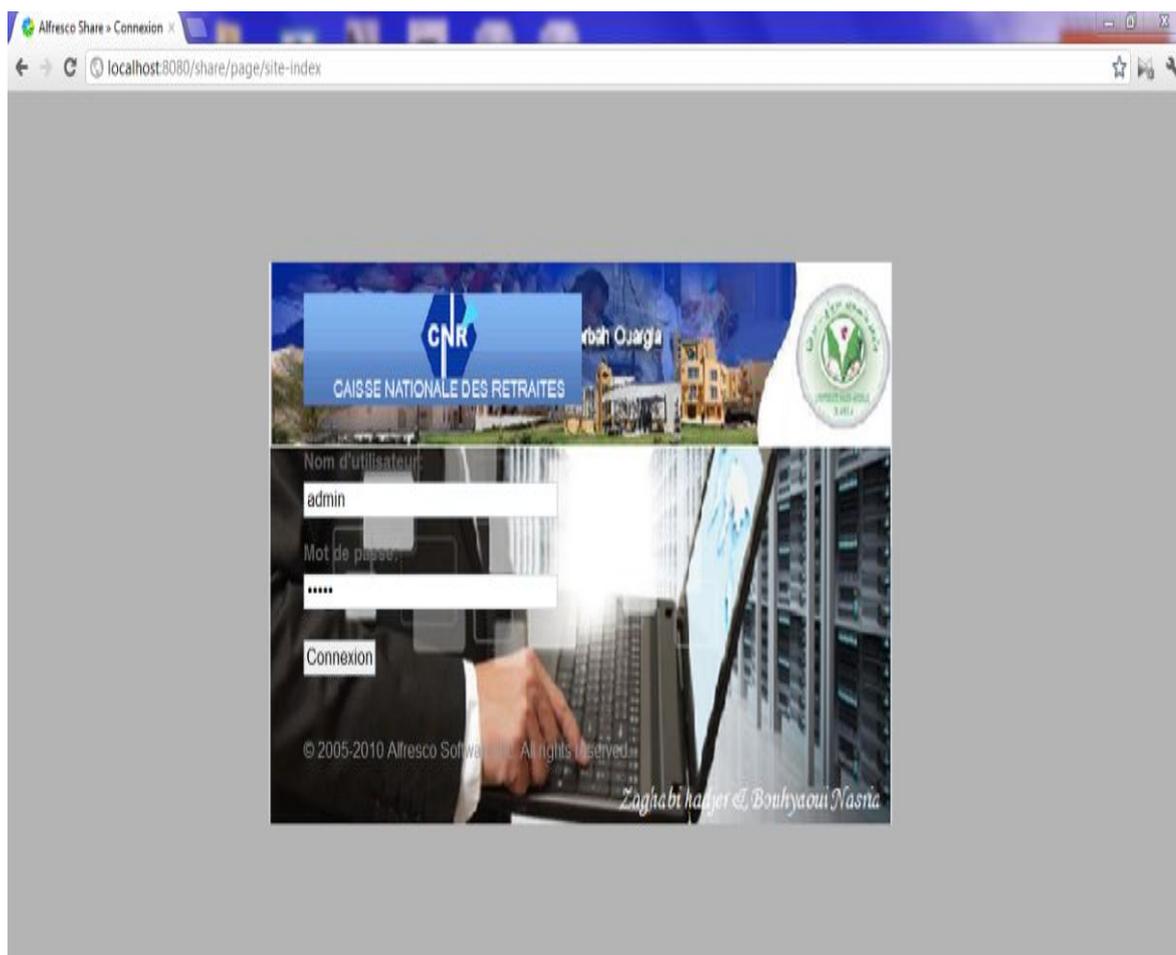
[http://\[serveur\]:8080/share](http://[serveur]:8080/share)

[http://\[serveur\]:8080/alfresco](http://[serveur]:8080/alfresco)

### 2.3.4. Interface graphique de la solution

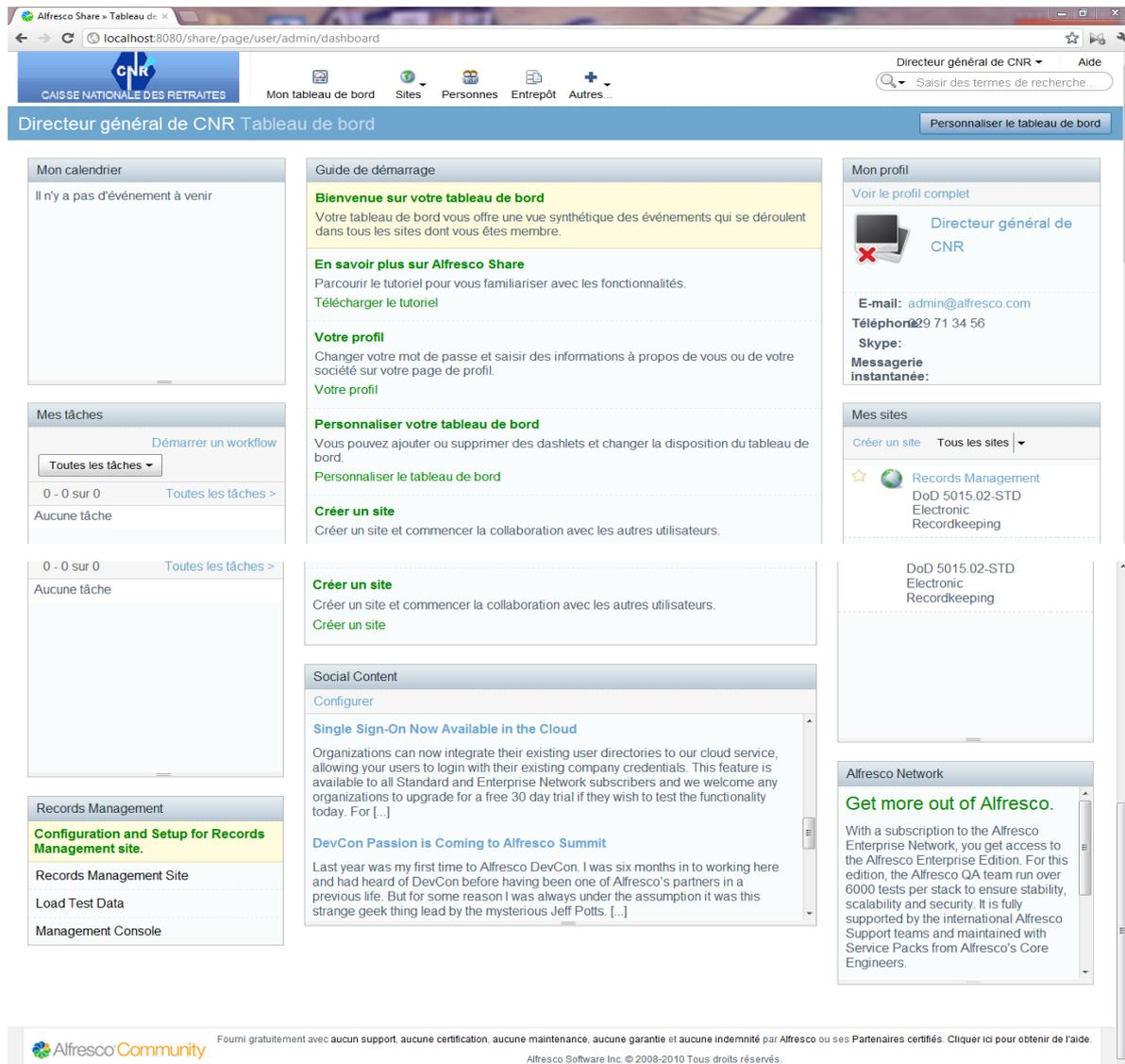
#### Authentification

La première page est la page d'authentification. En effet l'utilisateur doit tout d'abord s'authentifier avant d'accéder à la solution. Cette authentification est faite selon l'annuaire LDAP de CNR.



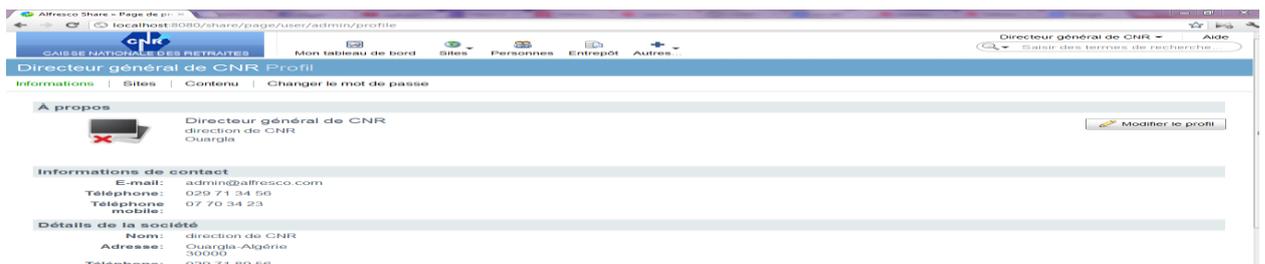
**Figure 38:**Page d'identification de l'application

**Tableau de bord :** une fois authentifié l'utilisateur accède à son tableau de bord qui se présente comme suit :



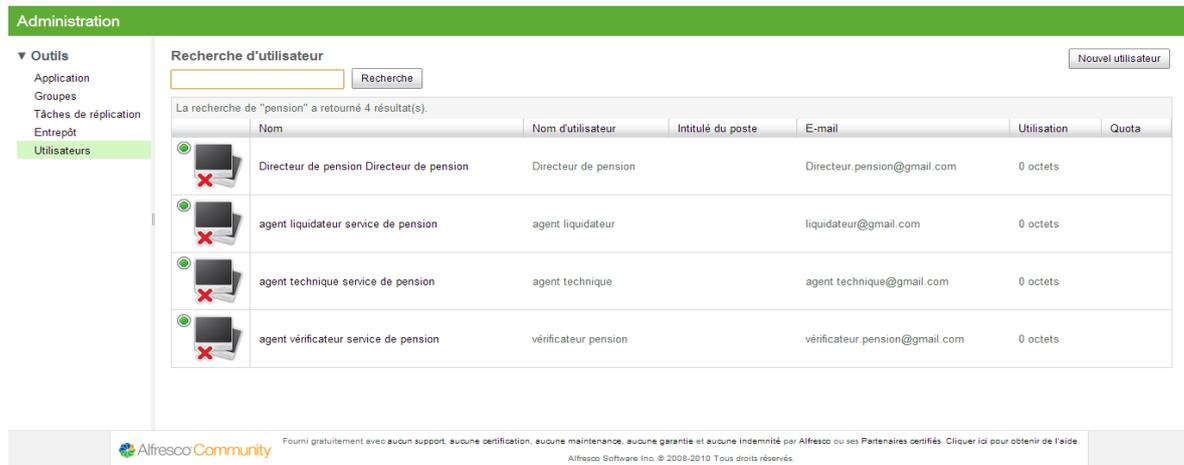
**Figure 39:** page de tableau de bord

**Le profil d'administrateur (Directeur général de CNR) ;**



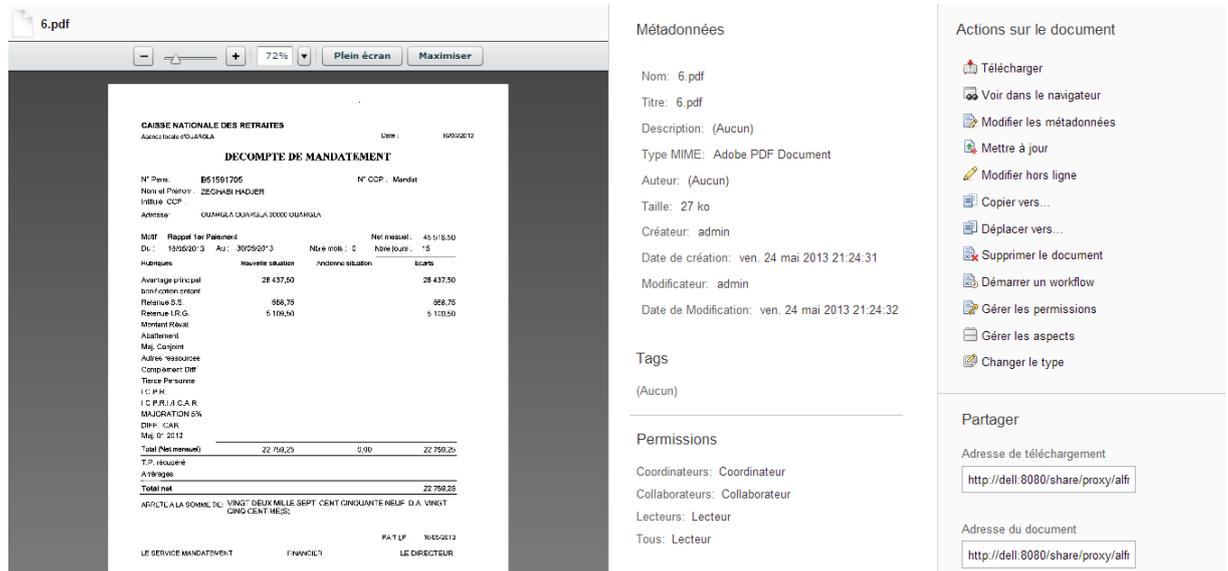
**Figure 40:** profil d'administrateur (Directeur général de CNR)

**Gestion des utilisateurs :** l'administrateur peut gérer les utilisateurs et les groupes à partir la page suivantes.



**Figure 41 :** utilisateurs de CNR

**Prévisualisation des documents :** Un document peut être pré-visualisé directement à partir de la GED avec toutes ces métadonnées ainsi que la possibilité d'ajouter un commentaire au document.



**Figure 42:** Prévisualisation des documents

**Métadonnées** : la modification des métadonnées.

#### Modifier les métadonnées

\* Champs requis

Nom: \*

Titre:

Description:

Type MIME:

Auteur:

Étiquettes:

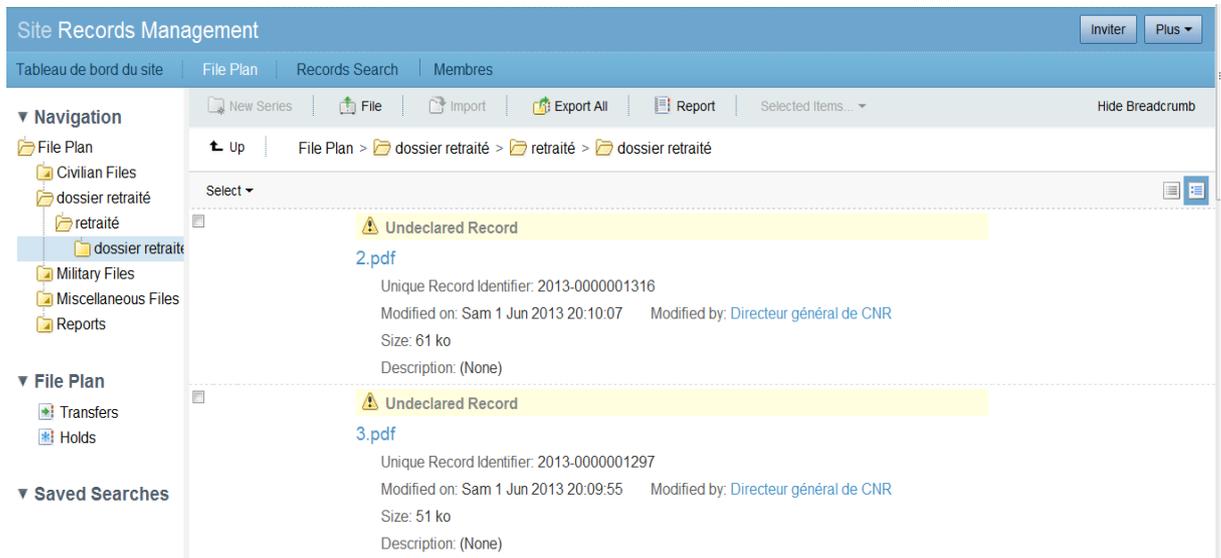
**Figure 43:** la modification des métadonnées

#### **Le module de Record Management :**

Le module de Record Management d'Alfresco dédié à l'archivage qui est caractérisé par un référentiel unique et simple. Tous les contenus sont recueillis, gérés et contrôlés au sein d'un seul référentiel et ainsi la gestion des archives directement via une interface web, qui simplifié la gestion pour les utilisateurs. Ce module fourni un modèle d'autorisations optimisé, basé sur les compétences et rôles des utilisateurs, des cycles de vie d'archivage très complets, basés sur des calendriers, la gestion de transferts complexes, la mémorisation des recherches effectuées pour optimiser la gestion et le recouvrement des documents. Ce document est « un document validé, non modifiable, qui est archivé à des fins de preuves et d'audit. On parle de documents « vivants », qu'on peut consulter régulièrement. D'archives intermédiaires qu'on peut ressortir **en cas** de litige pendant un DUA (durée d'utilité administrative) **et**, enfin, d'archives définitives de valeur historique ou patrimoniale ».[49]

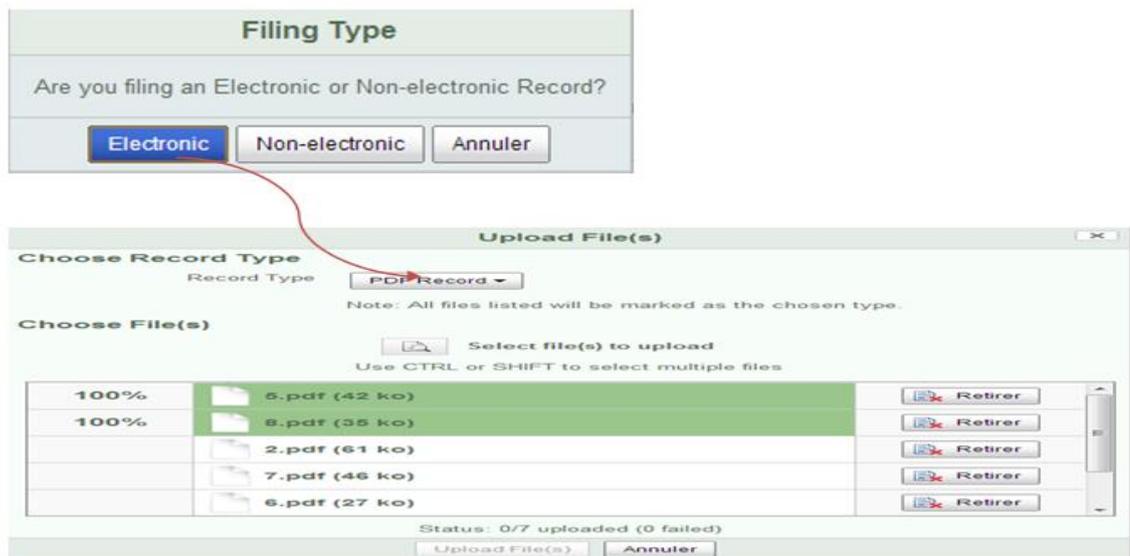
### La conception de File Plan

Record series, record category, et la hiérarchie record folder.



**Figure 44:**File Plan de record management

L'importation des fichiers électroniques et la spécification de l'emplacement des fichiers physiques non électroniques.



**Figure 45:**L'importation des fichiers électroniques

## La Déclaration des fichiers comme records.



Figure 46: La Déclaration des fichiers comme records

## Disposition schedule

« Disposition schedule » c'est des procédures requises pour la conservation des records dans le système de "records management" jusqu'à leur destruction ou leur transfert vers un autre emplacement.

Préparation de "disposition schedule" pour les "record" au sein d'une catégorie "category".

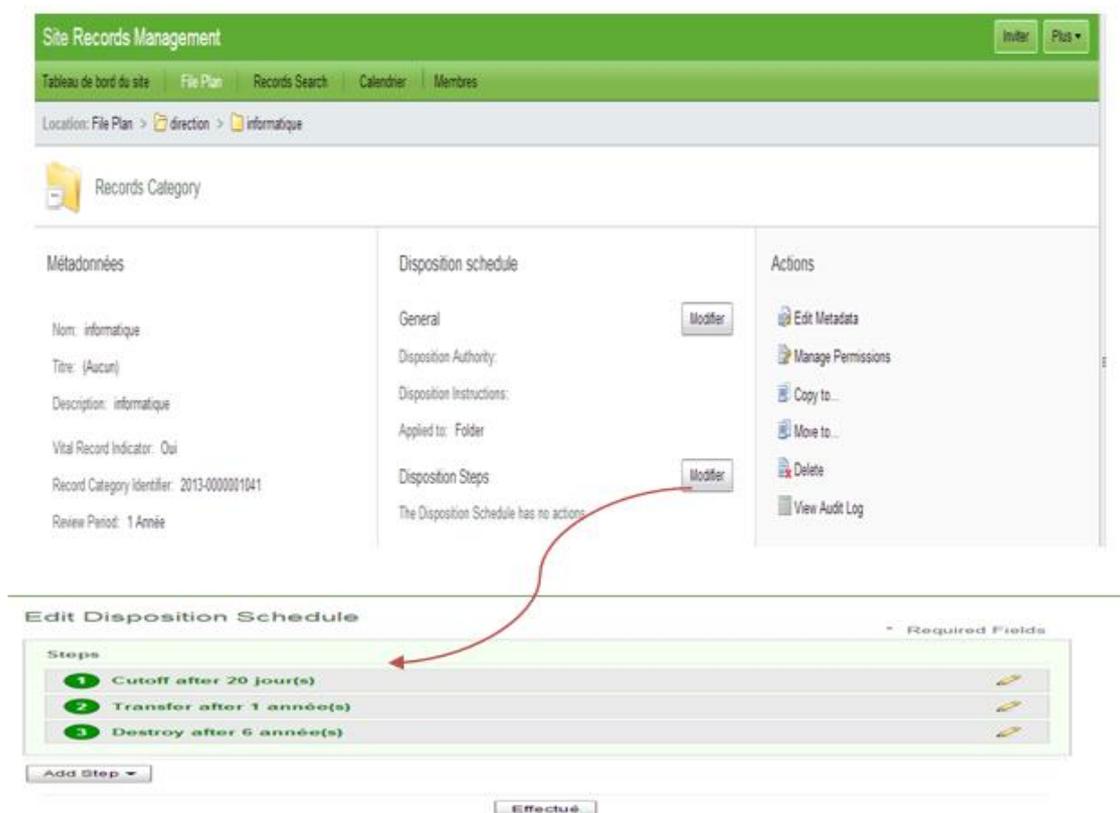


Figure 47: Disposition schedule

**Conclusion**

Dans ce dernier chapitre, nous avons montré les différentes étapes utilisées pour le déroulement du développement, adaptation et configuration de la solution choisie (Alfresco) au sein du CNR (Caisse Nationale des Retraites) de Ouargla.

Tout d'abord et comme notre encadreur avait prédit, il s'est avéré très utile et très bénéfique d'avoir choisi de mettre une partie de nos connaissances et résultats au sein d'un milieu professionnel. Nous avons également eu l'occasion de faire profiter une entreprise locale avec notre étude et nos conclusions qui auraient pu coûter trop cher à l'entreprise.

La plateforme « Alfresco » semble être très intéressante puisqu'elle dispose d'une interface utilisateur conviviale et de beaucoup de fonctionnalités avancées.

# *Conclusion Général et Perspectives*

Aujourd'hui, les documents numériques sont de plus en plus utilisés par les entreprises et les administrations à jour avec les nouvelles technologies. Ces documents numériques sont soit nativement le résultat de l'utilisation de l'informatique ou obtenus suite à un processus de dématérialisation. Par conséquent, les entreprises doivent faire face au traitement d'un grand nombre de documents archivés en toute sécurité. Pour rester compétitives, elles doivent être capables de réagir rapidement à toute demande pour trouver un document archivé. Les solutions de « Gestion Electronique des Documents »- GED –sont des outils indispensables pour répondre à cette exigence. En effet, la GED permet de centraliser rapidement, et en quelques clics, l'information. Le document est facilement accessible par la suite. Il peut aussi être partagé par plusieurs personnes au même moment. La GED garantit également la sécurité des données et crée un dynamisme de travail collaboratif. Le souci avec ce type de système réside essentiellement dans les supports de stockage ; il convient de veiller particulièrement à ceux-ci afin de garantir la pérennité des données et des documents.

Malgré les difficultés rencontrées durant ce projet et en particulier le fait de traiter un sujet d'actualité avec peu de documentation, nous sommes parvenues à nous familiariser avec ces nouvelles technologies liées au GED.

Nous pensons que l'objectif initial de ce projet de fin d'études, qui était la réalisation d'une étude comparative des solutions de stockage et d'archivage en local et sur le « Cloud » – Optimisation et Intégration d'une solution OPEN SOURCE, est atteint. A l'exception et par manque de temps de ne pas arriver à intégrer la solution sur le Cloud.

L'intégration de la solution Open Source « Alfresco » au sein de la Caisse Nationale des Retraites (CNR) de Ouargla a été très bénéfique. En effet, le stage au CNR nous a permis de se rapprocher du milieu professionnel et faire profiter une administration locale de nos travaux.

Il est clairement montré que la solution « Alfresco » est très intéressante. Il sera plus facile et encore plus bénéfique que d'autres personnes poursuivent notre travail et faire une étude de cas complète au CNR ou chez toute autre administration similaire, tout en ajoutant l'implémentation de la solution sur le « Cloud ».

# Bibliographie

## *Livres électroniques*

- [1] **AbderrazakMkadmi**, document pdf disponible à l'adresse URL :  
<http://abderrazakmkadmi.free.fr> /Cours destiné aux Etudiants de l'Institut supérieur de documentation Licence : Documentation Bibliothéconomie et Archivistique 2<sup>ème</sup>, **2009**.
- [2] **Michel Bolognesi**, mémoire « *archivage de données SAP* », informatique, systèmes d'information, 17 septembre 2010.
- [3] **livre blanc** « *progiciel de gestion d'archivage Mixtes* » enjeux ; apports et critères d'archivage.
- [4] « *Principes et exigences fonctionnelles pour l'archivage dans un environnement électronique* », Conseil international des Archives 2008
- [5] **Marie-Anne CHABIN, Eric CAPRIOLI, /Jean-Marc RIETSCH** « *l'archivage électronique à l'usage du dirigeant* »
- [6] **Instance responsable : Unité Innovation et préservation** « *formats de fichier adaptés a l'archivage normes et standards normes et standards pour l'archivage de documents Numériques/archives fédérales* »suisses, Juillet **2007**.
- [7] **LUDOVIC Desaubry**,mémoire, « La dématérialisation des dossiers documentaires», 16 novembre 2009
- [8] **CHOPPY Thomas, Consultant ECM /Nicolas RICHTON** « Smile », livre blanc « *GED open source* »
- [9] « *Mandat archivage électronique Archivage électronique Guide de bonnes pratiques Fiches annexes* », **Version du 13/09/2012**
- [10] **BOLOGNESI Michel**, mémoire : spécialité : informatique /option : informatique, systèmes d'information, « *Archivage de données SAP* »/ 17 septembre 2010
- [11] **AGNIESZKA Tona**, Mémoire « *Comment choisir une solution de gestion d'archives* »Master 2 Gestion de l'Information et de la Documentation en Entreprise, 30 septembre 2010,

- [12] **Gardarin Georges**, georges.gardarin.free.fr « Cours\_BD/TD1\_Gestion\_des\_disques »
- [13] « *Archivage, stockage, protection et récupération des données de l'entreprise* » 29 route de la Borde - 78110 Le Vésinet [France/storage@decisionmakersit.com](mailto:France/storage@decisionmakersit.com)
- [14] **BLANC Xavier et TAHIR Mohamed**, « *Les solutions de stockage de grande capacité* »
- [15] **Edissyum & Aliocha Iordanoff, Semaweb** « *Le cloud computing?* » Conférence du 17 janvier 2013
- [16] **Tuteur : M. Lucas Nussbaum/Groupe : Vincent kherbache, Mohamed Moussalih, Yannick** « *Cloud Computing* », Kuhn, Allan Lefort, IUT Charlemagne 2009/2010.
- [17] **Livre blanc** « *Sécuriser les données sur Google Drive Le guide des entreprises souhaitant garder leurs données en sécurité sur Google Drive* ».
- [18] **BONNEAU Céline**, mémoire « - *gestion et traitement de l'information spécialisée Mise en place d'une Gestion Électronique de Documents à la SACD participation à l'élaboration du cahier des charges* », Diplôme national de master Domaine - sciences humaines et sociales Mention - sciences de l'information et des bibliothèques Spécialité
- [19] **Mme ROLAND** (tutrice professionnelle) & **M. DEVELEY Lille 3 : M. DAVIGNON** (tuteur universitaire), « *mise en place d'une méthodologie de gestion documentaire au sein d'une entreprise Master 2 Information, Communication, Culture & Documentation* », Stage de fin d'année, **2010-2011**.
- [20] **RAÏS Nadia**, « *La dématérialisation des dossiers documentaires : les enjeux et les techniques étude de faisabilité réalisée pour le centre de documentation des Arts Décoratifs* »
- [21] **P.MIREBEAU** « *dossier GED (Gestion Electronique des Documents)* », **PRESENTATION GED - 2012**
- [22] **DESBETES Géraldine, LEROY Lucile**, « *La Gestion Electronique des documents, Typologie des systèmes d'information* », Mars 2008.
- [23] **Conseil international des Archives** « *Principes et exigences fonctionnelles pour l'archivage dans un environnement électronique* », **2008/ISBN : 978-2-918004-00-4**.
- [24] **GED GUIDE** utilisateurs Référencement de ce guide « *utilisation GED-CRI-G-DOCU* », université de pau et des pays de l'adour centre de ressources informatiques Date de publication /**2013**

- [25] **RAIS Bouchra**, « *Etude et mis en place d'une solution de gestion électronique de documents open source* », Ecole National Supérieure d'Informatique et Analyse des Systèmes, année 2010/2011.
- [26] « *Les avantages de la GED* » <http://ngominh.zevillage.org/files/GED.pdf>
- [27] **GUILLAUME Superczynski**, livre « *Alfresco 3.4 Travail collaboratif et GED avec la plate forme Share* », Editions ENI, 12sept.2011, pages 12  
[http://books.google.dz/books/about/Alfresco\\_3\\_4.html](http://books.google.dz/books/about/Alfresco_3_4.html)
- [28] **BEGAULT Agathe - LACOUR Delphine - SAINTON Etienne - SERIN Adrien** « *Rapport d'audit Alfresco Community* », Printemps 2011.

## webographie

- [29] « **Archivage** », <http://www.dlge.ch/competences/73-archivage.html> , consulté le 20/02/2013
- [30] « Archivage légal « à valeur probante » <http://www.locarchives.fr/nosprestations/dematerialisation/dematerialisation-de-processus>, consulté le 28/02/2013
- [31] « Dématérialisation de documents », <http://www.ged-de-france.fr/projetdetail-dematerialisation-4.htm>, 15/02/2013
- [32] « *Les images JPEG* », <http://www.buvetteetudiants.com/cours/administrator/html-css/jpeg-jpg-jpe-jfif.php>, consulté le 13/03/2013
- [33] « Stockage, le guide pratique »  
<http://stockage.comprendrechoisir.com/comprendre/entreprise-archivage.html>, consulté le 20/02/2013
- [34] « Les réseaux de Stockage » <http://www.numidit.dz/index.php/gestion-electronique-des-documents>, consulté le 13/03/2013
- [35] « DAS (Direct Attached Storage) »  
<http://www.maisondunumerique.com/produits/peripheriques/serveur-de-stockage/das>
- [36] « cloud-service », <https://www.dropbox.com> / <http://www.clubic.com/application-web/article-458202-2-cloud-service.html>, consulté le 22/05/2013
- [37] « iCloud », <http://www.aaai.fr/actualites/icloud-quest-ce-que-cest>, consulté le 20/06/2013
- [38] « SkyDrive », <http://febangers-micro.overblog.fr/pages/skydrive.html>, consulté le 20/04/2013

- [39] « ged-open-source.smile » [http://ged-open-source.smile.fr/LES\\_fonctionnalites/Metadonnees/](http://ged-open-source.smile.fr/LES_fonctionnalites/Metadonnees/)
- [40] « open source » [http://www.opensource.tn/fileadmin/Code\\_php/GED.web](http://www.opensource.tn/fileadmin/Code_php/GED.web), consulté le 26/06/2013
- [41] « Import / export en masse » <http://ged-open-source.smile.fr/LES-FONCTIONNALITES/Import-export>, consulté le 28/05/2013
- [42] « Gestion des profils », consulté le 26/04/2013  
[http://www.opensource.tn/fileadmin/Code\\_php/GED.web/AnnuaireEtSSO.html](http://www.opensource.tn/fileadmin/Code_php/GED.web/AnnuaireEtSSO.html)
- [43] « Annuaire et SSO », consulté le 28/04/2013  
[http://www.opensource.tn/fileadmin/Code\\_php/GED.web/AnnuaireEtSSO.html](http://www.opensource.tn/fileadmin/Code_php/GED.web/AnnuaireEtSSO.html)
- [44] « Single Sign On ». <http://www.savoirfairelinux.com/solutions/entreprise/ged>
- [45] « La GED en mode SaaS », consulté le 19/05/2013  
<http://www.cddoc.fr/logiciel-ged/ged-mode-saas.html>  
<http://www.cddoc.fr/hebergement-ged-saas-cloud>  
<http://www.opensource.tn/fr/open-source/thematique/>
- [46] « Alfresco », consulté le 22/05/2013  
<http://www.alfresco.com>  
<http://www.starxpert.fr>  
<http://www-igm.univ-mlv.fr/~dr/XPOSE2010/alfresco/alfresco.html#alfresco>  
<http://www.sollan.com/v6/index.php/Produits-Partenaires/Fonctionnalites-d->
- [47] « Knowledge Tree », consulté le 23/05/2013  
[http://www.opensource.tn/fileadmin/Code\\_php/GED.web/PrSention4.html](http://www.opensource.tn/fileadmin/Code_php/GED.web/PrSention4.html)  
<http://www.opensource.tn/actualit%C3%A9single/newspage/1/article/la-ged-en-mode-open-source>  
<http://referentiel.lecko.fr/knowledgetree/>
- [48] « open source Jahia », consulté le 30/05/2013  
[http://www.opensource.tn/fileadmin/Code\\_php/GED.web/PrSention3.html](http://www.opensource.tn/fileadmin/Code_php/GED.web/PrSention3.html)
- [49] « Connexion à open source Nuxeo EP », consulté le 26/05/2013  
<http://www.nuxeo.com/fr/produits/digital-asset-management>

- [50] « l'espace de travail collaboratif dans Nuexo », consulté le 29/05/2013  
<http://www.nuxeo.com/fr/produits/social-collaboration>
- [51] « La Caisse Nationale des Retraites (CNR) », <http://www.cnr-dz.com/>, consulté le 30/05/2013
- [52] « PostgreSQL » <http://www.commentcamarche.net/s/PostgreSQL+>, consulté le 30/05/2013

# *Annexes*

## **LEXIQUE DU MÉTIER**

### **Archivage**

Opération consistant à assurer la conservation d'un document, quel que soit son support, en vue d'une consultation ultérieure, à titre de preuve ou d'information.

### **Archivage électronique**

Opération qui consiste à stocker des documents sur des supports informatiques à des fins de conservation et de consultation ultérieure. La conservation de l'information doit offrir des garanties d'intégrité et de fiabilité dans le temps.

### **Archives**

Ensemble des documents, quels que soient leur date, leur forme et leur support matériel, produits ou reçus par toute personne physique ou morale ou par tout service ou organisme public ou privé, dans l'exercice de son activité .

### **Document d'archives**

Documents créés, reçus et préservés à titre de preuve et d'information par une personne physique ou morale dans l'exercice de ses obligations légales ou la conduite de son activité.

### **Archives courantes**

Dossiers ouverts ou récemment clos utilisés pour le traitement des affaires.

### **Archives définitives**

Documents conservés sans limitation de durée.

### **Archives intermédiaires**

Documents qui ont cessé d'être considérés comme archives courantes, et ne peuvent, en raison de leur intérêt administratif, faire l'objet de tri et d'élimination.

### **Archives privées**

Ensemble des documents qui n'ont pas le caractère d'archives publiques.

### **Archives publiques**

Documents qui procèdent de l'activité de l'Etat, des collectivités locales, des établissements et entreprises publics ou bien d'organismes de droit privé chargés de la gestion des services publics ou d'une mission de service public ou bien encore des minutes et répertoires des officiers publics ou ministériels ainsi que des fonds et collections acquis à titre divers.

## **Archivistique**

Discipline qui étudie la formation, le tri, l'exploitation et la conservation des ensembles organiques d'archives (dossiers - séries - fonds) et par suite les méthodes et les techniques nécessaires à leur traitement.

## **Communication**

Opération matérielle et administrative de sortie d'une archive d'un centre d'archives vers un utilisateur.

## **Conservation**

Stockage et gestion d'archives dans des magasins adaptés aux conditions de sécurité requises selon la nature des archives.

## **Document**

Ensemble d'informations cohérentes possédant un format et un support.

## **Dossier**

Ensemble des documents réunis pour la conduite et le traitement d'une affaire.

## **Fonds d'archives**

Ensemble de documents de toute nature qu'une personne physique ou morale a automatiquement et organiquement réunis dans l'exercice de ses activités et conservés en vue d'une utilisation éventuelle.

## **GEIDE**

Gestion Electronique d'Information et de Documents pour l'Entreprise.

## **Indexation**

Opération qui consiste à analyser et enregistrer sous une forme structurée le contenu d'un document, afin d'obtenir des points d'accès permettant le repérage des documents et/ou des informations.

## **Knowledge Management**(ou gestion des connaissances)

Discipline de management qui considère les connaissances, la mémoire et les savoir-faire de l'entreprise comme un capital. Dans un système de knowledge management, les connaissances sont identifiées, diffusées et stockées de manière évolutive et au profit de la performance générale de l'entreprise.

## **LAD**

(Lecture Automatique de Documents) Application de conversion des formulaires papier en données caractères et images.

Les applications de LAD reconnaissent les marques, les caractères imprimés, les caractères manuscrits pré-casés, les caractères manuscrits en caractères d'imprimerie, les mots manuscrits.

## **Métadonnées**

Données décrivant le contexte, le contenu et la structure des documents ainsi que leur gestion dans le temps

## **Numérisation**

Opération d'acquisition de documents en vue de leur transformation sous forme d'images numériques.

## **Plan d'archivage**

Etat typologique et parfois descriptif des dossiers et documents d'un organisme, classés par fonds, donnant pour chacun d'eux les règles de conservation et de communication.

## **Plan de classement**

Système qui préside à l'organisation des archives courantes dans les bureaux.

## **Records Management**

Ensemble des techniques d'organisation et outils fédérant la gestion administrative, logistique et informatique des processus de création, utilisation, maintenance, recherche et élimination des documents produits ou reçus par l'entreprise.

## **Signature électronique**

La signature numérique consiste en une suite d'opérations mathématiques sur un document électronique, afin de permettre à une personne d'assurer à la fois l'intégrité et l'authenticité d'un document.

## **Tableau de gestion**

Outil utilisé dans la gestion des archives et définissant, pour chaque type de document, la durée de conservation à appliquer en fonction de la valeur du document ainsi que les critères de sélection des archives définitives.

## Versement

Opération matérielle et administrative de transfert d'archives d'un service versant vers le centre d'archives. **Remarque** : Les définitions proposées ci-dessus sont, pour partie, extraites de la norme NF Z40-350 sur les « prestations en archivage et gestion externalisée de documents sur support papier » et de l'ouvrage « Ecole Nationale des Chartes - AFNOR; Dictionnaire des archives : de l'archivage aux systèmes d'information; AFNOR; 1991 ».

(1) Norme NF Z 42-013 : Archivage électronique : Spécifications relatives à la conception et à l'exploitation de systèmes informatiques en vue d'assurer la conservation et l'intégrité des documents stockés dans ces systèmes / Association Française de Normalisation. —AFNOR : Décembre 2001, 34 p.

(2) Norme internationale ISO 15489-1 : Information et documentation — « Records management » / ISO. — ISO : 15 septembre 2001, 19p.

## Les 5 grandes catégories de GED existantes sont

- **La GED administrative** permet de numériser puis de classer les documents administratifs (factures, fiches techniques, formulaires, devis ...).
- **La GED bureautique** regroupe l'offre de progiciels de travail collaboratif permettant d'échanger des documents, de les lire dans leur format d'origine (Word, Excel, Powerpoint, Outlook ...).
- **La GED COLD** (Computer Output on Laser Disc) qui permet d'archiver sous une forme électronique les états produits par l'informatique d'un organisme (relevés de compte, factures, etc).
- **La GED technique** ou GED métier qui concerne la manipulation de documents dont le format et le contenu sont propres à un métier (plans, schémas etc..).
- **La GED documentaire** consiste à indexer un grand nombre de fichiers numériques aux formats les plus divers (texte, image, ..) selon des critères définis par et pour l'organisme. Ce mode de gestion suppose de mettre en place de puissants moteurs de recherche afin que chaque fonction de l'organisme puisse exploiter efficacement les multiples ressources mises à sa disposition. Cette catégorie de GED a pour objectif la diffusion à distance des documents, et recouvre des applications variées : la fourniture

électronique de documents, la consultation de fonds documentaires numérisés, les dossiers documentaires (accessibles à distance et de manière partagée), les banques d'images, etc.

**SAE (Système d'Archivage Electronique)/ERMS (Electronic Records Management System) :** Système de gestion des documents numériques qui s'applique aux archives intermédiaires et définitives. L'objectif d'un SAE est de garantir la pérennité et la fiabilité des informations archivées.

**Thésaurus- thesaurus :** « Liste organisée de termes normalisés (descripteurs et non-descripteurs) servant à l'indexation de documents et des questions dans un système documentaire ». (Vocabulaire de la documentation, 3e édition, ADBS Editions, 2004).

### **Format des documents électroniques**

Come défini précédemment, un document numérique est un ensemble de données organisées selon une structure stable associée à des règles de mise en forme permettant une lisibilité partagée entre son concepteur et ses lecteurs. Ces règles sont régies par des formats de fichiers. La lisibilité des formats de fichiers est liée au matériel physique, au système d'exploitation et au logiciel. Il est possible de distinguer les types de format que l'on peut rencontrer du plus « ouvert » au plus « fermé » :

- **Standard :** lorsqu'ils sont produits par un organisme de normalisation et que leur utilisation est libre.
- **Ouvert :** quand les spécifications sont libres.
- **Propriétaire :** lorsqu'ils sont définis par une entreprise privée et que leur utilisation est soumise à des droits.
- **Fermé :** quand les spécifications sont tenues secrètes par leur propriétaire.

Les formats fichier peuvent, pour différentes raisons, disparaître au cours du cycle d'archivage du document. Pour cette raison, il est nécessaire de pouvoir accéder aux spécifications dudit format et ainsi pouvoir écrire au besoin un programme destiné à l'interpréter. Ce sont les raisons pour lesquelles, dans le cadre d'archivage, il est fortement recommandé d'utiliser des formats standards ou ouverts tel que XML ou PDD/A.

### **XML (*eXtensible Markup Language*)**

XML est un langage informatique de balisage générique. Il sert essentiellement à stocker/transférer des données au format texte Unicode8 structurées en champs arborescents, Sa structure arborescente permet de modéliser la majorité des problèmes informatiques.

## **PDF:**

Le PDF est un langage de représentation de documents, riche en fonctionnalités qui n'a cessé de s'étoffer au travers de ses versions successives. Les documents PDFs peuvent être simples ou complexes, structurés ou non structurés. Ils peuvent contenir du texte, des images, des graphiques et même des contenus multimédia (documents sonores ou images animées). Les annotations, les métadonnées, les liens hypertextes, les signets ainsi que la signature électronique sont également supportés par ce format de fichier.

## **PDF/A**

Le format PDF/A est défini par la norme ISO 19005. Il s'agit d'un format adapté pour la conservation permanente et la diffusion de documents. Il préserve la mise en page, les polices de caractères et la mise en forme. De plus, il s'agit d'un standard ouvert et libre de droits. La raison d'être du PDF/A est d'assurer la conservation à long terme de documents.

## **Le JPEG (Joint Photographic Experts Group)**

*Les images JPEG* ont l'extension ".jpg", ".jpeg", ".jpe" ou ".jfif", les quatre notations sont justes. Le mot **JPEG** veut dire **Joint Photographic Expert Group** qui représente le nom d'un groupe d'experts qui s'est chargé de créer un format de compression pour les images photographiques. **Le format d'images JPEG** est très adapté aux photographies ou images volumineuses qui contiennent des millions de couleurs telles qu'un fond d'écran, *capture d'écran vidéo* etc.

## **TIFF (Tagged Image File Format)**

Convient pour: données d'image (données d'images raster). Le format TIFF sauvegarde des images raster (graphisme pixel ou bitmap), une succession de pixels sous forme de tableau (éléments d'image), auxquelles sont attribuées des couleurs.

## **Groupware**

Ensemble d'outils logiciels permettant à des utilisateurs travaillant ensemble de s'organiser à travers le réseau informatique, et facilitant ainsi le travail collaboratif à distance.

## **LDAP (Lightweight Directory Access Protocol)**

Protocole d'accès aux annuaires légers. Ce protocole standard permet la gestion (interrogation, modification, etc) des annuaires par l'intermédiaire des protocoles TCP/IP.

## **Le scanner «numériseur»:**

Les Scanners sont des périphériques d'acquisition des documents qui permettant la transformation des documents sous forme d'image numérique.

On distingue généralement trois catégories de scanners :

- **Les scanners à plats** permettant de numériser un document en le plaçant à plat contre une vitre. Il s'agit du type de scanner le plus courant.
- **Les scanners à main** possédant une taille réduite. Ces scanners doivent être déplacés manuellement (ou semi manuellement) sur le document, par bandes successives, afin de le numériser en entier.
- **Les scanners par défilement** faisant défiler le document devant une fente lumineuse fixe afin de le numériser, à la manière des télécopieurs (fax). Ce type de scanner est de plus en plus intégré dans des appareils tels que des imprimantes multifonctions.

## Caractéristiques d'un scanner

Un scanner est généralement caractérisé par les éléments suivants :

- **La résolution** d'une image est le nombre de pixels par unité de mesure, la plupart du temps exprimée en points par pouce (ou dpi en anglais, dots per inch). Une résolution élevée est l'un des éléments garantissant une reproduction fidèle de l'original et une impression de qualité .
- **Le format de document** : selon leur taille, les scanners sont capables d'accueillir différentes taille de documents, généralement A4 (21 x 29,7 cm), plus rarement A3 (29,7 x 42 cm).
- **Vitesse d'acquisition** : exprimée en *pages par minute (ppm)*, la vitesse d'acquisition représente la capacité du scanner à acquérir un grand nombre de pages par minute. La vitesse d'acquisition dépend du format du document ainsi que de la résolution choisie pour la numérisation.
- **Caractéristiques physiques** : d'autres éléments peuvent être pris en compte lors du choix d'un scanner :
  - L'encombrement, correspondant aux dimensions physiques du scanner.
  - Le poids.
  - La consommation électrique, exprimée en Watts (W).
  - Les températures de fonctionnement et de stockage.
  - Le niveau sonore. Un scanner peut s'avérer être très bruyant, ce qui peut constituer une nuisance non négligeable.
  - Les accessoires : Les pilotes et le manuel d'utilisation sont habituellement fourni, mais il faut s'assurer que les câbles de raccordement le sont également ou, dans le cas contraire, les acheter à part.

## Résumé

A l'ère du numérique et face au nombre croissant d'outils et de théories managériales diverses, la gestion documentaire est au cœur des problématiques soulevées par le monde de l'entreprise aujourd'hui. Afin d'améliorer la qualité de service et de simplifier les procédures du travail, les entreprises, les associations, les administrations et même les particuliers s'orientent de plus en plus vers les nouvelles technologies de l'information. A travers un cas concret, notre travail cherche à étudier un ensemble de solutions de gestion électronique des documents (GED) afin de sélectionner une solution adaptée pour la mettre en place dans une PME Algérienne. Nous avons pris l'établissement administratif de la Caisse Nationale de la Retraite (CNR) de Ouargla comme exemple. Cette étude a permis de conclure que « Alfresco » est la solution la plus adaptée et la plus accessible. L'intégration d'Alfresco au CNR a nécessité des configurations, des intégrations indispensables et le paramétrage de certaines fonctionnalités d'Alfresco. Pour se faire, un ensemble de technologies ont été utilisés, en l'occurrence : JEE, Tomcat, ...etc. Une description détaillée de ce travail est présentée dans ce mémoire.

**Mots Clés:** GED, Archivage, Stockage, local, Cloud

## ملخص

في العصر الرقمي ومع العدد المتزايد من الوسائل والنظريات الإدارية المختلفة، لقد أصبحت إدارة الوثائق في صميم القضايا التي أثارها عالم الشركات اليوم. ولتحسين نوعية الخدمة وتبسيط إجراءات العمل اتجهت الشركات والجمعيات والحكومات بل حتى الأفراد نحو التكنولوجيات الجديدة للمعلوماتية.

يهدف عملنا من خلال دراسة تطبيقية ملموسة لدراسة مجموعة من الحلول لإدارة الوثائق الإلكتروني GED. لتحديد حل مناسب لتطبيقه في PME (الشركات الصغيرة والمتوسطة) جزائرية، اخترنا مؤسسة إدارية متمثلة في الصندوق الوطني للتقاعد CNR كمثال تطبيقي. خلصت هذه الدراسة إلى أن « Alfresco » هو الحل الأنسب. لقد تطلب دمج « Alfresco » في CNR تهيئة وادماجات ضرورية و ضبط وظائف « Alfresco ». للقيام بذلك استخدمت مجموعة من التقنيات وهي JEE و Tomcat... الخ. وصف مفصل لهذا العمل في هذه المذكرة.

**الكلمات المفتاحية :** التسيير الرقمي للوثائق GED، أرشفه، تخزين، المحلية، السحابة

## Abstract

In the digital age and with the growing number of tools and managerial theories, Content Management is at the heart of the issues raised by the world of today's company. To improve the quality of service and to simplify work procedures, companies, associations, administrations and even individuals are moving more and more towards the new information technologies.

Through a concrete case, our work aims to study a set of solutions for electronic content management (ECM) to choose an adapted solution for setting up in an Algerian small company. We took the administrative establishment CNR -National Pension Fund- as an example. This study concluded that "Alfresco" is the best and most affordable solution. Integrating Alfresco in CNR required configurations, indispensable integrations and setting certain functions of Alfresco. To do so, a set of technologies were used, like: JEE, Tomcat ... etc. A detailed description of this work is presented in this report.

**Key words:** ECM, RM, back up, local, Cloud