

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE MOULOUD MAMMARI DE TIZI-OUZOU



FACULTE DU GENIE ELECTRIQUE ET D'INFORMATIQUE
DEPARTEMENT D'INFORMATIQUE

Mémoire de Fin d'Etudes de MASTER ACADEMIQUE

Domaine : **Mathématiques et Informatique**

Filière : **Informatique**

Spécialité : **Systemes informatiques**

Présenté par

IRNATENE Ouiza

AISSIOUENE Nabila

Thème

**Conception d'une plateforme de mise en
ligne de manuscrits arabes anciens
numérisées.**

Mémoire soutenu publiquement le 17/07/2016 devant le jury composé de :

Président : Mr HABET Mohand Said

Encadreur : Mr SOUALAH Mohand Ourabah

Examineur : Mr KERBICH Mohand

Examineur : Mme YESLI Yassmine

Remerciements

*Nous tenons à remercier et à exprimer notre
profonde*

*Gratitude à notre promoteur pour nous avoir
proposé ce thème et aidé au long de son
développement.*

*Comme nous remercions les membres de
jury d'avoir accepté de juger ce modeste
travail ainsi que toute personne ayant aidée
de près ou de loin.*

Dédicaces

Je dédie ce travail :

A mes parents :

Je vous remercie pour tout le soutien et l'amour que vous me portez depuis mon enfance et j'espère que votre bénédiction m'accompagne toujours.

Que ce modeste travail soit l'exaucement de vos vœux tant formulés, le fruit de vos innombrables sacrifices, bien que je ne vous en acquitterai jamais assez.

Puis Dieu, le très Haut, vous accorder santé, bonheur et longue vie.

A mes chers frères :

Ferhat, mahrez, massi et amirouche et surtout a toi mahrez je vous remercie pour ton soutien.

Je dédie ce travail a ma sœur zohra et bien sur a louiz aussi

A mes amis

A mes belles copines Lydi, Maya, Lydia, Nawel.

A Imade que je remercie pour son encouragement.

Nabila

Sommaire

Sommaire

Introduction générale

Chapitre I : présentation des manuscrits arabes anciens

I.1 Introduction.....	1
I.2 Définition d'un manuscrit.....	1
I.3 Caractéristiques des manuscrits arabes anciens.....	1
I.4 Disciplines traités par les manuscrits arabes.....	2
I.5 Etude descriptive des manuscrits.....	3
I.5.1 Les champs bibliographiques.....	3
I.5.2 Eléments descriptif identifiés par la codicologie.....	4
I.5.3.Eléments descriptifs identifiés par la paléographie.....	5
I.6 Protocole de description des manuscrits arabes.....	5
I.6.1 Le volume.....	6
I.6.2 Unité codicologique.....	7
I.6.3 Le texte.....	8
I.6.4 Les notes.....	9
I.7 Classification des manuscrits.....	9
I.7.1 L'aspect appartenance religieuse.....	10
I.7.2 L'aspect « nature ».....	10
I.8 La numérisation des manuscrits arabes anciens.....	11
I.8.1 Définition de la numérisation.....	11
I.8.2 Objectifs de la numérisation.....	12
I.8.3 Les mode de numérisation.....	13
I.8.4 Les étapes de la numérisation.....	13
Conclusion.....	14

Chapitre II : Le catalogage des manuscrits arabes anciens

II.1 Introduction.....	15
II.2 Généralités sur le catalogage.....	15
II.3 Le catalogage des manuscrits arabes.....	15
II.3.1L'école arabe.....	15
II.3.2 L'école orientaliste.....	16
II.4 Les métadonnées.....	16
II.4.1 Définition des métadonnées.....	16

II.4.2	Caractéristiques des métadonnées	17
II.3.3	Les métadonnées définies dans quelques bibliothèques arabes.....	17
II.3.4	Les métadonnées définies par l'IRHT	18
II.4.5	Les métadonnées définies par EL Saoud.....	19
II.4.6	Le nouveau protocole de catalogage	21
II.5	Encodage des métadonnées décrivant les manuscrits arabes anciens	24
II.5.1	Définition du standard	24
II.5.2	Formats d'encodage de métadonnées	24
II.5.2.1	Le format MARC	24
II.5.2.1.1	Définition.....	24
II.5.2.1.2	MARC 21	25
II.5.2.1.3	L'UNIMARC.....	26
II.5.2.2	L'EAD(Encoded Archival Description).....	27
II.5.2.2.1	Définition.....	27
II.5.2.2.2	La structure de l'EAD.....	28
II.5.2.3	Le projet MASTER	30
II.5.2.3.1	L'encodage des manuscrits dans le projet MASTER.....	30
II.5.2.4	Le format METS(Metadata Encoding and Transmission Standard)	31
II.5.2.4.1	Définition.....	31
II.5.2.4.2	La structure d'un documentMETS	31
II.5.2.4.3	Les éléments de METS.....	32
II.5.2.4.4	Quelques grands principes.....	33
II.5.2.4.5	Les pointeurs de METS	34
II.5.2.5	Le Dublin Core	34
II.5.2.5.1	Qu'est-ce que le Dublin Core?	34
II.5.2.5.2	Les quinze éléments du Dublin Core	34
II.5.2.6	La TEI (Text Encoding Initiative) :	36
II.5.2.6.1	Definition.....	36
II.5.2.6.2	Caractéristiques de la TEI	36
II.5.2.6.3	La TEI et la description des manuscrits	36
II.5.2.6.3	Principe de description des manuscrits.....	36
II.5.2.6.3.2	Les éléments de la TEI	37
II.5.2.6.3.3	Les élément de niveau de phrase	39
	Conclusion.....	41

Chapitre III : La technologie Java EE

III.1 Introduction	42
III.2 Définition	42
III.3 Objectif de Java EE.....	42
III.4 Fonctionnement interne.....	43
III.5 Les éléments de Java EE	43
III.6 L'architecture de Java EE	43
III.7 Outils de programmation.....	45
III.8 Les conteneurs.....	45
III.9 Le conteneur web	45
III.10 Le conteneur d'EJB.....	46
III.11 Les API de Java EE.....	46
III.10 Les services JEE.....	47
III.11 Java EE et la sécurité.....	48
III.12 les avantages d'utilisation Java EE	48
Conclusion.....	48

Chapitre IV : Analyse et Conception

IV.1 Introduction.....	49
IV.2 Objectifs de notre travaille.....	49
IV.4 Analyse et Conception	49
IV.4.1 Analyse.....	49
IV.4.1.1 Spécification des besions	49
IV.4.1.2 Identification des acteurs	50
IV.4.1.3 Spécification des taches.	50
IV.4.1.4 Spécification des scénarios	52
IV.4.1.5 Diagramme représentatif.....	55
IV.4.1.6 Les diagrammes de cas d'utilisation	50
IV.5 Conception	60
IV.5.1 Elaboration des diagrammes de séquences	60
IV.5.2 Elaboration des diagrammes d'activités	66
IV.5.3 Elaboration des diagrammes de classe.....	69
IV.5.4 Conception de modèle de donnée	70
Conclusion.....	77

Chapitre V : Implémentation et réalisation

V.1	Introduction	78
V.2	Chiox technologique.....	78
V.2.1	L'architecture MVC	78
V.3	Présentation des langages de programmation	79
V.3.1	Le langage Java EE	79
V.3.2	Les JSP	80
V.3.3	Les Servlets	80
V.3.4	Le langage JavaScript	80
V.3.5	Le langage SQL.....	81
V.3.6	Le langage HTML	81
V.4	Les outils de développement	81
V.4.1	NetBeans version 8.1.....	82
V.4.2	WampServer.....	82
V.4.3	Le serveur Apache.....	82
V.4.4	Apache Tomcat.....	84
V.4.5	L'interface graphyque phpMyAdmin.....	84
V.4.6	Serveur MySql.....	85
V.5	présentation de quelque interfaces de notre application.....	86
	Conclusion.....	89
	Conclusion générale	90
	Annexe	91
	bibliographie.....	96

Liste des Figures

Figure I.1 : Les manuscrits arabes anciens.....	1
Figure I.2 :Le colophon en forme triangulaire	4
Figure I.3 : les différentes décore d'un manuscrits arabes anciens.....	5
Figure I.4 :Décore de section	6
Figure II.1 :Structure générale d'un outil de recherche encodé en EAD	28
Figure II.2 :Structure de l'élément <archdesc>	29
Figure II.3: Structure de l'élément <did>	30
Figure II.4 :La structure d'un document	32
Figure II.5 :Les éléments de métadonnées Dublin.....	34
Figure III.1:L'architecture de Java EE.....	44
Figure IV.7:Diagramme de contexte	55
Figure IV.8 :Diagramme de cas d'utilisation général	57
Figure IV.9 :Diagramme de cas d'utilisation « administrateur»	58
Figure IV.10 :Diagramme de cas d'utilisation « Catalogueur»	58
Figure IV.11 :Diagramme de cas d'utilisation «recherche adhérent»	59
Figure IV.12 :Diagramme de cas d'utilisation «Visiteur»	59
Figure IV.13 :diagramme de séquence du cas d'utilisation «Authentification».....	61
Figure IV.14 :diagramme de séquence du cas d'utilisation «Ajouter une notice»	62
Figure IV.15 :diagramme de séquence du cas d'utilisation «modifier le mot de passe».....	63
Figure IV.16 :diagramme de séquence du cas d'utilisation «Inscription».....	64
Figure IV.17 :diagramme de séquence du cas d'utilisation « Recherche un manuscrit ».....	65
Figure IV.18 :Diagramme d'activité de cas d'utilisation « authentification ».....	66
Figure IV.19 :Diagramme d'activité de cas d'utilisation « Modification de mot de passe »..	67
Figure IV.20 :Diagramme d'activité de cas d'utilisation « rechercher un manuscrit »	67
Figure IV.21 :Diagramme d'activité de cas d'utilisation « Gestion des adhérents ».....	68
Figure IV.22 :Diagramme d'activité de cas d'utilisation « Gestion des notices ».....	68
Figure IV.23 :Diagramme de classe générale.	69
Figure V.1:Interface NetBeans version 8.1.....	82
Figure V.2:Panneau de gestion de serveur.....	83
Figure V.3:Interface phpMyAdmin.....	85
Figure V.4 :Interface page d'accueil.....	86
Figure V.5:Interface administrateur.....	87

Figure V.6:Interface d'ajout d'un catalogueur.....	87
Figure V.7:Interface phpMyAdmin.....	87
Figure V.8:Insertion d'une notice.	88
Figure V.9:Biote de reception.	89

Liste des tableaux

Tableau II.1 Protocole de description des manuscrits arabes.....	24
Tableau III.1 : API JEE.....	47
Tableau IV.1 : tableau des cas d'utilisation	51
Tableau IV.2 : Spécification des scénarios de « Visiteur ».....	52
Tableau IV.3 : Spécification des scénarios de « Adhérent ».....	52
Tableau IV.4 : Spécification des scénarios de « Catalogueur ».....	53
Tableau IV.5 : Spécification des scénarios de « Administrateur »	54
Tableau IV.6 : Spécification des scénarios de « Responsable d'orientation »	54
Tableau IV.3 : Spécification des scénarios de « Adhérent ».....	52
Tableau IV.4 : Spécification des scénarios de « Catalogueur ».....	53

Introduction Générale

Introduction générale

Un nombre de documents hérités de nos parents et ancêtres sont conservés dans musées, bibliothèques et archives, parmi ces documents nous citons les manuscrits. Aujourd'hui ces derniers ils ont profité de cette révolution scientifique, surtout dans leurs conservations.

Le manuscrit est un document ancien, rare et précieux qui a une valeur historique indéniable, ainsi nous pouvons confronter notre passé avec leur présent et remettre en cause notre connaissance de l'histoire. Alors que les documents parmi les plus fragiles nécessite d'être conservés dans des conditions de température et d'humidité particuliers afin de les préservé aux générations future.

Les manipulations dues à la lecture détériorent inévitablement les documents. Malgré les précautions prises, les supports sont fragiles et souffrent d'être trop souvent maniés, ce qui impose leur conservation dans des bonnes conditions.

En conséquence, les moyens de conservation du manuscrit ont pris une autre dimension, notamment grâce à la naissance et la mise en service de la technique de numérisation. Avec ce dernier, de nouveaux moyens ont rendu plus facile et plus rapide la sauvegarde, la préservation et la diffusion des informations contenues dans ces documents.

L'objectif de notre projet, travail sur la numérisation des manuscrits arabes anciens pour donner un accès à distance de ce patrimoine à une échelle internationale et de facilité la possibilité de la recherche d'information d'une manière automatique.

Pour ce faire, nous avons organisé notre mémoire en cinq chapitres :

- Le premier chapitre porte sur la présentation des manuscrits arabes anciens.
- Le deuxième chapitre est abordé sur le catalogage des manuscrits arabes anciens. Il débute par une présentation des différents types du catalogage des manuscrits arabes anciens dans plusieurs bibliothèques et institut. Il se poursuit par une présentation de métadonnées. La suite de chapitre décrit les formats d'encodage de métadonnées.
- Le troisième chapitre est consacré pour la présentation de langage de programmation java EE.

- Le quatrième chapitre porte sur les différentes étapes de l'analyse et conception de notre application.
- Le cinquième chapitre est consacré à l'implémentation et la réalisation du notre projet.

Ce travail se termine par une conclusion générale, ainsi qu'une partie d'annexe.

Chapitre I

Présentation des manuscrits arabes anciens

I.1 Introduction :

Les manuscrits arabes sont des documents rare et fragile constituant une source d'information inestimable. Le traitement automatique de ces documents en vue de leur restauration, indexation et exploitation offre un avantage certain. Cependant, elles sont confronte à de nombreuse difficulté due au mauvaise état de conservation. Donc il doit les préservé au longtemps possible afin d'être consulté par la génération future.

I.2 Définition d'un manuscrite :

Un manuscrite en latin manu scriptum. Est, littéralement, un texte écrit à la main sur un support que ce soit par son auteur (manuscrite autographe) ou par un copiste avant la découverte de l'imprimerie.

Chaque manuscrit est un témoin unique. Produit dans des conditions artisanales a l'aide de matériaux rares et couteux qu'il est possible d'étudier de différentes façons.

Les manuscrits sont d'abord des livres qui contiennent une ou plusieurs ouvres intéressantes par leur contenu. Leur copie, leur diffusion, leur appartenance à une bibliothèque ou à une institution témoignent aussi de l'état des connaissances et de la circulation du savoir à une époque et en un lieu donnés.

En outre, le manuscrit est un document archéologique, un objet qui peut être étudié sous l'angle de sa production artisanale, de ses caractéristiques matérielles et de son décor. Ceci ne concerne pas les seuls manuscrits arabes, mais tous les manuscrits quelle que soit l'aire culturelle à laquelle ils appartiennent [1].



Figure I.1 : exemple des manuscrits arabes anciens

I.3 Caractéristique des manuscrits arabes anciens :

Les manuscrits arabes, comme les manuscrits de différentes langues, ont des caractéristiques communes mais chacun de ces manuscrits a sa propre identité.

Parfois le même élément qui est en commun avec les autres diffère dans son contenu. Ainsi, les informations mentionnées dans un colophon sont différentes entre les manuscrits arabo-islamiques, les manuscrits arabo-chrétiens et les manuscrits en langue latine. En plus, les manuscrits arabes diffèrent selon leurs régions (manuscrits orientaux ou maghrébins), l'appartenance religieuse de la communauté qui l'on a produits (manuscrits arabo-islamiques et arabo-chrétiens) et la période de leur achèvement (création).

Les caractéristiques communes aux manuscrits arabes peuvent être résumées comme suit [2]:

- début des manuscrits au verso du premier feuillet, alors que le recto est réservé à l'inscription du nom de l'auteur, au commanditaire de l'œuvre et parfois au cachet.
- le début de texte peut être accompagné d'un décor particulier et représente souvent, le début de chaque chapitre, section ou sourate quand il s'agit d'une œuvre coranique.
- le texte est écrit en longues lignes, à l'exception du texte poétique.
- des règles d'usage de l'encre sont observées : la couleur rouge est souvent utilisée pour l'écriture des noms propre, des nombre et des citations.
- l'usage du texte encadré dans les manuscrits coraniques.
- présence du texte dans les marges.

I.4 Les disciplines traitées par les manuscrits arabes :

Les manuscrits traitent en générale les domaines suivants [3]:

- l'histoire (*Tarik*) en particulier celle de l'époque précoloniale.
- la théologie musulmane (*Tawhidi*).
- l'astrologie (*ilm al-falak*).
- la littérature arabe (*al-adab al- arabi*).
- la poésie religieuse en langues africaines et arabe (*unshudadiniyya*).
- des traités en sciences occultes (*sirr*).
- la dissertation en droit (*tahrir fi 'l-fiqh*).
- la mystique musulmane (*tasawwuf*).
- la médecine locale (*al-ḥibbal-mahalli*).
- la pharmacopée (*saydala*).

- des traités politiques (*siyasa*).

Par ailleurs, nombreux sont les manuscrits généraux qui traitent des thèmes très divers, ce qui rend difficile leur classification thématique.

I.5 Etude descriptive des manuscrits :

Le manuscrit est un témoin non seulement d'un savoir faire et d'une culture médiévale, mais aussi d'un héritage antique possédant sa propre histoire, sa propre archéologie. Il fait partie de divers ensemble et contexte intellectuels et matériels qui constituent l'histoire du savoir et de transmission, alors, le manuscrit est loin d'être coursier passif de l'information.

Le manuscrit est un objet qui peut être approché selon divers aspects :

- Aspect codicologiques (formes et support, encre, cahier, mise en page, reliure,.....),
- Aspect paléographique (étude des écritures et décoratifs),
- Histoire de manuscrit.

Autant d'éléments qui permettant bien souvent de dater ou de localiser un document anonyme ou mal identifié.

I.5.1 Les champs bibliographiques :

L'ensemble des éléments descriptifs d'un manuscrit est désigné par « champ bibliographique ».

L'identité de ces éléments sera très utile pour la création d'un format de description électronique, propre aux manuscrits arabes. Parmi les éléments d'identification nous retrouvons principalement [3]:

A) la mention de responsabilité :

L'écriture manuelle du manuscrit fait apparaître une nouvelle mention de responsabilité qui est celle du "copiste", pouvant être mentionnée à la suite du nom de l'auteur. Le nom de l'auteur, ainsi que le nom du copiste, figurent parfois au recto du 1er feuillet des manuscrits arabes de luxe, quelquefois entourés par des décors. Dans les manuscrits « ordinaires », le nom de l'auteur figure au début du texte et/ou sur la page de titre, celui du copiste au colophon.

B) le nom du propriétaire :

Le nom de propriétaire ou possesseur du manuscrit est également un indice important pour les chercheurs désirant suivre le développement historique du manuscrit.

Le propriétaire peut être une personne, une bibliothèque ou une institution. Son nom figure soit écrit au 1er recto, soit sous forme d'un cachet ou d'un médaillon ou dans le colophon ou encore sur une page de garde.

C) le titre du manuscrit : Le titre de l'ouvrage ou le titre de la série est l'élément principal d'identification qui dans la majorité des cas, est présenté en page de titre. Mais dans les manuscrits, il y a aussi "l'incipit" (c'est à dire la première phrase du texte) qui est un élément de recherche et d'identification très important, ainsi que "l'explicit" (la dernière phrase du texte manuscrit). Normalement, le titre du manuscrit se trouve accompagné du nom de l'auteur inscrit sur le recto de la 1ère page. Mais, le titre et le nom de l'auteur peuvent aussi apparaître tous deux dans le colophon. Souvent l'auteur indique dans le texte le titre qu'il a donné à son œuvre.

I.5.2 Eléments descriptifs identifiés par la codicologie :

A) Le colophon : Le colophon est placé à la fin d'un texte sur un livre ou un manuscrit. Il contient des indications sur le nom du scribe ou du copiste, ainsi que la date à laquelle la copie a été faite et autres renseignements sur le paratexte. Parfois il indique aussi le nombre d'exemplaires imprimés et le temps que le copiste a passé à copier l'ouvrage. Le colophon prend souvent la forme d'un triangle.

Les informations contenues dans le colophon sont presque les mêmes à celle qu'on trouve dans les mentions d'éditions des ouvrages imprimés (la date, le titre, lieu d'éditions, nom éditeur) .



Figure I.2 : Exemple de colophon en forme triangulaire

B) La reliure : Ensemble des opérations qui donnent à un ouvrage sa forme définitive et par les quelles on l'habille d'une couverture rigide ou souple pour le protéger et le décorer . La reliure participe à l'histoire des techniques, à l'histoire du livre, à l'histoire des bibliothèques. Enfin, dater la constitution d'un recueil peut être de la plus haute importance pour comprendre

la tradition manuscrite d'un texte. La reliure utilise au moins deux pièces de cuivre, une pour couvrir la planche et le dos du livre et l'autre qui, collée en bas du côté intérieur de la planche, constitue une couverture qui protège les côtés du bloc de texte [2].

La reliure des manuscrits arabes se caractérise par la présence d'un recouvrement, pièce de cuir qui recouvre la tranche de gouttière et vient s'insérer sous la plate supérieure.

C) Décore : Le décor d'un manuscrit sert en premier lieu à articuler le texte et permet au lecteur de se repérer. Au-delà, il joue d'innombrables rôles : didactique, symbolique, historique et purement esthétique. Le décor possède une histoire propre de ses formes et de ses significations. Que ce soit sous forme peinte ou dessinée, il apporte des indices de datation et de localisation du manuscrit [4].



Figure I.3 : exemple sur les différentes décorations d'un manuscrit arabe ancien

Il existe plusieurs types de décor [2]:

1- Décor de la reliure : La reliure des ouvrages est parfois décorée par des doublures et des miniatures spéciales. Le motif le plus utilisé dans les manuscrits anciens, et que l'on trouve sur la reliure des manuscrits de Sana au Yémen et de Kairouan, est le panneau central.

2- Décor de texte : Cette partie de la description est importante pour les chercheurs qui s'intéressent à l'étude esthétique des manuscrits. L'enluminure est considérée comme une des gloires de l'art islamique. Autant que la calligraphie, l'enluminure est considérée par les arabo-musulmans comme un moyen d'exprimer leurs talents artistiques. Les principaux types d'enluminure rencontrés dans les manuscrits islamiques sont :

- ✚ Le décor en pleine-page.
- ✚ Les décors de section.
- ✚ Les colophonnés.
- ✚ Les signes de division du texte.
- ✚ Les encadrements.
- ✚ Le décor des reliures, et des doublures.



Figure I.4 : Décore de section

D) Organisation : un manuscrit peut être composé de plusieurs cahier de parchemin ou de papier, on parle de volume importance de l'observation du volume permet de déterminer si le manuscrit est homogène ou composite, si les cahiers sont reliés en ordre ou dans le désordre. Les signes de numérotation et de renvoi utilisé dans les manuscrits permettent de se repérer dans le manuscrit [5]. Parmi les styles de numérotation nous distingue la foliotation, la signature et la réclame [6] :

- ❖ **La foliotation :** Numérotation des feuilles d'un manuscrit ou d'un registre.
- ❖ **La signature :** chiffres, lettres ou signes utilisés pour indiquer le numéro d'ordre des cahiers ou des feuillets dans les manuscrits.
- ❖ **La réclame :** la réclame correspond au premier mot de chaque cahier annoncé à la fin du cahier précédent. Elle permet au couseur et à la relieur de vérifier la bonne succession des feuillets et des cahiers.

E) Les genres et les couleurs d'encre :

Peu d'études ont été faites sur l'utilisation de l'encre dans les manuscrits arabes, mais on peut classer les couleurs utilisées en deux grands ensembles:

- 1- L'encre brune : utilisée prioritairement dans les occidents musulmans à l'époque ancienne.
- 2- L'encre noire largement utilisée dans toutes les régions.

F) Type du support : Plusieurs type de support ont été utilisés pour rédiger le manuscrit arabe : le papyrus, le parchemin et le papier.

I.5.3 Eléments descriptifs identifiés par la paléographie :

La paléographie prend en charge l'étude des écritures anciennes et trace les règles à suivre dans la lecture de celles-ci. Son domaine, en principe, peut s'étendre à tous les genres de documents, à tous les pays, à tous les temps.

I.6 Protocole de description de manuscrits arabes :

Le protocole de description de manuscrits arabes est devisé en quatre volets : étude de volume, étude de codicologique, étude du texte, et les éléments variables [3].

I.6.1 Le volume

Un manuscrit est une œuvre qui peut être formée de plusieurs volumes. Chaque volume est décrit comme suit :

A- la cote : La cote d un manuscrit est bien souvent représenté par un numéro d'ordre séquentiel. Toutefois, une codification particulière est préposée par le protocole de L'IRHT. il est définit par : Ville-Dépôt- [fonds] - [langue/matière] - N° d'ordre et éventuellement la ou les cotes antérieures.

B- libellé générale : Le libellé d un manuscrit est définit par :

- ❖ **La matière :** Décrit la matière de base de fabrication d un manuscrit.
- ❖ **La détention des pages :** La mesure des pages est généralement le millimètre.
- ❖ **Nombre total de folios :** Le papier étant un produit rare, les auteurs de transcrivaient sur les deux faces d une feuille, on parle alors de folio. Le protocole prévoit d indiquer les billets ou les feuilles volantes.

C- Numérotation : la numérotation est représentée en générale sous forme de réclame. Par ailleurs, on retrouve parfois le foliotage (pagination) réalisé par le chiffre indien, arabes voire en bagad et description de la liaison vers la copie microfilmée.

D- Reliure : La reliure est partie intégrante du codex comme objet archéologique ce titre elle est définie par :

- Le Type (à rabat ou sans rabat).
- datation par siècle.
- Nature de la reliure (Originale ou non).
- Dimension des plats en millimètre.
- épaisseur du volume fermé.
- Ais (bois, carton).
- Couvrure : matière (cuir, étoffe, papier marbré, etc.), et couleur.
- décoration (estampage des plats et contre plats, dans les différentes zones).
- Etat de conservation. observations diverses.

I.6.2 Unité codicologique :

La description de l'unité codicologique se résume à la description du papier des cahiers, de l'écriture et éventuellement de la copie.

A- Papier :

- Type de papier (Filigrané ou non et Description des filigranes).
- Datation de papier.
- Mesure de la largeur de vingt(20) vergeures.
- Orientation des vergeures : pliage, format de la feuille et origine du papier.
- Disposition (simple ou groupée) et écartement des fils de chaînette.

B- Cahiers : Le cahier est un groupe de feuillets obtenus par pliage d'une feuille de parchemin ou de papier ou par la réunion de plusieurs pliées. Les cahiers sont décrits par les éléments suivants :

- Nature des cahiers (binions, ternions, etc.)
- Collation des cahiers (nombre de bifeuill et du cahier, nombre de cahiers, somme des folios).
- Anomalie de composition (localiser les feuillets manquants, feuillets ajoutés, l'ordre de lecture.
- Numéros de cahier.
- Marques de milieu de cahiers.
- La réclame: le protocole définit la réclame par sa présence ou son absence dans le cahier et le lieu de présence (dans tous les feuillets, ou en fin de cahier, au milieu,.....etc.).

C- Préparation de la page avant l'écriture : Avant la transcription sur le manuscrit, une préparation préliminaire et adéquate est réalisée. La préparation consiste à l'introduction de décor ou de tracés de lignes. Elle est décrite par :

- La surface écrite.
- Instrument de réglure (pointe sèche, *mistara*, ou mixte).
- Indication et description des piqûres.

D- Écriture : La calligraphie se présente comme un art dans la civilisation arabe. En effet, si les dessins et les gravures se retrouvent plus au moins réprimées, la calligraphie prend une dimension plus importante. Ainsi, plusieurs styles calligraphiques se distinguent les uns des autres par la forme et le caractère. L'écriture se trouve décrite par l'élément suivant :

- Style calligraphique : Le degré de cursivité, la Couleur d'encre, la taille, la Particularités d'écriture, la proportion des lettres, la ligature particulière et l'inclinaison, la taille de calame, etc.
- Description des usages orthographiques : cette élément est décrit par la présence ou l'absence des Signes de vocalisation et *sukūn*, la présence ou l'absence de la Ponctuation des lettres, Usage de *sadda*, règles suivies pour le siège de hamza (omission ou non de hamza Sur *-alifen* début de mot. *alif*, *wāw*, *nabra* à l'intérieur des mots). Points sous le *yāfinal*, Confusions emphatiques et inter dentales pouvant renseigner sur les origines régionales des copistes.
- Organisation du manuscrit : Description de l'articulation des textes et des chapitres entre eux et de la hiérarchie des subdivisions (titres, rubricages, sur lignements).
- Corrections de copiste : Raturage, grattage. Omissions, rectifications, collations, additions marginales et sur linéaires.
- Colophon : Transcription et Identification (lieux, noms, interprétation des dates conversion des dates).

I.6.3 Le texte :

Le protocole préconise une description du contenu textuel du manuscrit sans pour autant s'intéresser au contenu proprement dit. La description du texte se fait en générale en trois niveaux hiérarchique, que nous décrivons dans ce qui suit :

A- Numérotation du texte : La numérotation se fait selon l'ordre d'apparition dans le volume.

B- Intitulé : Dans l'intitulé, il s'agit de mentionner *le titre normalisé de l'œuvre*, la reconstitution totale ou partielle *du nom de l'auteur*. Les anépigraphe (absence de l'auteur et titre), ou de pseudépigraphe (l'intitulé, texte ou auteur, ne correspond pas à la vérité) doivent être indiqués.

L'intitulé sera décrit en mentionnant *les titres alternatifs*, la ou les *localisations du titre* dans le volume (page de titre, frontispice, incipit, colophon, tranche) et citer *la formulation*. Par exemple : *Titre au frontispice. Sur la tranche inférieure, intitulé: ma ād Ibn Sīnā.*

C- matière scientifique : Appelée encore genre, il s'agit de déterminer le thème sur lequel se porte le texte selon la division traditionnelle des sciences arabo-musulmanes (*fiqh, tafsir, adabkalam, médecine, astronomie, etc.*).

D- résumé : Le résumé contiendra une description contextuelle du volume, tout en notant les mots clés.

E- Références : Il s'agit de reporter les références des différentes éditions récentes et antérieures du texte.

F-contenu de texte : Un intérêt particulier est porté sur la nature de l'œuvre. En effet, en cas d'une copie, la description portera d'abord sur la collation avec l'édition originale (contrôle du nombre de ligne par page). On s'intéressera aussi aux divergences (omissions, additions de séquence, différences de divisions en chapitre). Dans le cas d'un texte inédit, on prévoit de reporter le sommaire du texte, relever les têtes de chapitres en respectant la terminologie (*bāb, faṣl, qawl*).

G- Incipit : Les trois séquences, basmala, hamdala et entrée en matière.

H- Explicit : Dernier mot de texte avant le colophon.

I- Auteur : Reporter les références bibliographiques de l'auteur, sa date de décès.

I.6.4 Notes :

Dans ce volet le protocole reprendra toutes les inscriptions hors copie au niveau codicologique ou du volume. L'hétérogénéité de la structure des manuscrits oblige les spécialistes à réserver un élément qui portera les spécificités non communes. On parle parfois, d'élément variable : c'est le rôle que joue l'élément "Notes". Il décrit par les éléments suivants :

A- Notes de lecture : Gloses, attestations de lecture, attestations de collation, certificats d'audition (*samā*) et de transmission (*iğāza*).

B-Notes diverses : Marque de possession, inscription de *dawaqf*, table des matières, vers, carrés magiques, formule propitiatoire, invocations, exercices d'écriture, essais de plume, prix de vente.

I.7 Classification des manuscrits arabes :

Etant témoins de leurs ères, il est donc intéressant de pouvoir ranger les manuscrits suivant des critères communs, et pour un meilleur classement on peut introduire deux aspects : la région, l'appartenance religieuse et nature.

I.7.1 L'aspect appartenance religieuse :

Le deuxième aspect est l'appartenance religieuse qui l'on a produit. Les manuscrits arabo-chrétiens, sont les premiers apparus, ensuite on trouve les manuscrits arabo-islamiques surtout à partir du 13^e siècle [5].

A- les manuscrits arabo-chrétiens :

Les Eglises chrétiennes du Proche et Moyen-Orient qui ont commencé à s'organiser, à partir du 3^e siècle, à la faveur de la christianisation du pourtour de la méditerranée.

C'est qu'à partir du septième siècle que les manuscrits chrétiens écrits en langue arabe sont apparus, à cause de l'influence des grecs et des byzantins sur les copistes chrétiens, ces derniers ont rédigé en différentes langues : grecque, le copte, etc.

B- les manuscrits arabo-islamiques:

Les premières enluminures de manuscrits islamiques datent du 11^e siècle, mais ce n'est qu'à la fin du 13^e siècle qu'ils ont connu une véritable progression.

La plus longue et la plus célèbre des dynasties musulmanes est les Abbassides, se sont les protecteurs des sciences et de la religion. Ils sont ainsi les premiers souverains musulmans à la tête d'une véritable civilisation islamique, ne contenant pas d'imposer la civilisation arabe aux pays conquis. Sous leur règne, Bagdad devient le centre de recherches théologique, industriel et commercial, l'Empire islamique atteint son apogée sur les plans économique et culturel grâce au niveau de développement acquis. La plupart des manuscrits islamiques retournent à cette époque.

I.7.2 L'aspect « nature » :

C'est un classement proposé par le docteur Fadhl Djamil Kolib et Fouad Mohammed Khali Abid dans leur livre :

A- Le manuscrit mère: C'est celui écrit par l'auteur lui-même. Cette catégorie comporte toutes les caractéristiques du manuscrit arabe. Les auteurs arabes mettaient leurs œuvres manuscrites (manuscrit mère) dans des endroits considérés comme centre d'études et de culture (DAAR EL Khalifa), enfin de le mettre des correcteurs pour la révision, des copistes pour la reproduction d'autres exemplaires de manuscrits, et en fin, pour les collations.

A- Le manuscrit rapporté : C'est une copie conforme au manuscrit mère, ils sont manipulés de la même façon.

B- Le manuscrit ambigu : C'est une sorte de manuscrit incomplet et désordonné, qui a des folios perdus, surtout les premières feuilles qui contiennent le nom de l'auteur et l'adresse, ou bien, il contient des parties répétées.

C- Le manuscrit par étape : L'auteur l'édite périodiquement, il produit sa première œuvre, il l'a diffusée, et à chaque nouvelle édition il ajoute de nouvelles connaissances.

D- Les manuscrits photographiés ou microfilmés : Dans plusieurs études concernant les manuscrits du point de vue artistique, on trouve un bon nombre de manuscrits numérisés ou microfilmés.

E- Les manuscrits sous forme de collections : Il y a beaucoup de manuscrits sous forme d'ensemble ou de collections qui contiennent souvent des écrits, des fragments ou des lettres.

I.8 La numérisation des manuscrits arabes :

Le passage au numérique est devenu une priorité et souvent même une nécessité dans le paysage actuel de la recherche et de sa patrimonialisation, dans le but de préserver ou d'élargir l'accès aux documents, en particulier aux documents rares ou patrimoniaux tel que les manuscrits.

Les projets de numérisation de fonds liés au patrimoine culturel des différents pays sont de plus en plus en vue de par le monde. Ce genre de projets n'est plus l'apanage des instances nationales ou internationales officielles (bibliothèques nationales ou organismes culturels).

Les manuscrits ont profité de cette révolution technologique, surtout du côté moyens de conservation et méthodes de consultation.

La transformation des fonds de bibliothèques en documents électroniques par la technique de numérisation.

I.8.1 Définition de la numérisation :

La numérisation est un acte qui consiste à créer une copie en mode image d'un document physique existant. Ces documents numérisés ainsi générés ne sont pas à confondre avec les documents dits « informatiques » créés dès l'origine sur ordinateur. On utilise le terme « numérique » du fait que, par cette procédure, des données analogiques sont traduites en langage binaire (une suite de 1 et de 0) par un numériseur (scanner). Cette « traduction » implique qu'il existe de bonnes et de mauvaises numérisations. La qualité de la numérisation se détermine alors en fonction de l'usage qui en est défini, et de la portabilité sur le réseau.

I.8.2 Les objectifs de numérisation :

La numérisation peut avoir plusieurs objectifs [7] :

- Préserver un document fragile et offrir un outil de valorisation.
- Donner un accès distant ou partagé à des documents peu accessibles ou uniques et exploiter leurs contenus.
- La numérisation permet de conservé un document original.
- La numérisation couplée à un system documentaire offre en outre des possibilités de recherche et d'exploitation nouvelles.
- Un chercheur peut il rapprocher sur son écran d ordinateur des textes ou des images éparpillés dans divers documents qu'il devait jusqu' alors consulter séparément.
- Peut offrir une alternative à la production de documents papier, comme les revues de presse ou les dossiers documentaires, qui posent d énormes problèmes de diffusion et de mise à jour.

I.8.3 Les modes de numérisation :[8]**A- Le mode image :**

On appelle « *mode image* » le fait de numériser un document en faisant une image, une photographie de chaque page. On obtient un fac-similé : exactement semblable à l'ouvrage original, il respecte l'organisation des informations sur la page et restitue fidèlement son contenu. On distingue différentes modes d images qui sont comme suit :

- **Le mode bitonale** : Image noire et blanc, fichier non volumineux. Ce mode convient aux documents contrastés.
- **Le mode niveau de gris** : contrairement au codage bitonal qui utilise deux niveaux de codification (noire ou le blanc).
- **Le mode couleur** : trois couleurs de base dont le rouge, le vert et le bleu sont utilisées et codifiées selon le principe de niveau de gris.

B- Le mode texte :

Le « *mode texte* » ne prend en compte la mise en page que de façon très limitée, il s'intéresse surtout à la retranscription du contenu de l'ouvrage en lettres et en mots, exactement comme si on utilisait un traitement de texte. Dès lors, on peut faire des recherches dans le contenu du document, copier des citations.

I.4 Les étapes de numérisation : les étapes de numérisation sont comme suit [5] :

I.4.1 Prétraitement et binarisation :

Le prétraitement consiste à améliorer la qualité des images en éliminant les défauts dus à l'éclairage et au processus d'acquisition. Ces défauts proviennent notamment de l'inclinaison du document sur la vitre du scanner, de la courbure du texte sur les bords due à l'épaisseur du livre, d'un éclairage non homogène, de la fonction de transfert du capteur, de contours de l'écriture flous. Le processus de vieillissement se traduit sur l'image par des trous, des taches d'humidité, de l'encre du verso apparaissant sur le recto, des traits partiellement effacés produisant une écriture fragmentée lors de la binarisation.

Les prétraitements peuvent être effectués à des fins de visualisation et/ou en vue des traitements ultérieurs de structuration ou de reconnaissance. Ils utilisent des opérations sur l'image telle que les modifications d'histogramme et les filtrages pour corriger la luminosité, réduire les bruits.

La binarisation consiste à transformer les images en image binaires, elle produit deux classes de pixels, qui sont représentées par des pixels noirs et des pixels blancs, avec différents niveaux de gris.

I.8.4.2 Segmentation d'image :

La segmentation est une opération de traitement d'images qui a pour but de rassembler des pixels entre eux suivant des critères prédéfinis. Les pixels sont ainsi regroupés en régions, qui constituent un pavage ou une partition de l'image. Il peut s'agir par exemple de séparer les objets du fond. Si le nombre de classes est égal à deux, elle est appelée aussi binarisation.

I.8.4.3 La reconnaissance :

La reconnaissance consiste en l'extraction des métadonnées prédéfinies dans les images numérisées, binarisées et segmentées des manuscrits anciens en langue arabe.

La reconnaissance de manuscrits consiste en l'étude et l'analyse des manuscrits numérisés.

Pour ce fait, l'extraction automatique de métadonnées est un outil très utilisé dans la classification des manuscrits. Elle se base sur des outils de traitement d'images (filtrage, binarisation et segmentation) et de vision artificielle.

Les trois étapes du processus de numérisation présentent le moyen de la reconnaissance automatique de la structure des manuscrits arabes, incluent la possibilité de reconnaître des mots, des caractères, des dessins et des schémas à l'intérieur du texte.

Plusieurs obstacles de caractères arabes et l'état physique actuel des manuscrits arabes.

Conclusion :

Dans cette partie nous avons étudié la description générale de manuscrits arabes du point de vue codicologique, paléographique et historique afin d'élaborer un mode d'accès électronique aux manuscrits arabes numérisés en mode image.

Par la suite, nous avons passé à la numérisation des manuscrits arabes anciens, qui se propose comme une solution d'aide à préservation, à la conservation, à l'accès et à la valorisation de documents fragile et rares.

En définitive, seuls les catalogages des manuscrits numérisés peuvent permettre d'exploiter les potentialités de diffusion des versions numériques, il faut les rendre accessibles et communiquer au public l'ensemble des informations qu'ils souhaitent recevoir en regard des manuscrits arabes anciens numérisés.

Dans le chapitre suivant nous allons procéder au chapitre du catalogage des manuscrits arabes anciens.

Chapitre II

Catalogage des

manuscripts arabes anciens

II.1 Introduction :

Un document manuscrit implique des conditions de conservation des plus drastiques pour que celui-ci puisse être consulté le plus longtemps possible, afin de garantir la disponibilité de ces documents il faut qu'elle possède d'un catalogage.

II.2 Généralités sur le catalogage:

Le catalogage consiste à analyser le document en tant que support, fait partie des deux opérations intellectuelles nécessaire à la constitution d'un catalogue : catalogage (Description physique) et indexation (Description intellectuelle).

Le catalogue est destiné à identifier les documents (décrits dans des notices) et à faciliter leur recherche ou leur localisation.

Il existe différents types de catalogues :

- Un catalogue local décrit les documents d'une seule bibliothèque.
- Un catalogue commun référence les documents de plusieurs bibliothèques relevant d'un même organisme, comme une bibliothèque centrale et des bibliothèques de quartier.
- Un catalogue collectif décrit les collections de plusieurs bibliothèques indépendantes.
- Un catalogue virtuel est un outil permettant d'interroger plusieurs catalogues simultanément.

L'informatisation des catalogues des bibliothèques a commencé dans les années 1960. Les catalogues informatisés permettent de créer de nouveaux accès, comme la recherche portant sur tous les mots de la notice (recherche plein-texte). Ils permettent également de construire des équations de recherche en utilisant des opérateurs booléens.

Le catalogue interactif permet de passer à la consultation de livres en ligne, de documents rares, fragiles, ou non-empruntables tels que les manuscrits. Grâce aux technologies Web, le catalogue interactif permet à l'internaute de feuilleter directement en ligne le document.

II.3 Le catalogage des manuscrits arabes :

Le catalogue est le référent bibliographique unique pour les documents numérisés. C'est de celui-ci que l'on extrait les notices décrivant les documents numérisés [9].

Le catalogage des manuscrits arabes anciens se trouve confronté à deux approches issues de deux écoles différentes:

-L'école arabe, d'un côté, qui à son tour se trouve répartie en école du Machriq et celle du Maghreb.

-L'école orientaliste, d'un autre côté, représentée quant à elle par des catalogueurs occidentaux qui ont pris l'initiative de répertorier les manuscrits arabes.

II.3.1 L'école arabe :

Les travaux de catalogage des manuscrits arabes datent de l'époque Abasside, considérée comme l'âge d'or de la civilisation arabo-islamique. En effet, deux principaux auteurs, Ibn Nadim et Hadji Khalifa, ont réalisé deux catalogues bibliographiques. Le premier nommé "*El Fihrist*" d'Ibn Anadim réalisé en 377 de l'ère du Hégire et le second nommé "*Kachf Adhonoun aan Asami Al Koutoub Wa Alfounoun*" de Mustapha Ben Abdellah, connu sous le nom de Hadji Khalifa, décédé en 1067 de l'hégire laissant son œuvre inachevée. Toutefois, le travail fut achevé par Ismaïl Pacha Ibn Mohamed Amine Ibn El Amir Salim El Babani en 1330 de l'hégire [10].

II.3.2 L'école orientaliste :

Les orientalistes ont joué un rôle très important dans le domaine du catalogage des manuscrits arabes.

Plusieurs catalogues sont réalisés par les orientalistes et on le trouve dans différentes bibliothèques que ce soit dans les bibliothèques arabes ou dans des bibliothèques étrangères, principalement européennes. Nous citons entre autre quelques références :

-L'histoire de la littérature arabe, une bibliographie des littératures arabes depuis la première manifestation de la littérature arabe jusqu'en 1900, réalisé par Carl Brockelmann.

-En Europe, le premier catalogue de manuscrits arabes existant dans la bibliothèque Royale à Copenhague en 1980 par Jacob George Christian Adler.

-Le catalogue des notices pour la collection de manuscrits de la Bibliothèque Nationale à Paris réalise en 1848, par l'orientaliste sicilien Michel Amari.

Les premiers catalogues des manuscrits arabes sont réalisés par les orientalistes, qui ont joué un rôle très important dans les recherches de littérature arabe.

Les travaux de catalogage de manuscrits arabes réalisés par les arabes, eux-mêmes, datent de l'époque Abbasside, qui est considérée comme l'âge d'or de la civilisation arabo-islamique. Pour cela, les bibliothèques ont commencé à acquérir des ressources sous forme numérique, sont apparus de nouveaux niveaux de granularité permettant de les décrire et d'y donner accès. Les reproductions numériques ou les documents nés numériques rendent désormais possibles des niveaux d'accès plus fines, ces différents niveaux d'accès nécessitent que soient créées des métadonnées de même niveau de granularité [2].

II.4 Les métadonnées :

II.4.1 Définition des métadonnées :

Une métadonnée est littéralement une donnée sur une donnée .Plus précisément, c'est un ensemble structuré de données décrivant une ressource quelconque.

Les métadonnées sont de l'information structurée qui décrit, explique, localise ou facilite autrement l'obtention, l'utilisation ou la gestion d'une ressource d'information.

II.4.2 Caractéristiques des métadonnées :

La métadonnée est utilisée pour Faciliter la recherche des documents et gérer les restrictions d'utilisation.

Le manuscrit est un document d'une valeur inestimable (du point de vue support, contenu et rareté), sa diffusion a entraîné la nécessité de catalogage respectant des recommandations pour le décrire d'une manière détaillées, tout en utilisant [11]:

- Métadonnées descriptives, décrivent une ressource à des fins de découverte et d'identification (auteur, titre, mot-clé.....) pour permettre la recherche et la récupération.

- Métadonnées structurelles des différents éléments qui composent le manuscrit. la relation physique (folio, reliure,...), la relation logique entre ces éléments (numérotation, volume, cahier...).

Elles servent à se repérer dans un document complexe et à avoir une vue d'ensemble du contenu, accéder à un niveau de granularité plus fin.

On a deux types de métadonnées :

- internes : contenues dans la ressource elle-même.
- externe : contenues dans une notice séparée de la ressource.

Les métadonnées permettent de nommer, désigner, décrire, ou de catégoriser les divers ressources pédagogiques d'une manière interprétable et compréhensible par les machines comme par les humains, afin de les rendre identifiables, exploitables et partageables. Elles rendent ainsi possible la gestion des bases documentaires (identification, archivage, indexation recherche...).

II.4.3 Les métadonnées définies par quelques bibliothèques arabes :

Après étude de différents protocoles de catalogage dans différents instituts et bibliothèques des pays arabes, tels que la bibliothèque centrale de Ben Ghazi en Libye, la bibliothèque Hasan Hosni Abdelwahab en Tunisie, la bibliothèque du Qatar, l'université du Koweït, l'institut des recherches islamiques en Arabie saoudite et quelques bibliothèques au Yémen. Nous remarquons qu'ils ne respectent aucune méthode de description de manuscrits, la notice bibliographique se caractérise par une description formelle des documents et sous une forme réduite.

Les notices des catalogues arabes se ressemblent dans les éléments descriptifs suivants :

- Le titre du manuscrit ;
- L'auteur (nom et date de décès) ;
- Le copiste (nom et date de copie) ;
- Description physique (nombre de folios et les dimensions) ;
- La langue d'écriture ;
- Le colophon ;

- Observations et bibliographie.

Mais elles se différencient par la présence de quelques métadonnées dans l'une et leur absence dans les autres, telles que :

- Le numéro du manuscrit ;
- L'état du manuscrit ;
- La date d'édition ;
- L'encre ;
- Le lieu de conservation ;

De ce fait, nous pouvons la qualifier de simple fiche rédigée par les bibliothécaires, et non pas de notice documentaire qui doit fournir le maximum de renseignements possibles sur le manuscrit.

Ces catalogues produits sont traditionnels, sous formes de livres imprimés basés sur la classification selon des critères prédéfinis, par ordre alphabétique, auteurs ou par collections de thèmes.

II.4.4 Les métadonnées définies par l'IRHT :

IRHT : Institut de recherche et d'histoire des textes se consacre à la recherche fondamentale sur les manuscrits médiévaux et les imprimés anciens dans les principales langues de culture du pourtour méditerranéen. Il a pour but d'établir une plateforme de catalogage des manuscrits médiévaux d'Europe occidentale et du monde méditerranéen.

L'originalité de l'IRHT est de rassembler, au sein d'une même équipe, des spécialistes de disciplines diverses mais complémentaires qui se proposent d'étudier l'histoire du livre sous deux angles principaux : En tant qu'objet avec ses aspects codicologiques et paléographiques.

Les champs qui encapsulent les métadonnées définies par IRHT sont [2]:

- La forme
- Support
- Encre
- Cahier
- La mise en page
- Reliure
- Écriture
- Décoration
- Histoire
- Contenu

- Etat du manuscrit

Le protocole d'IRHT est consacré pour la recherche fondamentale sur des manuscrits médiévaux. Il étudie le manuscrit en considérant son histoire, son aspect paléographique et son aspect codicologique. Mais, il se retrouve incomplet pour décrire des manuscrits arabes anciens.

II.4.5 Les métadonnées définies par la fondation EL Saoud :

EL Saoud : Le centre de documentation et des études islamiques est constituée sous forme d'une association de droit marocain à but non lucratif. Elle a été créée à l'initiative du prince Abdallah Ibn Abdul-Aziz, en vue de fournir aux chercheurs les moyens les plus appropriés en documentation et en support de la recherche scientifique [12].

La fondation El Saoud, en 1987, a mis en oeuvre un ensemble de métadonnées descriptives, pour le catalogage des manuscrits arabes. Elles sont regroupées dans sept champs différents, chaque champ cerne un certain nombre de métadonnées :

- Identification ;
- Contenu ;
- Description première ;
- Description physique du manuscrit ;
- Additifs ;
- Notice supplémentaire et traductions ;
- Observation.

La description du manuscrit arabe ancien par le modèle El Saoud n'est pas très détaillée et ne prend pas en considération des éléments qui peuvent être très utiles pour les chercheurs du domaine.

II.4.7 Le nouveau protocole de catalogage : il est issu de la fusion des protocoles cités précédemment [10] :

Auteur :

- Nom complet de l'auteur :
 - ✓ Ism.
 - ✓ Lakab.
 - ✓ Kunya.
 - ✓ Nisba.
- La naissance et le décès de l'auteur :

- ✓ Date. (système hégirien et grégorien).
- ✓ Siècle.
- ✓ Intervalle.
- ✓ Lieu. (pays, région).

Description physique :

- Etat du manuscrit. (complet ou pas).
- Epaisseur du manuscrit. (en cm).
- Matière de la couverture. (cuir, papier, carton...).
- Dimension de la couverture. (en cm).
- Matière des folios.
- Dimensions des folios.
- La reliure :
 - ✓ Technique (plat, couverture, dos)
 - ✓ Matière.
 - ✓ Emplacement.
 - ✓ Motif (rabat,...).
 - ✓ Artisan (nom).
- Restauration :
 - ✓ Date de restauration (système hégirien et grégorien).
 - ✓ Lieu de restauration (pays, région).
 - ✓ Responsable (nom).

Description logique :

- Nom du manuscrit.
- Autres noms du manuscrit.
- Responsable.
- Titre de la couverture.
- Nombre de copies.
- Numéro du volume :
 - ✓ Système de numérotation (foliation, pagination).
 - ✓ Signature.
 - ✓ Réclame.
 - ✓ Cahier :
 - numéro du cahier.

- sujet.
- Début.
- page :
 - Numéro de la page.
 - Titre de la page.
- Nombre de pages
- fin.
- Nombre de volumes

Contenu :

- Le thème.
- La langue, le dialecte.
- Analyse du texte :
 - Sujet(s) (médecine, histoire, religion...).
 - Responsable (l'auteur, possesseur, bibliothèque).
 - Explications.
 - Lectures.
 - Confrontation (Collations)
 - Ecoute.
 - Autorisation.
- Incipit.
- Explicite.
- Colophon (existe ou pas).
- Intitulé de début et de fin.
- Résumé (existe ou non).

Ecriture et décoration :

- Ecriture :
 - ✓ Style d'écriture (kufique, cursive...).
 - ✓ Qualité de l'écriture (lisible, mauvaise...).
 - ✓ La couleur de l'encre : -noire pour le texte original.
-rouge pour l'explication du texte.
 - ✓ Outil de l'écriture.
 - ✓ Mise en page :
 - Pique /réglure.

- Nombre de lignes.
- Nombre de colonnes (cas de texte poétique)
- Surface écrite.
- Manchette.

- Décoration :
 - Description du décor :
 - les couleurs du décor.
 - Décor du texte.
 - Décor de la page.
 - Dessins explicatifs (cartes, tableau...).
 - Sceaux (marque, cachet).
 - Localisation du dessin (entête, pied de page ...).

Copiste :

- Nom complet du copiste :
 - ✓ Nom.
 - ✓ Prénom.
- Date de naissance et de décès (système hégirien et grégorien).
- Lieu de naissance et de décès. (pays, région).
- Nom de la copie.
- Copier pour (Destination : personnes physiques ou logiques)
- Date de la copie. (système hégirien et grégorien).
- Lieu de la copie. (pays, région).

Additions :

- Annotation marginales.
- Annotations sur les bordures.
- Responsables.
- Emplacement dans le manuscrit.
- Critiques.
- Corrections.
- Correcteurs (responsables).
- Folios détachés de la copie.
- Copie du manuscrit dans d'autres bibliothèques.

Histoire et conservations :

- Numéro de conservation.
- Ancien nom du manuscrit.
- Ancien numéro du manuscrit.
- Lieu d'origine (pays, région).
- Possesseurs successifs (noms des personnes, institut...).
- Lieu de conservation :(pays, région, institution, collection...).
- Etat de conservation.
- Bibliographie :
 - ✓ Edition du texte :
 - Date (dans le système hégirien et grégorien).
 - Lieu (pays, région).
 - Etude référençant le manuscrit.
 - ✓ Traduction :
 - langue de traduction.
 - Référence traduction.

Tableau II.1 : Protocole de description des manuscrits arabes**II.5 Encodage des métadonnées décrivant les manuscrits arabes anciens :**

Les bibliothèques ont une longue expertise en matière de gestion et de classement de leurs collections, par le biais du catalogage, qui sert également à donner accès aux ressources. Les catalogues sur fiches proposaient un accès principal, des entrées secondaires, une Cote et des éléments de classification. Les bibliothèques utilisent désormais d'autres ensembles d'éléments, notamment des normes d'encodage de métadonnées, dont le MARC (Machine-Readable Cataloguing), EAD (Encoded Archival Description), METS (Meta data Encoding and Transmission Standard), TEI (Text Encoding Archival), etc.

Ainsi, plusieurs formats proposent leur propre modèle de représentation des métadonnées descriptives des manuscrits Arabes anciens numérisés, mais qu'il est le modèle de catalogage à travers lequel il serait possible de définir des métadonnées concises et représentatives du manuscrit arabe numérisé ?

II.5.1 Définition du standard :

D'après les organismes chargés de standardiser, le standard technique est créé à partir des spécifications élaborées par des groupes de travail (tel que le Dublin Core) ; les spécifications sont étudiées, analysées et ont l'authenticité de standard après qu'un consensus soit obtenu lors d'une réunion annuelle de l'IEEE (Institute of Electrical and Electronics Engineers).

Le standard n'est qu'un ensemble de recommandations provenant d'un groupe représentatif d'utilisateurs qui peuvent être des sociétés, des groupements de professionnels ou des institutions.

II.5.2 Formats d'encodage de métadonnées :**II.5.2 Le format MARC :**

De nos jours, il est impossible de lire une revue de bibliothéconomie, d'assister à une conférence pour bibliothécaires voire d'échanger de manière informelle avec un confrère du domaine sans entendre parler du format MARC, des notices MARC ou de compatibilité avec la norme MARC.

II.5.2.1 Définition :

MARC est un acronyme constitué à partir de l'expression anglaise "MACHINE-Readable Cataloguing", qui se traduit par catalogage ordinaire, dit plus couramment catalogage lisible par machine.

Le format MARC né à la bibliothèque du Congrès des États-Unis au milieu des années 1960, il définit une structure logique des données saisies dans un système informatique de gestion de bibliothèque. Le format a subi des modifications selon les besoins des bibliothèques et des pays [13]. Il a connu plusieurs versions nationales comme :

- USMARC.
- CANADAMARC : le format MARC au CANADA. MARC21.
- UNIMARC.
- UKMARC.

Les deux formats MARC particulièrement intéressants reconnus par IFLA (International Federation of Library Associations) sont : MARC 21 et UNIMARC.

II.5.2.1.2 MARC 21 :

Le format MARC d'origine a évolué jusqu'à devenir la norme MARC 21, il est maintenu par la bibliothèque du congrès. Aujourd'hui utilisée par la plupart des bibliothèques pour les notices bibliographiques, les notices d'autorité et les notices sur les fonds relatives aux différents systèmes de bibliothèque [14].

A- Les notices MARC :

MARC utilise deux types de notices de catalogage :

A.1 La notice MARC bibliographique :

Contiennent l'information traditionnelle d'une fiche catalographique et l'information à propos d'un livre, d'un périodique, d'un enregistrement audio, d'une cassette vidéo, etc.

A.2 La notice MARC d'autorité :

Les notices MARC d'autorité servent à la normalisation et la présentation des noms de personne, de personnes morales (par exemple, les sociétés, les entreprises, les institutions, etc.), des collectivités, des titres et des sujets. Elles fournissent ainsi un contrôle d'autorité, c'est-à-dire une forme reconnue de désignation de ces entités que l'on peut utiliser pour accéder aux notices bibliographiques à partir d'un nom.

Une notice d'autorité est constituée de trois éléments principaux [15]:

1) La Vedette :

Est la forme normalisée et « faisant autorité » d'un nom, d'un sujet ou d'un titre qui servira de point d'accès aux notices bibliographiques. L'emploi de vedettes de nom et de sujet uniformisées facilite le repérage simultané des notices qui les partagent.

2) Les Renvois :

Sont des indications qui permettent à l'utilisateur de repérer une variante de la forme établie du nom ou du sujet (désignées par le terme « Rappel de renvoi 'voir' ») ou diverses formes de vedettes autorisées associées à la vedette de la notice (désignées par le terme « Rappel de renvoi 'voir aussi' »). Les renvois inscrits dans les notices MARC se rapportent à la vedette d'autorité.

3) Les Notes :

Contiennent de l'information d'ordre général à propos des vedettes normalisées ou des renseignements plus spécialisés, comme des citations d'une source de données consultée où des indications sur la vedette ont pu être trouvées ou non.

II.5.2.1.3 UNIMARC :

UNIMARC (acronyme pour Universel MARC) est le format développé par l'IFLA pour permettre l'échange de l'information bibliographique informatisée et servir d'interface entre les formats MARC nationaux. Sa première édition remonte à 1977 [16].

II.5.2.1.3.1 Les métadonnées en UNIMARC :

Le format UNIMARC contient 10 blocs et chaque bloc contient des zones ou champs qui sont représenté comme suit [17] :

- 0- Bloc des numéros d'identification ISBN, ISSN, numéro dans le catalogue local**
- 1- Bloc des informations codées :** Contient des éléments de données de longueur fixe (le plus souvent codés) décrivant divers aspects de la notice ou des données (Dates, langues, pays).
- 2- Bloc des notes :** Contient des notes (ex : le contenu), destinées à l'affichage public, expliquant les relations qui existent entre la vedette de la notice (2XX) et d'autres vedettes, et contribuant à identifier l'entité décrite dans la notice d'autorité.
- 3- Bloc des indices de classification :** Contient les indices de classification relatifs à la vedette de la notice.
- 4- Bloc des liens :** Collection, histoire...
- 5- Bloc des titres associés :** Titre uniforme, titre parallèle...
- 6- Bloc de l'indexation matière :** Nom commun, classification...
- 7- Bloc des responsabilités intellectuelles :** Auteur principal personne physique et collectivité
- 8- Bloc des informations sur les sources :** Contient la mention des sources de la notice ainsi que des notes du catalogueur concernant les données, non destinées à l'affichage public .
- 9- Bloc à usage national :** Contient des données propres à l'établissement créateur de la notice.

II.5.2.1.3.2 L'usage de l'UNIMARC en Algérie : [18]

Le CERIST possède une bibliothèque riche en documentation scientifique et technique qu'il a entièrement informatisée en s'appuyant sur des normes et standards internationalement reconnus.

En matière de catalogage, c'est le format UNIMARC qui a été adopté pour traiter le fonds documentaire de la bibliothèque et de l'atelier de microfiches.

Le fonds du CERIST est mis en ligne depuis 1998, à leur tour les bibliothèques universitaires algériennes ont tendance à adopter le format UNIMARC.

L'usage d'UNIMARC en Algérie souffre de plusieurs maux résumés en deux points essentiels:

- L'absence d'un fichier d'autorité algérien.
- Le manque d'expérience du personnel en matière d'UNIMARC.

Le CERIST compte dans le futur proche mettre en place un catalogue collectif national sous UNIMARC. Par ailleurs, une équipe de recherche du département de bibliothéconomie de l'Université d'Alger entreprend depuis février 2010 un projet de mise en œuvre d'un fichier d'autorité algérien : noms géographiques, noms arabes anciens et collectivités algériennes.

II.5.2.2 L'EAD (Encoded Archival Description) :

II.5.2.2.1 Définition :

L'EAD est la DTD (Document Type Définition) utilisée pour décrire des fonds d'archives, des collections de manuscrits et plus généralement tous ensemble hiérarchisé de documents [19].

La DTD EAD est née à la bibliothèque de l'université de Berkeley en 1993 pour remplacer le standard MARC-AMC (Archival and Manuscript Control), elle est maintenue par la bibliothèque de congrès.

L'EAD facilite l'échange et la portabilité des instruments de recherche, il est possible d'utiliser le même document encodé pour des restitutions multiples [20].

II.5.2.2.2 La structure de l'EAD :

Un document encodé en EAD comporte, sous l'élément racine <ead>, trois éléments

Principaux :

- Un en-tête (<eadheader>).
- Une page de titre (<frontmatter>).
- Une description archivistique avec attribut de niveau de description (<archdesc>).

A- <eadheader> :

L'élément En tête EAD <eadheader>, il est directement inspiré par la TEI (Text Encoding Initiative). Cet élément contient des informations bibliographiques et descriptives sur l'instrument de recherche (titre, auteur, date, etc.).

La structure générale d'un outil de recherche encodé en EAD peut se présenter comme suit :

```

<eadheader>
  <eadid> Identifiant EAD
  <filedesc> Description du fichier
    <titlestmt> Mention de titre
    <titleproper> Titre propre de l'inventaire
    <subtitle> sous-titre de l'inventaire
    <author> auteur
    <editionstmt> Mention d'édition
    <publicationstmt> Mention de publication
    <seriestmt> Mention de collection
    <notestmt> Mention de note
  <profiledesc> Description du profil
    <creation> Informations sur la création de l'inventaire
    <langusage> Langue utilisée
    <descrules> Règles de description
  <revisiondesc> Descriptions des révisions
</eadheader>

```

Figure II.1 : Structure générale d'un outil de recherche encodé en EAD

B- <frontmatter> :

L'élément Préliminaires <frontmatter> il donne des informations sur la création, la publication ou l'utilisation de l'instrument de recherche lui-même et non sur les documents en cours de description. Cela peut être une page de titre, une préface, une dédicace et des instructions relatives à l'utilisation de l'instrument de recherche. Dans le sous-élément facultatif page de titre <titlepage> de <frontmatter>, on peut répéter des informations choisies dans l'élément En-tête EAD <eadheader>, puis les autres parties liminaires.

C- <archdesc> :

Cet élément englobe toutes les informations décrivant un corpus documentaire dans son ensemble et recouvrant les différents niveaux hiérarchiques de cette description. L'élément <archdesc> comprend des informations de haut niveau : identification du fonds (élément <did>), biographie/histoire institutionnelle (élément <bioghist>) ; puis on décrit les parties du fonds dans l'élément <dsc> Description des sous-composants, en commençant par le niveau hiérarchique le plus élevé et en allant si on le souhaite jusqu'au niveau de la pièce.

La structure de l'élément <archdesc> peut être décrite comme suit :

```
<archdesc>
  <did> Identification et description
  <bioghist> Biographie/histoire institutionnelle
  <custodhist> Historique de la conservation
  <acqinfo> Informations sur les modalités d'entrée
  <accessrestrict> Conditions d'accès
  <userrestrict> Conditions d'utilisation
  <arrangement> Classement
  <bibliography> Bibliographie
  <scopecontent> Présentation du contenu
  <controlaccess> Point d'accès contrôlé
  <dsc> Description des composants
  <c> niveaux de description définis grâce à un attribut
</archdesc>
```

Figure II.2: Structure de l'élément <archdesc>

L'élément <did> représente l'identification de l'unité documentaire. Il peut être décrit comme suit :

```
<did>
  <origination> Origine
  <unittitle> Intitulé de l'unité documentaire
  <unitdate> Date de l'unité documentaire
  <physdesc> Description physique
  <repository> Organisme responsable de l'accès intellectuel
  <unitid> Identifiant de l'unité documentaire
  <langmaterial> Langue de l'unité documentaire
  <physloc> Localisation physique
</did>
```

Figure II.3: Structure de l'élément <did>

II.5.2.3 Le projet MASTER (Manuscript Access through Standards for Electronic Records):

Le projet MASTER est un projet européen inauguré en le 1er janvier 1999, a pour but de créer un système générique, suffisamment souple et robuste pour permettre son application dans différents domaines de la description du manuscrit et la technique choisie atteindre ce but plus ambitieux et fondé sur les normes international SGML et XML.

Un groupe de travail dirigé par TEI a été chargé de développer une DTD (définition de type Document) spéciale pour la description de manuscrit qui serait compatible avec l'encodage concernant le texte électronique crée par la TEI (Texte Encoding Initiative) [2].

II.5.2.3.1 L'encodage des manuscrits dans le projet MASTER :

Le projet MASTER a pour but de créer un système générique, suffisamment souple et robuste pour permettre son application dans différents domaines de la description du manuscrit. La solution proposé par ce projet est de développer une DTD (définition de type Document) spéciale pour la description de manuscrit qui serait compatible avec l'encodage concernant le texte électronique crée par la TEI (Texte Encoding Initiative).

Permettant essentiellement au sous élément <msDescription> de se produire dans un ordre spécifié.

L'élément <msDescription> peut contenir jusqu'à sept sous éléments différents qui sont les suivants, dont seul le premier est obligatoire:

1. <msIdentifier> : (manuscript identifier) cet élément englobe huit éléments qui permettent l'identification du manuscrit ou d'un fragment de manuscrit, qui sont mentionnés comme suivant : <country>, <region>, <settlement>, <institution>, <repository>, <collection>, <idno> et <altName>.

2. <msHeading> : C'est un élément facultatif.

3. <msSummary> : Élément facultatif, qui peut contenir des éléments tels que <author>, <title>, <origPlace>, <origDate>, etc.

4. <msContents> : Élément facultatif, une liste détaillée du contenu intellectuel d'un manuscrit ou partie d'un manuscrit.

5. <physdesc> : Cet élément regroupe des informations sur l'aspect physique d'un manuscrit.

6. <additional> : élément facultatif, qui regroupe des éléments additifs au manuscrit.

7. **<msPart>** : Nous pouvons avoir cet élément plusieurs fois, il contient la description d'un fragment du manuscrit.

II.5.2.4 Le format METS :

II.5.2.4 Définition :

Le modèle METS (Meta data Encoding and Transmission Standard) est un exemple de format conteneur permettant de rassembler en un fichier unique des données du type objet numérique et des métadonnées.

Le format METS est initié par la Fédération des Bibliothèques numériques américaine-DLF (Digital Library Federation) et maintenu par la bibliothèque du Congrès.

Le schéma XML METS permet l'encodage des métadonnées descriptives, de gestion et de structure des objets numériques [21].

II.5.2.4.2 La Structure d'un document METS :

METS se décompose en sept sections, dont une seule est obligatoire, *structmap* décrivant la structure du document et chacune de ces sections comprend différents types de métadonnées.

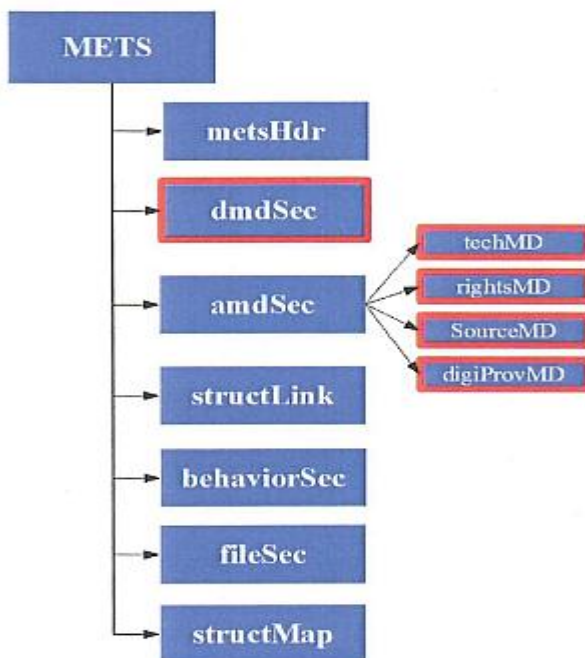


Figure II.4 : La structure d'un document

II.5.2.4.3 Les éléments de METS :

Un fichier METS est structuré en sept sections, pouvant comprendre un ou plusieurs groupes de métadonnées, dont une seule (structMap) est obligatoire [22] :

-**Header** <metsHdr> : Cette section facultative et répétable contient les métadonnées concernant le document METS lui-même telles que le créateur, le logiciel d'édition employé.

– **Descriptive metadata** <dmdSec> : Contient les métadonnées descriptives de l'objet. Des schémas externes comme Dublin Core simples et qualifiés, MARCXML, MODS, EAD, peuvent être utilisés. Cette section est facultative et répétable.

– **Administrative metadata** <amdSec> : Contient les métadonnées de gestion correspondant au document numérique. Elle contient quatre sous section :

1-sourceMD: donne des informations sur la source dont le document numérique décrit est issu.

2-techMD : donne des informations techniques sur le document dans son ensemble (identifiant,version...).

3-digiproMD : donne des informations sur l'historique du document numérique, depuis sa création jusqu'au moment présent en passant par toutes les opérations qu'il a pu subir.

4-rightsMD : donne des informations sur le statut juridique du document.

– **File Section** <fileSec> : elle liste les fichiers composants de l'objet numérique, les sous-éléments fileGrp répétables permettent l'enregistrement de plusieurs versions d'un même objet.

– **Structural Map** <structMap> : contient les informations concernant la hiérarchie de structure de l'objet et lie cette hiérarchie avec les fichiers. Cette section est la seule obligatoire.

– **Structural Links** <structLink> : renseigne les hyperliens entre composantes de l'objet, par exemple des hyperliens entre une page et une autre si le document décrit est un site web.

– **Behavior** <behaviorSec> : Cette section facultative et répétable contient les informations concernant les programmes nécessaires à la lecture.

II.5.2.4.4 Les pointeurs de METS :

METS utilise un système de pointeurs afin de mettre en relation des éléments de métadonnées et des fichiers entre eux. Plusieurs types de pointeurs sont proposés [23] :

- **Xlink:** permet de pointer vers un bloc à l'extérieur du document METS.
- **Filepointer:** interne au document METS et permet de pointer vers l'identifiant d'un fichier.
- **Metspointer:** externe au document METS et permet de lier plusieurs documents METS entre eux.
- **Area:** élément intégré dans la carte de structure et permet de pointer vers une partie de fichier.

II.5.2.5 Le Dublin Core :

II.5.2.5.1 Qu'est-ce que le Dublin Core ?

Le Dublin Core est un schéma de métadonnées générique qui permet de décrire des ressources numériques ou physiques et d'établir des relations avec d'autres ressources. Il comprend officiellement 15 éléments de description formels (titre, créateur, éditeur), thématiques (sujet, description, langue...) et relatifs à la propriété intellectuelle.

Le Dublin Core est employé par l'Organisation mondiale de la santé, ainsi que d'autres organisations intergouvernementales.

II.5.2.5.2 Les quinze éléments du Dublin Core :

Le Dublin Core est un ensemble de 15 éléments de métadonnées ayant trait comme suit [24] :

- au contenu: Title, Description, Subject, Source, Coverage, Type, Relation.
- à la Propriété intellectuelle: Creator, Contributor, Publisher, Rights.
- à la Version: Date, Format, Identifier, Language.

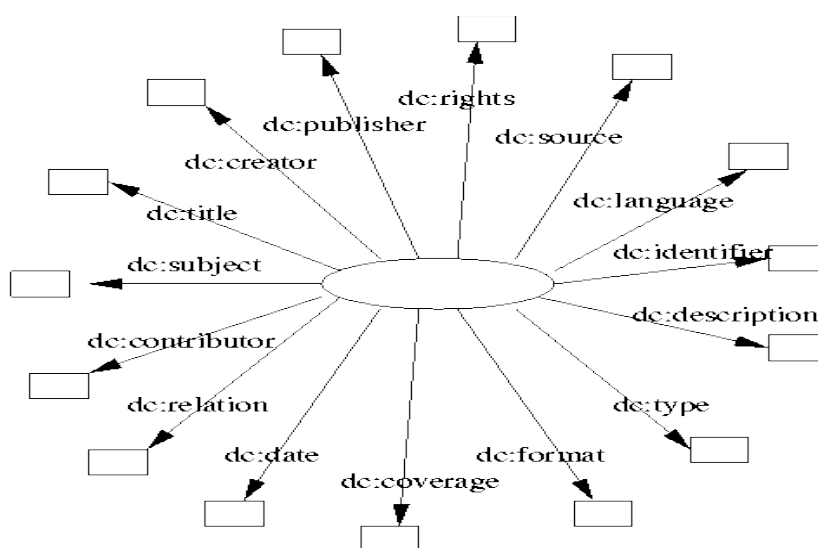


Figure II.5 : les éléments de métadonnées Dublin

1. titre/title : Le nom donné à la ressource.

2. Description/description: Une description du contenu de la ressource, peut inclure un résumé, une table des matières, une référence à une représentation graphique du contenu.

3. Sujet/subject : Le sujet du contenu de la ressource.

4. source/source : Une référence à une ressource dont dérive-la présente ressource. La présente ressource peut être dérivée de la ressource source en partie ou en totalité. La pratique recommandée est de référencer la ressource au moyen d'une chaîne ou d'un numéro conformes à un système d'identification formelle.

5. couverture/coverage : L'étendue ou la portée du contenu de la ressource. La couverture comprendra habituellement une position dans l'espace (le nom ou les coordonnées géographiques d'un lieu), une période temporelle (le nom d'une époque, une date ou un ensemble de dates).

6. type/type : La nature ou le genre du contenu de la ressource.

7. relation/relation : Une référence à une ressource apparentée. La pratique recommandée est de référencer la ressource au moyen d'une chaîne ou d'un numéro conformes à un système d'identification formelle.

8. créateur/creator : Une entité principalement responsable de la fabrication du contenu de la ressource. Comme exemples de créateurs une personne, une organisation ou un service.

9. contributeur/contributor : Une entité responsable de contributions au contenu de la ressource. Comme exemples de contributeurs, une personne, une organisation ou un service.

10. éditeur/publisher : L'entité responsable de la diffusion de la ressource, dans sa forme actuelle, tels, un département universitaire, une entreprise.

11. gestion des droits/rights : informations à propos des droits détenus à l'intérieur et sur la ressource. L'élément rights contiendra typiquement une déclaration de gestion des droits de la ressource, ou référencera un service fournissant ce renseignement. Les informations légales comprennent souvent les droits de propriété intellectuelle, les droits d'auteur et divers droits patrimoniaux.

12. date/date : Une date associée à un événement dans le cycle de vie de la ressource.

13. format/format : La matérialisation physique ou digitale de la ressource.

14. identifiant de la ressource/identifier : Une référence univoque de la ressource dans un contexte donné. La pratique recommandée est d'identifier la ressource au moyen d'une chaîne ou d'un numéro conformes à un système d'identification formelle.

15. langue/language : La langue du contenu intellectuel de la ressource.

II.5.2.5.3 L'inconvénient de Dublin Core :

L'inconvénient de Dublin Core est le manque de clarté dans le mécanisme de la Qualification. Le mécanisme de la qualification est l'élément qui enrichit la description Dans Dublin Core.

Ce manque de clarté a conduit à des interprétations locales différentes, ce qui a rendu Difficile l'interopérabilité de Dublin Core.

II.5.2.6 La TEI (Text Encoding Initiative) :

II.5.2.6.1 Définition :

La TEI (Text Encoding Initiative) est un projet universitaire pluridisciplinaire visant à uniformiser autant que possible le codage de documents en vue de leur échange et de leur publication en ligne ou hors ligne. Il s'agit d'un format de codage de documents dit «structuré»: il a besoin d'un langage, XML, pour aider à la saisie d'un texte en lui donnant une structure compatible à la fois avec les exigences des différentes communautés qui l'utilisent (bibliothécaires, historiens, philo-sophes, spécialistes de littérature, linguistes, amateurs) et avec les possibilités des outils de consultation [25].

La TEI (Text Encoding Initiative) est une communauté académique internationale dans le champ des humanités numériques visant à définir des recommandations pour l'encodage de documents textuels

La TEI a donné naissance à plusieurs applications parmi lesquels nous citons [26] :

- La TEI Lite : est une version simplifiée des *Recommandations* de la « TextEncoding Initiative » (TEI) qui s'adressent à tous ceux qui souhaitent échanger des informations stockées sous forme électronique. Elles mettent l'accent sur l'échange des données textuelles, notamment pour les sciences humaines et les études sur les textes littéraires, mais d'autres types de données comme les images et les sons sont également pris en compte.
- TEI-ms : la TEI manuscript description, utilisée pour la description des manuscrits médiévaux.

- La TEI Guideline for Electronic Text Encoding : pour définir et documenter un langage de balisage pour représenter les caractéristiques structurelles des textes.

II.5.2.6.2 Caractéristiques de la TEI :

La TEI est un format d'encodage standard qui comprend des métadonnées, repose sur le langage XML.

La TEI sert à faciliter la création, l'échange, et l'intégration des données textuelles informatisées de toute sorte de texte, de toutes les langues et de toute origine temporelle ou culturelle. La TEI s'adresse également aux débutants, cherchant des solutions bien connues et consensuelles et aux experts, cherchant à créer de nouvelles solutions [27].

II.5.2.6.3 La TEI est la description des manuscrits :

II.5.2.6.3.1 Principe de description des manuscrits :

La TEI-ms propose deux méthodes de description de manuscrits:[10]

- La première, dite simple, consiste en une série de paragraphes imbriqués dans des identificateurs définissant un aspect particulier du manuscrit. Les paragraphes sont constitués de petites phrases décrivant le manuscrit.
- La seconde, dite complexe, définit une notice avec une structure de balisage plus détaillée, décrivant le manuscrit d'une manière précise et complète.

La première méthode est simple à utiliser par le catalogueur, elle lui offre un environnement familier de description de manuscrits. Quant à la seconde, la description des manuscrits se fait d'une manière délicate, par ce qu'elle exige le respect de la grammaire de la TEI_ms.

II.5.2.6.3.2 Les éléments de la TEI :

L'élément<msDesc> normalement apparaît à l'intérieur de l'élément<sourceDesc> de l'entête d'un document conforme à la TEI, dans lequel le document codé est une représentation numérique d'un certain manuscrit original. <msDesc> (description du manuscrit) : contient une description d'un manuscrit unique identifiable et fournisse des informations plus détaillées dans un certain nombre de rubriques, que nous décrivons comme suit [28] :

- < **msIdentifier**> : (identificateur de manuscrit) contient les informations nécessaires pour identifier le manuscrit. Il contient les éléments suivants :

<country> (pays), <region> (région), <settlement> (établissement), <institution> (institution), <repository> (dépôt).

<msIdentifier> dispose des autres éléments afin de fournir différents moyens pour identifier le manuscrit qui sont les suivants : <collection> (collection), <idno> (identifiant), <altIdentifier> (identifiant alternatif), <msName> (nom du manuscrit).

- <head > : cet élément contient tout type de position, par exemple, le titre d'une section, ou la tête d'une liste, un glossaire, la description du manuscrit, etc. Nous retrouvons en général dans cet élément les éléments de niveau de phrase : <title>, <name>, <date>.

Ou des éléments spécialisés tels que : <origPlace> (place d'origine), <origDate> (date d'origine)

- < msContents > :(contenu du manuscrit) décrit le contenu intellectuel d'un manuscrit ou d'une partie de manuscrit, soit comme une série de paragraphes ou comme une série d'éléments manuscrits structurés. Il contient trois éléments de base qui sont : <msContents>, <msItem> et <msItemStruct>.

L'élément <msItem> contient les éléments suivants : <author>, <respStmt>, <title>, <rubric>, <incipit>, <quote>, <msItem>, <msItemStruct>, <explicit>, <finalRubric>, <colophon>, <decoNote>, <listBibl>, <bibl>, <filiation>, <note>, <textLang>

- < physdesc > (description physique) : contient une description physique complète d'un manuscrit ou d'une partie de manuscrit, éventuellement subdivisé en utilisant des éléments plus spécialisés :

-Description de l'objet : constitué par les éléments relatifs au support : <supportDesc>, <support>, <extent>, <collation>, <foliation> et <condition> et des éléments relatifs à la mise en page <layout> et <layoutDesc>. Les éléments relatifs au support sont formés par des éléments de niveau de phrase combinés avec des éléments <p>.

-Description de l'écriture, de la décoration et autres annotations : Ce deuxième aspect de la description physique du manuscrit est décrit par les

éléments : <handDesc>, <handNote>, <decoDesc>, <musicNotation> et <additions>.

-Reliure et matière : Le troisième volet de la description physique du manuscrit est relatif à la reliure et la matière de fabrication du manuscrit. Il est défini par les éléments : <bindingDesc>, <binding>, <sealDesc>, <seal> et <accMat>.

- **<history>** : L'élément <history> définit l'historique du manuscrit ou d'une partie de manuscrit. Il contient les éléments suivants : <origine>, <provenance> et <acquisition>.
- **<additional>** : combine des informations bibliographiques sur un manuscrit, ou des copies de substitution de celui-ci avec des informations de conservation ou d'administration. Il est défini par trois éléments de base :
 - **<adminInfo>** : définit par <availability>, <custodialHist> et <recordHist> (défini par les éléments <source> et <change>).
 - **<surrogates>** : Permet de fournir les références sur d'éventuels microfilms du manuscrit.
 - **<listbibl>** : Contient la liste des citations bibliographiques du manuscrit.
- **<msPart>** (partie manuscrite) : l'élément <msPart> permettrait la description de chaque manuscrit comme un objet à part, d'une manière indépendante des autres volumes du manuscrit lorsque ce dernier est composé de plusieurs manuscrits
- **<msFrag>** (fragment de manuscrit) : contient des informations sur un fragment d'un manuscrit dispersé, tenue comme une seule unité ou liée dans un manuscrit plus grand.

II.5.2.6.3.3 Les éléments de niveau de phrase :

La TEI possède des éléments descriptifs pour décrire le contenu des manuscrits au niveau d'un paragraphe : <p>

Lorsque l'élément <msDesc> est utilisé au sein d'une description du manuscrit, de nombreux autres éléments de niveau phrase sont disponibles :

1-Origine :

Les éléments suivants peuvent être utilisés pour fournir des informations sur les origines de tout aspect d'un manuscrit:

- **<origDate>** (date de la création) : contient tout forme de date, Utilisée pour déterminer la date d'origine d'un manuscrit ou une partie d'un manuscrit.
- **<origPlace>** (lieu de création) : Contient tout forme de nom de lieu, utilisé pour déterminer le lieu d'origine pour un manuscrit ou une partie d'un manuscrit.

Les éléments <origDate>et <origPlace>sont spécialisés pour toutes formes existantes de la date et les noms a des places.

2- Timbre et filigranes :

Deux autres éléments sont fournis pour marquer d'autres éléments décoratifs caractéristiques des feuilles manuscrites et fixations:

- **<Watermark>** : contient un mot ou une phrase décrivant un filigrane ou un dispositif similaire.
- **<stamp>** : contient un mot ou une phrase décrivant un timbre ou un dispositif similaire.

3- Dimensions :

L'élément <dimensions> contient des spécifications de dimension, il peut être utilisé pour spécifier la taille de certains aspects du manuscrit, donc il peut être considéré comme une forme spécialisée de la TEI qui est dans l'élément < mesure>.

Les éléments suivants sont disponibles dans l'élément <dimensions> :

- **<height>** (hauteur) : élément décrivant la hauteur de la couverture du manuscrit, de ces folios ou de la surface écrite.
- **<width>** (largeur) : élément décrivant la largeur de la couverture du manuscrit, de ces folios ou de la surface écrite.
- **<thickness>** (épaisseur) : élément décrivant l'épaisseur du manuscrit.

4- Références à des emplacements dans un manuscrit :

L'élément<locus> et l'élément de regroupement <locusGrp>sont des formes spécialisées de l'élément<ref>, utilisé pour indiquer un emplacement, ou une séquence d'emplacements, dans un manuscrit.

<Locus> contient des attributs suivants :

- From (à partir de) spécifie le point de départ de l'emplacement sous une forme normalisée.
- To(à) spécifie le point final de l'emplacement sous une forme normalisée.
- Schema(régime) identifie le système de feuilletage en termes de laquelle l'emplacement est spécifié.

L'élément < locus> est utilisé pour faire référence à un seul emplacement dans un manuscrit, généralement pour spécifier l'emplacement occupé par l'élément dans lequel il apparaît. Si, par exemple, il est utilisé comme premier composant d'un élément <msItem> ou <msItemStruct>, ou un quelconque des éléments spécifique qui apparaissent dans la description, il est étendu pour spécifier l'emplacement(ou les emplacements) de ce point dans le manuscrit décrit.

5- Les noms des personnes, des lieux, et des organisations :

L'élément standard <name> de la TEI, peut être utilisé pour identifier les noms en toutes sortes qui se produisant dans une description.

6- Les mots clés, les signatures, Secundo Folio :

L'élément <catchwords> décrit le système utilisé afin d'assurer un ordre des folios du manuscrits, typiquement en insérant des annotations en bas de page.

L'élément <signateurs>est utilisé, d'une manier similaire, pour décrire un système dans lequel les feuilles sont marquées progressivement afin de faciliter l'arrangement, cependant l'élément <secFol> est utilisé pour enregistrer une phrase d'identification prise à partir d'un point connu spécifique dans un codex.

Conclusion :

Dans ce chapitre, nous avons présenté différents types du catalogage des manuscrits arabes anciens dans plusieurs bibliothèques et quelques institutions, et nous avons mis en place un protocole pour le catalogage en fusionnant les métadonnées constituant les notices bibliographiques de chacune d'elles, afin d'aboutir à une notice aussi complète que possible. Par la suite, nous avons passé à l'étude des formats d'encodage de métadonnées, et nous avons découvert l'existence d'un grand nombre de format d'encodage de métadonnées (METS, EAD, DUBLIN CORE, TEI...).

Chapitre III

Java EE

III.1 Introduction :

Dans le mode actuel de l'informatique et des technologies de l'information et de la communication, les applications sont de plus en plus complexes, et doivent être réalisées en un minimum de temps et de cout et portable à travers les différents systèmes utilisées. Sun Microsystems a proposée pour cela la plate-forme java Entreprise Edition ou java EE, crée dans le but de faciliter la construction d'applications distribuées, et a doté le langage java d'une plateforme logicielle fiable pour répondre aux besoins des entreprises.

III.2 Définition de Java EE :

Java Entreprise Edition(JEE) est une spécification pour le langage de programmation java de Sun destiné aux applications d'entreprise. Java EE offre une plate-forme de développement en langage Java pour les applications distribuées à plusieurs niveaux. On parle généralement de « plate-forme Java EE » pour désigner l'ensemble constitué des services API (Application Programming Interface) offerts et de l'infrastructure d'exécution. Dans la mesure ou Java EE s'appuie entièrement sur Java, il bénéficie des avantages de ce langage, en particulier une bonne portabilité et une maintenance du code. De plus, L'architecture Java EE repose sur des composants distincts, interchangeables et distribués, ce qui signifie notamment qu'un système reposants sur Java EE peut posséder des mécanismes de haute-disponibilité, afin de garantir une bonne qualité de service et que la maintenance des applications est facilité [29].

III.3 Objectifs Java EE :

Les objectifs de Java EE sont comme suit : [30]

- Faciliter le développement de nouvelles applications à base de composants
- Intégration avec les systèmes d'information existants
- Support pour les applications « critiques » de l'entreprise
- Disponibilité, tolérance aux pannes, montée en charge, sécurité ...
- une architecture d'applications basée sur les composants qui permet un découpage de l'application et donc une séparation des rôles lors du développement

III.4 Fonctionnement interne :

Le langage Java, sur lequel les bibliothèques Java EE sont utilisées, met à disposition un compilateur et une machine virtuelle (JVM – Java Virtual Machine) qui se charge de créer un environnement standard pour le lancement de l'application sur tout type de système opérationnel. Le compilateur compile le code source et produit le byte code, soit un code intermédiaire qui sera ensuite lu par la machine virtuelle Java. Chaque système opérationnel majeur possède une JVM expressément codée [31].

III.5 Les éléments de Java EE : [31]

- Une plate-forme (Java EE Platform), pour héberger et exécuter les applications, incluant outre Java SE des bibliothèques logicielles additionnelles du Java Développement Kit (JDK). Elle est basée sur des spécifications donc les projets sont portables sur n'importe quel serveur d'applications conforme (Tomcat, JBoss, WebSphere...).
- Une suite de tests (Java EE Compatibility Test Suite) pour vérifier la compatibilité.
- Une réalisation de référence (Java EE Reference Implementation) dénommée GlassFish.
- Un catalogue de bonnes pratiques (Java EE BluePrints).

III.6 L'architecture de Java EE : [32]

J2EE ajoute nombreuses couches de niveau entreprise au-dessus de la plate-forme J2SE - Java Standard Edition. Chaque couche est conçue pour supporter une différente technologie de développement.

- **Technologie web application:** technologies liées à la production des interfaces web dynamiques, par exemple JSP (Java Servlet Pages) et servlet
- **Technologie entreprise application:** technologies plus directement liées à la logique de business : EJB (Enterprise Java Bean), JavaMail, JMS (Java Message Service), JTA (Java Transaction), etc.
- **Technologie web services:** technologies utiles au développement des applications adhérentes au paradigme SOA (Service Oriented Architecture) : web services,

JAX-WS (Java API for XML-based web services), JAX-RPC (Java API for XML-Based RPC).

- **Technologie management and security** : technologies liées à la gestion de la technologie d'entreprise afin de réaliser l'accès et l'échange d'information entre machines et services distribués : JAAS (Java Authentication and Authorization Service), JCA (Java Connector Architecture).

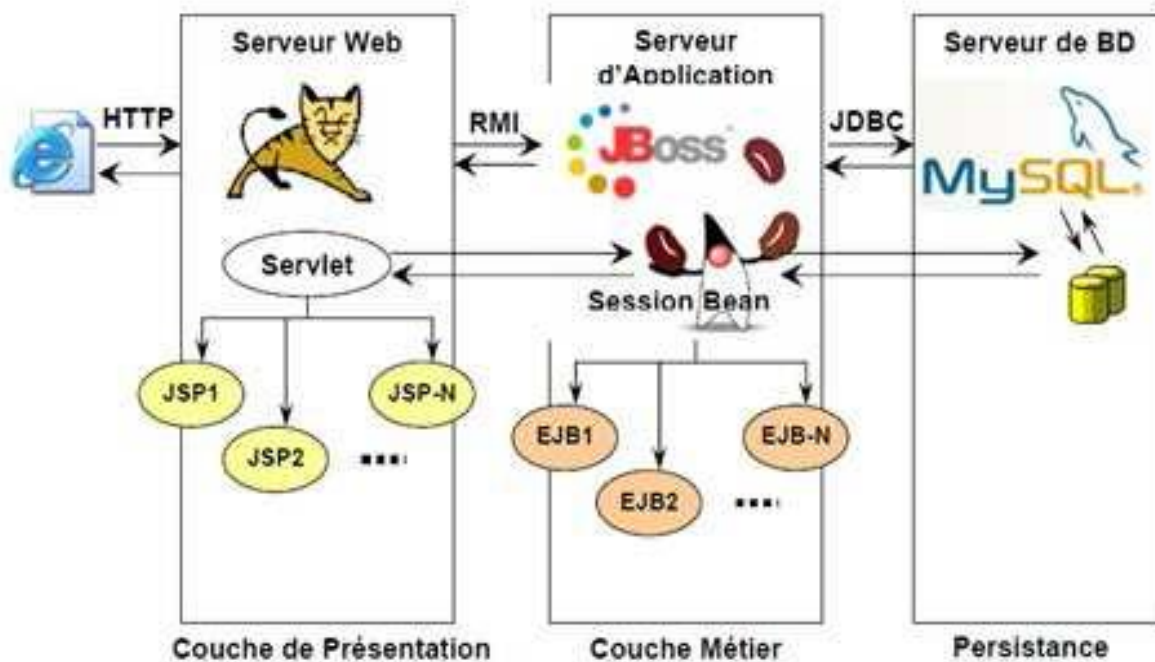


Figure III.1: L'architecture de Java EE

Pour expliquer l'utilisation de ces technologies on peut imaginer que les technologies entreprise sont utilisées pour gérer l'accès aux données (généralement une ou plusieurs data base), les technologies web application sont utilisées pour montrer les données aux utilisateurs génériques. Dans un contexte Business to Business, les technologies **Web service** seront utilisées pour échanger les informations avec les partenaires commerciales et les technologies de gestion gèrent tous les processus informationnels assurant la sécurité des transactions. Une application JavaEE est composée principalement de trois couches (pour cela on dit application 3-tiers ou n-tiers) :

Couche présentation : contient les vues que l'utilisateur vas interroger soit pour afficher des données ou envoyer des données.

Couche métier : la couche la plus lourde normalement les classes de traitement de données (le langage de l'application).

Couche de persistance(ou accès aux données) : la couche basse, c'est elle qui va stocker nos données (texte, XML, base de données, Excel).

III.7 Outils de programmation : [32]

Plusieurs environnements de développement intégrés (IDE) sont disponibles pour la plate-forme J2EE. Parmi les logiciels libres on rappelle Eclipse de IBM, qui met à disposition, grâce au système des plug-ins, un riche environnement de développement. Encore on retrouve NetBean de la Sun que, à différence d'Eclipse, a été créé exclusivement pour faciliter le développement d'applications J2EE.

III.8 Les conteneurs :

Les conteneurs assurent la gestion du cycle de vie des composants qui s'exécutent en eux. Les conteneurs fournissent des services qui peuvent être utilisés par les applications lors de leur exécution.

Il existe plusieurs conteneurs définis par Java EE [39] :

- **Conteneur web** : pour exécuter les servlets et les JSP
- **Conteneur d'EJB** : pour exécuter les EJB
- **Conteneur client** : pour exécuter des applications standalone sur les postes qui utilisent des composants Java EE

III.8.1 Le conteneur web : [33]

Le conteneur web est une implémentation des spécifications servlets et par extension des spécifications des JSP.

- **JSP : Java Server Page**

JSP est une application pour le développement de pages Web incluant du contenu dynamique. Contrairement à une page HTML qui ne contient que de contenu statique qui reste par définition toujours le même, JSP peut changer selon l'identité du visiteur, etc.

Une page JSP contient des balises standards, comme du HTML (ou du WML, XML...), comme toute page web normale. Portant, une page JSP contient aussi des éléments JSP spécifiques (scriptlets), permettent au serveur l'insertion dynamique de contenu (contenu de BDD, préférences du visiteur...). Lorsqu'un utilisateur accède à une page JSP, le serveur exécute les éléments JSP, fusionne les résultats avec les parties statiques de la page, et envoie

le tout au navigateur. JSP définit des outils standards utiles pour n'importe quelle application web comme l'accès à des composants JavaBeans, les contrôles d'accès entre pages, le partage d'information entre requêtes, pages et utilisateurs.

- **Servlets :**

Technologie Java utilisé pour effectuées des traitements coté serveur en réponse aux requêtes provenant en général de poste clients distants. Bien que les Servlets puissent répondre à n'importe quel type de requête, elles sont généralement employées pour répondre à des requêtes de type HTTP et qui permettent de retourner dynamiquement des pages HTML.

L'utilisation de servlets se fait par le biais d'un conteneur de servlets (framework) côté serveur. Celui-ci constitue l'environnement d'exécution de la servlet et lui permet de persister entre les requêtes des clients. L'API définit les relations entre le conteneur et la servlet. Le conteneur reçoit la requête du client, et sélectionne la servlet qui aura à la traiter. Le conteneur fournit également tout un ensemble de services standard pour simplifier la gestion des requêtes et des sessions.

III.8.2 Le conteneur d'EJB :

- **EJB ou Enterprise JavaBeans :**

Est une architecture de composants logiciels côté serveur pour la plateforme de développement Java EE. Cette architecture propose un cadre pour créer des composants distribués (c'est-à-dire déployés sur des serveurs distants) écrit en langage de programmation Java hébergés au sein d'un serveur applicatif permettant de représenter des données (EJB dit entité), de proposer des services avec ou sans conservation d'état entre les appels (EJB dit session), ou encore d'accomplir des tâches de manière asynchrone (EJB dit message). Tous les EJB peuvent évoluer dans un contexte transactionnel. C'est le serveur applicatif qui a en charge la création, la destruction, la passivation ou l'activation de ses composants en fonction des besoins.

III.9 Les API de JEE :

API (Application Programming Interface) (Interface Applicative de Programmation). Les API sont un moyen efficace de faire communiquer entre elles deux applications. J2EE regroupe un ensemble d'API pour le développement d'applications d'entreprise. Le tableau suivant représente quelque API ainsi que leur rôles [33]:

API	Rôle
Entreprise Java Bean (EJB)	Composants serveurs contenant la logique métier
Servlets	Composants basés sur le concept C/S pour ajouter des fonctionnalités à un serveur. Pour le moment, principalement utilisé pour étendre un serveur web
Java Server Pages (JSP)	
Java Naming and Directory Interface (JNDI)	Accès aux services de nommage et aux annuaires d'entreprises
Java Database Connectivity (JDBC)	Accès aux bases de données. J2EE intègre une extension de cette API
Java Transaction API (JTA) Java Transaction Service (JTS)	Support des transactions
Java Server Faces (JSF)	

Tableau III.1 : API JEE

III.10 Les services JEE : [34]

➤ Services d'infrastructures :

- JDBC (Java Database Connectivity) : API d'accès aux SGBDs
- JNDI (Java Naming and Directory Interface) : service de noms (annuaire) de référencement des objets
- JTA/JTS (Java Transaction API/Service) : service de gestion des transactions distribuées
- JCA (J2EE Connector Architecture) est une API de connexion au système d'information de l'entreprise.
- JMX (Java Management Extension) fournit des extensions permettant de développer des applications web de supervision d'applications.

➤ Services de communication :

- JAAS (Java Authentication and Authorization Service) est une API de gestion de l'authentification et des droits d'accès.

- Java Mail est une API permettant l'envoi de courrier électronique.
- JMS (Java Message Service) service de gestion des messages asynchrones (appelées MOM pour Middleware Object Message) entre applications.
- RMI-IIOP est une API permettant la communication synchrone entre objets.

III.11 Java EE et la sécurité : [35]

Vu la complexité des ces serveurs d'application il est très difficile d'assurer la protection contre les menaces et les intrusions dans ce type d'architecture, donc il faut faire attention aux risques que peut produire un serveur mal configuré et essayer de fermer les portes d'entrées que peuvent laisser de telles serveurs pour diminuer les risques d'attaque et protéger ce type d'application.

III.12 Les avantages d'utiliser Java EE :

L'utilisation de Java EE pour développer et exécuter une application représente plusieurs avantages :

- Une architecture d'application basée sur les composants qui permet un découpage de l'application et donc une séparation des rôles lors du développement.
- La possibilité de s'interfacer avec le système d'information existant grâce à de nombreuses API : JDBC, JNDI, JMS, JCA...
- La possibilité de choisir les outils de développement et les serveurs d'applications utilisés qu'ils soient commerciaux ou libres.

Conclusion :

Au cours de ce chapitre, nous avons présenté le fonctionnement, les éléments, l'architecture java EE, ainsi que les services de la Java EE, les conteneurs et l'environnement d'exécution.

Dans le chapitre qui suit nous allons présenter la phase d'analyse et conception.

Chapitre VI

Analyse et Conception

IV.1 Introduction :

Pour aboutir à une meilleure organisation et une bonne maîtrise du travail et donc arriver à déployer des meilleures applications, il est nécessaire de suivre une démarche méthodologique rigoureuse. Pour cela le choix d'une méthode de conception est d'une très grande importance.

Dans ce chapitre, nous allons présenter la démarche de modélisation pour le développement de notre application pour la mise en ligne des manuscrits arabes anciens numérisés. Pour se faire, nous introduisons un ensemble de concepts proposés par le langage UML (Unified Modeling Language) qui sera un support pour l'analyse et la spécification des besoins et aussi pour la conception de cette application.

IV.2 Objectifs de notre travail :

L'objectif principal de notre travail consiste à mettre en place une plate-forme de mise en ligne de manuscrits arabes anciens numérisés qui permet de :

- Donner un accès à distance aux manuscrits arabes anciens à une échelle internationale.
- Le catalogue des manuscrits arabes numérisés.
- Permettre l'accès à des documents qui, souvent ne sont pas consultable pour des raisons de conservation.

IV.3 Description de l'UML : [annexe]**IV.4 Analyse et conception :****IV.4.1 Analyse :**

Cette activité commence par la spécification des besoins de notre application, et s'enchaîne par l'identification des acteurs, et vient ensuite la définition de toutes les tâches et les scénarios appropriés.

IV.4.1.1 Spécification des besoins :

Cette étape permet de définir les besoins et les objectifs du système à concevoir ainsi que les différents cas d'utilisation existants. Le modèle résultant de cette étape nous donne le diagramme de cas d'utilisation de chacun de ces derniers.

IV.4.1.2 Identification des acteurs :

- ✚ **Définition** : un acteur représente un rôle joué par une personne ou une chose qui interagit avec le système, autrement dit c'est un objet actif qui utilise les fonctions du Système.

Dans le cas de notre application, nous avons sept acteurs qui interagissent avec le système :

- ❖ **Le visiteur** : toute personne visite le site.
- ❖ **L'adhérent** : toute personne inscrite à la bibliothèque et pouvant bénéficier de ses services de notre application.
- ❖ **Le catalogueur** : toute personne chargée de la gestion de la base de documentaire.
- ❖ **Le responsable d'orientation** : c'est la personne qui permet d'établir les accès vers le système aux divers utilisateurs externes à la bibliothèque (adhérents et visiteurs).
- ❖ **Le producteur** : toute personne chargé de versement de nouveaux manuscrits ou notices dans le système.
- ❖ **Le numériseur** : toute personne chargée de la numérisation des manuscrits.
- ❖ **L'administrateur** : est la personne chargée de la gestion administrative.

IV.4.1.3 Spécification des taches :

- ✚ **Définition** : Une tache est l'ensemble des différents fonctions qui peut accéder un acteur bien spécifique. Les acteurs définis précédemment effectuent un certain nombre de taches, ces taches sont résumées dans le tableau ci-dessous :

Acteur	Taches
Visiteur	T01 : Se connecter au site. T02 : Chercher des manuscrits. T03 : Naviguer dans le site. T04 :S'inscrire.
Adhérent	T05 : Se connecter au site. T06 :S'authentifier. T07 : Chercher des manuscrits. T08 : Se déconnecter.
Catalogueur	T09 : Se connecter au site. T10 :S'authentifier.

	<p>T11 : Chercher des manuscrits.</p> <p>T12 : Gérer les notices.</p> <p>T13 : Se déconnecter.</p>
Administrateur	<p>T14 : Se connecter au site.</p> <p>T15 : S'authentifier.</p> <p>T16 : Chercher des manuscrits.</p> <p>T17 : Gérer les catalogueurs.</p> <p>T18 : Gérer les adhérents.</p> <p>T19 : Changer le mot de passe.</p> <p>T20 : Se déconnecter.</p>
Producteur	<p>T21 : Se connecter au site.</p> <p>T22 : S'authentifier.</p> <p>T23 : Valider le versement du manuscrit.</p> <p>T24 : Demande de numérisation.</p> <p>T25 : Etablissement des statistiques.</p> <p>T26 : Se déconnecter.</p>
Responsable d'orientation	<p>T27 : Se connecter au site.</p> <p>T28 : S'authentifier.</p> <p>T29 : Gérer les visiteurs et les adhérents.</p> <p>T30 : Gérer l'accès ver le système.</p> <p>T31: Réception des commandes.</p> <p>T31 : Se déconnecter</p>
Numériseur	<p>T32 : Se connecter au site.</p> <p>T33 : S'authentifier.</p> <p>T34 : Numériser les manuscrits.</p> <p>T35 : Corriger les manuscrits.</p> <p>T36: Se déconnecter</p>

Tableau IV.1 : tableau des cas d'utilisation

IV.4.1.4 Spécification des scénarios :

✚ **Définition de scénario :** Un scénario est une description narrative de comment est-ce que le système pourra être utilisé. Les scénarios doivent être décrits par les utilisateurs eux même, chacune des tâches effectuées par un ou plusieurs acteurs sera décrite par un ensemble de scénarios.

Les tableaux si dessous résumant les différents scénarios par tâches :

❖ Le visiteur :

Tache	Scénarios
T01 : Se connecter au site.	S01 : Saisir l'URL du site dans le navigateur.
T02 : Chercher des manuscrits.	S02 : Choisir la langue de recherche, S03 : Lancement de la recherche. S04 : Consulter les résultats
T03 : Naviguer dans le site.	S05 : sélectionner un lien. S06 : suivre le lien.
T04 :S'inscrire.	S07 : sélectionner le lien « s'inscrire ». S08 : remplir le formulaire d'inscription à envoyer.

Tableau IV.2 : Spécification des scénarios de « Visiteur »

❖ L'adhérent :

Tache	Scénarios
T05 : Se connecter au site.	S09 : Saisir l'URL du site dans le navigateur.
T06 :S'authentifier.	S10 : Saisir login et le mot de passe. S11 : cliquer sur connexion.
T07 : Chercher des manuscrits.	S12 : Choisir la langue de recherche. S13 : Saisir la requête. S14 : Lancement de la recherche. S15 : Consulter les documents. S16 : Télécharger le support PDF.
T08 : Se déconnecter.	S17 : Cliquer sur le lien « déconnexion ».

Tableau IV.3 : Spécification des scénarios de « Adhérent »

❖ Le catalogueur :

Tache	Scénarios
T09 : Se connecter au site.	S18 : Saisir l'URL du site dans le navigateur.
T10 :S'authentifier.	S19 : Saisir login et le mot de passe. S20 : cliquer sur connexion.
T11 : Chercher des manuscrits.	S21 : Choisir la langue de recherche. S22 : Saisir la requête. S23 : Lacement de la recherche. S24 : Télécharger le support PDF.
T12 : Gérer les notices.	S25 : Choisir la langue du catalogage. S26 : Ajouter des notices, S27 : Modifier des notices, S28 : Supprimer des notices.
T13 : Se déconnecter.	S29 : Cliquer sur le lien « déconnexion ».

Tableau IV.4 : Spécification des scénarios de « Catalogueur »

❖ L'administrateur :

Tache	Scénarios
T14 : Se connecter au site.	S30 : Saisir l'URL du site dans le navigateur.
T15 :S'authentifier.	S31 : Saisir login et le mot de passe. S32 : cliquer sur connexion.
T16 : Chercher des manuscrits.	S33 : Choisir la langue de recherche. S34 : Lacement de la recherche. S35 : Consulter les résultats.
T17 : Gérer les catalogueurs.	S36 : Ajouter des catalogueurs. S37 : Modifier des catalogueurs. S38 : Supprimer des catalogueurs. S39 : Afficher liste des catalogueurs.
T18 : Gérer les adhérents.	S40 : Ajouter des adhérents. S41 : Modifier des adhérents. S42 : Supprimer des adhérents. S43 : Afficher liste des adhérents.
T19 : Changer le mot de passe.	S44 : Cliquer sur le lien «changer le mot de passe». S45 :remplir le formulaire de changement.

	S46: valider le changement.
T20 : Se déconnecter.	S47: Cliquer sur le lien « déconnexion ».

Tableau IV.5 : Spécification des scénarios de « Administrateur »

❖ **Responsable d'orientation :**

Tache	Scénarios
T27 : Se connecter au site.	S48: Saisir l'URL du site dans le navigateur.
T28 : S'authentifier.	S49 : Saisir login et le mot de passe. S50 : cliquer sur connexion.
T29 : Gérer les visiteurs / adhérents.	S51: Ajouter des visiteurs/adhérents. S52: Modifier des visiteurs/adhérents. S53: Supprimer des visiteurs/adhérents.
T30 : Gérer l'accès ver le système.	S54: arrêt d'accès ver le système. S55 : prolongation d'accès ver le système.
T31: Réception des commandes.	S56 : En clique sur le lien « réception des commandes ».
T32 : Se déconnecter	S57: Cliquer sur le lien « déconnexion ».

Tableau IV.6: Spécification des scénarios de « Responsable d'orientation »

IV.4.1.5 Diagramme représentatif :**Le diagramme de contexte de l'application :**

Le diagramme de contexte permet de spécifier le nombre d'instances d'acteurs connectés à ce système à un moment donné. La définition des besoins fonctionnels des acteurs présentée dans le paragraphe, nous permet de dégager le diagramme de contexte suivant :

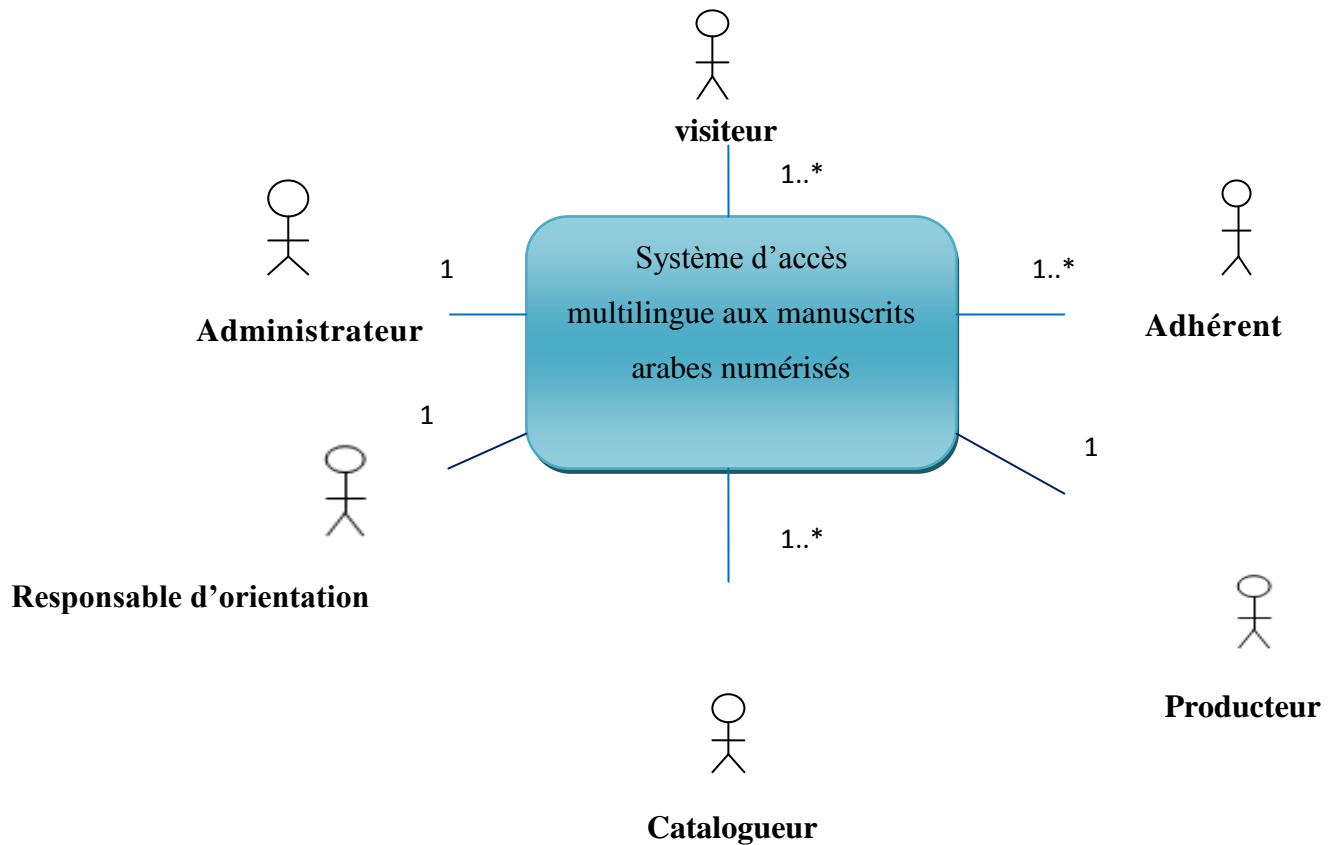


Figure VI.7 : Diagramme de contexte

IV.4.1.6 Les diagrammes de cas d'utilisation :

UML définit une notation graphique pour représenter les cas d'utilisation. Cette notation est appelée diagramme de cas d'utilisation qui fait intervenir les acteurs et les cas d'utilisation eux-mêmes. Il est utilisé pour donner une vision globale du comportement fonctionnel de l'application modélisée. Il existe deux types de relation standard entre cas d'utilisation qui sont proposés par UML [36] :

- ❖ **La relation « include »** : Une relation inclusion d'un cas d'utilisation A par rapport à un cas d'utilisation B, signifie qu'une instance de A contient le comportement décrit dans B, le cas d'utilisation A ne peut pas être utilisé seul.
- ❖ **Relation « extend »** : Une relation d'extension d'un cas d'utilisation A par rapport à un cas d'utilisation B, signifie qu'une instance de A peut être étendue par le comportement décrit dans B.

Le diagramme de cas d'utilisation générale de notre système est le suivant :

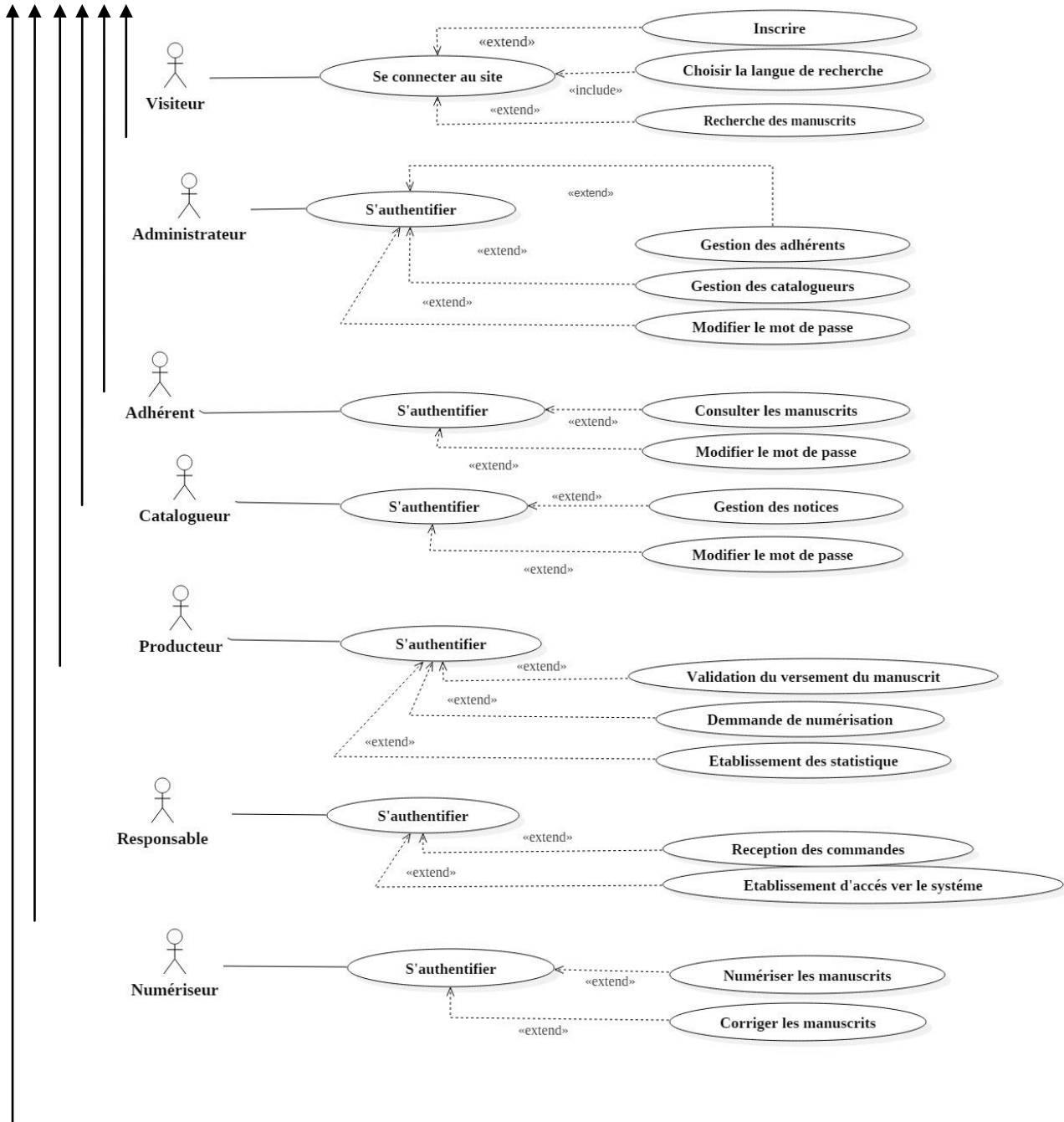


Figure IV.8: Diagramme de cas d'utilisation général

Quelque diagramme des cas d'utilisation détaillé de notre système sont:

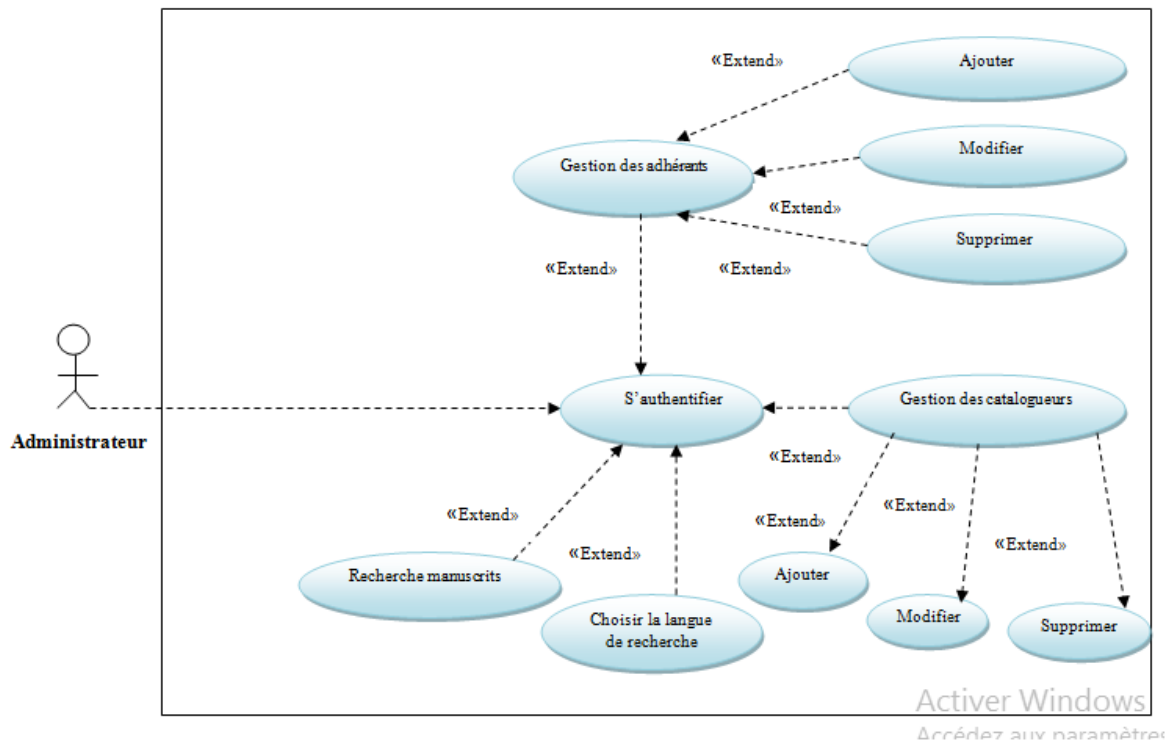


Figure IV.9: Diagramme de cas d'utilisation « administrateur »

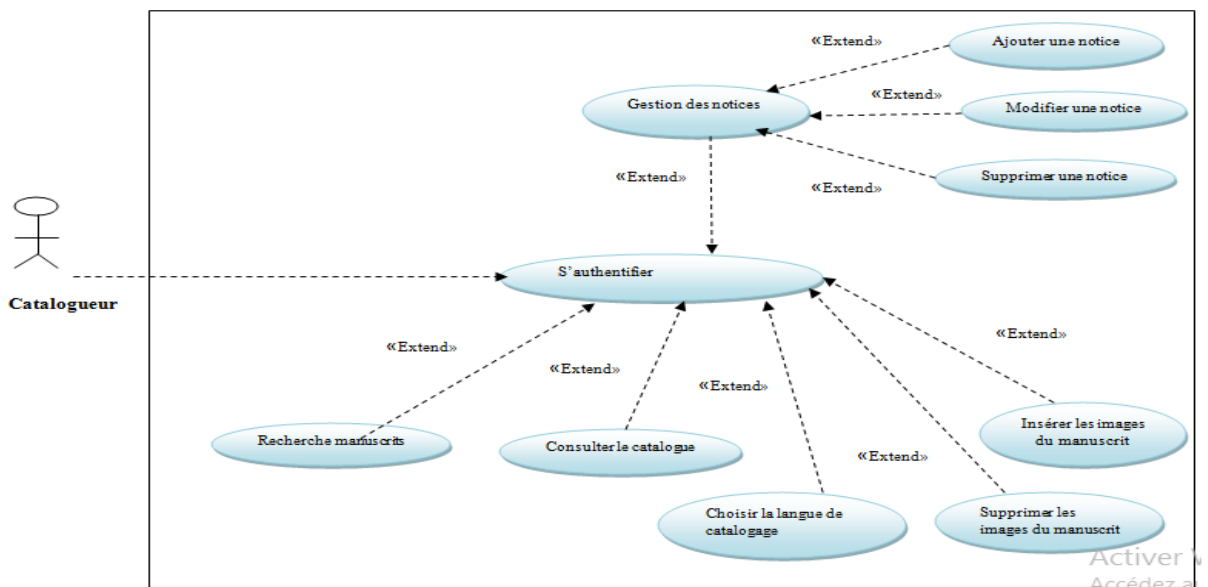


Figure IV.10: Diagramme de cas d'utilisation « Catalogueur »

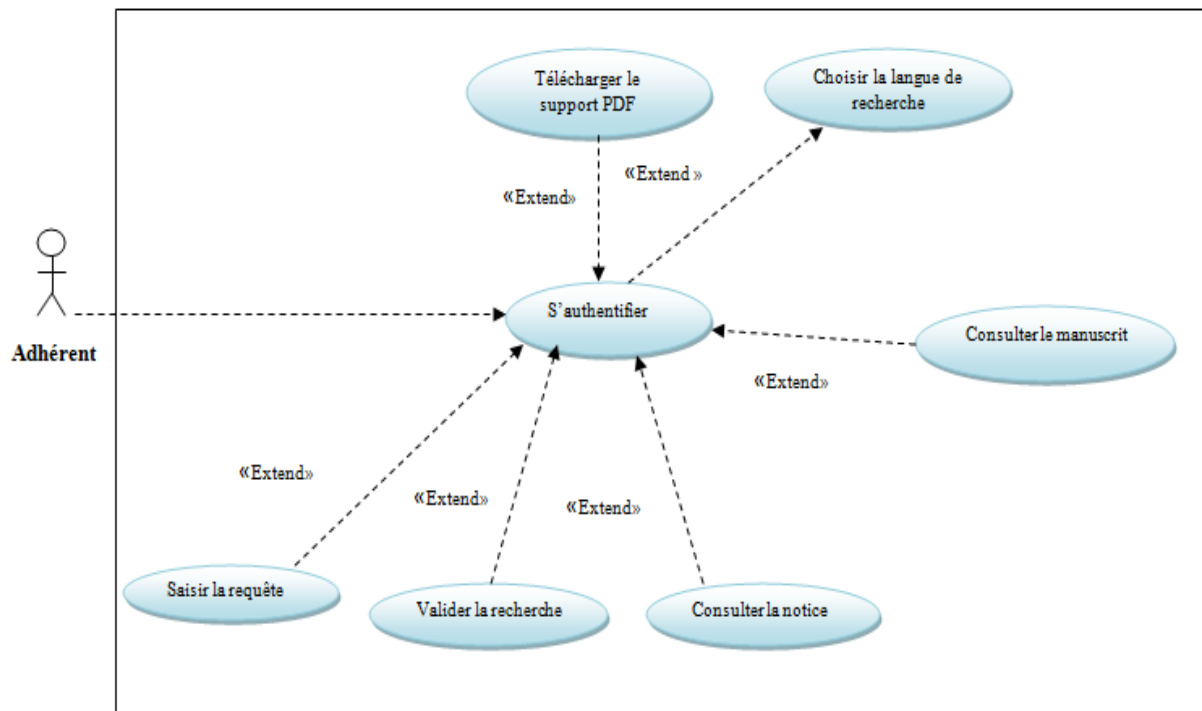


Figure IV.11 : Diagramme de cas d'utilisation «recherche adhérent»

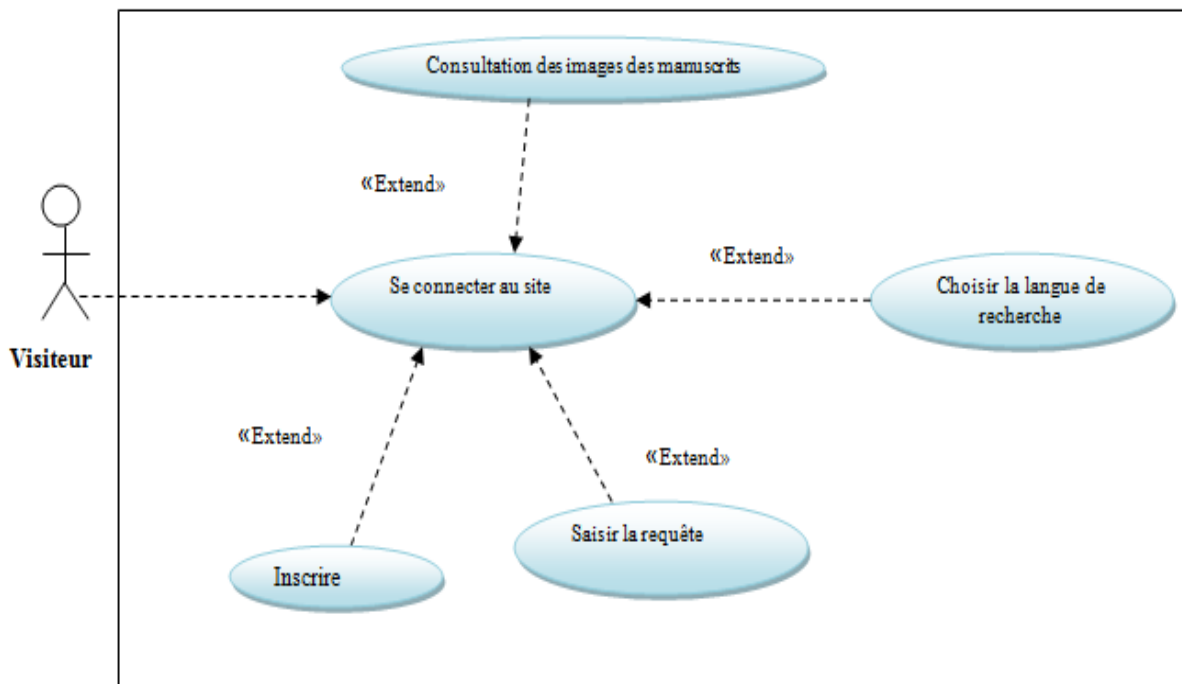


Figure IV.12: Diagramme de cas d'utilisation «Visiteur»

IV.5 Conception :

La conception est la phase la plus complexe du projet, elle est en grande partie un processus d'affinement du modèle d'analyse, en intégrant les besoins non fonctionnels du système et les contraintes d'architecture, elle affine le modèle d'analyse jusqu'à l'obtention d'une structure qui puisse être codée.

Dans cette phase une nouvelle vue du modèle fait son apparition, elle exprime les modules et les exécutable physiques sans aller à la réalisation concrète du système.

IV.5.1 Elaboration des Diagrammes de séquences

- ✚ **Définition :** Les diagrammes des séquences permettent de représenter les interactions entre objet selon un point de vue temporel .l'accent est mis sur la chronologie des envois de messages [37].
- ❖ **Scénario :** une liste d'actions qui décrivent une interaction entre un acteur et le système.
- ❖ **Interaction :** Un comportement qui comprend un ensemble de messages échangés par un ensemble d'objets dans un certain contexte pour accomplir une certaine tâche.
- ❖ **Message :** Un message est une transmission d'information unidirectionnelle entre deux objets, l'objet émetteur et l'objet récepteur.

Nous présentons dans ce qui suit quelques diagrammes de séquence pour les cas d'utilisations précédemment définis :

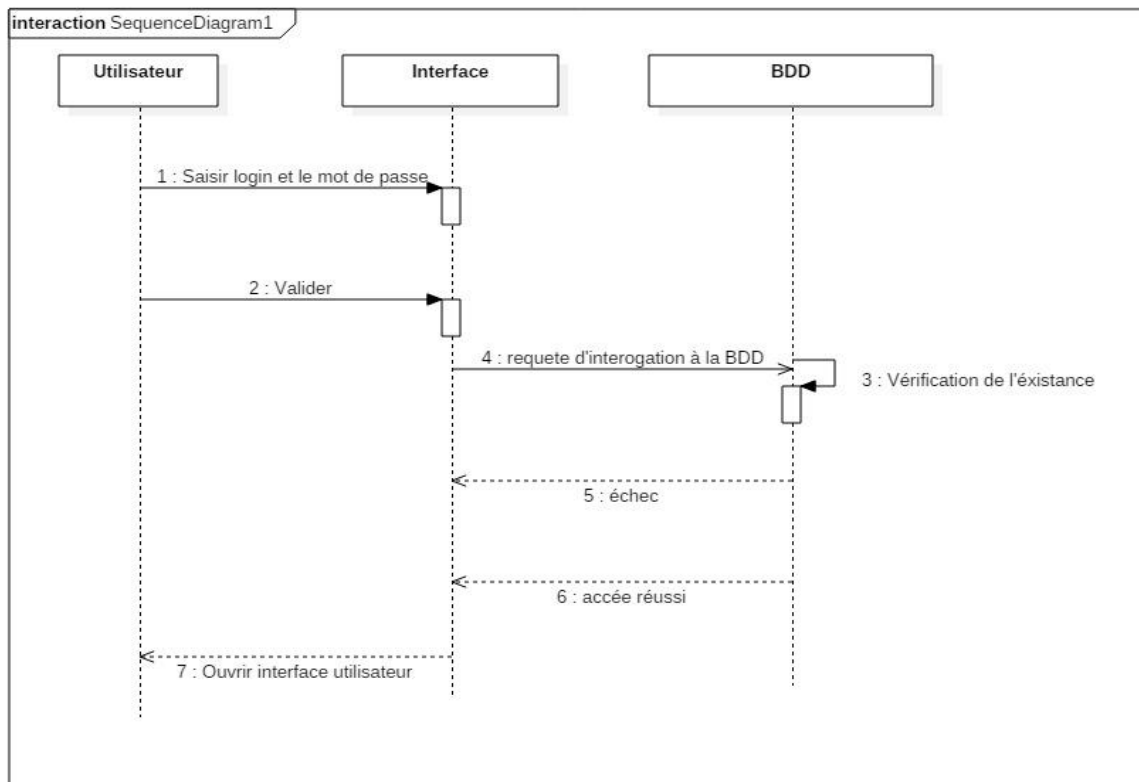


Figure IV.13 : diagramme de séquence du cas d'utilisation «Authentification»

Description textuelle :

1. l'utilisateur entre son login et mot de passe.
2. l'utilisateur valide.
3. le système interroge la BDD, vérifier les informations de l'utilisateur et retourne un message de confirmation si non il retourne une page d'erreur.
4. le système ouvre une interface utilisateur si le login et le mot de passe sont corrects.

Remarque : l'utilisateur peut être un adhérent, un catalogueur, un administrateur, un producteur, un numériser, un responsable d'orientation.

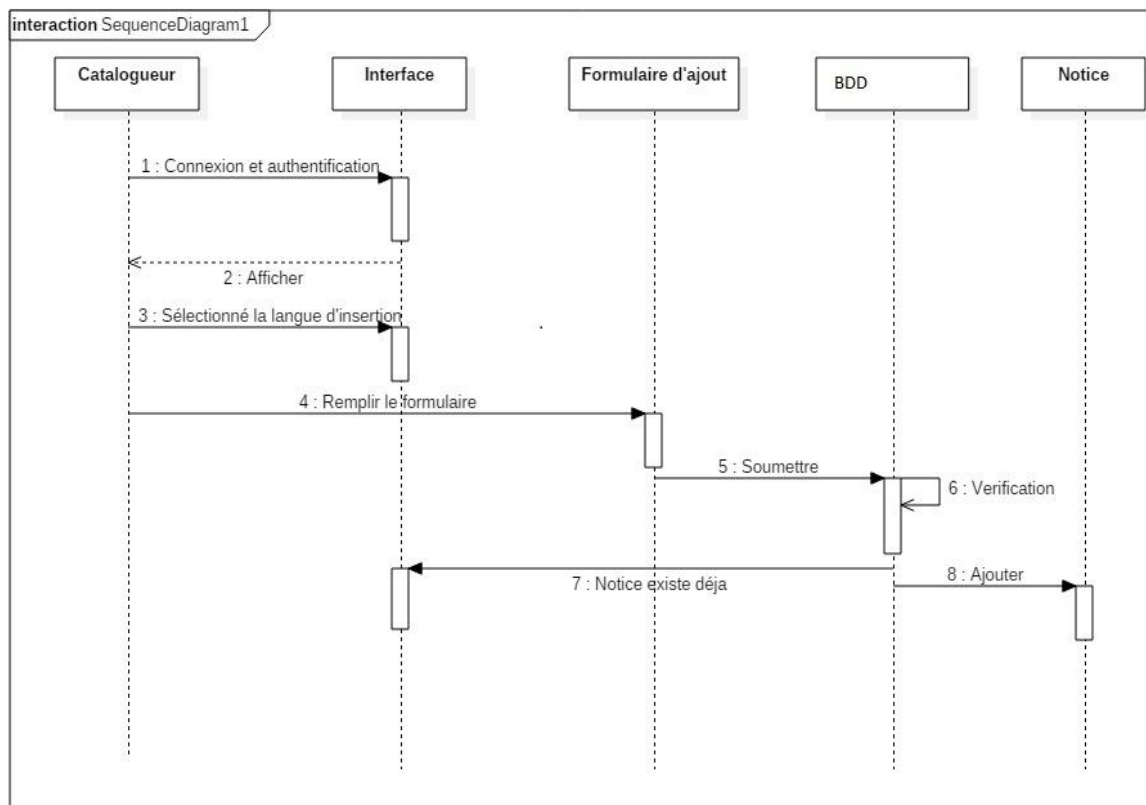


Figure IV.14: diagramme de séquence du cas d'utilisation «Ajouter une notice»

Description textuelle :

1. le catalogueur s'authentifie à son interface.
2. le système affiche l'interface principale.
3. le catalogueur sélectionne la langue d'insertion.
4. le catalogueur remplit le formulaire d'ajout d'une notice.
5. le système soumet les informations saisies dans le formulaire à la BDD, vérifie les informations et retourne un message de confirmation d'ajout si non il retourne une page d'erreur si la notice existe déjà.

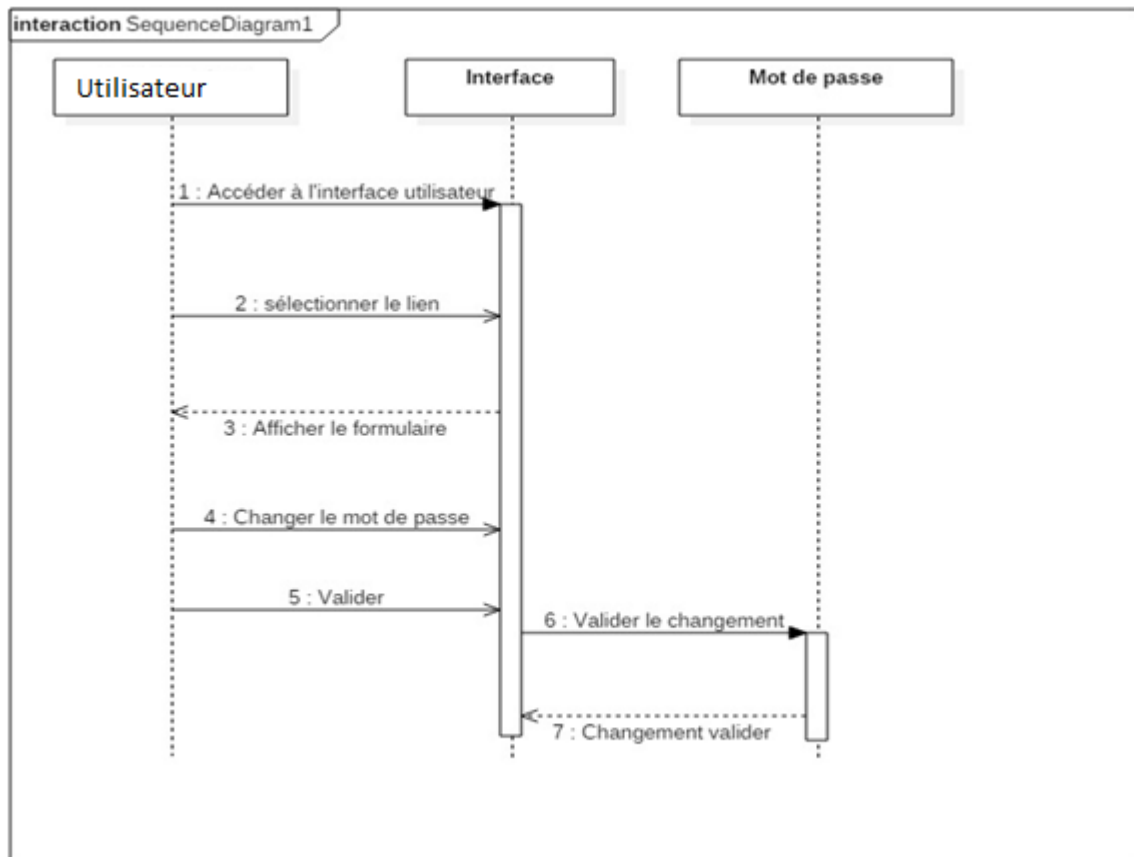


Figure IV.15 : diagramme de séquence du cas d'utilisation «**modifier le mot de passe**»

Description textuelle :

1. l'utilisateur accède à son l'interface.
2. l'utilisateur sélectionne le menu changement de mot de passe.
3. le système affiche un formulaire pour changer le mot de passe.
4. l'utilisateur remplir le formulaire.
5. l'utilisateur valide.
6. le système affiche un message de confirmation.

Remarque : l'utilisateur peut être un adhérent, un catalogueur, un administrateur, un producteur, un numériser, un responsable d'orientation.

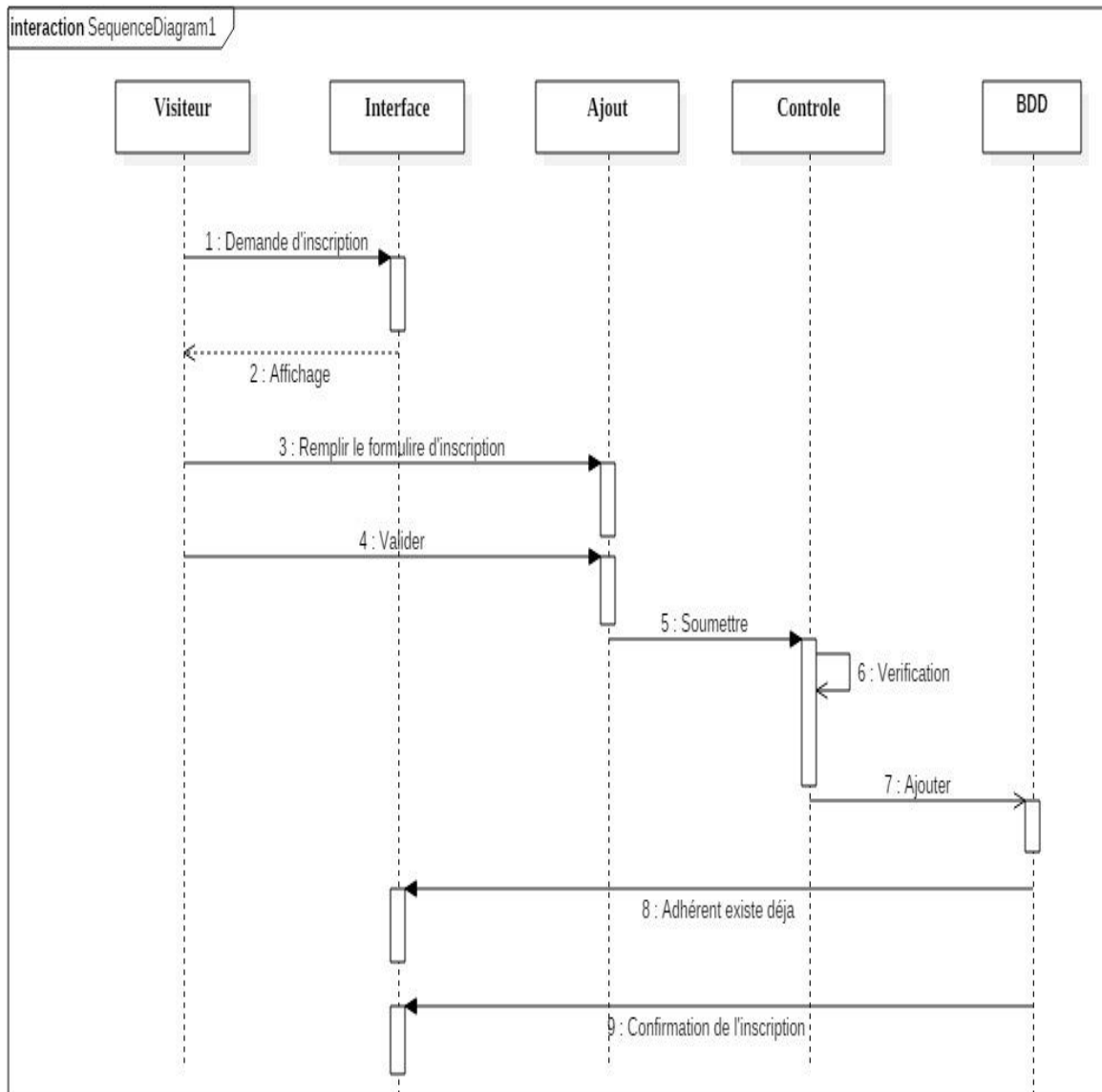


Figure IV.16 : diagramme de séquence du cas d'utilisation «Inscription»

Description textuelle :

1. le visiteur demande le formulaire d'inscription.
2. le formulaire s'affiche.
3. le visiteur remplit le formulaire.
4. une vérification de l'existence de l'adhérent dans la base se lance.
5. si l'adhérent existe déjà un message d'erreur s'affiche.
6. si c'est un nouvel adhérent, confirmation de l'inscription s'affiche.

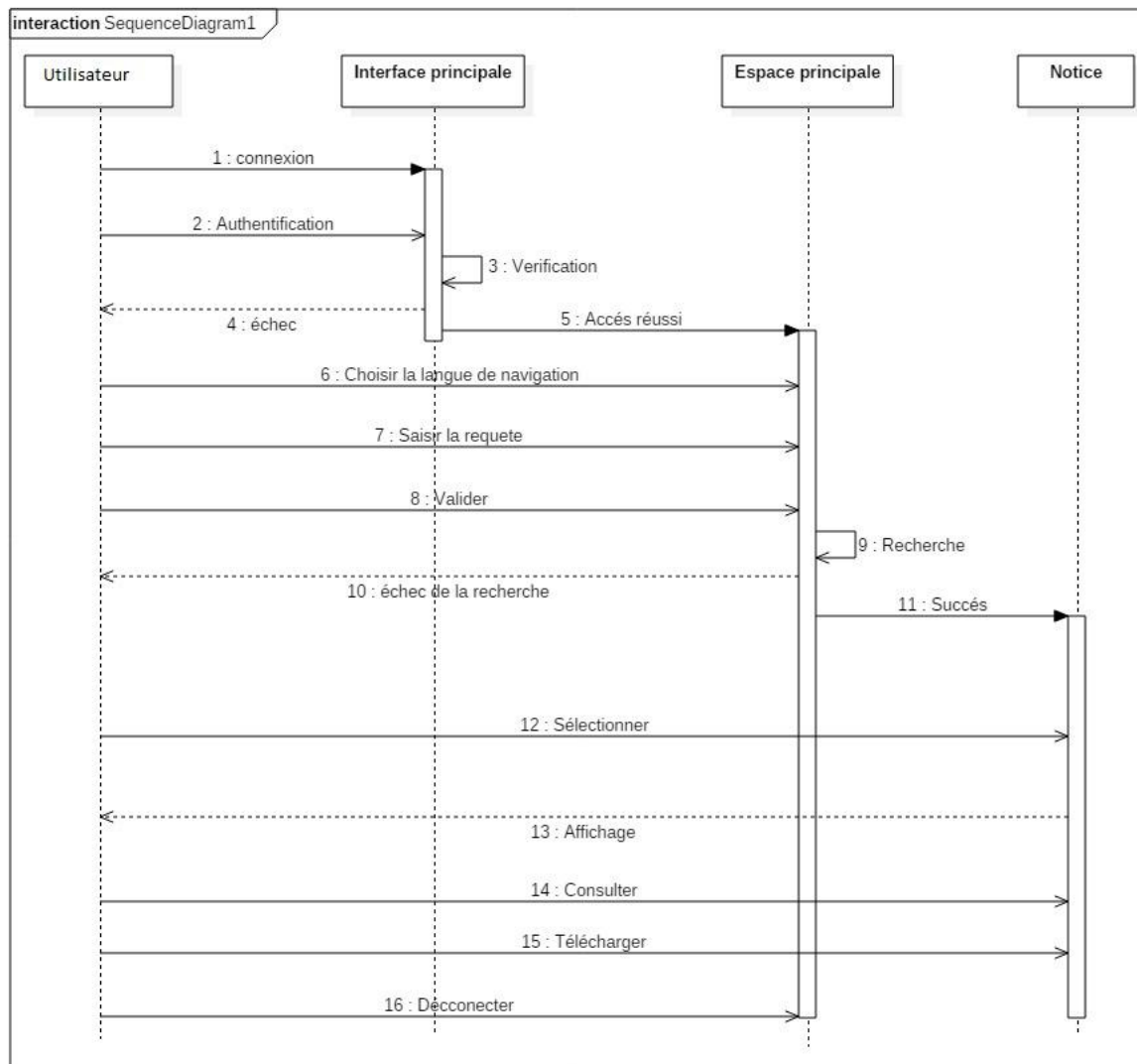


Figure IV.17: diagramme de séquence du cas d'utilisation « Recherche un manuscrit »

Description textuelle :

1. l'utilisateur se connecter et s'authentifier à son interface.
2. le système vérifier les informations de l'utilisateur et retourne un message de confirmation si non il retourne une page d'erreur.
3. le système ouvre une interface utilisateur si le login et le mot de passe sont corrects.
4. l'utilisateur choisir la langue de recherche puis saisir la requête.
5. l'utilisateur valide.
6. une vérification de la recherche se lance.
7. l'utilisateur sélectionne la notice si la recherche est réussie.
8. le système affiche les notices.
9. l'utilisateur consulte les notices puis les télécharger.

10. l'utilisateur quitte son espace.

Remarque : l'utilisateur peut être un adhérent, un catalogueur, un administrateur, un producteur, un numériseur, un responsable d'orientation.

IV.5.2 Elaboration des Diagrammes d'activités :

✚ **Définition :** Le diagramme d'activité est un diagramme comportemental d'UML, permettant de représenter le déclenchement d'événement en fonction des états du système et de modéliser des comportements parallélisables [38].

On représente ci-dessous une description de quelques diagrammes d'activités :

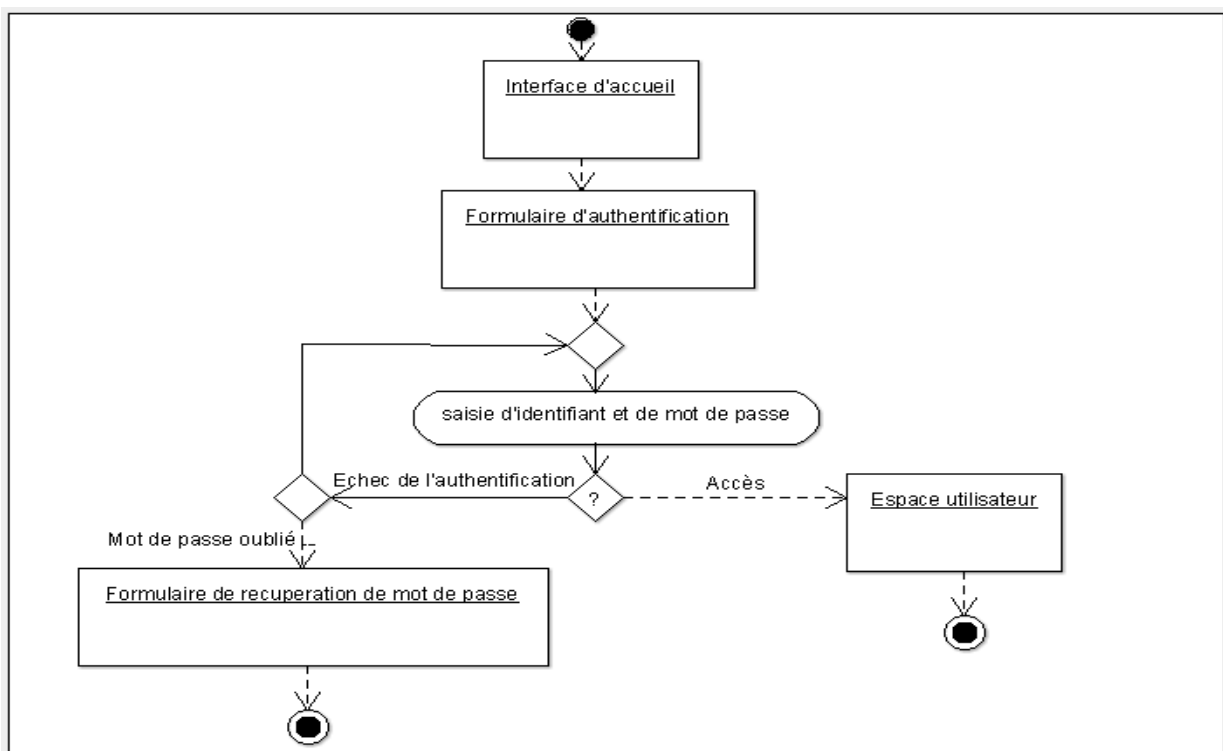


Figure IV.18 : Diagramme d'activité de cas d'utilisation « authentification »

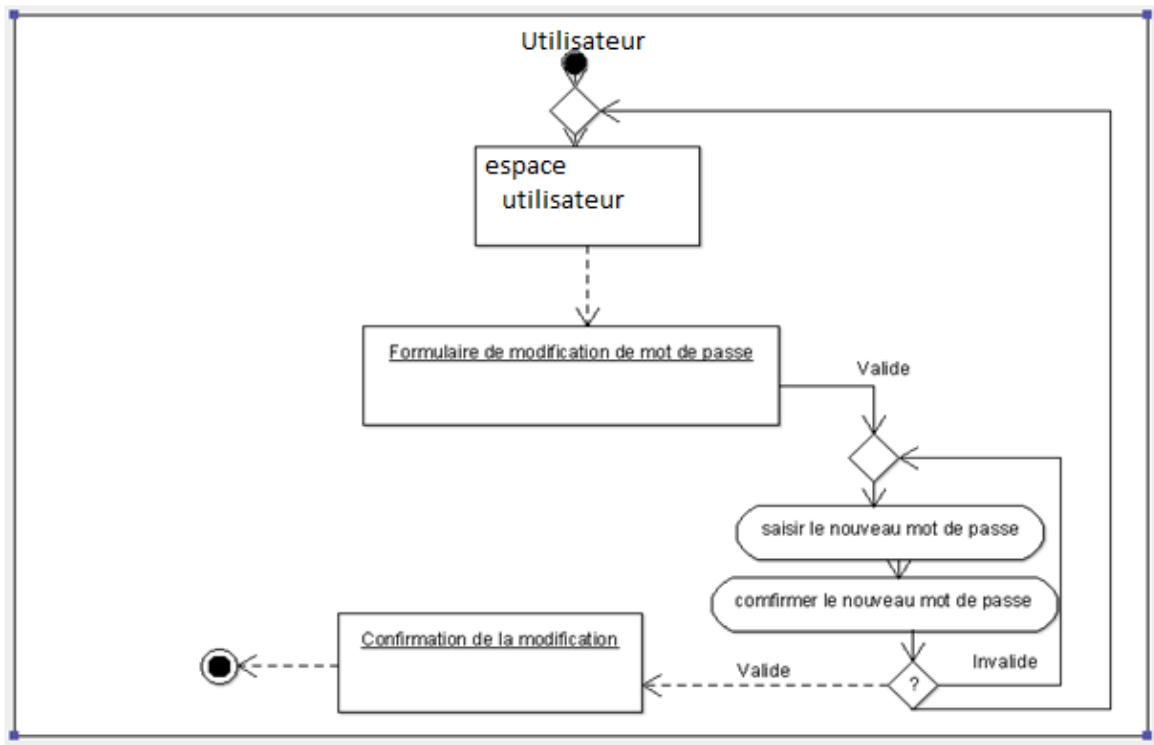


Figure IV.19: Diagramme d'activité de cas d'utilisation « **Modification de mot de passe** »

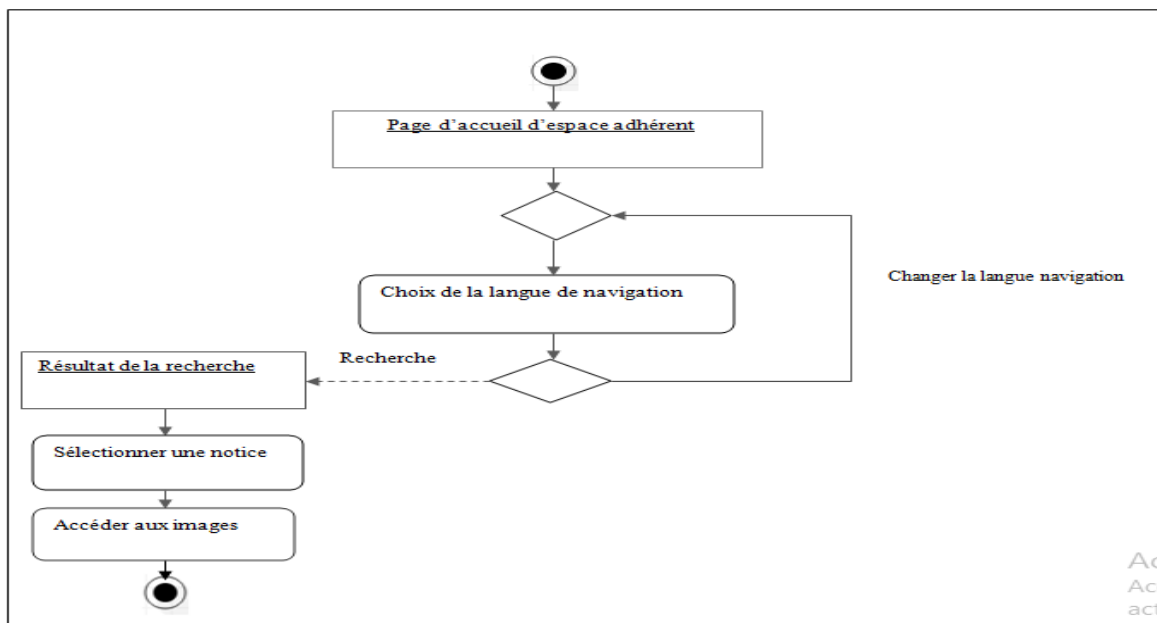


Figure IV.20: Diagramme d'activité de cas d'utilisation « **rechercher un manuscrit** »

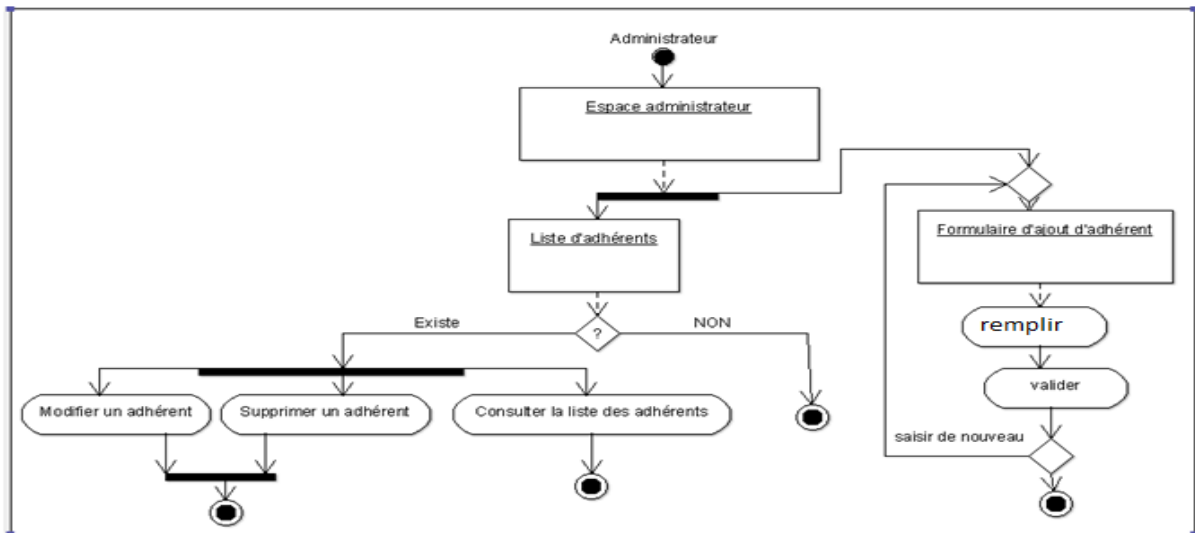


Figure IV.21 : Diagramme d'activité de cas d'utilisation « Gestion des adhérents »

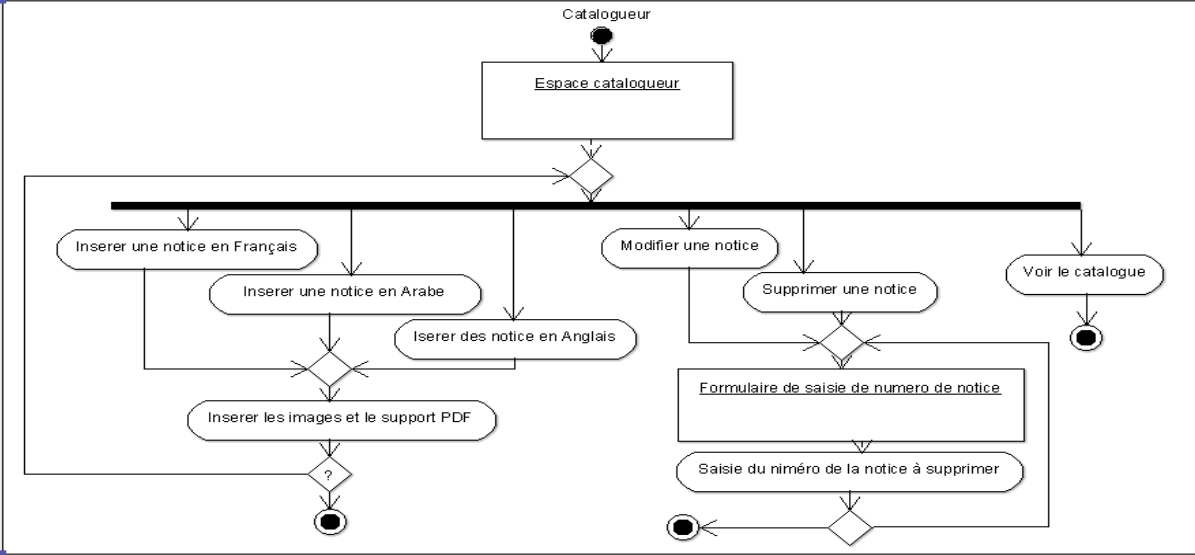


Figure IV.22 : Diagramme d'activité de cas d'utilisation « Gestion des notices »

IV.5.3 Elaboration des Diagrammes de classe :

🔧 Définition : Le diagramme de classe est un diagramme structurel qui exprime d'une manière générale la structure statique d'un système en termes de classe ainsi que leurs relations.

IV.5.4 Conception de modèle de donné :

Après avoir modélisé le site avec les différents diagrammes vu précédemment (d'utilisation, de séquence, d'activité et de classe général), nous nous intéressons maintenant aux données nécessaires pour le fonctionnement de l'application. Pour les obtenir, nous suivons de près le déroulement de chaque cas d'utilisation et les diagrammes des classes détaillés qui nous renseignent sur les données nécessaires, qui seront stockées dans les tables de la base de données.

❖ Table administrateur :

Nom de champ	Type de donnée	Taille	Description	Observation
Id_admin	Int	10	Identificateur de l'administrateur	Primary key
Login	Varchar	100	Login de l'administrateur	-
Pass	Varchar	100	Mot de passe de l'administrateur	-

❖ Table catalogueur :

Nom de champ	Type de donnée	Taille	Description	Observation
Idcatalogueur	Int	11	Identificateur de catalogueur	Primary key
Nomcatalogueur	Varchar	254	Nom catalogueur	-
Prenomcatalogueur	Varchar	254	Prénom catalogueur	-
Datecatalogage	varchar	20	Date de naissance du catalogueur	-
Numero_manuscrit	varchar	20	Numéro du manuscrit	Foreign key

❖ Table adhérent :

Nom de champ	Type de donnée	Taille	Description	Observation
Num_carte	Int	10	Numéro de carte d'adhésion	Primary key
Passerword	Varchar	10	Mot de passe adhérent	-
Nom	text	-	Nom adhérent	-
Prénom	text	-	Prénom adhérent	-
Date_naiss	Date	-	Date de naissance de l'adhérent	-
Adresse	text	-	Adresse de l'adhérent	-
Adresse_mail	Varchar	20	Adresse mail catalogueur	-

❖ Table responsable :

Nom de champ	Type de donnée	Taille	Description	Observation
Id_res	Int	11	Identificateur de responsable	Primary key
nom	Varchar	100	Nom de responsable	-
prenom	Varchar	100	Prénom de responsable	-
email	Varchar	100	Email de responsable	-
login	Varchar	100	Login de responsable	-
pass	Varchar	100	Mot de passe de responsable	-

❖ Table producteur :

Nom de champ	Type de donnée	Taille	Description	Observation
Id_prod	Int	11	Identificateur de producteur	Primary key
nom	Varchar	100	Nom de producteur	-
prenom	Varchar	100	Prénom de producteur	-
email	Varchar	100	Email de producteur	-
login	Varchar	100	Login de producteur	-
pass	Varchar	100	Mot de passe de producteur	-

❖ Table numériseur :

Nom de champ	Type de donnée	Taille	Description	Observation
Id_numér	Int	11	Identificateur de numériseur	Primary key
nom	Varchar	100	Nom de numériseur	-
prenom	Varchar	100	Prénom de numériseur	-
email	Varchar	100	Email de numériseur	-
login	Varchar	100	Login de numériseur	-
pass	Varchar	100	Mot de passe de numériseur	-

❖ Table visiteur :

Nom de champ	Type de donnée	Taille	Description	Observation
numcart	Varchar	10	Numéro de la carte	Primary Key
dateinscription	Date	-	Date d'inscription	-
datefin	Date	-	Date fin	-
nom	Varchar	100	Nom de visiteur	-
prenom	Varchar	100	Prénom de visiteur	-
etude	Varchar	254	Etude de visiteur	-
specialite	Varchar	100	Spécialité de visiteur	-
fonction	Varchar	100	Fonction de visiteur	-
adresse	Varchar	254	Adresse de visiteur	-

❖ Table notice :

Nom de champ	Type de donnée	Taille	Description	Clé(s)
Numeromanuscrit	varchar	50	numéro de la notice	Primary key
Numerosecondaire	varchar	50	Numéro secondaire de la notice	-
Numérosequentiel	varchar	50	Numéro séquentiel de la notice	-
Numerovolume	varchar	50	numéro de volume de la notice	-
Premierpage	Varchar	50	Premier page de la notice	-
Dernierpage	Varchar	50	Dernier page de la notice	-
Numerodinscription	varchar	50	Numéro d'inscription	-
Numeropartie	varchar	50	Numéro de partie de la notice	-
Numerofolios	varchar	50	Numéro des folios de la notice	-
Etatmanuscrit	text		Etat du manuscrit	-

Foliosdetachees	text		Folios des taches	-
Reliure	text		Reliure de la notice	-
Rabat	text	-	Rabat de la notice	-
Nombredecopie	varchar	50	Nombre de copie de la notice	-
Incipit	text	-	Incipit de la notice	-
Explicit	text	-	Explicit de la notice	-
Langue	Varchar	100	Langue de la notice	-
Typeecriture	Varchar	100	Type de l'écriture de la notice	-
Typeencre	varchar	254	Type d'encre de la notice	-
Couleurencre	varchar	254	Couleur d'encre de la notice	-
Resume	varchar	10	Résumé de la notice	-
Dessins	text	-	dessins	-
Corrections	text	-	Corrections de la notice	-
Consultation	text	-	Consultation de la notice	-
Ecoutes	text	-	Ecoutes de la notice	-
Arrets	text	-	Arrêts	-
Aviores	text	-	Avoirs	-
Inspections	text	-	Inspections de la notice	-
Editionimprimees	text	-	Édition de l'imprimé de la notice	-
Traduction	text	-	Traduction du la notice	-
referencebibliographique	text	-	Référence bibliographique de la notice	-
Conservation	text	-	Conservation de la notice	-
Urne	text	-	Urne de la notice	-
Remarques	text	-	Remarques	-

❖ **Table message :**

Nom de champ	Type de donnée	Taille	Description	Observation
Id_mesg	Int	11	Numéro de message	Primary key
prenom	varchar	200	Prénom de l'expéditeur	-
nom	varchar	100	Nom de l'expéditeur	-
email	varchar	100	Email de l'expéditeur	-
sujet	varchar	200	Sujet de message	-
message	Text		Email de l'expéditeur	-
date	bigint	20	Date message	-

❖ **Table sujet :**

Nom de champ	Type de donnée	Taille	Description	Observation
Id_sujet	Int	11	Identificateur du sujet	Primary key
Sujet	varchar	255	sujet	-
Sujettrans	varchar	255	Sujet transcription	-
Numéromanuscrit	Varchar	255	Numéro du manuscrit	Foreign key

❖ **Table titre :**

Nom de champ	Type de donnée	Taille	Description	Observation
Id_titre	Int	11	Identificateur du titre	Primary key
Titre	varchar	254	titre	
Titretrans	text		Titre transcription	
Numeromanuscrit	varchar	20	Numéro du manuscrit	Foreign key

❖ **Table auteur :**

Nom de champ	Type de donnée	Taille	Description	Observation
Idauteur	Int	11	Identificateur du l'auteur	Primary key
Nomauteur	varchar	25	Nom de l'auteur	-
nomtran	Varchar	254	Nom transcription de l'auteur	-
deceshegire	Varchar	254	Décès hégire de l'auteur	-
decesgregorien	Varchar	254	Décès grégorien de l'auteur	-
numeromanus	varchar	20	Numéro de manuscrit	Foreign key

❖ **Table valider :**

Nom de champ	Type de donnée	Taille	Description	Observation
numeromanuscrit	varchar	50	Numéro du manuscrit	Foreign Key
valider	varchar	50	valider	-
AR	varchar	10	arabe	-
FR	varchar	10	français	-
AN	varchar	10	anglais	-
biblioinsert	int	11	Insert un manuscrit	-
dateinsert	date	-	Date insertion	-
bibliosupp	int	-	Supprimer un manuscrit	-
datesupp	date	-	Date	-

❖ **Table image :**

Nom de champ	Type de donnée	Taille	Description	Observation
numeromanuscrit	varchar	254	Numéro de manuscrit	Foriegne Key
lienpdf	varchar	254	Lien PDF du l'image	-
Vise 1	varchar	254	Visionneuse du l'image	-

Conclusion :

A l'issue de ce chapitre nous avons introduit les objectifs principaux de notre application. Pour les atteindre, nous avons proposé une solution basée sur une analyse et une conception selon l'approche orientée objet, et cela en s'aidant des différents outils de modélisation du langage UML, qui facilite la maîtrise et la production de logiciel.

Après avoir analysé la problématique en profondeur et effectué la conception de notre système, nous sommes désormais en mesure d'entamer l'étape d'implémentation de notre application qui sera présentée dans le chapitre qui suit.

Chapitre V

Implémentation et Réalisation

V.1 Introduction :

Après avoir présenté dans le chapitre précédent les différentes étapes de la conception de notre système, nous allons présenter dans ce chapitre le choix technologique, les outils qui ont servi à la réalisation de notre travail et fonctionnalités de notre application à travers des exemples réelle de ces différentes interfaces.

V.2 choix technologique :

La première question qui s'est posée pour la réalisation de la partie technique du projet a été de choisir une architecture pour l'application.

V.2.1 L'architecture MVC (Modèle-Vue-Contrôleur) :

Plusieurs architectures sont proposées pour les projets informatiques de plus en plus complexes à réaliser mais surtout à maintenir, une bonne organisation du code est fortement recommandée afin de faciliter ces deux aspects. Dans cette optique, nous avons opté l'architecture MVC pour Modèle-Vue-Contrôleur qui est devenue très courante pour la réalisation des gros projets.

L'architecture Modèle-Vue-Contrôleur est une manière d'organiser son code et une méthode de développement professionnel d'application logicielle. Ce modèle d'architecture impose la séparation entre les données, en l'occurrence celles du traitement de l'information et de sa mise en forme, ce qui donne trios parties fondamentales dans l'application finale : le modèle, la vue et le contrôleur :

- Les données(Modèle).
- L'interface home-machine(Vue).
- La logique de contrôle(Contrôleur).

Grossièrement, cela permet une séparation entre les traitements de données et la présentation.

➤ Le modèle :

Le Modèle représente le comportement de l'application : traitement des données, interactions avec la base de données, etc. Il décrit les données manipulées par l'application et définit les méthodes d'accès. Dans une application web, le modèle est souvent le code qui permet de faire de requête à la base de données (récupérer, insérer et mettre à jour des informations de la base de données).

➤ **La vue :**

La Vue correspond à l'interface qui verra l'utilisateur et avec laquelle il interagira. Les résultats renvoyés par le modèle sont dénués de toute présentation mais sont représentés par les vues. Plusieurs vues peuvent afficher les informations d'un même modèle. Elle est conçue en html. La vue n'effectue aucun traitement, elle se contente d'afficher les résultats des traitements effectués par le modèle, et de permettre à l'utilisateur d'interagir avec elles.

➤ **Le contrôleur :**

Il est juste la passerelle entre le modèle, l'interface graphique et l'utilisateur qui prend en charge la gestion des événements de synchronisation pour mettre à jour la vue ou le modèle et effectuer la synchronisation entre eux. Il n'effectue aucun autre traitement, ne modifier aucune donnée, il analyse la requête du client et se contente d'appeler le modèle adéquat et renvoyer la vue correspondant à la demande.

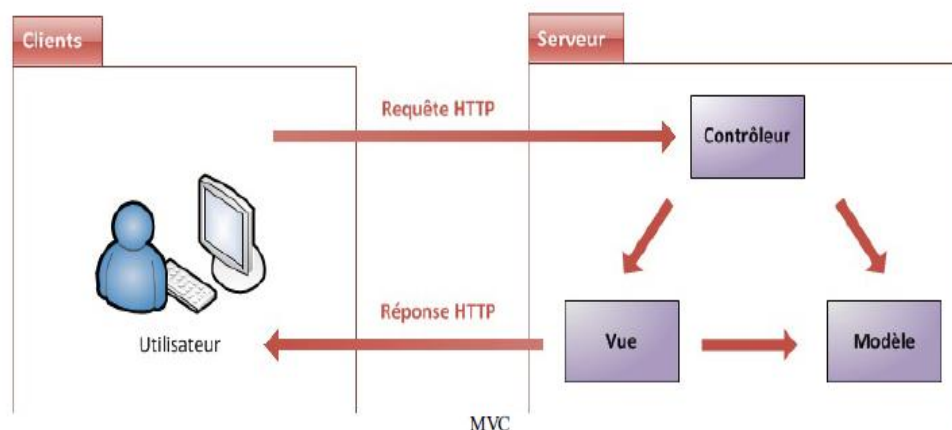


Figure V.1 :L'architecture MVC

V.3 Présentation des langages de programmation utilisés :

V.3.1 Le langage java EE :

Java Enterprise Edition, ou Java EE (anciennement J2EE), est une spécification pour la technique Java d'Oracle plus particulièrement destinée aux applications d'entreprise. Ces applications sont considérées dans une approche multi-niveaux. Dans ce but, toute implémentation de cette spécification contient un ensemble d'extensions au *framework* Java standard (JSE, *Java Standard Edition*) afin de faciliter notamment la création d'applications

répartie [39]. L'utilisation de J2EE pour développer et exécuter une application offre plusieurs avantages :

- une architecture d'applications basée sur les composants qui permet un découpage de l'application et donc une séparation des rôles lors du développement
- la possibilité de s'interfacer avec le système d'information existant grâce à de nombreuses API : JDBC, JNDI, JMS, JCA ...
- la possibilité de choisir les outils de développement et le ou les serveurs d'applications utilisés qu'ils soient commerciaux ou libres.

V.3.2 Les JSP :

Les JSP (Java Server Pages), une technologie Java qui permet la génération de pages web dynamiques. La technologie JSP permet de séparer la présentation sous forme de code HTML et les traitements sous formes de classes Java définissant une servlet. Ceci est d'autant plus facile que les JSP définissent une syntaxe particulière permettant d'appeler une servlet et d'insérer le résultat de son traitement dans la page HTML dynamiquement [40].

V.3.3 Les servlets :

Une servlet est un programme qui s'exécute côté serveur en tant qu'extension du serveur. Elle reçoit une requête du client, elle effectue des traitements et renvoie le résultat. La liaison entre la servlet et le client peut être directe ou passer par un intermédiaire comme par exemple un serveur http.

V.3.4 Le langage Java Script :



JavaScript est langage de scripts incorporé aux balises HTML s'exécutant dans le navigateur sans faire appel aux ressources du serveur, permettant d'effectuer des contrôles de saisie pour valider les formulaires et aussi d'améliorer la présentation. Il est plus simple à mettre en œuvre, car c'est du code que vous ajouterez à votre page écrite en HTML [41].

V.3.5 Le langage SQL(Structured query language) :

C'est un langage permettant d'interroger les bases de données d'une manière simple. Il est doté d'une syntaxe particulière que l'on doit respecter pour que la communication avec la base soit crédible. Son succès est essentiellement à sa simplicité, et à son attachement à la stratégie de SGBD.

❖ Quelques avantages de MYSQL :

Les principaux concurrents de MYSQL sont PostgreSQL, Microsoft SQL Server et Oracle.

Par rapport à ces produits MYSQL possède plusieurs avantages majeurs, dont les suivants :

- Des performances élevées.
- Sa simplicité de configuration et d'apprentissage.
- Un coût réduit.
- Sa portabilité.
- L'accessibilité de son code source.
- La disponibilité du support.

V.3.6 Le langage HTML (Hyper Text Markup Language) :

Il a fait son apparition dès 1991 lors du lancement du web. Son rôle est de gérer et organiser le contenu. C'est donc en HTML que vous écrivez ce qui doit être affiché sur la page : du texte, des liens, des images.....

Le langage HTML permet notamment la lecture de documents sur internet à partir de machines différentes, grâce au protocole http, permettant d'accéder via le réseau à des documents repérés par une adresse unique, appelée URL.

V.3.7 Le langage CSS (Cascading Style Sheets) :

C'est un langage qui permet d'arranger le contenu des pages HTML et de définir la présentation : couleurs, image de fond, marge, taille du texte....le principe des feuilles de styles consiste à regrouper dans un même document des caractéristiques de mise en forme associées à des groupes d'éléments. Il suffit de définir par un nom, un ensemble de définition

et de caractéristiques de mise en forme, et de l'appeler pour l'appliquer a un texte. Les feuilles de styles ont été mises au point afin de récompenser les manques du langage HTML.

V.4 Les outils de développement :

Les outils de développement utilisés sont :

V.4.1 NetBeans version 8.1 :

NetBeans est un environnement de développement intégré(EDI) à code ouvert (open source) permettant d'écrire, compiler, déboguer, déployer des programmes. En plus NetBeans permet également de supporter différents autres langages, comme Python, C, C++, JavaScript, XML, PHP... . Il comprend toutes les caractéristiques d'un éditeur graphique d'interfaces et de pages Web. Conçu en Java, NetBeans est disponible sous Windows, Linux, Solaris.

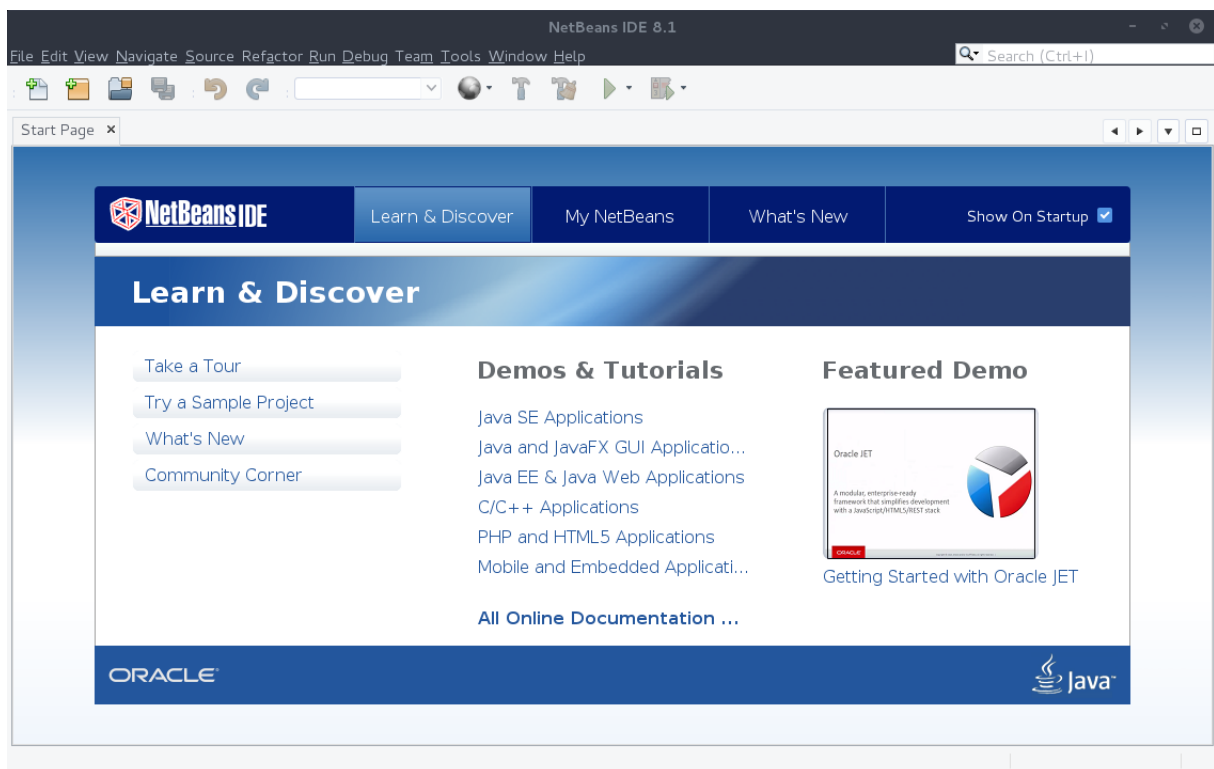


Figure V.2 : interface NetBeans version 8.1

V.4.2 WampServer :



WampServer est une plate-forme de développement Web sous Windows. Il permet de développer des applications Web dynamiques à l'aide du serveur Apache, du langage de scripts PHP et d'une base de données MySQL. Il possède également PHPMyAdmin pour gérer plus facilement les bases de données.

WampServer s'installe facilement et son utilisation très intuitive permet de le configurer très rapidement (sans toucher aux fichiers de configuration). Une fois la base installée, WampServer offre la possibilité d'ajouter autant de versions d'Apache, MySQL et PHP que l'on souhaite.

WampServer dispose également d'un "TrayIcon" permettant de gérer et configurer simplement les serveurs.

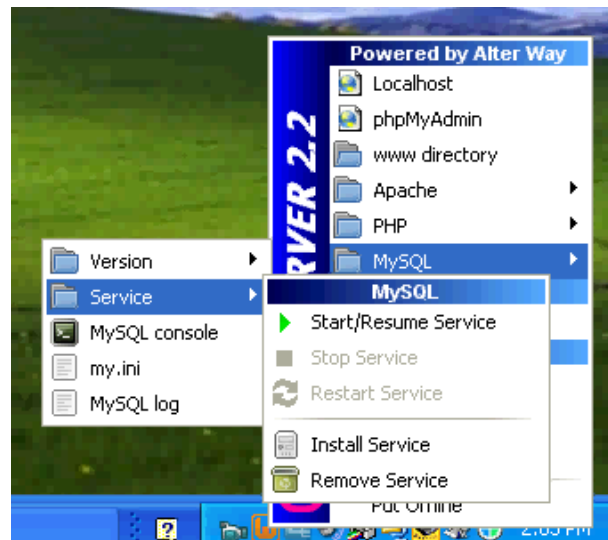


Figure V.3: Panneau de gestion du serveur

Les fonctionnalités de WampServer sont très complètes et très intuitives, via un click gauche sur l'icône de WampServer, on peut notamment :

- Gérer les services d'Apache et MySQL.
- Passer en mode online/offline (accessible à tous ou limité à localhost).
- Installer et changer des versions d'Apache, MySQL et PHP.
- Gérer les paramètres de configuration des serveurs.
- Accéder aux fichiers de configuration Créer des alias.

V.4.3 Le serveur Apache :



Apache HTTP Server, communément appelé Apache, est un serveur http en Open Source produit par Apache Software Fondation. Il est, à l'heure actuelle, le serveur le plus utilisé sur le web.

Différentes fonctionnalités sont implémentées comme la possibilité d'utiliser un seul serveur Internet pour héberger plusieurs sites, l'utilisation des langages interprétés Perl, PHP et Python, sauvegarde des accès dans un fichier log (statistiques), htaccess pour la protection des répertoires et l'URL-Rewriting, Le serveur apache est:

- Disponible sur pratiquement toutes les plates formes (Unix, linux, WindowsNT et Windows95/98).
- Un niveau élevé de performance et des exigences matérielles modestes.
- Gratuit.
- Robuste et sécurisé.
- Son développement est actif.

Dans notre cas apache est intégré à IDE NetBeans 8.1 dans sa version 8.0.20, qui s'appelle Apache Tomcat 8.0.20.

V.4.4 Apache Tomcat :

Apache Tomcat est un conteneur Web libre de Servlets et JSP java EE. Il implémente les spécifications des servlets et des JSP du Java Community Process, est paramétrable par des fichiers XML et de propriétés, et inclut des outils pour la configuration et la gestion .Il comporte également un serveur HTTP.

Tomcat ne constitue qu'un conteneur Web, et non un serveur Web à part entière: il (gère spécifiquement les servlets et les JSP (un compilateur Jasper compilant les pages JSP pour en faire des servlets).IL peut être également parfois désigné comme moteur de servlet.

Tomcat est en réalité souvent employé en combinaison avec un serveur Web Apache ou d'autres serveurs Web. Tomcat a été écrit en langage java, il peut donc s'exécuter via la machine virtuelle java sur n'importe quel système d'exploitation la supportant.

V.4.5 L'interface graphique PhpMyAdmin :

Il s'agit de l'une des plus célèbres interfaces pour gérer une base de données MySQL sur un serveur PHP. De nombreux hébergeurs, qu'ils soient gratuits ou payants, le proposent ce qui permet à l'utilisateur de ne pas avoir à l'installer.

Intégralement développé en PHP, phpMyadmin permet une administration aisée des bases de données MySQL. C'est un outil libre et gratuit développé par la communauté des programmeurs libres.

Cette interface pratique permet d'exécuter, très facilement et sans grandes connaissances dans le domaine des bases de données, de nombreuses requêtes comme la création de tables de données, les insertions, les mises à jour, les suppressions, les modifications de structure de la base de données. Ce système est très pratique pour sauvegarder une base de données sous forme de fichier .SQL et ainsi transférer facilement ces données. De plus celui-ci accepte la formulation de requêtes SQL directement en langage SQL, cela permet de tester ses requêtes par exemple lors de la création d'un site.

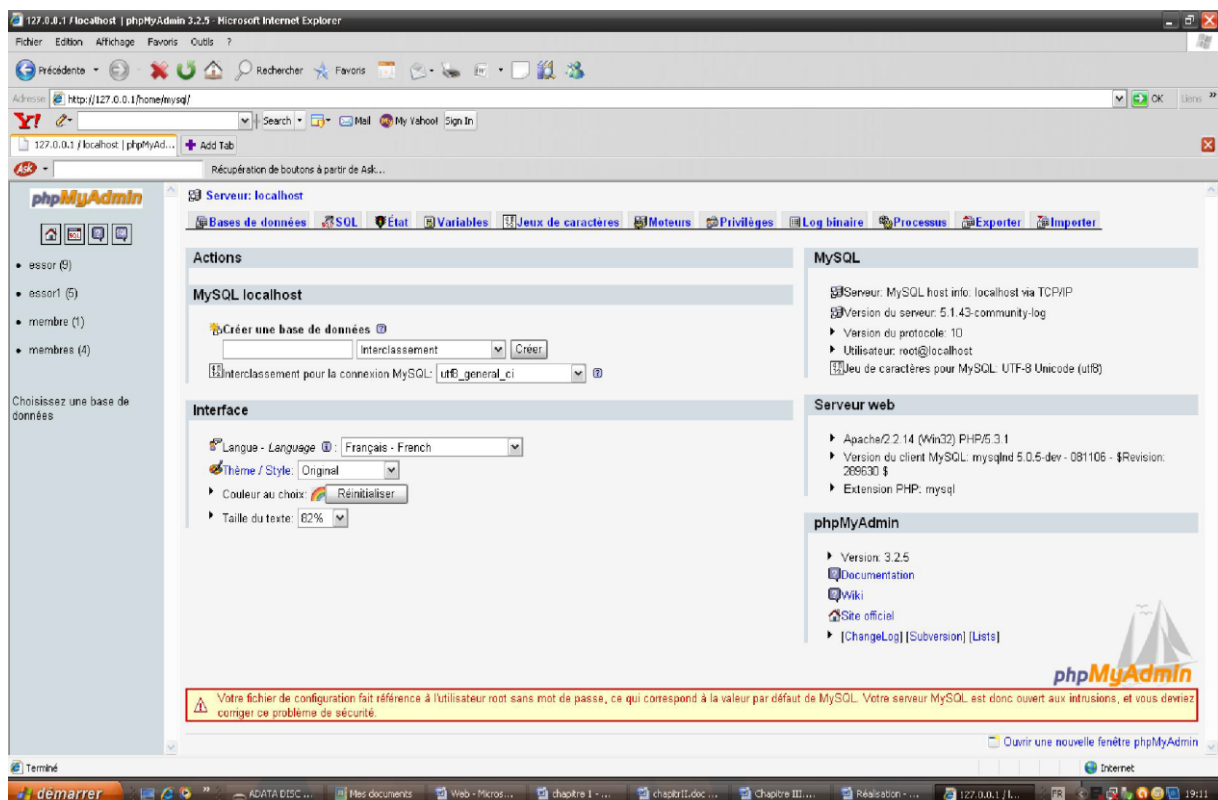


Figure V.4: interface PhpMyAdmin

V.4.6 Serveur MySQL :



MySQL est un véritable serveur de base de données SQL multiutilisateur. Le langage SQL (Structured Query Language) étant le langage le plus populaire de base de données dans le monde. Il en prend la syntaxe mais n'en conserve pas toute la puissance puisque de nombreuses fonctionnalités de SQL n'apparaissent pas dans MYSQL (sélection imbriquées, clés étrangère...). SQL est un langage standardisé qui rend facile le stockage, la mise à jour et l'accès à l'information. Il est suffisamment rapide et flexible pour gérer des historiques et des images.

Un serveur de bases de données stocke les données dans des tables séparées plutôt que de tout rassembler dans une seule table. Cela améliore la rapidité et la souplesse de l'ensemble. Les tables sont reliées par des relations définies, qui rendent possible la combinaison de données entre plusieurs tables durant une requête.

MySQL est probablement la base de données disponible en freeware la plus utilisée. Réputé pour sa rapidité.

V.5 Présentation de quelque interface de notre application :

➤ Page d'accueil :

C'est la page principe de l'application où tous les acteurs du système doivent s'authentifier pour accéder à leur espaces,et aussi un espace de recherche (recherche réduite) pour toute personne désirant visiter le site.



Figure V.5 : Interface page d'accueil

➤ **L'espace administrateur :**

C'est la page d'accueil de l'administrateur, où il peut effectuer différentes tâches.



Figure V.6 : Interface administrateur

➤ **L'espace producteur :**

C'est la page d'accueil de producteur ou il peut effectuer différentes tâches, telle que la validation des manuscrits, établissement des statistique et demande de numérisation.

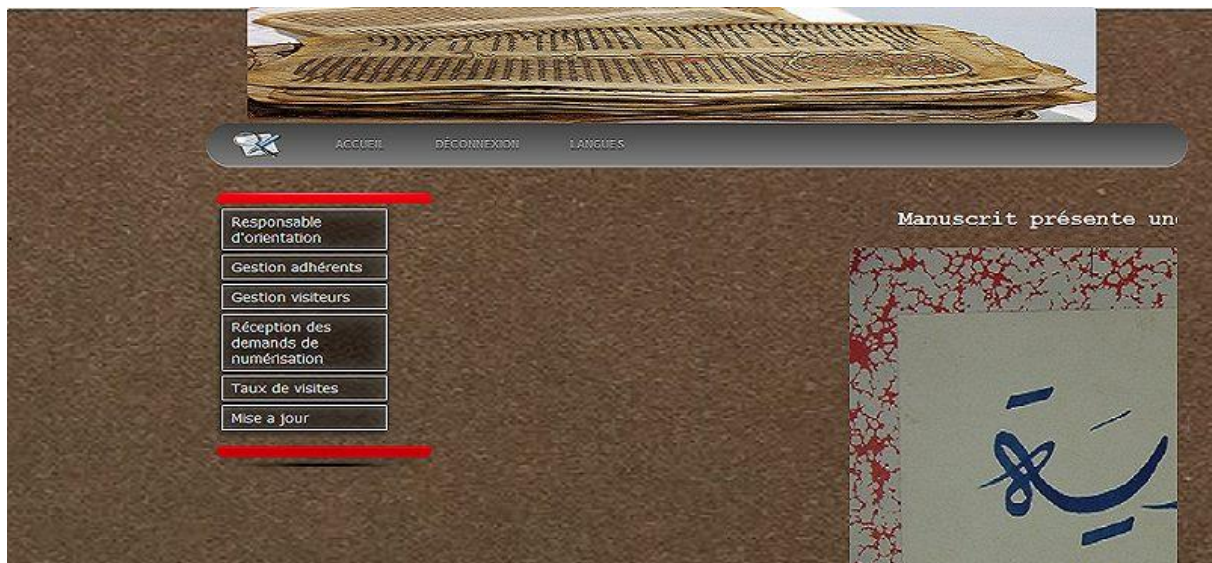


Figure V.7: Interface producteur

➤ **Page d’ajout d’un catalogueur :**

Cette page représente un formulaire d’insertion de nouveaux catalogueurs.

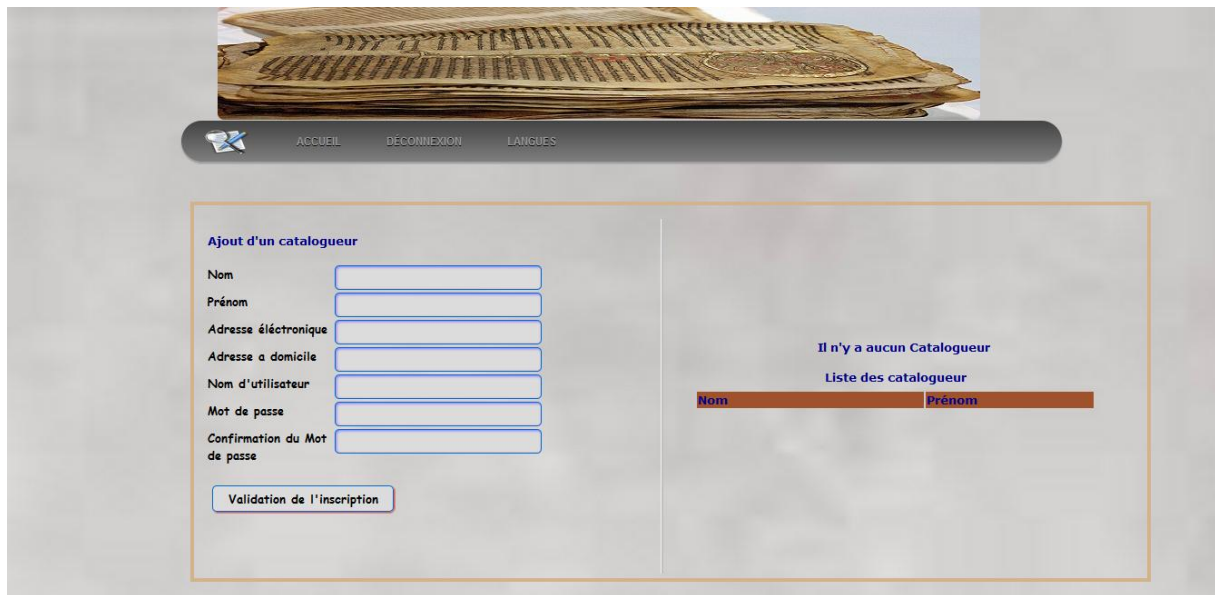


Figure V.8:interface d’ajout d’un catalogueur.

➤ **Page insertion notice :**

Cette page représente l’interface d’insertion des nouvelles notices par un catalogueur.

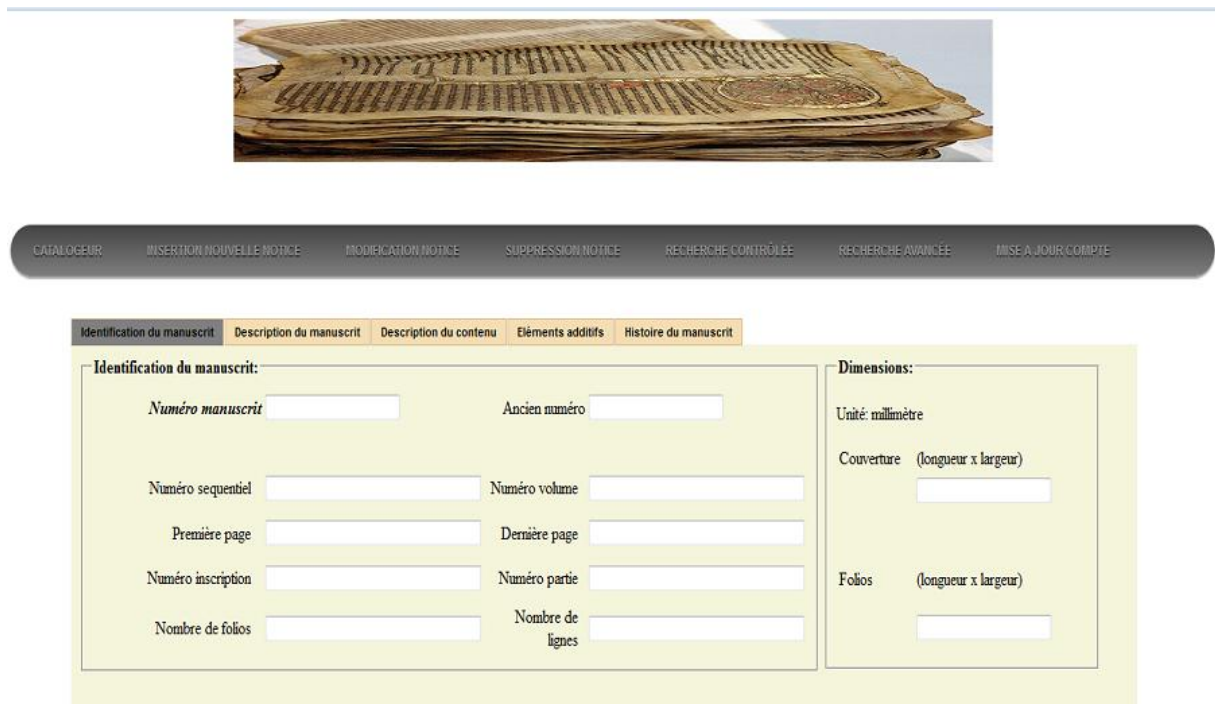


Figure V.9 : insertion d’une notice

➤ **Boite de réception :**

L'interface présente la liste des messages dans la boîte de réception.

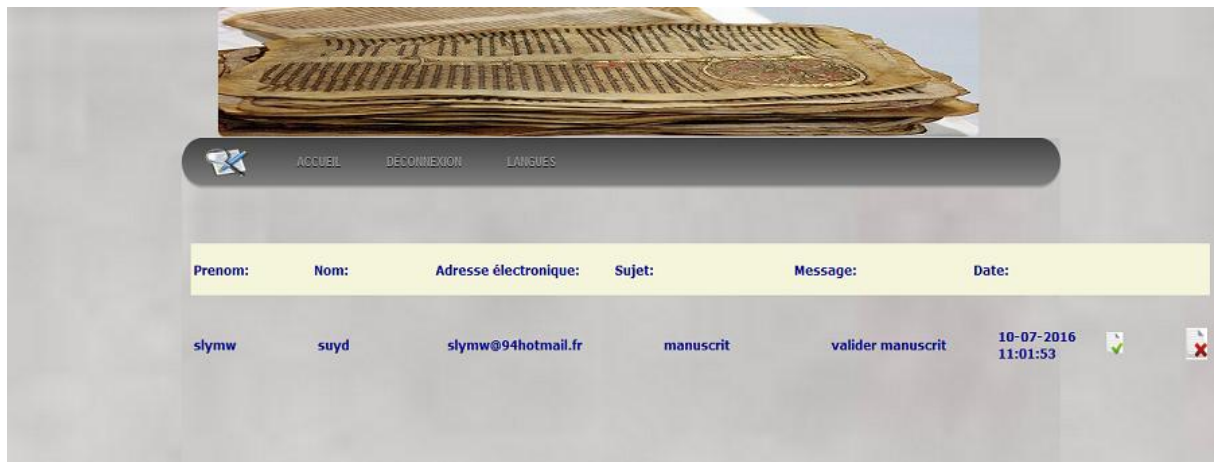


Figure V.10 : interface de boite de réception.

Conclusion :

Dans ce chapitre nous avons, au premier lieu, présenté les différentes outils et langages que nous avons utilisé pour implémenter notre application. Par la suite, nous avons présenté quelques interfaces de notre application.

Conclusion Générale

Conclusion générale

L'objectif de notre projet de fin d'étude était de concevoir et implémenter une application web sur la mise en ligne de manuscrit arabes anciens numérisés.

Le point de départ de la réalisation de ce projet était une étude théorique sur les manuscrits arabes anciens.

Par la suite, nous nous sommes intéressés à l'analyse et la spécification des besoins qui nous a permis de distinguer les différents acteurs interagissant avec l'application visée.

L'objectif de la partie suivante était la conception détaillée, dans laquelle nous avons fixé la structure globale de l'application.

Le dernier volet de notre projet était la partie réalisation qui a été consacrée à la présentation des outils du travail et les interfaces les plus significatives de notre application.

La réalisation de notre projet nous a donné l'occasion d'acquérir des nouvelles connaissances et d'approfondir sur le développement des applications, on citera les langages de développement Java EE, Java Script, SQL avec des outils de développement à titre d'exemple nous citons Netbeans, WampServeur, Apache Tomcat. Ainsi les accès distants aux bases de données.

Enfin, nous estimons avoir réalisé l'objectif tracé au début de notre projet, à savoir la mise en place d'un système pour accéder aux informations de manuscrits arabes anciens et de permettre aux possesseurs des manuscrits arabes anciens de cataloguer leurs fonds de manuscrit.

Nous espérons que notre travail sera de grande utilité pour les générations à venir.

Bibliographie

Bibliographie

[1] : [IRHT 2006]. Livret du stage d'initiation au manuscrit médiéval, réalisé par l'Institut de Recherche en Histoire et du Texte (IRHT), Lyon – France. Département des manuscrits arabes - disponible sur <http://www.irht.cnrs.fr/stage/descriptionmanuscrit.htm>

[2] : [Hala Kaileh 2004].Thèse pour l'obtention du grade de Doctorat l'université de Lumière Lyon II. L'accès à distance aux manuscrits arabes numérisés en mode image disponible sur <http://theses.univlyon2.fr>.

[3] : [I.R.S.H. 1995].Document publié par l'I.R.S.H., Niamey, Le Département de Manuscrits Arabes et Ajamis de l'Institut de Recherche en Sciences Humaines, Université Abdou Moumouni, Niamey (Niger)-disponible sur <http://www.org.uib.no/smi/sa/07/7Niamey>

[4]: La décoration du manuscrit - les carnets de IRHT. Disponible sur <http://irht.hypotheses.org/751>

[5]: mémoire ingénieur (Accès en ligne aux manuscrits Arabes en mode contrôlé) promotion 2010/2011.Réalisé par M^{elle} Taleb Symouna et Mr Kouchi Ismail.

[6] : http://www.univ-montp3.fr/incunables_et_manuscrits.html

[7] : les fiches d'Interbibly-la numérisation en bibliothèque et en archives disponible sur <http://www.interbibly.fr>

[8] : <http://mediadix.u-paris10.fr/cours/Bibliographie/BB/Collectionsnumeriques>

[9] : le catalogage, description des documents. Disponible sur <http://www.math.u-psud.fr/~biblio/ecole/support/Cours-Catalogage>.

[10] : [Soualah 2008]. Soualah Mohammed Ou Rabah Thèse pour l'obtention du grade de Magister à l'I.N.I.Numérisation des manuscrits arabes et accès multilingue.

[11]: <https://www.library.cornell.edu/preservation/tutorial/metadata/table5-1.html>

[12] : http://www.fondation.org.ma/fondation1_fr/presentation.html

[13] : <http://www.collectionscanada.gc.ca>

[14] : <http://www.marc21.ca/M21/040010-239>

[15] : comprendre les notices d'autorités MARC : catalogage ordinolingue disponible sur www.marc21.ca/M21/040010-238-f.html

[16] : unimarc-Bibliothèque nationale de France disponible sur http://www.bnf.fr/fr/professionnels/f_um/s.format_unimarc_notices_bibliographie.html

Bibliographie

- [17] : [Françoise Leresche, 2004] Bibliothèque nationale de France, Agence bibliographique nationale, Bureau de normalisation documentaire.
- [18] : Habchi, Khaled. L'usage d'UNIMARC en Tunisie. Disponible sur <http://www.ifla.org/past-wlic/2010/148-habchi-fr>
- [19] : faire un répertoire ou inventaire simple en EAD, group d'experts AFNOR, octobre 2005. Disponible sur <http://www.enssib.fr/bibliotheque-numerique>
- [20] : l'EAD et l'EAC, deux DTD pour la description archivistique, archive de France, présentation group PIN, mai 2004. Disponible sur <http://pin.association-aristote.fr>
- [21] : mets metadata encoding and transmission standard .Disponible sur <https://archivengines.wordpress.com>
- [22] : <https://www.loc.gov/standards/mets/METSOverview.v2.html>
- [23] : <http://www.academia.edu/591323>
- [24] : La bibliothèque nationale de France, disponible sur : http://www.bnf.fr/fr/professionnels/formats_catalogage/a.f_dublin_core.html
- [25] : Manuel d'encodage TEI - Renaissance et temps modernes. Disponible sur : http://www.bvh.univ-tours.fr/XML-TEI/manuelTEIrenaissance3_2012.
- [26] : la même source que [5].
- [27] : la bibliothèque numérique de l'enssib. Disponible sur : <http://www.enssib.fr/bibliotheque-numerique/documents/1335-tei-text-encoding-initiative>
- [28] : Manuscript Description - The TEI Guidelines. Disponible sur: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html>.
- [29] : www.comment-ça-marche.net
- [30] : www.gibello.com/cours/JavaEE
- [31] : mémoire ingénieur (conception et réalisation d'une application java ee sur la gestion des achats locaux et importation) promotion 2014/2015. Réalisé par Mlle MOKHBI Ouiza et Mlle TOUBAL Samira
- [32] : <http://lig-membres.imag.fr/plumejeaud/NFE107-fichesLecture>
- [33] : [Jean-Michel Doudoux]. Disponible sur: www.jdoudoux.fr/java/dej/chap-j2ee-java-ee.htm
- [34] : Mémoire promotion 2012/2013 réalisé par : Amrane Samia et Amrane Hassina (conception et réalisation d'une application web avec J2EE pour la gestion de scolarité).
- [35] : Mémoire ingénieur (conception et réalisation d'une application client/serveur). Promotion 2009/2010.

Bibliographie

[36] : Le guide d'utilisation UML de Grady Bouch

[37] : Définition et caractéristique d'UML, 2008

[38] : Livre d'UML et SQL de Christian Soutous

[39] : <http://www.jmdoudoux.fr/java/dej/chap-jsp.htm>

[40] : <http://www.jmdoudoux.fr/java/dej/chap-servlet.htm>

[41] : Wazir Dino,cours JavaScript,TS IRIS-LEGT Louis-Modeste Leroy-evreux,2002.

Disponible sur : info.arqendra.net

Annexes

I. Introduction

La conception de toute solution logicielle doit être traitée d'une manière précise et détaillée, précédée d'une analyse profonde et bien réfléchie, car elle est le reflet du futur système avant même sa réalisation. Dans le but d'avoir une meilleure analyse et de rendre la conception de notre application plus complète, nous avons opté le langage UML qui permet de bien représenter l'aspect statique et dynamique d'une application par une série de diagramme qu'il offre.

I.1 Présentation du langage UML

I.1.1 Origine et définition :

UML (Unified Modeling Language, "langage de modélisation objet unifié") est né au milieu des années 90 par la consolidation de plusieurs méthodes dont « Booch » et « OMT ».

Le succès a été immédiat et UML est aujourd'hui universellement accepté et supporté par l'ensemble des outils de développement.

UML est essentiellement un langage qui facilite la représentation et la compréhension des solutions objets :

- ❖ **Sa notation graphique** permet d'exprimer visuellement une solution objet, ce qui facilite la comparaison et l'évaluation de solutions.
- ❖ **Sa notation formelle** limite les ambiguïtés et les incompréhensions.
- ❖ **Son indépendance** par rapport aux langages de programmation, aux domaines d'application et aux processus, en fait un langage universel.

I.1.2 Pourquoi le langage UML ?

Le choix d'UML se justifie à partir de nombreux points forts qu'il offre, parmi eux on peut citer les trois suivants :

- Pour maîtriser la complexité technique grâce au formalisme et l'abstraction fournis par les concepts UML.
- Pour faciliter le dialogue entre le métier et la technique et ainsi garantir une prise en compte optimum des besoins utilisateurs.
- Modéliser tout les squelettes de son système et le générer par la suite dans le langage désiré, est une des principales fonctionnalités offertes.

I.1.3 Présentation des diagrammes UML :

Un modèle est une représentation simplifiée d'un problème. UML permet d'exprimer les modèles objets à travers un ensemble de diagrammes. Ces derniers sont des moyens de description des objets ainsi que des liens qui les relient.

Un diagramme est une représentation graphique qui s'intéresse à un aspect précis du modèle. UML offre 13 types de diagrammes. Chaque type de diagramme offre une vue d'un système. Combinés, les différents types de diagrammes offrent une vue complète du système.

Les digrammes UML peuvent êtres classés sous deux grandes catégories :

- Les diagrammes de vue statiques,
- Les diagrammes de vue dynamique.

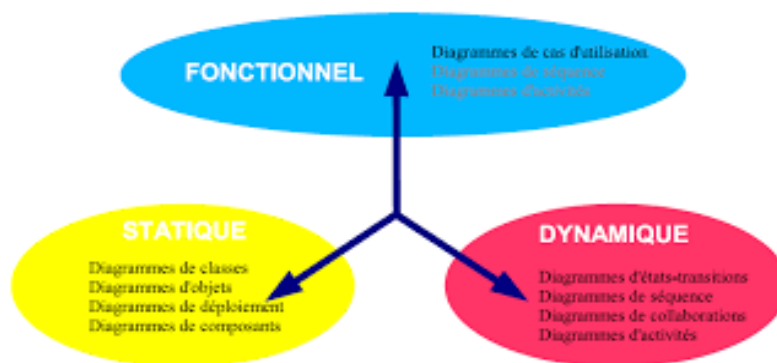


Figure 1 : les diagrammes de l'UML et leurs vue

a. Les diagrammes de vue statique :

- **Les diagrammes de classes** : sont sans doute les diagrammes les plus utilisés d'UML. Ils décrivent les types des objets qui composent un système et les différents types de relations statiques qui existent entre eux.



Activer Windows

Figure 2: représentation graphique de diagramme de classe.

Annexe

- **Les diagrammes d'objets** : représente un instantané des objets d'un système à un moment donné. Ces diagrammes sont souvent appelés diagrammes d'instances car ils représentent des instances et non pas des classes.
- **Les diagrammes de composants** : Décrivent l'architecture interne d'une classe. Ils sont très utilisés pour représenter l'architecture physique et statique d'une application en termes de module (fichiers, sources, bibliothèques, exécutables,...) pour montrer la mise en œuvre physique des modèles statiques avec l'environnement de développement.
- **Les diagrammes de déploiement** : représentent l'agencement physique d'un système montrant sur quels composants matériels les différents composants logiciels s'exécutent. Les principaux éléments sont des nœuds connectés par des voies de communication. Un nœud représentera une unité qui va héberger un logiciel, un équipement matériel ou ordinateur, un environnement d'exécution, un système d'exploitation,...

b. Les diagrammes de vue dynamique :

- **Les diagrammes d'activités** : décrivent le comportement d'une méthode, le déroulement d'un cas d'utilisation, les enchaînements d'activités. Une activité désigne une suite d'actions. Le passage d'une action vers une autre est matérialisé par une transition. Les transitions sont déclenchées par la fin d'une action et provoquent le début immédiat d'une autre (elles sont automatiques).

Annexe

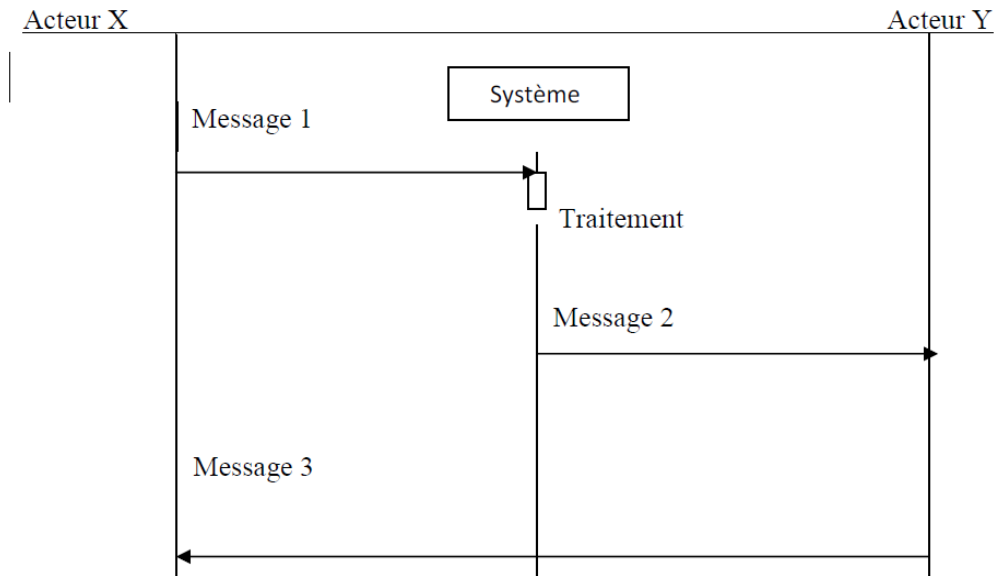


Figure 3 : représentation graphique de diagramme d'activité.

- **Les diagrammes de cas d'utilisation :** les cas d'utilisation (en anglais use cases) ont été introduits par Ivar Jacobson dans sa méthode OOSE (Oriented Object Software Engineering). Ils constituent une technique qui permet de déterminer les besoins des utilisateurs et de capturer les exigences fonctionnelles d'un système. En d'autres termes, ils décrivent le comportement d'un système du point de vue de ses utilisateurs. Ils décrivent les interactions entre les utilisateurs d'un système et le système lui-même. Et un diagramme de cas d'utilisation permet de représenter graphiquement les cas d'utilisation. Le fait qu'un acteur déclenche un cas d'utilisation est représenté par une flèche entre ces deux derniers.

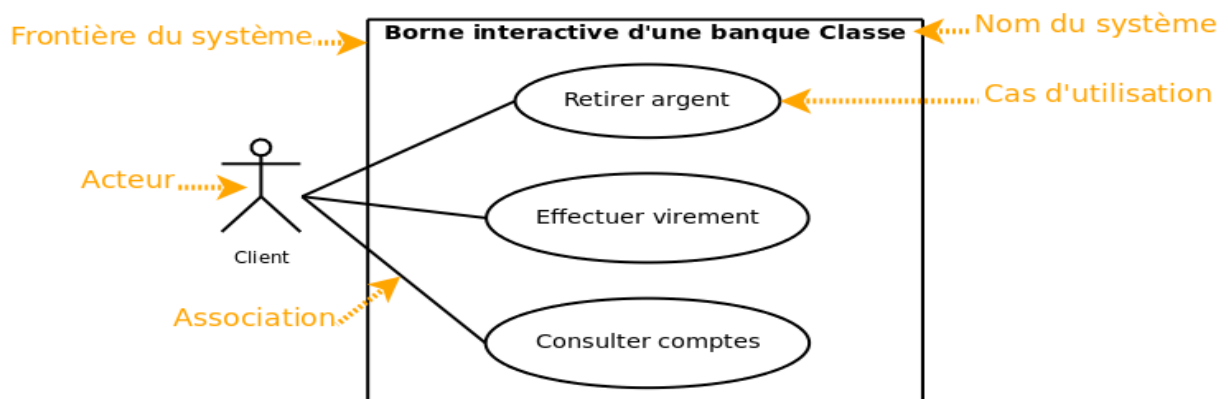


Figure 4: représentation graphique de diagramme

Annexe

- **Les diagrammes de séquence** : permettent de représenter les interactions entre objets selon un point de vue temporel. L'accent est mis sur la chronologie des envois de messages.
- **Les diagrammes de machine d'état** : appelés diagrammes d'état transitions dans UML 1.x, permet de décrire le comportement d'un objet durant son cycle de vie.

Conclusion :

Comme UML n'impose pas de méthode de travail particulière, il peut être intégré à n'importe quel processus de développement logiciel de manière transparente. UML est une sorte de boîte à outils, qui permet d'améliorer progressivement vos méthodes de travail, tout en préservant vos modes de fonctionnement.