

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique



Université Larbi Tébessi – Tébessa -

Faculté des Sciences Exactes et Sciences
de la Nature et de la Vie

Département des Mathématiques et Informatique



Laboratoire de Mathématiques, d'Informatique et Systèmes (LAMIS)

Thèse

Présentée et soutenue publiquement par

Mohamed Hemili

En vue de l'obtention du diplôme de

Doctorat

Intitulée :

***Méthodologie pour un E-library Decision Support
System basé Cloud et Data Warehouse.***

Spécialité : **Informatique**

Option : **Systèmes d'information coopératifs**

Devant le jury composé de :

BENDJENNA Hakim	Professeur	Université Larbi Tébessi, Tébessa	Président
DERDOUR Makhlof	Professeur	Université Larbi Tébessi, Tébessa	Examineur
MAAROUK Toufik	MCA	Université Abbès Laghrour, Khenchela	Examineur
BENMERZOUG Djamel	MCA	Université Abdelhamid Mehri Constantine 2	Examineur
LAOUAR Mohamed Ridha	Professeur	Université Larbi Tébessi, Tébessa	Directeur de thèse

*Méthodologie pour un E-library Decision Support
System basé Cloud et Data Warehouse.*

Mohamed Hemili



Université Larbi Tébessi



Laboratoire LAMIS

Dédicace

Je dédie ce travail,

A mes très chers parents,

A toute ma famille,

A tous mes amis,

A mon encadreur Pr. LAOUAR Mohamed Ridda

A tous mes chers.

Remerciements

C'est avec un réel plaisir et un grand enthousiasme que je me livre à la rédaction de cette page. Bien plus que le point final du manuscrit scientifique, cette page constitue l'opportunité de m'accorder une réflexion sur une période de ma vie très riche en émotions.

Les travaux présentés dans ce mémoire ont été menés au sein du Laboratoire Laboratoire de Mathématiques, d'Informatique et Systèmes (LAMIS) de l'Université Larbi Tébessi de Tébessa.

*Premièrement je tiens à remercier mon directeur de thèse le Professeur **LAOUAR Mohamed Ridda** pour la confiance qu'il m'a accordée et ses encouragements tout au long de la préparation de cette thèse.*

*Je tiens également ma gratitude à Monsieur **Sean B. EOM**, Professeur à l'Université Southeast Missouri en USA d'avoir accepté de répondre à mes questions sur le sujet d'étude, je le remercie pour son immense aide, pour son attention de tout instant sur mes travaux et pour ses conseils avisés. Je suis ravie d'avoir travaillé en sa compagnie.*

*Je voudrai exprimer mes remerciements à tous **les membres de laboratoire LAMIS** pour le support qui m'a permis de réaliser et de finir ce travail dans les meilleures conditions.*

*Je remercie le plus sincèrement possible les membres de jury : **Pr. BENDJENNA Hakim, Pr. DERDOUR Makhoulf, Dr. MAAROUK Toufik et Dr. BENMARZOUG Djamel** qui ont accepté d'évaluer ce travail et de fournir des critiques pertinentes pour leur attention portée à cette thèse et le temps qu'ils ont consacré à son évaluation.*

Finalement je voudrais remercier tous mes amis qui m'ont soutenu et encouragé durant ces cinq années de préparation.

Résumé

Aujourd'hui, les bibliothèques numériques académiques (BNAs) représentent une très importante source d'information pour la communauté académique. Les chercheurs, les étudiants, et les autres utilisateurs académiques vont toujours répondre à leurs besoins d'informations en utilisant les différents services de bibliothèques numériques académiques. Par conséquent, le premier but de ces bibliothèques est de satisfaire les besoins d'informations de ses utilisateurs à travers ses services offerts. Cependant, avec l'augmentation du nombre des utilisateurs des bibliothèques numériques académiques, leurs attentes et leurs besoins d'information, et aussi le nombre des travaux publiés, la gestion des bibliothèques numériques devient de plus en plus difficile. Les gestionnaires des bibliothèques numériques doivent prendre des décisions effectives pour une bonne gestion de ces bibliothèques, afin d'atteindre ce but, ils sont obligés d'analyser un grand nombre de sources de données, des processus et des volumes de données hétérogènes.

Les bibliothèques numériques sont des profitables organisations qui visent à gagner des profits pour continuer leur travail. Elles sont obligées de satisfaire leurs utilisateurs, par offrir des services de qualité. Les utilisateurs des bibliothèques numériques académiques demandent aux bibliothèques d'offrir un certain niveau de qualité de service, tel qu'une collection riche, un temps de réponse acceptable, et une interface de site web facile à utiliser ...etc. Les anciennes méthodes de développement et d'amélioration des services ne peuvent plus satisfaire ces utilisateurs. Le problème des anciennes méthodes est que ces sources de données hétérogènes ne sont que partiellement utilisées pour les processus de prise de décision en raison de la grande variété de formats, normes et technologies, ainsi que le manque de méthodes efficaces d'intégration. Récemment, plusieurs travaux de recherche ont été proposés pour le but d'aider les gestionnaires des bibliothèques numériques dans les différents processus de gestion. Dans cette thèse on s'intéresse aux problèmes de développement et de gestion concernant les collections des bibliothèques numériques académiques. Ce travail a pour objectif l'analyse et la conception d'un système d'aide à la décision pour le développement et la gestion des collections bibliothèques numériques académiques.

Mots-clés: *Les bibliothèques numériques académiques, développement de collection, gestion de collection, système d'aide à la décision, aide à la décision multicritère, Data Warehouse, Cloud Computing.*

Abstract

Today, academic digital libraries represent a very important source of information for the academic community. Researchers, students, and other academic users will always respond to their information needs by using the academic digital library services. Therefore, the primary purpose of these libraries is to satisfy the information needs of its users through its offered services. However, with the increasing in the number of users of academic digital libraries, their expectations and their information needs, and the increasing number of published works, the management of digital libraries becomes more and more difficult. Digital library managers (librarians) must make effective decisions for a good management of these libraries, in order to achieve that goal they are forced to analyze a large number of heterogeneous data sources, processes and large data volumes.

Digital libraries are profitable organizations that aim to earn profits to continue their work. They are obliged to satisfy their users, by offering good quality services. Users of academic digital libraries require these libraries to maintain a certain level of quality of service, such as a rich collection, acceptable response time, and an easy-to-use website interface ... etc. With their limited budgets, the decision making process at the management level becomes very critical. The old ways of developing and improving services can no longer satisfy these users. The problem with old methods is that these heterogeneous data sources are only partially used for decision-making processes because of the wide variety of formats, standards and technologies, as well as the lack of effective methods of integration. Recently, several research projects have been proposed for the purpose of helping digital library managers in the various management processes. This thesis gives a brief summary of the state of the art on the problems of management of digital libraries. In this thesis we are interested in management problems concerning the collections of academic digital libraries. This thesis aims to analyze and design a decision support system for the development and management of academic digital libraries collections.

Keywords: Academic digital libraries, collection developpement, collection management, decision support system, multi-criteria decision aid, data warehousing, cloud computing.

الملخص

أصبحت المكتبات الرقمية الأكاديمية اليوم تمثل مصدرًا مهمًا جدًا للمعلومات بالنسبة للمجتمع الأكاديمي. إذ يعتمد الباحثون والطلاب والمستخدمون الأكاديميون الآخرون في تلبية احتياجاتهم من المعلومات بشكل أساسي على استغلال الخدمات المختلفة التي توفرها المكتبات الرقمية الأكاديمية، وبالتالي فإن الغرض الأساسي من هذه المكتبات هو تلبية احتياجات المستخدمين من المعلومات من خلال الخدمات التي تقدمها. ومع ذلك، مع زيادة عدد مستخدمي المكتبات الرقمية الأكاديمية، ارتفع سقف توقعاتهم من هذه المكتبات، الزيادة في احتياجاتهم من المعلومات، وكذلك الزيادة في عدد المصادر المنشورة، أصبحت إدارة المكتبات الرقمية أكثر صعوبة. إذ أنه يجب على أمناء أو مدراء هاته المكتبات الرقمية اتخاذ قرارات فعالة من أجل إدارة فعالة لهذه المكتبات، من أجل تحقيق هذا الهدف، يضطر هؤلاء الأمناء إلى تحليل عدد كبير من مصادر البيانات غير المتجانسة و عدد كبير من العمليات وأحجام كبيرة من البيانات.

إن المكتبات الرقمية باعتبارها منظمات ربحية تهدف إلى كسب الأرباح لمواصلة عملها. و منه فهاته المكتبات ملزمة بإرضاء مستخدميها من خلال توفير خدمات عالية الجودة. يتوقع مستخدمو المكتبات الرقمية الأكاديمية من هذه المكتبات أن تحافظ على مستوى معين من جودة الخدمة، مثل مجموعة مصادر غنية. وقت استجابة مقبول، وواجهة ويب سهلة الاستخدام ... إلخ. نظرا للميزات المحدودة المتوفرة لدى هاته المكتبات، تصبح عملية صنع القرار على مستوى الإدارة بالغة الأهمية. في وقتنا الحالي لم تعد الطرق القديمة لتطوير وتحسين الخدمات كافية لتوفير خدمات ترضي المستخدمين. تكمن مشكلة الطرق القديمة في أن مصادر البيانات غير المتجانسة المستخدمة في صناعة القرار تستخدم جزئيًا فقط، وذلك نظرًا للتنوع الكبير في الأشكال والمعايير والتقنيات الذي تتميز به هاذه المصادر، فضلاً عن عدم وجود طرق فعالة لدمج هذه الأشكال والمعايير والتقنيات. في الآونة الأخيرة، تم اقتراح العديد من المشاريع البحثية بغرض مساعدة مديري المكتبات الرقمية في مختلف عمليات الإدارة. تقدم هذه الرسالة ملخصًا موجزًا عن أحدث ما توصلت إليه الأبحاث الموجهة لحل مشكلات إدارة المكتبات الرقمية. في هذه الأطروحة نحن مهتمون بمشاكل الإدارة المتعلقة بمجموعة المصادر في المكتبات الرقمية الأكاديمية. تهدف هذه الرسالة إلى تحليل وتصميم نظام لدعم اتخاذ القرار لتطوير وإدارة مجموعات المصادر في المكتبات الرقمية الأكاديمية بحيث يساعد مديري هذا النوع من المكتبات الرقمية على أداء مهامهم بفعالية أكبر.

الكلمات المفتاحية: المكتبات الرقمية الأكاديمية، تطوير مجموعة المصادر، إدارة مجموعة المصادر، نظام دعم القرار، دعم القرار المتعدد المعايير، مستودع البيانات، الحوسبة السحابية.

Table des Figures

Figure 2.1 Répartition des termes (Digital library, Electronic library, Virtual library)	27
Figure 2.2 L'augmentation de Nombre de Livres Electroniques Publiés	30
Figure 2.3 Processus de Développement de Collection	31
Figure 2.4 Exemple de formulaire de demande de ressource numérique interne en ligne	35
Figure 2.5 Cycle de Vie de Gestion de Collection	38
Figure 3.1 Les Composants des Systèmes D'aide à la Décision	45
Figure 3.2 Caractéristiques de Data Warehouse	57
Figure 3.3 Les Modèles de Livraison de Services du Cloud	60
Figure 3.4 Les Modèles de Déploiement Du Cloud	61
Figure 5.1 Les Etapes du DSS	80
Figure 5.2 Cycle de Vie des Ressources Electroniques	81
Figure 5.3 Architecture générale du système proposé	82
Figure 5.4 Obtenez des requêtes à partir de l'application de recherche	85
Figure 5.5 Les Etapes d'Analyse d'un Fichier Log	86
Figure 5.6 Algorithme d'Analyse du Fichier Log	86
Figure 5.7 Capture de l'interface du DSS	90
Figure 5.8 Le rapport principal du DSS	91
Figure 5.9 Table d'évaluation	91
Figure 5.10 Classement totale	92
Figure 5.11 BarChart d'évaluation	93
Figure 5.12 Data Warehouse pour CDDSS	94
Figure 5.13 Architecture des Data Marts	95
Figure 5.14 Schéma des Données dans le Data Warehouse	96

Table des Figures

Figure 5.15 Architecture du DSS basé Cloud pour les Bibliothèque Numériques	98
Figure 5.16 Les Etapes de Créer une VM en Utilisant Amazon EC2	99
Figure 5.17 Amazon AMI permet de Créer des VMs avec plateformes	100

Liste des Tableaux

Tableau 3.1 Comparaison entre les anciens travaux	64
Tableau 4.1 Comparaison des travaux	69
Tableau 4.2 Les Méthodes d'Evaluation des Collections	72
Tableau 4.3 Les critères de Sélection	74
Tableau 5.1 Préférences d'utilisateur	87
Tableau 5.2 Préférences d'évaluation	88
Tableau 5.3 Services de Déploiement d'une Application Java	97

Liste des Abréviations

<i>AHP</i>	<i>Analytic Hierarchy Process</i>
<i>BiblioML</i>	<i>Bibliographic Markup Language</i>
<i>BN</i>	<i>Bibliothèque Numérique</i>
<i>BNA</i>	<i>Bibliothèque Numérique Académique</i>
<i>BP</i>	<i>Business Process</i>
<i>BU</i>	<i>Bibliothèques Universitaires</i>
<i>CC</i>	<i>Cloud Computing</i>
<i>CD-ROM</i>	<i>Compact Disc Read-Only Memory</i>
<i>DDA</i>	<i>Demand-Driven Acquisition</i>
<i>DL</i>	<i>Digital Library</i>
<i>DLF</i>	<i>Digital Library Federation</i>
<i>DM</i>	<i>Decision Maker</i>
<i>DOI</i>	<i>Digital Object Identifier</i>
<i>DRM</i>	<i>Digital Rights Management</i>
<i>DSS</i>	<i>Decision Support System</i>
<i>DW</i>	<i>Data Warehousing</i>
<i>DAG</i>	<i>Direct Access Graphe</i>
<i>EBL</i>	<i>EBook Library</i>
<i>EBSCO</i>	<i>Elton B. Stephens COmpany</i>
<i>ED</i>	<i>Entreposage de Données</i>
<i>ELECTRE</i>	<i>ELimination Et Choix Traduisant la REalité</i>
<i>ENSSIB</i>	<i>Ecole Nationale Supérieure des Sciences de l'information et des Bibliothèques</i>
<i>ETL</i>	<i>Extract Transform Load</i>
<i>GBN</i>	<i>Gestionnaire de Bibliothèque Numérique</i>
<i>HTML</i>	<i>HyperText Markup Language</i>
<i>IaaS</i>	<i>Infrastructure as a Service</i>
<i>IFLA</i>	<i>International Federation of Library Associations</i>
<i>IN</i>	<i>Informatique en Nuage</i>
<i>IR</i>	<i>Information Retrieval</i>

Liste des Abréviations

<i>ISBN</i>	<i>International Standard Book Number</i>
<i>KDD</i>	<i>Knowledge Discovery in Databases</i>
<i>KM</i>	<i>Knowledge Management</i>
<i>KW</i>	<i>Key Word</i>
<i>MARC</i>	<i>MAchine-Readable Cataloging</i>
<i>MCD A</i>	<i>Multi-Criteria Decision Aid</i>
<i>MCDSS</i>	<i>Multi-Criteria Decision Support System</i>
<i>MySQL</i>	<i>My Structured Query Language</i>
<i>NISO</i>	<i>National Information Standards Organization</i>
<i>OPAC</i>	<i>Online Public Access Catalogs</i>
<i>PaaS</i>	<i>Platform as a Service</i>
<i>PDA</i>	<i>Patron-Driven Acquisition</i>
<i>PDF</i>	<i>Portable Document Format</i>
<i>PROMETHEE</i>	<i>Preference Ranking Organization METHod for Enrichment Evaluation</i>
<i>RSSI</i>	<i>Received Signal Strength Indicator</i>
<i>SaaS</i>	<i>Software-as-a-Service</i>
<i>SAW</i>	<i>Simple Additive Weighting</i>
<i>SI</i>	<i>Système d'Information</i>
<i>SIAD</i>	<i>Systèmes Interactifs d'Aide à la Décision</i>
<i>TOPSIS</i>	<i>Technique for Order of Preference by Similarity to Ideal Solution</i>
<i>UNESCO</i>	<i>United Nations Educational, Scientific and Cultural Organization</i>

Table des matières

<i>Dédicace</i>	<i>iv</i>
<i>Remerciements</i>	<i>v</i>
<i>Résumé</i>	<i>vi</i>
<i>Abstract</i>	<i>vii</i>
<i>المخلص</i>	<i>viii</i>
<i>Table des Figures</i>	<i>ix</i>
<i>Liste des Tableaux</i>	<i>xi</i>
<i>Liste des Abréviations</i>	<i>xii</i>
<i>Table des matières</i>	<i>xiv</i>
<i>Chapitre 1 Introduction Générale</i>	<i>18</i>
1.1 <i>La gestion des collections des bibliothèques numériques</i>	<i>19</i>
1.2 <i>Définition de la problématique</i>	<i>20</i>
1.3 <i>Présentation des contributions</i>	<i>21</i>
1.4 <i>Organisation de la thèse</i>	<i>22</i>
<i>Chapitre 2 La Gestion des Bibliothèques Numériques</i>	<i>24</i>
2.1 <i>Introduction</i>	<i>26</i>
2.2 <i>Définition des concepts de base des bibliothèques numériques</i>	<i>26</i>
2.2.1 <i>Bibliothèque numérique, électronique, ou virtuelle ?</i>	<i>26</i>
2.2.2 <i>Définition d'une bibliothèque numérique</i>	<i>27</i>
2.2.3 <i>Définition de collection</i>	<i>27</i>
2.3 <i>Les concepts liés aux bibliothèques numériques</i>	<i>28</i>
2.3.1 <i>Service de bibliothèque numérique</i>	<i>28</i>
2.3.2 <i>Ressources numérique</i>	<i>28</i>
2.3.3 <i>Live numérique</i>	<i>28</i>
2.3.4 <i>Journal numérique</i>	<i>28</i>
2.3.5 <i>Les gestionnaires de bibliothèque numérique</i>	<i>28</i>
2.4 <i>La gestion de bibliothèque numérique</i>	<i>29</i>
2.5 <i>Le développement de collection</i>	<i>31</i>
2.5.1 <i>Collectages d'information sur les ressources</i>	<i>31</i>
2.5.2 <i>Catégorisation des ressources</i>	<i>32</i>

Table des Matières

2.5.3	<i>Evaluation de ressources</i>	33
2.5.4	<i>Le processus de sélection</i>	33
2.5.5	<i>Abonnement ou acquisition</i>	34
	i. <i>Décision d'achat ou d'abonnement</i>	34
	ii. <i>Préparation de la commande et options d'acquisitions</i>	34
	iii. <i>Sélection du fournisseur approprié</i>	36
2.5.6	<i>Construction de collection</i>	37
2.6	<i>La gestion de collection</i>	37
2.6.1	<i>Collectage des données d'utilisation</i>	38
2.6.2	<i>Evaluation des ressources</i>	38
2.6.3	<i>Sélection de ressources</i>	39
2.6.4	<i>Désabonnement ou renouvellement</i>	39
2.6.5	<i>Mise à jour du catalogue</i>	40
2.7	<i>Défis liés au développement et à la gestion des collections</i>	40
2.7.1	<i>Détermination des besoins d'information des utilisateurs</i>	40
2.7.2	<i>L'évaluation de la collection</i>	40
2.7.3	<i>Sélection des ressources numériques</i>	40
2.7.4	<i>L'analyse des données hétérogènes</i>	41
2.7.5	<i>La puissance de calcul</i>	41
2.8	<i>Conclusion</i>	41
	Chapitre 3 DSS pour les Bibliothèques Numériques	42
3.1	<i>Introduction</i>	44
3.2	<i>L'aide à la décision</i>	44
	3.2.1 <i>Définitions de l'aide à la décision</i>	44
	3.2.2 <i>Objectif de l'aide à la décision</i>	45
	3.2.3 <i>Pourquoi l'aide à la décision</i>	46
	3.2.4 <i>Système d'aide à la décision (DSS)</i>	47
	3.2.5 <i>Systèmes interactifs d'aide à la décision</i>	47
	3.2.6 <i>L'aide à la décision multicritère</i>	48
	3.2.1 <i>Les méthodes d'aide à la décision multicritères</i>	48
3.3	<i>Les requêtes de recherche des utilisateurs des BNAs: Les traces informatiques</i>	51
	3.3.1 <i>Définitions</i>	51

Table des Matières

3.3.2	<i>Des Traces informatiques, pourquoi faire ?</i>	51
3.3.3	<i>Objectifs des systèmes de traçage</i>	51
3.3.4	<i>Utilisation des traces enregistrées</i>	52
3.3.5	<i>Types des traces</i>	52
3.3.6	<i>Finalités des traces</i>	53
3.3.7	<i>Cycle de vie et qualité des traces</i>	54
3.4	<i>Data Warehouse (DW)</i>	56
3.4.1	<i>Définitions de Data Warehouse</i>	56
3.4.2	<i>Caractéristiques de Data Warehouse</i>	56
3.4.3	<i>Méthodes de conception de Data Warehouse</i>	57
3.4.4	<i>Objective de Data Warehouse</i>	58
3.5	<i>Cloud Computing (CC)</i>	58
3.5.1	<i>Caractéristiques du CC</i>	59
	<i>Service-libre à la demande</i>	59
	<i>Large accès au réseau</i>	59
	<i>Mise en commun des ressources</i>	59
	<i>Élasticité rapide</i>	59
	<i>Service mesuré</i>	59
3.5.2	<i>Modèles de livraison de service</i>	59
	<i>Le modèle Software-as-a-Service (SaaS)</i>	60
	<i>Le modèle Platform as a Service (PaaS)</i>	60
	<i>Le modèle Infrastructure as a Service (IaaS)</i>	60
3.5.3	<i>Modèles de déploiement du Cloud</i>	61
1.	<i>Cloud privé</i>	61
2.	<i>Cloud communautaire</i>	61
3.	<i>Cloud public</i>	61
4.	<i>Cloud hybride</i>	62
3.6	<i>Travaux connexes</i>	62
3.6.1	<i>Les approches générales pour l'évaluation des BNs</i>	62
3.6.2	<i>Analyse comparative des approches existantes</i>	63
3.7	<i>Conclusion</i>	65
	Chapitre 4 Utilisation de MCDA pour l'Evaluation de Collection	66
4.1	<i>Introduction</i>	67

Table des Matières

4.2	<i>Prise de décision dans la gestion de collection</i>	68
4.3	<i>L'évaluation de collection</i>	70
4.3.1	<i>L'évaluation de collection en tant qu'outil de gestion</i>	70
4.3.2	<i>Les buts de l'évaluation de collection</i>	71
4.4	<i>Techniques d'évaluation de collection</i>	72
4.5	<i>Critères d'évaluation des collections</i>	73
4.6	<i>Approche multicritères basée sur la méthode PROMETHEE II pour l'évaluation des collections</i>	74
4.7	<i>Conclusion</i>	76
Chapitre 5 Un Système d'Aide à la Décision pour Gérer le Développement des Collections		78
5.1	<i>Introduction</i>	79
5.2	<i>Évaluation et sélection des ressources numériques</i>	80
5.3	<i>DSS pour soutenir le développement de la collection dirigé par les demandes des utilisateurs (CDDSS)</i>	82
5.3.1	<i>Critères considérées par le DSS</i>	83
5.3.2	<i>Analyse de fichier log pour les bibliothèques numériques académiques</i> ..	85
5.3.3	<i>Phase 1: extraction des préférences des utilisateurs à partir des requêtes de recherche</i>	87
5.3.4	<i>Phase 2: Détermination des poids des critères</i>	87
5.3.5	<i>Phase 3: Evaluation et classement des ressources numériques</i>	88
5.4	<i>L'implémentation d'un prototype du CDDSS</i>	89
5.5	<i>Data Warehouse pour CDDSS</i>	94
5.5.1.	<i>Les composantes de l'architecture proposée</i>	94
5.5.2.	<i>Schéma des données</i>	95
5.6	<i>Proposition d'une Architecture Cloud pour le CDDSS développé</i>	96
5.6.1.	<i>Logiciel de gestion Cloud</i>	96
5.6.2.	<i>Les composantes de l'architecture Cloud</i>	97
5.7	<i>Conclusion</i>	100
Chapitre 6 Conclusion Générale		102
6.1	<i>Contributions</i>	103
6.2	<i>Perspectives</i>	104
Liste des publications et communications		105
Bibliographie		106

Chapitre 1

Introduction Générale

Contenu

<i>1.1</i>	<i>La gestion des collections des bibliothèques numériques</i>	<i>19</i>
<i>1.2</i>	<i>Définition de la problématique</i>	<i>20</i>
<i>1.3</i>	<i>Présentation des contributions</i>	<i>21</i>
<i>1.4</i>	<i>Organisation de la thèse</i>	<i>22</i>

Chapitre 1. Introduction Générale

Ce premier chapitre de cette thèse présente une introduction générale sur le contexte de la thèse qui est le domaine de bibliothèques numériques et leurs gestions. Puis, le besoin des systèmes d'aide à la décision et des technologies d'entreposage de données et de l'informatique en nuage est présenté et motivé. La problématique de recherche traitée dans cette thèse est exposée avec des contributions qui répondent aux problèmes. Enfin, l'organisation de la thèse est présentée.

1.1 La gestion des collections des bibliothèques numériques

Dans les dernières années, avec les développements dans la technologie de Web, le nombre de sources d'information et le volume de l'information disponible en ligne a augmenté dramatiquement (Waller, 2009). Le Web aussi donne la possibilité d'accéder à un énorme volume de contenu varié. Les créateurs de contenu peuvent facilement fournir ses travaux sur internet, ce qui permet aux internautes d'accéder à ces contenus de n'importe quel endroit et à tout moment. Par conséquent, comme toutes les autres sources d'information, le nombre de bibliothèques numériques aussi augmente rapidement sur internet et en conséquence le volume de collections et documents disponibles en ligne devient énorme (Arms, 2001). Cette disponibilité de contenu scientifique ou académique permet aux enseignants, étudiants, et chercheurs de satisfaire leurs besoins d'information, et les donne multiple choix de sources d'information, ce qui les permettra de choisir la meilleure option tenu en compte la qualité de sources et leur situation économique (Fox, 2016; Kavulya, 2015).

Cependant, malgré les avantages évidents du Web sur les demandeurs d'information, les changements apportées par le Web sur la disponibilité de l'information, cause beaucoup de défis pour les bibliothèques numériques, et oblige les bibliothèques numériques en générale et les bibliothèques numériques académiques en particulier à repenser leurs qualité de services, leurs méthodes de gestions, et leurs dépense afin de se repositionner en tant qu'instituts de références pour la diffusion des savoirs . Pour faire face à cette situation et réagir avec la concurrence croissante, les bibliothèques numériques doivent s'adapter à ces changements par l'amélioration de ses services offertes aux utilisateurs, et d'autre côté minimiser leurs dépenses par augmenter la qualité de gestion de ses ressources (R. A. Kadir, W. A. Dollah, F. A. Saaid, & S. Diljit, 2009). Par ailleurs, les bibliothèques académiques doivent grader leurs utilisateurs et même attirer des nouveaux utilisateurs pour continuer à fonctionner. Afin d'atteindre cet objectif, satisfaire les utilisateurs est un élément clé, offrir des informations fiables et pertinentes est le premier rôle des bibliothèques numériques académiques et en même temps c'est leur chemin pour garder une place principale dans le marché de l'information.

Par suite, la collection de la bibliothèque numérique définir la situation de cette bibliothèque par rapport aux utilisateurs. Une collection riche veut dire que les utilisateurs sont satisfaits, par contre si la collection est pauvre, elle n'offre pas les informations demandées par les utilisateurs cause des problèmes sérieuses à son bibliothèque (Johnson, 2014).

Une collection de bibliothèque numérique consiste tous les documents numériques possède ou avoir un accès. La gestion des collections des bibliothèques numériques académiques contient tous les taches reliées aux modifications survenues aux collections, comme la sélection des ressources numériques, la détermination de la politique de sélection, la détermination des besoins des utilisateurs, la gestion du budget

...etc. les gestionnaires des bibliothèques numériques sont demandés d'accomplir les différentes tâches de gestion des collections en créant un équilibre entre le développement et l'amélioration des collections et les conditions économiques de ces processus (Johnson, 2014). Dans les processus de gestion, les gestionnaires faire face plusieurs problèmes, on mentionne, le volume de donnée, le nombre des sources de données, et des processus à analyser.

Aujourd'hui plus que jamais, la gestion de collection exige aux gestionnaires des bibliothèques numériques académiques de prendre des décisions critiques. Ces décisions doivent être soumises à des conditions de budget limité de ces bibliothèques, les besoins des utilisateurs, et le nombre énorme des travaux publiés chaque jour. Pour prendre des décisions appropriées à ces collections, les gestionnaires doivent déterminer les besoins des utilisateurs, évaluer leurs collections et évaluer les documents numériques dans ces collections, et évaluer les ressources numériques dans le marché. Une fois ces informations sont collectées, les gestionnaires peuvent les utiliser pour prendre des différentes décisions d'acquisitions, d'abonnement, de désabonnement, et de suppression, ...etc.

1.2 Définition de la problématique

Avec les demandes continues d'informations des utilisateurs des BNAs, et leurs attentes élevées d'une collection riche qui satisfait leurs besoins, et de bonnes qualité des services offerts. Les gestionnaires des bibliothèques numériques travaillent de façon continue sur l'amélioration de la qualité des collections tout en augmentant la disponibilité de l'information, et diminue la présence de l'information inutile, et l'amélioration des services ...etc.

L'amélioration de la collection de bibliothèques numériques académiques a pris une place principale dans la gestion des bibliothèques numériques. Il existe de nombreuses motivations pour améliorer la collection parmi ces motivations : garder les utilisateurs de la bibliothèque numérique académique par satisfaire leurs besoins d'informations, attirer de nouveaux utilisateurs, réaliser des profits pour les bibliothèques, ...etc.

La bonne gestion de collection de bibliothèque numérique veut une bonne évaluation de la collection d'une bibliothèque numérique, évaluation des ressources dans cette collection, et l'évaluation des ressources numériques dans le marché sont fait avec plusieurs critères

- **Problème principal**

Les différentes tâches de gestion de collection ont besoin de sélectionner des ressources numériques. Dans cette thèse nous nous intéressons à répondre à la question principale : Quels sont les meilleures ressources numériques à acquérir, abonner, désabonner, ou enlever ?

- **Sous problèmes**

Toutefois, le principal défi que rencontre les gestionnaires est comment déterminer les besoins de ces utilisateurs. Surtout que ces besoins sont intégrés dans des données volumineuses.

Chapitre 1. Introduction Générale

Pour avoir la possibilité de traiter et utiliser ces données, les gestionnaires sont obligés d'extraire ces données à partir des sources hétérogènes, de transformer leurs formats, et de les enregistrer dans des bases de données avec des formats standards. Mais, le problème major ici, est le manque des méthodes et d'outils efficaces pour accomplir cette tâche. Donc, le défi ici, c'est comment intégrer les données.

Autre problème, c'est comment utiliser ces informations pour évaluer la collection et les ressources numériques importantes pour les utilisateurs. Ces processus d'évaluation et de sélection doivent être réalisés avec la prise en considération de plusieurs critères contradictoires.

1.3 Présentation des contributions

Plusieurs travaux sont proposés pour le but d'offrir une assistance aux gestionnaires des bibliothèques numériques dans leurs différentes tâches de gestion. La plupart de ces travaux ont concentré sur l'évaluation des bibliothèques numériques, et de la qualité de ces services. Ces travaux fournissent aux gestionnaires une vue générale sur la qualité de ces services. Mais, ils ne donnent pas des propositions sur les solutions possibles en cas d'une évaluation négative de ces services.

Le but de cette thèse est de proposer une nouvelle approche pour la gestion des collections des bibliothèques numériques académiques. Cette approche va aider les gestionnaires de ces bibliothèques à évaluer et sélectionner les ressources numériques selon les besoins d'information de ces utilisateurs pour accomplir ses différentes tâches de gestion des collections.

Pour atteindre cet objectif on a proposé des contributions pour résoudre les problèmes reliés à la gestion des collections.

Contribution pour le problème de sélection des ressources numériques

Le problème de sélection des ressources numériques pour l'acquisition, l'abonnement, ou le désabonnement est très difficile, sa difficulté vient de la complexité de déterminer les besoins d'information des utilisateurs, et l'évaluation de ces ressources avec plusieurs critères contradictoires. Actuellement, les gestionnaires des bibliothèques numériques ne possèdent pas d'outils ou de méthodes efficaces pour évaluer et sélectionner ces ressources (Hemili & Laouar, 2018a).

Dans l'objectif de résoudre ce problème, nous avons proposé un système d'aide à la décision pour aider les gestionnaires à sélectionner les ressources numériques (Mohamed, Mohamed, & Sean, 2019). Ce système analyse les fichiers logs des bibliothèques numériques pour extraire les vrais besoins d'information des utilisateurs. Une fois les besoins des utilisateurs sont déterminés, le système utilise la méthode d'analyse multicritère PROMETHE II pour évaluer et classer les ressources selon les besoins d'information déterminés. Enfin, le système fournit ces évaluations et classements des ressources aux gestionnaires pour les aider à sélectionner les ressources appropriées.

Chapitre 1. Introduction Générale

Contribution pour l'évaluation des collections des bibliothèques numériques académiques

En raison du budget limité, l'allocation du budget pour améliorer les services des bibliothèques numériques devient un processus très important. Les gestionnaires doivent être plus conscients dans la sélection des services qui ont besoins plus d'amélioration. Pour cette raison, l'évaluation des collections des bibliothèques numériques est un élément clé pour une bonne allocation du budget.

Les gestionnaires rencontrent un problème principal dans l'évaluation des collections qui est le nombre des critères impliqués, et la contradiction entre ces critères. Ce problème rend l'évaluation très compliquée aux gestionnaires. On a proposé une approche basée sur l'aide à la décision multicritère pour résoudre ce problème. Cette approche utilise les critères impliqués dans l'évaluation des ressources numériques pour les évaluer, les valeurs de ces critères sont extraites des données de circulations des ressources, après on évalue la collection et les sous collections pour déterminer les parties qui ont le plus besoin d'améliorations.

Contribution pour les problèmes de manipulation de données

La gestion des collections des bibliothèques numériques académiques a besoin de l'utilisation des données qui sont stockées dans des sources hétérogènes, et elles sont des normes et formats différentes. Aussi cette gestion a besoin de consommation élevée de puissance de calcul pour manipuler ces données, comme on le sait les méthodes d'aide à la décision multicritère veulent beaucoup de comparaisons sous de nombreux critères.

Pour cela, on a proposé l'utilisation de deux technologies qui sont la technologie d'entreposage de données et la technologie de l'informatique en nuage. Pour résoudre le problème de sources de données hétérogène, normes, et formats des données différentes on a proposé une architecture d'entrepôt de données pour stocker les données en une seule base de données et sous le même format et norme, ce qui rend son utilisation plus facile et efficace.

Un autre problème faire face aux gestionnaire qui la puissance de calcul nécessaire pour traiter les données. Dans le but de résoudre ce problème, on a proposé une architecture d'un Cloud pour déployer le DSS et le Data Warehouse, cette architecture nous permet de profiter de la puissance de calcul du Cloud Computing.

1.4 Organisation de la thèse

Cette thèse est constituée de cinq chapitres, ce présent premier chapitre a présenté le contexte global de la thèse, la problématique à traiter, et les contributions proposées.

- ✓ **Chapitre 2** donne les définitions des concepts des bibliothèques numériques, et il présente le rôle joué par ces bibliothèques dans la communauté académique. Il focalise sur la gestion et le développement des collections des bibliothèques académiques et les défis qui faire face les gestionnaires de ces collections.
- ✓ **Chapitre 3** présente un état de l'art sur les systèmes d'aide à la décision, la notion de trace informatique, entreposage de données, l'informatique en nuage, et leurs rôles dans la résolution des différents problèmes de gestion. Et une étude comparative sur les travaux de recherches réalisés dans ce contexte.

Chapitre 1. Introduction Générale

- ✓ **Chapitre 4** expose l'utilisation de MCDA pour l'évaluation des collections des bibliothèques numériques académiques.
- ✓ **Chapitre 5** La contribution principale de la thèse est décrite dans ce chapitre. La contribution est un système d'aide à la décision basé sur la méthode PROMETHEE II pour gérer le développement des collections en fonction de la demande des utilisateurs dans les bibliothèques numériques universitaires. Aussi, ce chapitre présente l'architecture Data Warehouse proposée et l'architecture Cloud proposée.
- ✓ **Chapitre 6** présente une conclusion générale et quelques perspectives.

Chapitre 2

La Gestion des Bibliothèques Numériques

Contenu

2.1	<i>Introduction</i>	26
2.2	<i>Définition des concepts de base des bibliothèques numériques</i>	26
2.2.1	<i>Bibliothèque numérique, électronique, ou virtuelle ?</i>	26
2.2.2	<i>Définition d'une bibliothèque numérique</i>	27
2.2.3	<i>Définition de collection</i>	27
2.3	<i>Les concepts liés aux bibliothèques numériques</i>	28
2.3.1	<i>Service de bibliothèque numérique</i>	28
2.3.2	<i>Ressources numérique</i>	28
2.3.3	<i>Live numérique</i>	28
2.3.4	<i>Journal numérique</i>	28
2.3.5	<i>Les gestionnaires de bibliothèque numérique</i>	28
2.4	<i>La gestion de bibliothèque numérique</i>	29
2.5	<i>Le développement de collection</i>	31
2.5.1	<i>Collectages d'information sur les ressources</i>	31
2.5.2	<i>Catégorisation des ressources</i>	32
2.5.3	<i>Evaluation de ressources</i>	33
2.5.4	<i>Le processus de sélection</i>	33
2.5.5	<i>Abonnement ou acquisition</i>	34
i.	<i>Décision d'achat ou d'abonnement</i>	34
ii.	<i>Préparation de la commande et options d'acquisitions</i>	34
iii.	<i>Sélection du fournisseur approprié</i>	36
2.5.6	<i>Construction de collection</i>	37
2.6	<i>La gestion de collection</i>	37
2.6.1	<i>Collectage des données d'utilisation</i>	38
2.6.2	<i>Evaluation des ressources</i>	38
2.6.3	<i>Sélection de ressources</i>	39
2.6.4	<i>Désabonnement ou renouvellement</i>	39
2.6.5	<i>Mise à jour du catalogue</i>	40

Chapitre 2. La Gestion des Bibliothèques Numériques

2.7	<i>Défis liés au développement et à la gestion des collections</i>	40
2.7.1	<i>Détermination des besoins d'information des utilisateurs</i>	40
2.7.2	<i>L'évaluation de la collection</i>	40
2.7.3	<i>Sélection des ressources numériques</i>	40
2.7.4	<i>L'analyse des données hétérogènes</i>	41
2.7.5	<i>La puissance de calcul</i>	41
2.8	<i>Conclusion</i>	41

2.1 Introduction

Aujourd'hui, l'information est disponible sur Internet pour tout public. Les personnes peuvent accéder à n'importe quel type d'information facilement, des informations peuvent être créées par des amateurs, à cet effet, la qualité de ces informations varie.

Cependant, les bibliothèques numériques ont pris une place croissante en tant que source d'information de qualité. Elles offrent des ressources d'information plus sûres et plus organisées. Les bibliothèques numériques peuvent être destinées à tout public, ou à un public spécifique.

Les bibliothèques numériques académiques sont des bibliothèques numériques destinées aux publics universitaires et académiques. Elles jouent un rôle principal dans la diffusion des savoirs entre chercheurs, enseignants, étudiants, et tout autre personnel académique. Ainsi, elles jouent un rôle important dans l'apprentissage et la recherche scientifique par fournir des ressources d'information fiable, sûr, et garantir l'accessibilité à ces ressources de n'importe quel endroit et à tout moment.

La gestion des bibliothèques numériques rencontre plusieurs problèmes tels que la gestion des services de bibliothèque numérique, le recouvrement des coûts ressources numériques, les problèmes de droit d'auteur, la détermination des besoins des utilisateurs et la sélection des ressources numériques pour les différentes activités de développement et de gestion de collections.

Les objectifs de ce chapitre sont donc d'une part de préciser le contexte dans lequel s'inscrivent nos travaux de recherche et d'autre part de présenter quelques problèmes rencontrés par les gestionnaires des bibliothèques numériques dans les activités de gestion de ces bibliothèques, celle de l'analyse de gros volume de données hétérogènes pour extraire les informations nécessaires dans le processus de prise de décision.

2.2 Définition des concepts de base des bibliothèques numériques

Ce chapitre explore des définitions et des concepts sur les bibliothèques numériques. Des concepts de base nécessitent d'être clarifiés tels qu'une collection de bibliothèques numériques, un gestionnaire de collection et les ressources numériques.

2.2.1 Bibliothèque numérique, électronique, ou virtuelle ?

Ces termes ont été utilisés pour expliquer l'utilisation de la technologie numérique au sein des bibliothèques : acquérir, stocker, conserver et diffuser leurs contenus à des utilisateurs distants.

Dans cette thèse, nous adoptons le terme « bibliothèque numérique » ou bien « digital library (DL) » en anglais, parce que d'une part, nous pensons que c'est le terme le plus adapté pour désigner les contenus numériques d'une bibliothèque, et d'autre part, c'est désormais le terme le plus utilisé dans la littérature en informatique pour désigner les bibliothèques qui diffusent leur contenu sur le web voire Figure 2.1.

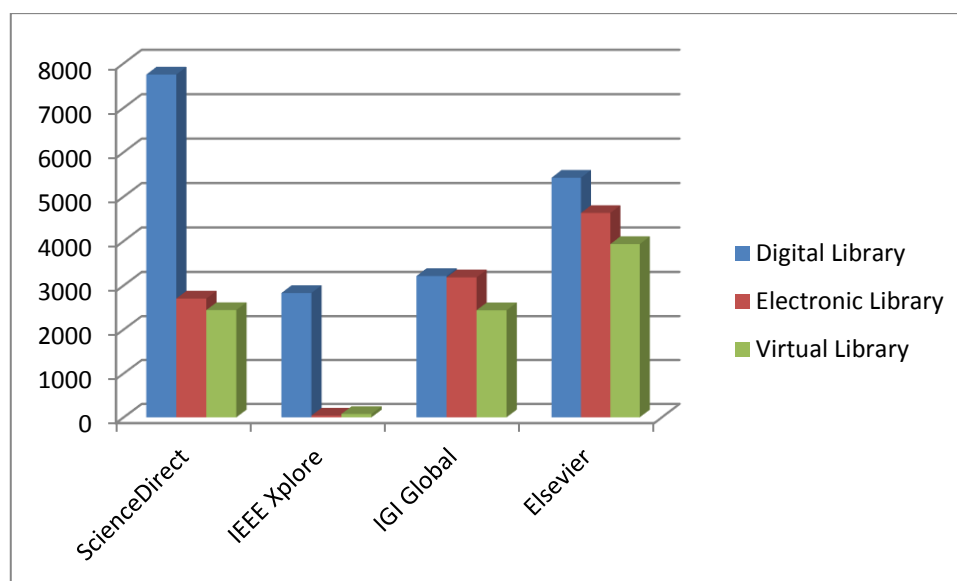


Figure 2.1 Répartition des termes (Digital library, Electronic library, Virtual library) (elsevier.com, 2019; ieeeexplore.ieee.org, 2019; igi-global.com, 2019; sciencedirect.com, 2019)

2.2.2 Définition d'une bibliothèque numérique

Dans la littérature, de nombreuses définitions ont été proposées pour définir les bibliothèques numériques. La façon de décrire une bibliothèque numérique diffère selon la communauté de la recherche qu'elle l'étudie (Carrigan, 1996; Cleveland, 1998). Ci-dessous trois définitions présentées pour clarifier la notion de bibliothèque numérique.

Définition 1 : *Les bibliothèques numériques sont des organisations qui fournissent des ressources, y compris le personnel spécialisé, pour sélectionner, structurer, offrir un accès intellectuel pour interpréter, distribuer, préserver l'intégrité de documents sous forme numérique, et assurer la maintenance des collections d'œuvres numériques, afin qu'ils soient facilement et économiquement utilisables par une communauté définie ou un ensemble de communauté (DLF, 1998).*

Définition 2 : *Les bibliothèques numériques proposent de véritables collections numériques, selon une politique documentaire déterminée. Elles sont alimentées soit par des opérations de numérisation (documents patrimoniaux ou non), soit par des documents nativement numériques. Les contenus sont organisés pour en faciliter la consultation (ENSSIB, 2012).*

Définition 3 : *La bibliothèque numérique est une collection ciblée d'objets numériques, y compris le texte, la vidéo et l'audio, ainsi que des méthodes d'accès, de récupération, de sélection, d'organisation et d'entretien de la collection (Witten, Bainbridge, & Nichols, 2009)*

2.2.3 Définition de collection

Une collection dans une bibliothèque en générale est constituée des matériaux de tous les formats et les genres qu'une bibliothèque possède ou auquel elle fournit un accès à distance, par achat ou location (Johnson, 2018).

Basé sur la définition de Johnson d'une collection de bibliothèque de façon générale, on peut définir une collection d'une bibliothèque numérique comme l'ensemble de

Chapitre 2. La Gestion des Bibliothèques Numériques

ressources numériques qui possède cette bibliothèque, et elle fournit un accès à distance à ces ressources.

2.3 Les concepts reliés aux bibliothèques numériques

Les principaux concepts utilisés dans le contexte des bibliothèques numériques sont les suivants :

2.3.1 Service de bibliothèque numérique

Un service désigne une fonctionnalité qui répond aux besoins des utilisateurs des bibliothèques numériques.

Les bibliothèques numériques peuvent potentiellement prendre en charge une gamme de services de bibliothèque traditionnels et non traditionnels. Bien que les projets de recherche avancés sur la DL exploitent les possibilités d'un monde exclusivement numérique, les bibliothèques numériques fonctionnelles prennent en charge pour la plupart des fonctions qui ressemblent beaucoup à leurs homologues de brique et de mortier, des formes légèrement différentes. Une DL fourni aux utilisateurs plusieurs services tel que la recherche et récupération d'informations, Formation des utilisateurs et la Facilitation de l'apprentissage et activités de collaboration (Schwartz, 2000).

2.3.2 Ressources numérique

Les ressources d'information numériques sont les sources d'information offertes sous un forme (Dadzie, 2005). Les ressources numériques peuvent inclure une source d'informations, telle qu'une base de données, livre numérique ou journal numérique en ligne / hors ligne, fourni via un réseau, tel qu'une zone locale réseau intranet, ou Internet (Ryan, McClure, & Bertot, 2001).

2.3.3 Live numérique

C'est un document numérique, sous licence ou non, dans lesquels le texte interrogeable est prédominant, et qui peuvent être vus comme une analogie avec un livre imprimé (monographie). L'utilisation de livres numériques dépend dans de nombreux cas d'un appareil dédié et / ou d'un lecteur spécial ou d'un logiciel de visualisation (NISO, 2005).

2.3.4 Journal numérique

Les revues numériques, également appelées e-journal en anglais, périodiques numériques, sont des revues académiques publiés sur Internet et accessibles à distant (Tenopir & King, 2000).

La majorité des journaux en ligne s'appuient sur des formats HTML et PDF parfois sont utilisés indépendamment. Un faible nombre font appel à Microsoft Word pour la publication de textes.

2.3.5 Les gestionnaires de bibliothèque numérique

Un gestionnaire de bibliothèques numériques (bibliothécaire, ou librarian en anglais) désigne d'une manière générale une personne à laquelle confiées des tâches de gestion,

Chapitre 2. La Gestion des Bibliothèques Numériques

de collections documentaires, de tâches d'aide aux usagers, ainsi, des fonctions d'administration générale (gestion du personnel, informatique, finances, action culturelle) (Sreenivasulu, 2000).

Dans cette thèse nous allons utiliser gestionnaire au lieu de bibliothécaire ou librarian.

2.4 La gestion de bibliothèque numérique

La bibliothèque numérique est une bibliothèque basée sur les besoins de l'utilisateur. Une bibliothèque connecte entièrement les utilisateurs aux informations dont ils ont besoin numériquement, quelle que soit leur source.

La gestion d'une bibliothèque numérique diffère de la gestion d'une bibliothèque traditionnelle de plusieurs façons, voici quelques facteurs qui font la différence:

- La distance par rapport aux utilisateurs ;
- Les mécanismes spécifiques de fourniture des produits/services ;
- La technologie et l'organisation qui doivent être planifiés et gérés différemment (Powell, 1994).

La gestion d'une bibliothèque numérique implique plusieurs tâches à faire tels que :

- Le recouvrement des coûts des ressources ;
- La gestion de droit d'auteur ;
- La formation des gestionnaires ;
- Planification de l'infrastructure: une bibliothèque numérique a besoin de composants informatiques, de mobilier, d'espace, etc.
- Planification des ressources d'information ;
- La maîtrise de méthodes et techniques d'accès à l'information ;
- La maîtrise de méthodes à adopter pour le développement de collection.
- Gestion de ressources humaines ;
- Gestion du budget ;
- Gestion de collection ;
- Développement de collection (Kumbhar & Harake, 2015).

- **Le monde des ressources numériques**

Chaque année le nombre des publications augmente Figure 2.2, par conséquent, sélectionner parmi le grand nombre des ressources publiées peut sembler une tâche très difficile. Les gestionnaires des bibliothèques numériques rencontrent les contraintes de l'augmentation du coût des ressources et du grand nombre de ressources.

Chapitre 2. La Gestion des Bibliothèques Numériques

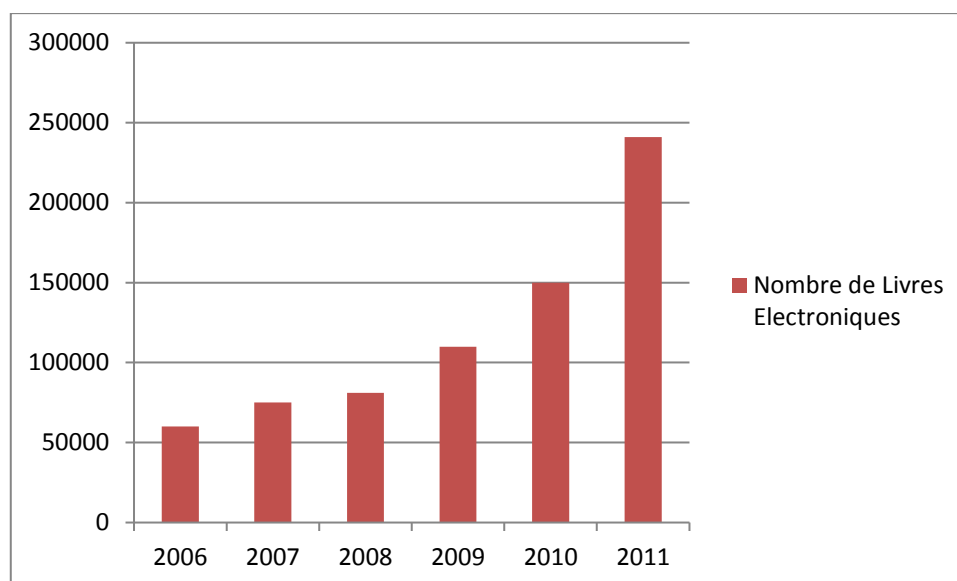


Figure 2.2 L'augmentation de Nombre de Livres Electroniques Publiés (Bowker, 2012)

L'UNESCO ne recueille plus de données sur le nombre de livres publiés par pays et par an, mais une entrée de Wikipédia, tirée de diverses sources, estime que plus de deux millions de titres sont publiés chaque année (Wikipedia, 2019).

Bowker a indiqué que la production de livres imprimés traditionnels aux États-Unis avait augmenté de 6% entre 2010 et 2011 pour atteindre un nombre projeté de 347 178 titres (Bowker, 2012). Même s'il s'agissait de la plus forte augmentation en plus de quatre ans, elle était due à la formidable croissance du marché de l'autoédition. Selon Bowker, aux États-Unis, les livres auto-publiés ont augmenté de 286% entre 2006 et 2011, chiffre qui ne reflète que les livres ISBN publiés (Bowker, 2012). Bien que des données sur les ventes de livres numériques soient disponibles, le nombre de nouveaux livres numériques publiés chaque année ne l'est pas. Le nombre de périodiques actifs continue également de croître. Lorsqu'il a été consulté en 2013, le site Ulrichsweb.com de Bowker a répertorié plus de 234 650 revues académiques et scientifiques (principalement des titres révisés par des pairs, y compris de nombreuses publications en accès libre), des magazines populaires, des journaux, des lettres d'information et d'autres types de périodiques en version imprimée et numérique autour du monde (Bowker, 2012).

Le coût des matériaux augmente de plus de plus de vingt-cinq ans et dépasse généralement le budget d'acquisition des bibliothèques des États-Unis. Entre 2007 et 2008, le prix des livres numériques a augmenté de 0,1% et le prix moyen des périodiques américains a augmenté de 8,0%. le prix moyen des ouvrages universitaires nord-américains publiés en 2010 a augmenté de 12,4% par rapport à 2009, par rapport à une augmentation de 1,5% de prix des livres numériques; pendant la même période, les livres à couverture rigide ont augmenté de 5,54% et le marché de masse livres de poche 1,94 pour cent (Board & Tafuri, 2017).

Avec l'augmentation de nombre de ressources numériques et de leurs prix, et avec les budgets limités des bibliothèques numériques, les bibliothèques subissent une pression constante pour fournir davantage de ressources numériques et font face à une confusion éventail d'options d'achat et de location de contenu numérique. Plus de matériaux parmi lesquels choisir, des prix à la hausse, des budgets limités et l'intérêt

Chapitre 2. La Gestion des Bibliothèques Numériques

des utilisateurs pour le contenu numérique signifient que les gestionnaires de collections doivent faire preuve de plus en plus de discrimination dans leurs choix.

Dans les sections suivantes on va exposer les activités impliquées dans les processus de développement et de gestion des collections des bibliothèques numériques.

2.5 Le développement de collection

Le développement d'une collection est le processus de construction ou de constitution d'une collection de bibliothèque en réponse aux priorités institutionnelles et aux besoins et intérêts de la communauté ou des utilisateurs (Carrigan, 1996; Johnson, 2014).

La sélection entre deux options ou plus fait partie de presque toutes les décisions prises par les gestionnaires des bibliothèques numériques pour la mise en œuvre des objectifs de développement et de gestion des collections, par conséquent le choix est l'essentiel dans le développement de la collection (Carrigan, 1996). La Figure 2.3 illustre les étapes principales du développement de collection.

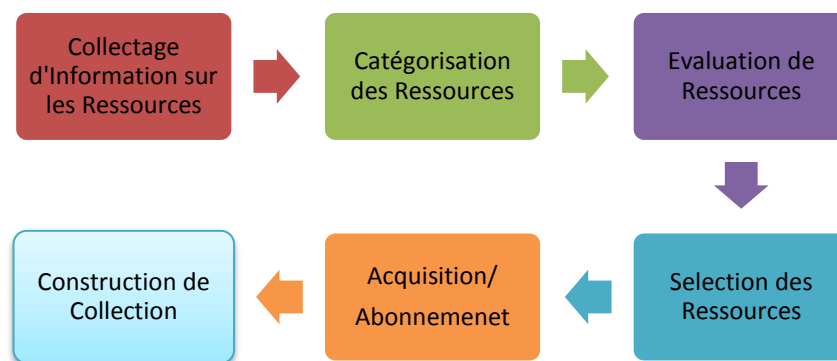


Figure 2.3 Processus de Développement de Collection (P. A. Kumar, 2017)

Dans les sous sections suivantes on va couvrir les activités de développement des collections.

2.5.1 Collectages d'information sur les ressources

Dans cette étape, les gestionnaires des bibliothèques numériques collectent les informations sur les ressources numériques. Les informations collectées seront utilisées ultérieurement pour identifier les ressources.

Il existe de nombreux outils et ressources pour aider les gestionnaires des bibliothèques numériques à trouver les informations factuelles de base sur les auteurs, les titres, les éditeurs et les sujets nécessaires pour identifier les éléments pouvant être sélectionnés. Les bibliographies et les listes peuvent être publiées par les bibliothèques, les éditeurs de bibliothèques, les systèmes scolaires, les associations professionnelles et les éditeurs commerciaux. Les bibliographies nationales et les listes de commerce sont des outils standards dans les bibliothèques depuis des décennies. Les bibliothèques consultent souvent les listes d'accession récentes, établies par d'autres bibliothèques. Les listes recommandées sont préparées par les associations de bibliothèques et d'autres associations professionnelles (Treviño & Treviño, 2006). Les bibliographies publiées

Chapitre 2. La Gestion des Bibliothèques Numériques

par les éditeurs commerciaux sont généralement disponibles sous forme de ressource en ligne (qui est constamment mise à jour).

Le résultat de cette étape sera utilisé pour catégoriser, évaluer, sélectionner, abonner ou acquérir des ressources.

2.5.2 Catégorisation des ressources

Une première étape fréquente dans la sélection des ressources consiste à les séparer en catégories, les bibliothèques peuvent attribuer à des responsables et personnels la tâche de catégorisation.

Plusieurs topologies de catégorisation ont été utilisées, tel que :

- **La catégorisation typique par le format :** par exemple, les enregistrements vidéo et audio et les autres ressources numériques. Le format indique souvent comment les ressources sont manipulées.
- **La catégorisation par le genre :** elle comporte les monographies, séries monographiques, manga et anime, magazines, thèses, partitions musicales, journaux, revues, manuels et documents gouvernementaux. Un seul genre peut être présenté dans plusieurs formats. Par exemple, les publications en série peuvent être acquises sous forme imprimée, microforme ou numérique (Al-Natsheh, Martinet, Muhlenbach, Rico, & Zighed, 2018).
- **La catégorisation par sujet :** Celles-ci peuvent être de grandes divisions (sciences humaines, sciences sociales, sciences), plus étroites (littérature, sociologie, ingénierie), ou très raffinée (littérature américaine, sciences sociales de la famille, génie chimique).
- **La catégorisation par langue :** dans laquelle ils sont produits ou la zone géographique dans laquelle ils sont publiés ou qu'ils couvrent. Ils peuvent être considérés en fonction de l'âge du lecteur auquel ils sont destinés: enfants, jeunes adultes. Ceux-ci peuvent également être subdivisés (par exemple, des livres d'images, des lecteurs précoces).
- **La catégorisation par type d'éditeur :** par exemple, les livres à couverture rigide pour adultes, les fictions commerciales et les livres de poche sur le marché de masse. Les bibliothèques universitaires et les bibliothèques de recherche peuvent faire la distinction entre ressources primaires (documents sources), secondaires (revues, résumés sur l'état de la technique, manuels, interprétations de sources primaires) et ressources tertiaires (reconditionnement de la littérature primaire dans les traitements populaires, annuelles, manuels et encyclopédies).

Les topologies guident comment les revues, les listes de publications et les introductions à la littérature sont organisées ou définies (Flynn, 1993).

De nombreuses bibliothèques fusionnent certaines catégories lors de l'attribution et de la gestion des responsabilités de sélection. Les bibliothèques universitaires peuvent utiliser une combinaison de spécialistes par matière ou discipline, de bibliothécaires pour l'étude de zones géographiques et de documents gouvernementaux.

Chapitre 2. La Gestion des Bibliothèques Numériques

Les bibliothécaires chacun avec des responsabilités de collecte associées. Les bibliothèques publiques peuvent classer par groupe de lecteurs, par domaine ou par une combinaison. Les bibliothèques scolaires peuvent penser en termes de fiction et de non fiction, d'âge du lecteur et de domaine (Database.com, 2013).

2.5.3 Evaluation de ressources

L'évaluation aide les gestionnaires des collections à décider si le titre doit être ajouté à la collection ou non. L'évaluation examine les qualités intrinsèques des ressources (P. A. Kumar, 2017), et les ressources sont considérées individuellement. L'évaluation considère la capacité des ressources à répondre aux besoins locaux. En pratique, un gestionnaire de collections ne devrait pas consacrer de temps à l'évaluation d'un élément si celui-ci ne convient pas à la communauté d'utilisateurs servie.

Dans l'évaluation, la ressource est considérée par rapport aux besoins des utilisateurs, à la collection existante, à la mission de la bibliothèque et aux obligations du consortium.

Les critères d'évaluation varient d'une ressource à une autre et entre les catégories de publications, mais incluent généralement plusieurs des considérations suivantes:

- ✓ Le sujet
- ✓ La langue
- ✓ La véracité
- ✓ Le style d'écriture (par exemple, bien écrit, facile à lire, aspects esthétiques)
- ✓ La réputation, références, ou autorité de l'auteur, de l'éditeur, éditeur, relecteurs fréquence à laquelle le titre est référencé dans des bibliographies ou des citations
- ✓ Lecture ou niveau utilisateur vers lequel le contenu est dirigé
- ✓ Fréquence des mises à jour ou des révisions
- ✓ Points d'accès (index, niveau de détail dans la table des matières, par exemple)
- ✓ Facilité d'utilisation
- ✓ Ressources externes qui indexent la publication
- ✓ Disponibilité du matériel nécessaire à l'audition ou au visionnage de documents audiovisuels Matériel
- ✓ Coût en relation avec la qualité de l'article et son utilisation projetée
- ✓ L'élément soutient-il le programme, les intérêts de la communauté, les spécialités du corps enseignant ou de l'enseignant ou les domaines de recherche en cours?

2.5.4 Le processus de sélection

La sélection est à la fois un art et une science. Il résulte d'une combinaison de connaissances, d'expérience et d'intuition. Les gestionnaires des collections expérimentées peuvent avoir du mal à expliquer exactement comment ils décident quoi acquérir ou non.

Rutledge et Swindler proposent un modèle mental qui attribue une valeur pondérée à chaque critère considéré (Rutledge & Swindler, 1987). Ils suggèrent qu'un gestionnaire des collections travaille à travers ce modèle mental et aboutit à l'une des trois conclusions suivantes:

- Le titre doit être ajouté à la collection ;
- Le titre devrait être ajouté ;

➤ Le titre pourrait être ajouté.

Williams explore le fonctionnement de l'esprit dans le processus de décision, en citant le rôle de la reconnaissance, «un processus de décision automatique ou délibératif selon lequel un signal est soumis à une sorte de test de familiarité et à une réponse affirmative ou négative (Williams, 1991)». Le gestionnaire détermine si l'élément est approprié et aide à répondre aux questions sur la pertinence du contenu et sur la connaissance de l'auteur, de l'éditeur ou du titre. Williams note que les capacités de reconnaissance sont renforcées par la construction fréquente, routinière et répétée des collections. La maîtrise passe par ces activités.

À cela s'ajoute une connaissance de la littérature dont le gestionnaire est responsable. Le gestionnaire qui maîtrise parfaitement ces éléments est en mesure de commencer la sélection.

La sélection se base sur le résultat de l'évaluation des ressources numériques.

2.5.5 Abonnement ou acquisition

Cette tâche passe par les étapes suivantes : (i) Décision d'achat ou d'abonnement, (ii) Préparation de la commande et options d'acquisitions ou d'abonnements, (iii) Sélection de fournisseurs

i. Décision d'achat ou d'abonnement

Une fois le gestionnaire a examiné toutes les questions d'évaluation, il est prêt à ajouter ou à rejeter la ressource à la collection. La décision d'ajouter une ressource est généralement considérée comme une décision d'achat ou d'abonnement, bien que la bibliothèque puisse payer pour posséder une ressource ou payer le droit d'accéder à une ressource ou à une collection de documents. Atkinson fait référence à l'univers des ressources non sélectionnés localement comme anti-collection. Il soutient que la sélection est «dans une large mesure, une série continue de décisions sur les ressources qui doivent être déplacés dans la collection» et suggère que la décision de sélection est relativement simple, car le gestionnaire n'a que deux options: ajouter ou ne pas ajouter (Shreeves, 2018).

Pour Atkinson, la limite entre l'acceptation et le rejet de ressource dépend principalement des ressources financières disponibles. Compte tenu du volume de documents publiés et de la nature limitée des budgets des bibliothèques, les gestionnaires seront toujours confrontés à des choix quant à ce qu'il ne faut pas ajouter, mais aussi à ce qu'ils doivent ajouter.

ii. Préparation de la commande et options d'acquisitions

Les processus et les systèmes de commande et d'obtention des documents de la bibliothèque après leur sélection sont appelés acquisition. L'acquisition de documents est étroitement liée au développement des collections, bien que dans la plupart des bibliothèques de moyenne et grande bibliothèques, la sélection et l'acquisition soient gérées par différentes personnes pouvant être situées dans différents départements ou unités de bibliothèques.

Les responsabilités liées aux acquisitions comprennent généralement la passation de commandes, voir Figure 2.4, la réclamation, l'annulation, la réception, le traitement de

Chapitre 2. La Gestion des Bibliothèques Numériques

factures et la préparation de demandes de propositions émanant de vendeurs monographiques et d'agents d'abonnement. Elles peuvent inclure le traitement de paiements. Les directives comptables classiques exigent que la responsabilité de l'approbation du paiement et du traitement du paiement soit attribuée à différentes personnes. La sélection et les acquisitions peuvent être traitées séparément dans des bibliothèques plus petites si le nombre de membres du personnel rend cela raisonnable. Cependant, la combinaison des fonctions de sélection et d'acquisition est courante. La facilité avec laquelle les gestionnaires peuvent travailler directement avec les bases de données en ligne des fournisseurs brouille la division traditionnelle du travail. Les gestionnaires peuvent passer des commandes directement en ligne dans le cadre du processus d'identification des ressources (Johnson, 2014).

Selector Name:	<input type="text"/>	
Budget Code:	Select code from drop-down menu <input type="button" value="v"/>	
Title(s): If this is a package purchase, indicate if you want "all titles" or "selected titles" only. If you wanted selected titles from the publisher, list all titles on this form.		
<input type="text"/>		
Publisher/Vendor: If you have been working with a particular sales representative, provide all relevant contact information.		
<input type="text"/>		
Contract Information: Is a contract with this publisher or vendor on file? Yes <input type="checkbox"/> No <input type="checkbox"/>		
Resource Type	Print Duplication	Timing
<input type="checkbox"/> Book	Do you want to retain an existing print subscription?	<input type="checkbox"/> Trial only
<input type="checkbox"/> DVD	<input type="checkbox"/> Yes	<input type="checkbox"/> One-time or back-file purchase
<input type="checkbox"/> Index	<input type="checkbox"/> No	<input type="checkbox"/> Subscription or standing order
<input type="checkbox"/> Journal		
<input type="checkbox"/> Streaming media		
Cost: If the resource is "free with print," provide the cost of the print subscription. Also note any one-time or ongoing costs.		
<input type="text"/>		
Additional Information: If unlimited access is not available, how many simultaneous users? Is extra software required, e.g., browser plug-ins, client program. If this is an e-book package, are MARC records available? If this is a trial, when should it start and end?		
<input type="text"/>		
<input type="button" value="Submit"/>	<input type="button" value="Reset"/>	

Figure 2.4 Exemple de formulaire de demande de ressource numérique interne en ligne (Johnson, 2014).

iii. Sélection du fournisseur approprié

Les bibliothèques peuvent passer des commandes auprès des fournisseurs et des agents, acheter auprès d'un détaillant, commander directement auprès d'éditeurs ou utiliser des agrégateurs. La sélection du fournisseur approprié implique plusieurs considérations et une évaluation minutieuse des avantages et des inconvénients. Les types de fournisseurs inclus :

- **Vendeurs et agents d'abonnement**

Les bibliothèques acquièrent des documents par l'intermédiaire d'un tiers en raison de leur capacité à regrouper les commandes, à recevoir des factures consolidées et à compter sur le tiers pour traiter les réclamations. Les vendeurs (parfois appelés grossistes, marchands de livres ou vendeurs de livres) peuvent fournir des livres imprimés, des livres numériques, de la musique, de l'audio et de la vidéo. Ils regroupent normalement les envois et offrent des rabais.

La différence entre les vendeurs de livres et les agents d'abonnement devient de moins en moins nette. Les agents d'abonnement vendent maintenant des livres numériques et les vendeurs de livres vendent collections de livres numériques. Bosch et ses collègues identifient le support suivant que les bibliothèques attendent des fournisseurs et des agents:

- **Détaillants**

L'objectif principal des détaillants est de vendre directement aux consommateurs. Les bibliothèques achètent parfois à des détaillants si elles ont besoin d'une ressource rapidement et le détaillant peut garantir une livraison rapide ou un membre du personnel de la bibliothèque peut se rendre dans un magasin local. Amazon.com est un détaillant fréquemment utilisé. Une des complications pour les institutions publiques dans leurs relations avec les détaillants est de s'assurer qu'aucune taxe n'est perçue. Les détaillants n'offrent généralement pas de réduction aux bibliothèques, mais le prix affiché (plus les frais d'expédition, le cas échéant) peut être justifié si le besoin est urgent.

- **Les éditeurs**

Pendant de nombreuses années, la plupart des éditeurs de livres n'ont pas considéré les bibliothèques comme leur principal marché. Le modèle commercial prédominant consistait à travailler avec des fournisseurs. Bien que les bibliothèques puissent également acheter directement auprès des éditeurs, elles ne bénéficiaient pas du rabais offert par les fournisseurs. Les avantages pour les éditeurs qui travaillent avec des fournisseurs sont similaires à ceux des bibliothèques: vous évitez de préparer des factures individuelles et d'expédier de petits colis. Les éditeurs de livres ont offert des rabais importants aux vendeurs, dont une partie a été transmise aux bibliothèques. Le passage au contenu numériques motivé les éditeurs à vendre activement aux bibliothèques en raison de la facilité de vente de paquets ou de paquets de contenu.

- **Les agrégateurs**

Les agrégateurs sont des tiers offrant un accès au texte intégral de périodiques, articles, livres ou supports publiés à l'origine par plusieurs éditeurs et fournissant un

Chapitre 2. La Gestion des Bibliothèques Numériques

accès en ligne au contenu via une interface ou une plate-forme commune. Agrégateurs telles que Ebook Library (EBL), EBSCO eBooks (anciennement Net Library), ebrary, OverDrive et la bibliothèque de nuages 3M hébergent le contenu de plusieurs éditeurs pour un accès en ligne.

- **Les échanges**

Certaines bibliothèques universitaires utilisent les échanges comme une forme de sélection en masse (Kristi Ekonen, 2006). Les échanges se font le plus souvent avec des partenaires étrangers et peuvent fournir des supports qui ne sont pas disponibles autrement ou de façon plus économique que l'achat direct. La bibliothèque fournit des publications locales à une bibliothèque ou à une institution partenaire étrangère, qui envoie ses publications de son propre pays à la bibliothèque. Les partenaires peuvent être des bibliothèques, des sociétés et associations savantes, des départements universitaires, des académies et des instituts de recherche.

- **Cadeaux et autres ressources gratuits**

Les cadeaux peuvent apporter des objets individuels ou une collection des ressources à la bibliothèque. Un cadeau est transféré volontairement sans compensation. Tout cadeau qui n'est pas en espèces s'appelle un cadeau en nature et peut être un bien, un service ou une propriété. Les cadeaux peuvent être acheminés vers la bibliothèque par un acte de donation, document juridique qui transfère le titre du donateur à la bibliothèque sans exiger de paiement; un acte de donation peut contenir des conditions auxquelles la bibliothèque doit se conformer (Association, 2019).

2.5.6 Construction de collection

Une fois des ressources sont abonnées ou achetées, les gestionnaires doivent les ajouter à la collection. Ainsi, ils intègrent ces ressources dans le catalogue de la bibliothèque numérique (Egan, Yearwood, & Kendrick, 2016).

2.6 La gestion de collection

La gestion de la collection est un terme générique qui couvre toutes les décisions prises après qu'une ressource fait partie de la collection. Ces décisions sont prises sur la base de plusieurs critères tel que le budget, les changements dans les communautés d'utilisateurs de la bibliothèque et les priorités de l'organisation mère. La Figure 2.5 expose les différentes tâches impliquées dans la gestion de collection.

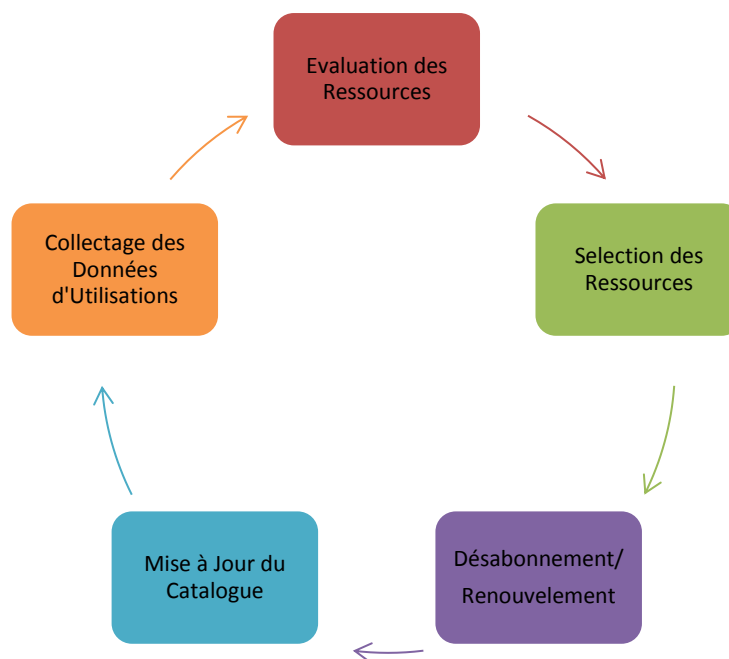


Figure 2.5 Cycle de Vie de Gestion de Collection (Shenton, 2003)

Dans ce qui suit nous présentons les étapes de gestion de collection, voir Figure 2.5.

2.6.1 Collectage des données d'utilisation

Les systèmes de gestion des bibliothèques numériques collectent constamment les données d'utilisation de ces services. Ils sauvegardent des statistiques sur la lecture, l'emprunt et la vente de ses ressources numériques. Ainsi, ils enregistrent toutes les transactions de ces utilisateurs avec ses services (Phillips & Kraemer, 2017).

Les données collectées seront utilisées plus tard pour des buts de gestion. Ils seront utilisés pour évaluer les services offerts, et pour l'amélioration des services.

2.6.2 Evaluation des ressources

Les bibliothèques numériques devraient évaluer régulièrement ces ressources numériques fournies. Cette évaluation continue est causée par le changement dans les prix des ressources, l'évolution avec le temps dans la pertinence du contenu abonné et le changement des besoins d'information des utilisateurs. La plupart des bibliothèques subissent une pression continue sur leurs budgets et doivent s'efforcer d'utiliser efficacement les fonds limités (Jareonruen & Tuamsuk, 2019).

L'évaluation de ressources numériques s'effectue sur la base de plusieurs critères, ces critères représentent le niveau d'utilisation des ressources. Dans le processus d'évaluation les gestionnaires exploitent les données d'utilisation collectées. L'International Federation of Library Associations and Institutions (IFLA) recommande les considérations suivantes pour l'évaluation de ressources:

- La ressource est-elle toujours pertinente pour les utilisateurs?
- L'usage augmente-t-il ou diminue-t-il par rapport aux années précédentes ou en comparaison avec des produits similaires?

Chapitre 2. La Gestion des Bibliothèques Numériques

- Comment le coût par utilisation de la ressource se compare-t-il à d'autres ressources?
- La ressource continue-t-elle de représenter une optimisation des ressources?
- Existe-t-il d'autres options d'accès (par exemple, paiement à la carte, contenu sélectionné ou un forfait) plus rentable?
- Le nombre d'utilisateurs simultanés est-il correctement défini?
- Comment l'utilisation du contenu actuel se compare-t-elle à celle du contenu du fichier précédent? (Weicher & Zhang, 2012)

2.6.3 Sélection de ressources

Une fois les ressources de la collection sont évaluées, les gestionnaires des bibliothèques numériques sélectionnent des ressources selon leurs évaluations pour désabonner ou renouveler à ces ressources (Shenton, 2003).

Les bibliothèques doivent déterminer dans quelle mesure les titres peuvent être désabonnés en une seule année.

2.6.4 Désabonnement ou renouvellement

Le processus de désabonnement des périodiques commence par un examen de ressources de la collection en parallèle avec autres fonctions de maintenance de la collection. Idéalement, les abonnements actifs sont examinés régulièrement afin de garantir que la collection continue de répondre aux besoins des utilisateurs et aux objectifs de la bibliothèque.

En réalité, l'identification des abonnements à désabonner est devenue une activité annuelle dans de nombreuses bibliothèques - peut-être la plupart - des vingt dernières années au moins, en raison de l'augmentation constante et rapide des prix dépassant les augmentations budgétaires. Bien que les bibliothèques universitaires et de recherche aient été plus durement touchées en raison de leur forte concentration de revues spécialisées onéreuses, toutes les bibliothèques ont connu une augmentation des coûts d'abonnement supérieure aux taux d'inflation nationaux et aux augmentations budgétaires (Phillips & Kraemer, 2017).

Les gestionnaires utilisent de nombreuses techniques pour rendre le processus de désabonnement aussi logique et défendable que possible. Chaque bibliothèque a besoin de politiques et de procédures pour guider les désabonnements et tenir les communautés d'utilisateurs informées et impliquées dans la mesure du raisonnable et du concret. Les mêmes critères (qualité et pertinence) qui guident la sélection d'un journal ou d'une autre ressource continue sont appliqués lors de son examen en vue de son désabonnement. L'utilisation est un critère déterminant. Les données peuvent être disponibles dans les statistiques de circulation et d'utilisation des ressources numériques, les demandes de prêt entre bibliothèques, les enquêtes auprès des utilisateurs ou les enregistrements d'utilisation interne (Carroll & Cummings, 2010; Thohira, Chambers, & Sprague, 2010).

Après la sélection des ressources les gestionnaires désabonnent de ses ressources. Dans le cas contraire ils renouvèlent l'abonnement dans les autres ressources.

Chapitre 2. La Gestion des Bibliothèques Numériques

2.6.5 Mise à jour du catalogue

Pour chaque désabonnement de ressources numériques les gestionnaires doit enlever ces ressources du catalogue (Shenton, 2003). Un catalogue à jour est une nécessité pour satisfaire les utilisateurs.

2.7 Défis liés au développement et à la gestion des collections

Aujourd'hui, les gestionnaires des bibliothèques numériques rencontrent des difficultés dans les activités de développement et de gestion des collections. Cette section comporte une discussion qui passe en revue certains problèmes que les gestionnaires doivent faire face, et qui constituent actuellement des thématiques de recherche très actives.

2.7.1 Détermination des besoins d'information des utilisateurs

Les besoins d'information des utilisateurs sont la base de développement des collections des bibliothèques numériques, par conséquent, la détermination de ses besoins est essentiel dans les différents processus de prise de décision concernant les collections des bibliothèques numériques (Hemili, Laouar, & Eom, 2019).

Le nombre augmenté des utilisateurs et la distinction entre ses besoins d'information rendent la tâche de détermination des besoins des utilisateurs est très difficile. Les gestionnaires des collections des bibliothèques numériques ont besoins d'analyser des grands volumes des données pour extraire les besoins des utilisateurs. Ces besoins d'information sont cachés dans des fichiers log (Armstrong & Lonsdale, 2005; Walters, 2013).

2.7.2 L'évaluation de la collection

L'évaluation de collection implique plusieurs critères qui représentent les données d'utilisation de ces collections.

Le manque de méthodes efficaces pour exploiter ces critères dans l'évaluation de collection, sous collection et ressources dans la collection cause beaucoup de problèmes aux gestionnaires.

2.7.3 Sélection des ressources numériques

Pour effectuer des tâches de gestion de la collection des bibliothèques numériques comme l'abonnement, le désabonnement, ou l'acquisition des ressources numériques la sélection est primordiale. Le processus de sélection des ressources numériques implique plusieurs critères. Ces critères sont basés sur les besoins de la faculté et les besoins des utilisateurs. Ce processus implique l'analyse des grands volumes de données, des processus, et des sources des données hétérogènes. Cette analyse représente des défis qui faire face les gestionnaires des BNAs (Hemili & Laouar, 2018a; Hemili et al., 2019).

La sélection, l'acquisition et la gestion de ressources numériques peuvent être complexes et coûteuses. Ces activités ne deviennent rentables que lorsque les bibliothèques peuvent tirer parti d'économies d'échelle - que lorsque les procédures nécessaires peuvent être routinières et déléguées au personnel d'appui (Armstrong &

Chapitre 2. La Gestion des Bibliothèques Numériques

Lonsdale, 2005). Le même argument peut être avancé du point de vue du client. Les efforts nécessaires pour trouver et utiliser des ressources numériques ne seront probablement utiles que si les clients ont une chance importante de trouver des titres de ressources numériques qui répondent à leurs besoins. Comme Robert Slater l'a observé, «une masse critique de livres numériques est nécessaire dans le domaine de besoin d'un client avant qu'il commence à accepter, attendre et utiliser régulièrement ces collections numériques» (Slater, 2010). Construire une collection solide de livres numériques est cependant difficile, car de nombreux titres académiques ne sont disponibles dans aucun format numérique (Dewan, 2012). Par exemple les titres des livres scolaires ne représentent qu'un dixième de tous les livres numériques (Himma & Just, 2007). En revanche, la romance et l'érotisme comptent parmi les genres les plus vendus (Himma & Just, 2007).

2.7.4 L'analyse des données hétérogènes

Toute bibliothèque numérique doit avoir un staff compétent. Ce personnel doit posséder également des connaissances techniques pour la manipulation d'équipements numériques ainsi que d'informations numériques. Dans les bibliothèques numériques, la gestion fait face au défi de garder son personnel professionnel et compétent, pour le but de tenir la bibliothèque à jour et de mettre en œuvre de nouvelles activités. Les professionnels de la bibliothèque doivent également constamment mettre à jour leurs propres connaissances et compétences pour s'adapter à l'environnement numérique en rapide évolution (Hobohm, 2011; Walters, 2013).

Le personnel des bibliothèques numériques doit analyser des données hétérogènes pour le but d'obtenir les informations nécessaires au processus de prise de décision. Le manque des outils pour intégrer ces données cause des difficultés aux gestionnaires (Ashcroft, 2011).

2.7.5 La puissance de calcul

Étant donné que presque toutes les évaluations dans la bibliothèque numérique impliquent beaucoup de calcul. Le nombre de comparaisons et le volume de données à analyser peuvent potentiellement augmenter le temps d'obtention de résultat et consomment les ressources du calcul (Eschenfelder, 2008).

2.8 Conclusion

Dans ce chapitre, nous avons exposé les notions de base liées au domaine des bibliothèques numériques. Nous avons présenté les éléments clés liées à la gestion des bibliothèques numériques incluant le développement des collections et la gestion des collections. Nous avons focalisé sur l'importance du processus de sélection des publications pour les différentes activités de développement et de gestion des collections. Le chapitre a été clôturé par un passage en revue sur les difficultés qui entravent le développement et de gestion des collections. Nous avons aussi soulevé quelques défis auxquels les gestionnaires de ces collections sont encore confrontés.

Le prochain chapitre sera consacré aux systèmes d'aide à la décision et d'autres technologies qui peuvent contribuer dans les processus de gestion des collections, à la discussion des travaux de recherche dédiés à l'amélioration de gestion des bibliothèques numériques.

Chapitre 3

DSS pour les Bibliothèques Numériques

Contenu

3.1	<i>Introduction</i>	44
3.2	<i>L'aide à la décision</i>	44
3.2.1	<i>Définitions de l'aide à la décision</i>	44
3.2.2	<i>Objectif de l'aide à la décision</i>	45
3.2.3	<i>Pourquoi l'aide à la décision</i>	46
3.2.4	<i>Système d'aide à la décision (DSS)</i>	47
3.2.5	<i>Systèmes interactifs d'aide à la décision</i>	47
3.2.6	<i>L'aide à la décision multicritère</i>	48
3.2.1	<i>Les méthodes d'aide à la décision multicritères</i>	48
3.3	<i>Les requêtes de recherche des utilisateurs des BNAs: Les traces informatiques</i>	51
3.3.1	<i>Définitions</i>	51
3.3.2	<i>Des Traces informatiques, pourquoi faire ?</i>	51
3.3.3	<i>Objectifs des systèmes de traçage</i>	51
3.3.4	<i>Utilisation des traces enregistrées</i>	52
3.3.5	<i>Types des traces</i>	52
3.3.6	<i>Finalités des traces</i>	53
3.3.7	<i>Cycle de vie et qualité des traces</i>	54
3.4	<i>Data Warehouse (DW)</i>	56
3.4.1	<i>Définitions de Data Warehouse</i>	56
3.4.2	<i>Caractéristiques de Data Warehouse</i>	56
3.4.3	<i>Méthodes de conception de Data Warehouse</i>	57
3.4.4	<i>Objective de Data Warehouse</i>	58
3.5	<i>Cloud Computing (CC)</i>	58
3.5.1	<i>Caractéristiques du CC</i>	59
	<i>Service-libre à la demande</i>	59
	<i>Large accès au réseau</i>	59
	<i>Mise en commun des ressources</i>	59

Chapitre 3. DSS pour les Bibliothèques Numériques

<i>Élasticité rapide</i>	59
<i>Service mesuré</i>	59
3.5.2 <i>Modèles de livraison de service</i>	59
<i>Le modèle Software-as-a-Service (SaaS)</i>	60
<i>Le modèle Platform as a Service (PaaS)</i>	60
<i>Le modèle Infrastructure as a Service (IaaS)</i>	60
3.5.3 <i>Modèles de déploiement du Cloud</i>	61
1. <i>Cloud privé</i>	61
2. <i>Cloud communautaire</i>	61
3. <i>Cloud public</i>	61
4. <i>Cloud hybride</i>	62
3.6 <i>Travaux connexes</i>	62
3.6.1 <i>Les approches générales pour l'évaluation des BNs</i>	62
3.6.2 <i>Analyse comparative des approches existantes</i>	63
3.7 <i>Conclusion</i>	65

3.1 Introduction

Le problème qui est à l'origine de cette thèse est la difficulté pour un gestionnaire de bibliothèque numérique académique (BNA) d'évaluer la collection de cette bibliothèque et d'évaluer les ressources numériques selon les besoins des utilisateurs afin de prendre des décisions de gestion pour développer ou améliorer ces collections. Un système d'aide à la décision (DSS) peut aider un gestionnaire dans le processus de détermination des besoins de ces utilisateurs, et dans l'exploitation de ces besoins pour prendre des décisions de développement et de gestion des collections ce qui permet aux BNA d'offrir des collections riches afin de satisfaire ses utilisateurs.

Le problème d'évaluation des bibliothèques numériques en générale a été étudié par plusieurs chercheurs de différents domaines. A cet égard, la question d'évaluation de ressources numériques et des collections des BNA est le cœur de l'évaluation des BNAs. La littérature existante est très riche en travaux qui ont tenté de relever les défis inhérents à cette problématique de recherche.

Ce chapitre présente les systèmes d'aide à la décision, en commençant par une description de ce qu'est l'aide à la décision et les différents systèmes qui existent pour arriver plus particulièrement aux systèmes qui nous intéressent. Ensuite nous exposons la notion de trace informatique. Aussi, nous verrons les technologies de Cloud Computing et de Data Warehousing qui peuvent être utilisé pour améliorer les performances de DSS. Enfin, ce chapitre présente une exploration, une étude et une analyse des travaux connexes qui traitent, d'une manière ou d'une autre, les questions d'évaluation des BNAs.

3.2 L'aide à la décision

On peut définir une décision comme le choix d'une action ou alternative basé sur des buts en prenant en considération un ou plusieurs critères. La détermination des données adéquates pour la prise de décision est une tâche difficile dans le domaine d'aide à la décision, alors il est indispensable d'employer des systèmes d'aide à la décision (DSS) pour l'évaluation d'un problème décisionnel, des différentes alternatives ainsi que leurs influences.

3.2.1 Définitions de l'aide à la décision

On peut définir l'aide à la décision comme un ensemble de méthodes et techniques, permettant à une personne confrontée à un problème de décision ou de choix de l'aider à prendre une décision. L'aide à la décision est utilisée dans des domaines très diversifiés tels que la finance, la gestion ou en médecine (Abdolmohammadi & Usoff, 2001; H. A. Simon, 1987; Zielke, Sitter, Rampp, Bohrer, & Rothmund, 2001). Les méthodes d'aide à la décision permettent de fournir des informations sur le problème (la question dans la Figure 3.1), des outils pour l'analyser ainsi que pour trier les informations pertinentes qui peuvent aider le décideur (le gestionnaire dans la Figure 3.1) dans le but de formuler plusieurs alternatives qui pourront être proposées à ce décideur (via l'interface utilisateur Figure 3.1). Ces étapes se retrouvent dans toutes les prises de décision que Bouyssou a définies comme étant un processus cognitif (Bouyssou, 1993).

Définition. *L'aide à la décision est l'activité de celui qui, prenant appui sur des modèles clairement explicités mais non nécessairement complètement formalisés, aide à obtenir des*

Chapitre 3. DSS pour les Bibliothèques Numériques

éléments de réponses aux questions que se pose un intervenant dans un processus de décision, éléments concourant à éclairer la décision et normalement à prescrire, ou simplement à favoriser, un comportement de nature à accroître la cohérence entre l'évolution du processus d'une part, les objectifs et le système de valeurs au service desquels cet intervenant se place d'autre part (Bouyssou, 1993).

Nous pouvons retrouver plusieurs types d'acteurs lors d'une prise de décision. La personne qui va prendre la décision est le décideur. C'est à lui que s'adresse ce processus d'aide à la décision dans lequel il est impliqué pour exprimer ses préférences. Un autre acteur important de ce processus est la personne d'étude (Bouyssou, 1993). C'est un expert qui va aider le décideur à comprendre le problème, lui donner des informations essentielles et lui expliquer les conséquences des différentes alternatives (choix) lorsque c'est possible. Généralement ces deux acteurs sont deux personnes distinctes. D'autres catégories d'acteurs peuvent intervenir pendant le processus comme le demandeur qui est à l'origine de la requête d'aide à la décision, les intervenants qui vont donner leurs avis en fonction de leurs préférences ou les agis qui regroupent toutes les personnes concernées par les conséquences de la décision.

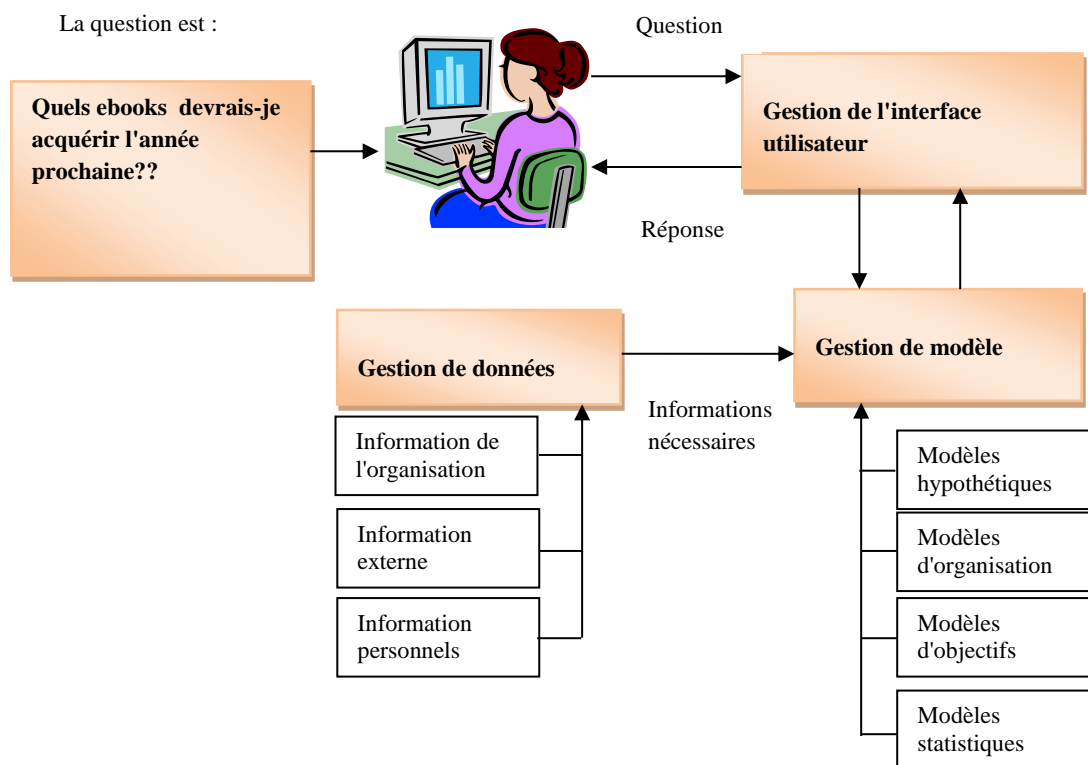


Figure 3.1 Les Composants des Systèmes D'aide à la Décision (Delgado, González, Miranda, Navarro, & Graverán, 2013)

3.2.2 Objectif de l'aide à la décision

L'objectif final de l'aide à la décision n'est donc pas de prendre la décision à la place du décideur mais bien de lui donner les informations importantes pour qu'il puisse les comprendre, lui donner des éléments de solution et les conséquences de chaque alternative afin qu'il puisse prendre la décision la plus appropriée (Fomba, 2018).. Il ne s'agit pas de décision optimale mais de décision satisfaisante. En effet Simon explique dans son principe de la rationalité limitée (H. Simon & MARCH, 1976) que certaines hypothèses admises à l'époque concernant la rationalité de certaines décisions n'étaient

Chapitre 3. DSS pour les Bibliothèques Numériques

pas valides. Il redéfinit la rationalité classique en rationalité limitée et énonce plusieurs hypothèses :

- L'accès à l'information est limité. Il n'est donc pas possible d'obtenir toutes les informations disponibles pour un problème donné (Fomba, 2018)..
- La capacité cognitive d'un individu est elle aussi limitée. Il n'est pas possible pour une personne d'analyser toutes les informations disponibles, de comprendre toutes les situations possibles et d'optimiser son choix (Fomba, 2018)..
- Les préférences d'un décideur peuvent évoluer pendant le processus de décision et les critères utilisés pour prendre une décision peuvent être contradictoires. Prenons l'exemple d'un gestionnaire d'une bibliothèque numérique voulant acheter ou acquérir un livre numérique. Le livre doit être agréable à lire par les utilisateurs mais aussi la moins cher possible, il est alors impossible de maximiser les deux critères. Le gestionnaire (décideur) devra faire un compromis entre les deux critères : lisibilité et prix (Fomba, 2018).

3.2.3 Pourquoi l'aide à la décision

Les bibliothèques numériques ont été confrontées à des défis importants au cours des dernières années. L'état actuel des bibliothèques numériques académiques, causé par des budgets limités, des développements technologiques rapides et d'une croissance étonnante du contenu numérique, expose l'importance de la compréhension de la gestion dans les bibliothèques numériques académiques (Meier, 2016). La prise de décision stratégique devient essentielle dans le processus de gestion des collections des bibliothèques numériques académiques. Néanmoins, ce type de processus de prise de décision dans les bibliothèques numériques académiques est extrêmement compliqué en raison du grand nombre de sources de données, de processus et du grand volume de données à analyser. Les sources de données typiques comprennent les systèmes de bibliothèque intégrés, les portails de bibliothèque et les OPAC, les systèmes de consortiums, les enquêtes sur la qualité et les systèmes de gestion des universités. Ces sources de données hétérogènes ne sont que partiellement utilisées pour les processus décisionnels en raison de la grande variété de formats, de normes et de technologies, ainsi que du manque de méthodes d'intégration efficaces.

Les systèmes traditionnels de gestion des bibliothèques numériques sont utilisés depuis des années pour répondre aux besoins des bibliothèques numériques. Cependant, les systèmes traditionnels remplissent des fonctions très limitées et ne peuvent pas promouvoir l'innovation et la création de connaissances pour la prise de décision stratégique. Dans ce sens, la gestion des connaissances (knowledge management (KM)) est devenue un outil puissant permettant aux bibliothèques numériques d'élargir leur rôle et leurs responsabilités dans des domaines où elles avaient peu d'impact dans le passé, telles que les décisions financières et la prise de décisions stratégiques (Hobohm, 2011; Townley, 2001).

Les systèmes d'aide à la décision (DSS) fondés sur la connaissance fournissent des informations importantes pour la prise de décision et l'amélioration des performances des bibliothèques numériques (Lai, Wang, Huang, & Kao, 2011). Un processus typique de découverte des connaissances est une séquence interactive d'étapes qui commence généralement par le nettoyage des données afin d'éliminer le bruit, les doublons et les

Chapitre 3. DSS pour les Bibliothèques Numériques

données incohérentes (J. Han, Kamber, & Pei, 2011). Ensuite, ces données nettoyées sont intégrées à partir de plusieurs sources et formats de données. Les données pertinentes sont sélectionnées et collectées à partir des bases de données en tant que données brutes à extraire; Ces données brutes sont ensuite transformées en formats appropriés qui peuvent être compris par d'autres outils, tels que des modèles d'exploration de données ou d'optimisation, et des techniques de filtration et d'agrégation appliquées pour intégrer les données de plusieurs tableaux dans un seul tableau. Des connaissances intéressantes sont extraites des données transformées (ou Découverte de connaissances dans des bases de données (Knowledge Discovery in Databases (KDD)), et ces informations sont analysées afin d'identifier des modèles réellement intéressants. Finalement, les connaissances stratégiques sont visualisées pour les gestionnaires afin de les aider à prendre des décisions.

Plusieurs DSSs ont été documentés dans la littérature pour appuyer la prise de décision dans les bibliothèques numériques; Cependant, la plupart d'entre eux se concentrent principalement sur des domaines généraux tels que l'allocation de budget, l'évaluation des performances de la bibliothèque numérique et l'analyse du comportement des utilisateurs. On sait peu de choses sur l'aide à la décision dans des domaines spécifiques, tels que l'évaluation de ressources numériques, l'évaluation de collection, la sélection des ressources numériques pour accomplir les différentes tâches de gestion et de développement de collection. On en sait encore moins sur les différentes perspectives de intervenants hétérogènes telles que les chercheurs, les universitaires, les gestionnaires, les bibliothécaires et les utilisateurs généraux (Y. Zhang, 2010). Les DSSs est donc un outil très important pour le succès de la prise de décision dans les BNAs

3.2.4 Système d'aide à la décision (DSS)

Un système d'aide à la décision (DSS) représente un mécanisme qui permet de soutenir et accompagner le décideur tout au long de son processus de prise de décision. En effet, un processus de décision ne peut pas être entièrement automatisable (Pomerol, 1992). Le DSS est aperçu comme étant un Système de Traitement de l'Information, son rôle est l'extraction et la fourniture des données nécessaires au décideur mais cela est insuffisant pour la caractérisation d'un DSS et le distinguer des autres systèmes d'information conventionnels. Il important de rajouter qu'un DSS peut être considéré comme un résolveur de problème (Pomerol, 1992).

3.2.5 Systèmes interactifs d'aide à la décision

L'expression Systèmes Interactifs d'Aide à la Décision (SIAD ou Decision Support System) a été définie au début des années soixante-dix par Gorry et Scott-Morton (Gorry & Scott Morton, 1971). De nombreuses définitions des SIAD ont été proposées depuis, comme par exemple celle de Keen et Scott-Morton (Keen, 1978) :

Définition. *Les systèmes interactifs d'aide à la décision associent les ressources intellectuelles d'une personne avec les capacités d'un ordinateur pour améliorer la qualité des décisions. Ce sont des systèmes informatiques d'assistance pour les décideurs qui doivent gérer des problèmes semi-structurés.*

Les SIAD combinent les modèles mathématiques pour analyser le comportement des décideurs et les outils informatiques pour leur interactivité et leurs techniques de visualisation. Grâce à la puissance de calculs des ordinateurs de nos jours et aux

Chapitre 3. DSS pour les Bibliothèques Numériques

avancées logicielles, un SIAD peut aider un décideur, en recréant un processus de décision, à gérer des problèmes de plus en plus complexes impliquant toujours plus d'information. L'interactivité entre le décideur et le système va permettre à ce dernier d'analyser le comportement du décideur et d'apprendre ses préférences. Cette interactivité va permettre au système de comprendre le point de vue du décideur afin de lui proposer une aide plus précise.

On peut voir l'aide à la décision de deux points de vue : une aide à la décision monocritère ou une aide à la décision multicritère. L'approche monocritère comme son nom l'indique ne prend en compte qu'un seul critère dans le processus de prise de décision. Par exemple une entreprise devant choisir un fournisseur pour acheter une pièce dont elle a besoin ne prendra en compte que le prix de la pièce. L'approche multicritère quant à elle utilise plusieurs critères pour prendre une décision. A cause de la nature complexe de notre problème, de nombreux critères peuvent rentrer en compte pour prendre une décision. Par exemple, je suis d'accord pour que ma famille accès à mes agendas à n'importe quel moment. Par contre, mes amis ne peuvent pas avoir accès à mon agenda professionnel et mes collègues ne peuvent avoir accès qu'à mon agenda professionnel et seulement durant les heures de travail. Dans ce cas-là, ma décision dépend de nombreux critères : la donnée, l'heure et la personne qui demande. Nous allons donc maintenant présenter l'aide à la décision multicritère.

3.2.6 L'aide à la décision multicritère

L'aide à la décision multicritère (Multi-Criteria Decision Analysis - MCDA) est un sous domaine de l'aide à la décision comprenant un grand nombre de méthodes permettant de résoudre des problèmes complexes qui ne pourraient pas être résolus avec une approche monocritère. Selon (Tzeng & Huang, 2011) "C'est seulement lorsqu'on est face à plusieurs attributs, objectifs, critères, fonctions, etc., que l'on peut parler d'aide à la décision et de sa théorie.

Lorsque les choix deviennent plus complexes et sont décrits par plusieurs objectifs, le problème de combiner ces aspects différents en une seule mesure d'utilité devient plus compliqué et moins réalisable."

Bien que ces méthodes ne soient pas toujours les plus efficaces sur des problèmes simples et bien structurés, elles permettent de modéliser des problèmes plus complexes en intégrant plusieurs paramètres de décision. Le but de ces méthodes reste le même qu'un système d'aide à la décision : aider un décideur et non pas le remplacer. Les solutions obtenues sont généralement des recommandations qui peuvent être acceptées ou refusées par le décideur.

Deux écoles existent en MCDA et suivent des méthodes différentes. L'école américaine avec des méthodes basées sur les fonctions d'utilité et l'école européenne principalement basée sur le concept de sur classement (Roy, 1985).

3.2.1 Les méthodes d'aide à la décision multicritères

Il existe deux grandes classes de méthodes d'aide à la décision multicritère. La première classe représente les méthodes d'aide à la décision multicritère qui utilise un attribut (critère) unique de synthèse, cette classe agrège les performances d'une action en un seul critère. La deuxième classe est celle des méthodes d'aide à la décision multicritère de sur-classement cette classe compare les différentes actions par paires.

Chapitre 3. DSS pour les Bibliothèques Numériques

Dans ce qui suit, nous présentons brièvement deux méthodes pour chaque classe. Pour la première classe nous présentons la méthode SMART et la méthode AHP, et pour la deuxième classe la méthode ELECTRE I et la méthode PROMETHEE.

i. La méthode SMART (Simple Multi Attribute Rating Technique)

Cette méthode utilise « la forme additive » pour agréger les évaluations sur les divers critères. En utilisant la « la forme additive », nous obtiendrons de bonnes approximations (Olson, 1996). Les phases de la méthode SMART sont les suivantes :

- **Phase 1** : trier les différents attributs (critères) du meilleur au moins bon selon leurs importances.
- **Phase 2** : définir les « poids » des attributs (critères).
- **Phase 3** : effectuer une normalisation des « coefficients d'importances » relative entre 0 et 1 en additionnant les « coefficients d'importance » puis en divisant chaque « poids » par cette somme.
- **Phase 4** : calculer la distance de chaque alternative sur chacun des attributs.
- **Phase 5** : définir la valeur de chacune des alternatives en utilisant la « somme pondérée »
- **Phase 6** : ranger les alternatives suivant un ordre décroissant.

ii. La méthode AHP (Analytic Hierarchy Process)

Cette méthode permet de représenter un problème décisionnel sous une forme hiérarchique tout en montrant les interconnexions qui existent entre les différents éléments de ce problème, par la suite elle procède à comparer paire par paire tous les éléments de la structure hiérarchique, et à la fin elle détermine les différentes priorités des alternatives (Saaty, 1988). Les phases de la méthode AHP sont les suivantes :

- **Phase 1** : effectuer une décomposition du problème décisionnel en une structure hiérarchique qui sera composée d'un ensemble d'éléments interconnectés. Au sommet de cette structure, on trouve le « but », et dans les sous niveaux, « les éléments qui contribuent pour arriver à ce but ». Le dernier niveau de la hiérarchie est celui des « alternatives (actions) ».
- **Phase 2** : comparer par paires les éléments de chaque niveau de la structure hiérarchique par rapport à un élément du niveau au-dessous. Cette phase construit un ensemble de tableaux de comparaisons où les valeurs de ces tableaux sont calculées en utilisant « la transformation des jugements en valeurs numériques selon l'échelle de Saaty ».
- **Phase 3** : définir l'importance liée aux éléments de la hiérarchie par le calcul des « Vecteurs propres » qui correspondent aux « valeurs propres maximales » des tableaux de comparaisons.
- **Phase 4** : procéder à la vérification de « la cohérence des jugements ».

- **Phase 5** : effectuer la « performance » lié à chaque alternative (action).

iii. La méthode ELECTRE I (Elimination Et Choix Traduisant la REalité)

Elle a été proposée par Roy en 1968. Cette méthode est considérée comme étant une problématique de choix ($P\alpha$). Elle est basée sur deux concepts : une condition de « concordance » qui impose que la plupart des attributs se dégagent avantagent l'alternative surclassant et une condition de « non discordance » qui impose le fait qu'il n'y ait pas assez d'influence, dans un des critères de la minorité, pour favoriser un sur-classement inverse (Schärlig, 1985).

Le but de la méthode ELECTRE I est l'obtention d'une division d'A en 2 sous-ensembles « N » et « A\N ». N représente l'ensemble des alternatives non surclassées donc, la meilleure alternative sera dans A\N.

Les étapes de la méthode Electre I sont les suivantes :

- **Phase 1** : on calcule tous les « indices de concordance ».
- **Phase 2** : on calcule « les indices de discordance »
- **Phase 3** : on construit les « liens de sur-classement »
- **Phase 4** : on étudie les « liens de sur-classement »

iv. La Méthode PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluation)

Les méthodes PROMETHEE ont été développées par Brans et Vincke en 1985 ; ces méthodes sont basées sur l'utilisation d'une fonction qui permet au décideur d'exprimer ces préférences pour une alternative par rapport à une autre alternative. Pour chacun des attributs (critères) (Brans & Vincke, 1985).

Les étapes de la méthode PROMETHEE I et la méthode PROMETHEE II sont les suivantes :

- **Phase 1** : choisi pour chacun des critères, une fonction d'indifférence parmi les six fonctions proposées par la méthode.
- **Phase 2** : calculer pour chaque couple d'alternatives (a_i, a_k) « la préférence globale » qui représente le degré de sur-classement.
- **Phase 3** : calculer les « flux sortants » et les « flux entrants » pour chaque alternative a_i .
- **Phase 4** : déterminer le flux total pour chaque alternative, par la suite on procède au classement des alternatives (actions).

3.3 Les requêtes de recherche des utilisateurs des BNAs: Les traces informatiques

Certaines applications ont, par nature, vocation à fournir à l'entreprise, par traçage informatique, le moyen de prouver la réception des ordres envoyés par la clientèle, l'accomplissement d'une prestation ou l'exécution d'une procédure. C'est la notion de trace.

3.3.1 Définitions

Le dictionnaire Robert donne la définition suivante : "Trace : suite d'empreintes ou de marques que laisse le passage d'un être ou d'un objet".

L'enregistrement systématique et temporaire d'un certain nombre d'informations caractérisant chaque transaction, appelées traces, est une contribution importante pour la sécurité et le contrôle des systèmes d'information et des moyens de communication.

Dans un système d'information, toute action d'un utilisateur ou d'un composant est susceptible de laisser des traces. Celles-ci se présentent sous la forme d'enregistrements stockés instantanément dans des fichiers.

On peut citer comme exemples d'actions traçables :

- Les connexions réussies à un système ;
- Les requêtes applicatives ;
- Les requêtes de recherche envoyées à un moteur de recherche ;
- Les accès aux serveurs de bases de données, aux services Intranet ou Internet ;
- Les actions de télémaintenance ;
- Les tentatives de prise de contrôle d'un équipement informatique ;
- Les tentatives d'intrusion dans le système d'information de l'entreprise ;

Etc.

Ces traces peuvent être utilisées soit :

- Immédiatement, par exemple par un système de détection d'intrusions ;
- De manière différée, par des équipes d'analyse et de recherche.

3.3.2 Des Traces informatiques, pourquoi faire ?

Une trace informatique est un ensemble d'informations qui ont été conservées et qui permettent de savoir ce qui est passé dans le système. Elles peuvent être issues d'un traitement effectué par un utilisateur ou encore d'un processus ou d'une application. Générée et inscrite sur le système où s'est passée l'action ou bien conservée sur un autre système, la trace informatique revêt des formes variées : ligne d'un fichier (chiffré ou en clair), événement d'un journal, entrée dans une base de données, information dans l'application elle-même, historique d'exécution (Charvériat, 2008).

3.3.3 Objectifs des systèmes de traçage

Un système de traçabilité répondra de fait à de nombreux objectifs :

- Conformité réglementaire ;

Chapitre 3. DSS pour les Bibliothèques Numériques

- Meilleure efficacité des processus ;
- Communication avec les fournisseurs et les clients ;
- Avantages commerciaux ;
- Avantages financiers (Exemple : réduction des intrants utilisés et des vols).

La traçabilité doit permettre aux acteurs opérant à tous les niveaux du processus et de la filière :

- De suivre les flux d'informations et d'activités.

Les traces constituent l'enregistrement de toutes les actions opérées par les différents intervenants sur le SI.

3.3.4 Utilisation des traces enregistrées

Récupérer des données effacées, démontrer que des contacts par e-mail ou par messagerie instantanée ont eu lieu, prouver l'existence de téléchargements illicites, montrer qu'un virus a été créé sur un poste ou établir qu'un site au contenu illicite a été visité... voici quelques-unes des tâches incombant aux experts de la recherche de traces informatiques. Activité encore émergente, celle-ci répond d'abord aux besoins des enquêtes judiciaires, mais aussi de plus en plus à ceux des entreprises. Une activité dans laquelle I-Tracing s'est spécialisée.

La richesse des traces est insoupçonnée et leur exploitation utile dans un large éventail de cas : lutte contre la fraude, e-discovery et gestion des preuves numériques, protection du patrimoine informationnel de l'entreprise, intelligence économique, conformité légale et réglementaire, exploitation et gestion des incidents, etc. Mais pour les obtenir, il faut utiliser des méthodes et des outils techniques adaptés s'inscrivant dans un cadre légal parfois complexe et en perpétuelle évolution (Charvériat, 2008).

Dans notre contexte, les traces seront utilisées pour déterminer les besoins des utilisateurs des BNAs.

3.3.5 Types des traces

Il existe plusieurs types des traces, à savoir :

1. Trace de profiling

Le profiling consiste à relever des indicateurs sur le comportement global du système pendant l'exécution. Typiquement ce sont des compteurs ou des variables statistiques qui reflètent le nombre d'appels de fonctions, le nombre et la proportion d'occurrence de certains événements, le temps d'exécution, etc. Le profiling ne fournit pas d'historique d'exécution mais donne des métriques la caractérisant. L'intrusion est bien moindre, la taille de la trace (du profil) est petite mais cela vient au coût d'un niveau d'information limité (Marangozova-Martin & Pagano, 2013).

2. Trace d'exécution

Les systèmes de gestion de BP permettent de retracer les exécutions d'instance. L'évènement pour chaque activité émise, son activation et d'achèvement sont signalés.

Chapitre 3. DSS pour les Bibliothèques Numériques

Pour une activité composée, les événements correspondant à son déclenchement interne sont signalés entre son activation et l'achèvement des événements.

Une trace d'exécution peut être considérée d'une manière abstraite comme un DAG (Direct Access Graphe) (imbriquée) qui contient des nœuds représentant l'activation et les événements de la fin des activités, et les bords qui décrivent leur flux. L'imbrication des DAG dans la trace découle naturellement de l'imbrication de la spécification. Des traces peuvent varier dans la quantité d'informations qu'ils enregistrent. En général, on peut distinguer trois familles de traces d'exécution. (i) des traces naïves qui fournissent un dossier complet de l'activation / événements achèvement de toutes les activités, (ii) des traces semi- naïves où les événements d'activation / d'achèvement sont tous enregistrés , mais peut-être avec qu'une information partielle sur leur activité d'origine , et (iii) traces sélectives où seul un sous-ensemble des événements est enregistré (Deutch & Milo, 2012).

3. Trace de sécurité

On entendra par traces de sécurité les traces générées par les équipements ou le système de sécurité.

4. Trace système (journal system)

On entendra par traces systèmes les traces disponibles sur les équipements directement liés au système d'exploitation. Ces traces existent en standard sur les équipements et peuvent être configurables par l'administrateur (Lathoud, 2002).

3.3.6 Finalités des traces

La notion de finalité est essentielle. Il s'agit de l'usage que l'entreprise compte faire de ses traces. Or, dans le cas de traces contenant des informations à caractère personnel, la Loi fait obligation à l'entreprise de déclarer toutes ces traces ainsi que les finalités.

A titre d'exemples, voici quelques finalités de l'usage des traces dans une entreprise :

- Surveiller le bon fonctionnement des composants du système d'information,
- Éluclider des incidents,
- Surveiller le respect des règles, le comportement des usagers,
- Détecter des situations anormales,
- Rassembler des éléments à valeur probantes,
- Imputer des actions frauduleuses,
- Contrôler les coûts,

Les traces sont donc utiles à toutes sortes de services de l'entreprise, notamment : les exploitants, les informaticiens, les services utilisateurs, les responsables d'unité, les services d'inspection ou d'enquête internes, etc.

Les traces sont également utilisées par la Police et la Justice. Les finalités des traces peuvent être regroupées en trois grandes rubriques :

1. La recherche de la qualité de service dont s'occupent principalement les exploitants informatiques et les responsables d'unités,

Chapitre 3. DSS pour les Bibliothèques Numériques

2. La protection du système d'information de l'entreprise, dont s'occupent principalement le personnel de sécurité,

3. Le domaine des enquêtes internes, administratives ou judiciaires menées par l'inspection interne de l'entreprise, les services de Police ou de Justice.

Utilité des traces d'exécution

Exploiter les traces d'exécution consiste à analyser les informations enregistrées afin de répondre à des questions sur le fonctionnement du système (Comment marche-t-il?), l'évaluation des performances (Le système est-il performant?), la validation (Respecte-t-il les spécifications?), le débogage (Où est l'erreur?) et l'optimisation (Comment pourrait-il marcher mieux ?) (Marangozova-Martin & Pagano, 2013).

Dans notre sujet on exploitera ces types de trace pour évaluer et sélectionner les ressources numériques afin de les utiliser dans la prise de décision dans les processus de développement et de gestion des collections des BNAs.

3.3.7 Cycle de vie et qualité des traces

Si les traces sont une aide à la conduite de l'exploitation, elles ont également des impacts sur l'architecture technique du SI (capacités de stockages, temps machine, . . . etc). Si les traces permettent de déjouer des menaces, elles engendrent également des risques notamment juridiques en cas d'atteinte présumée ou avérée aux libertés individuelles.

Auparavant, la limite de la gestion des traces venait des capacités de stockage. De plus en plus, la limite vient des méthodes d'utilisation des traces c'est-à-dire du recueil, de la sélection des traces pertinentes pour un objectif donné. Une politique de gestion des traces est donc nécessaire.

Les modalités de la gestion des traces sont directement liées aux utilisations prévues : les traces visant à détecter les dysfonctionnements d'un système ne sont pas collectées, traitées et archivées de la même manière que celles destinées à établir des éléments de preuve d'une transaction par exemple. La politique de gestion des traces est un élément de la politique de sécurité de l'entreprise. Elle y contribue aussi bien dans la définition des buts et des procédures que dans le contrôle de ces dernières.

Cycle de vie des traces

Le cycle de vie des traces constitue un enjeu pour l'entreprise dans la mesure où celle-ci doit appliquer des conditions strictes de collecte et d'exploitation pour pouvoir, le cas échéant, les utiliser comme éléments de preuve dans un contentieux.

Cependant, l'entreprise doit aussi garantir le « droit à l'oubli » qui figure dans la loi et donc organiser des processus de destruction des traces.

Les fonctions de collecte, d'archivage et de destruction peuvent être confiées, ensemble ou séparément, aux informaticiens, au Responsable des Services de la Sécurité de l'Information (RSSI) ou à une autre entité.

Qualité des traces

Les qualités d'une trace sont variées, et elles sont de deux ordres :

Chapitre 3. DSS pour les Bibliothèques Numériques

- Qualités intrinsèques ;
- Une trace doit d'abord être factuelle : C'est-à-dire qu'elle rend compte de faits ou d'événements, choisis pour leur signification ;
- Elle doit aussi être contextuée : C'est-à-dire qu'elle doit permettre de connaître l'environnement de la collecte : le système hôte, la version des logiciels, ...etc ;
- Une trace doit également être fidèle : D'éventuelles interruptions du traçage doivent être portées à la connaissance de l'exploitant informatique ou de la personne qui effectue l'analyse (critère de disponibilité des données) ;
- Une trace doit être fiable : C'est-à-dire qu'elle produit les mêmes enregistrements si l'événement se reproduit dans les mêmes conditions. Elle fournit des informations horodatées par rapport à base de temps commune (horloges synchronisées), en vue par exemple du rapprochement avec d'autres traces ;
- Une trace doit être inaltérable : Elle doit être protégée contre les altérations volontaires ou non, notamment l'effacement (critère d'Intégrité des données) ;
- Une trace doit être signifiante : L'événement doit être connu de l'éditeur du logiciel et explicable, il doit être interprétable par le client ;
- Une trace doit être pertinente : Un événement isolé ne doit pas être traité comme un événement répétitif ;
- Une trace doit être exploitable : Les délais de prétraitement et de traitement doivent être compatibles avec les besoins d'exploitation des traces. Les moyens mis en œuvre ne doivent pas être excessifs ou démesurément consommateurs de ressources ;
- Une trace doit faire l'objet d'un contrôle d'accès : (Confidentialité et Intégrité des données) (Lathoud, 2002);

3.3.8 L'analyse des traces pour la gestion des bibliothèques numériques

Les utilisateurs recherchent leurs informations dans différents types de systèmes IR en formulant des requêtes dans une zone de recherche. Les fonctions de recherche pour différents types d'environnements Web IR peuvent sembler similaires, mais le contenu peut être différent avec le champ de recherche et le bouton Exécuter (Wolfram, 2008). Pour analyser les patterns dans les requêtes de recherche d'informations des utilisateurs, les données du log des transactions sont utilisées dans de nombreuses études. Les données de fichier log aident les chercheurs à identifier et à comprendre les patterns de visites d'utilisateurs cachés et invisibles. Jansen (2006) a noté que les données stockées dans les logs de transaction des moteurs de recherche Web, des intranets et des sites Web peuvent fournir des informations précieuses sur le processus de recherche d'informations des utilisateurs en ligne. De nombreux chercheurs (Jansen, Booth, & Spink, 2008; Ozmutlu, Spink, & Ozmutlu, 2003; Tjondronegoro, Spink, & Jansen, 2009) ont procédé à une analyse des requêtes de journaux de transaction de sites Web. Wolfram (2008) a analysé les caractéristiques des requêtes dans une banque de données bibliographiques, un fichier OPAC, un moteur de recherche et un système de recherche spécialisé tel que HealthLink. Wang, Berry et Yang (2003) ont étudié les tendances et

Chapitre 3. DSS pour les Bibliothèques Numériques

les modèles de requêtes d'un site Web universitaire avec les données du log des transactions au cours d'une période de quatre ans (Wang, Berry, & Yang, 2003).

Cependant, peu d'analyses ont été effectuées sur l'analyse du log des transactions des requêtes des utilisateurs dans les bibliothèques numériques académiques, bien que des études aient été réalisées pour analyser les requêtes des utilisateurs en vue de la récupération d'information. Le but de cette étude est d'identifier et de comprendre les besoins d'informations des utilisateurs en utilisant les requêtes de recherche stockées dans les fichiers log des bibliothèques numériques académiques.

3.4 Data Warehouse (DW)

Un DW est une base de données analytique en lecture seule qui intègre toutes les informations provenant de diverses sources de données opérationnelles dont le but est de générer des rapports et d'analyser des données afin de soutenir le processus de prise de décision stratégique dans une entreprise.

3.4.1 Définitions de Data Warehouse

William H. Inmon, reconnu comme le père du «Data Warehouse» (DW), définit un DW comme une «collecte de données orientée sujet, intégrée, variant dans le temps et non volatile, à l'appui du processus décisionnel de la direction» (Inmon, 2005).

Selon Ralph (Kimball & Ross, 2011), autre figure prééminente de l'entreposage de données, un DW est un référentiel d'informations intégrées provenant de sources distribuées, autonomes, voire hétérogènes, spécifiquement structurées pour être analysées et consultées.

3.4.2 Caractéristiques de Data Warehouse

Il existe des fonctionnalités de base qui définissent les données dans le Data Warehouse, la Figure 3.2 expose les caractéristiques.

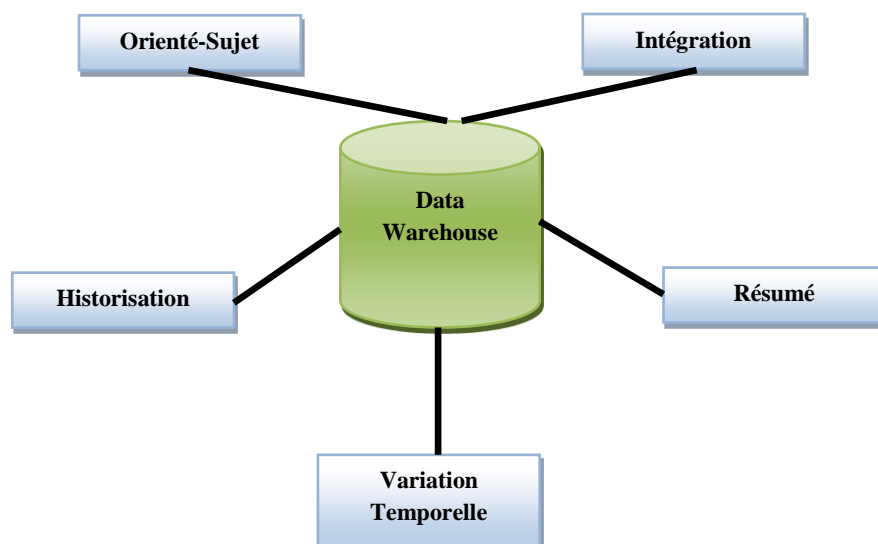


Figure 3.2 Caractéristiques de Data Warehouse (Saurabh & Nagpal, 2011)

- **Orienté-sujet** : Contrairement aux systèmes opérationnels, les données de Data Warehouse tournent autour de sujets de l'entreprise (normalisation de la base de données). L'orientation sujet peut être très utile pour la prise de décision.
- **Intégration** : Les données trouvées dans le Data Warehouse sont intégrées. Comme il provient de plusieurs systèmes opérationnels, toutes les incohérences doivent être supprimées. Les cohérences incluent les conventions de dénomination, la mesure des variables, les structures de codage, les attributs physiques des données, etc.
- **Variation temporelle** : Alors que les systèmes opérationnels reflètent les valeurs actuelles car ils prennent en charge les opérations quotidiennes, les données du Data Warehouse représentent des données sur un long horizon (jusqu'à 10 ans), ce qui signifie qu'elles stockent des données historiques. Il est principalement destiné à l'exploration de données et à la prévision (Inmon, 2005).
- **L'historisation (Non-volatilité)** : Les données du Data Warehouse sont en lecture seule, ce qui signifie qu'elles ne peuvent pas être mises à jour ou supprimées (Ponniah, 2011).
- **Résumé** : Dans le Data Warehouse, les données sont résumées à différents niveaux. L'utilisateur peut commencer à regarder le total des unités de vente d'un produit dans une région entière. L'utilisateur examine ensuite les états de cette région. Enfin, ils peuvent examiner les magasins individuels dans un certain état. Par conséquent, généralement, l'analyse commence à un niveau supérieur et descend vers des niveaux de détails inférieurs (Inmon, 2005).

3.4.3 Méthodes de conception de Data Warehouse

Il existe trois approches pour la conception de Data Warehouse :

- i. **Conception ascendante (Bottom-up)** : Dans l'approche ascendante, les data marts sont d'abord créés pour fournir des capacités d'analyse pour des processus métier spécifiques. Ces Data Marts peuvent ensuite être intégrés pour créer un Data Warehouse complet (Kimball & Ross, 2015).
- ii. **Conception descendante (Top-down)** : le DW est conçue à l'aide d'un modèle de données d'entreprise normalisé. Les données "atomiques", c'est-à-dire les données au plus haut niveau de détail, sont stockées dans le Data Warehouse. Des Data Marts contenant les données nécessaires à des processus métier ou des services spécifiques sont créés à partir de l'entrepôt de données (Sen & Sinha, 2005).
- iii. **Conception hybride** : un DW hybride élimine la redondance des données. Le DW fournit une source unique d'informations à partir de laquelle les data marts peuvent lire, fournissant une large gamme d'informations commerciales. L'architecture hybride permet de remplacer un DW par un référentiel de gestion des données maître dans lequel des informations opérationnelles et non statiques pourraient résider (Sen & Sinha, 2005).

3.4.4 Objective de Data Warehouse

L'objectif d'un DW est de consolider les informations stockées dans les différentes bases de données opérationnelles avec des informations provenant d'autres sources de données, souvent externes, et de les rendre disponibles pour l'analyse des données d'un point de vue gestionnaire (Singh, 1997). Les données stockées dans le DW sont des instantanés résultant de transformations de données, de contrôles de qualité et de l'intégration de données opérationnelles. Le principal avantage d'un DW est la possibilité d'avoir un accès interactif et immédiat aux informations stratégiques d'une entreprise. Les utilisateurs ayant un rôle de direction dans l'organisation effectuent leurs propres enquêtes et croisent les données, à l'aide d'outils spécialisés dotés d'interfaces graphiques, tels que l'exploration de données et les outils de traitement analytique en ligne. Les opérations dans un DW ne sont pas des transactions prédominantes, comme dans les bases de données opérationnelles, mais des requêtes complexes de jonction, de filtrage, de regroupement et d'agrégation de grandes quantités de données (Wrembel, 2006).

En raison des caractéristiques particulières d'un entrepôt de données, les stratégies de conception utilisées pour créer et gérer des bases de données opérationnelles ne sont généralement pas applicables à la conception d'entrepôts de données (Inmon, 2005; Kimball & Ross, 2011). La conception d'un DW est une tâche intrinsèquement complexe, coûteuse et longue. La création d'un DW implique l'extraction de données provenant de différentes sources de données, dans lesquelles de nombreux problèmes d'incohérence doivent être résolus (Wah, Peng, & Hok, 2007). L'extraction, le nettoyage et le stockage de données via des processus ETL (extraction, transformation, chargement) constituent également un problème complexe et fastidieux, car ce processus doit combiner les différentes sources de données et les convertir en un format uniforme, en excluant les éventuelles incohérences, redondances, etc. et incompatibilités (Nicholson, 2003). Par conséquent, pour que le résultat final réponde à ses besoins, la mise en œuvre réussie d'un DW nécessite, entre autres, un investissement important en temps et en énergie pour ceux qui en seront les utilisateurs, l'identification des objectifs commerciaux, la conception de l'architecture du DW et la mise en œuvre du DW. système (Curtis & Joshi, 1999). Cependant, ce n'est qu'en combinant et en reliant différentes sources de données que les gestionnaires de bibliothèque peuvent découvrir les informations stratégiques cachées qui peuvent aider à comprendre correctement les processus et les services dans la prise de décision.

3.5 Cloud Computing (CC)

Dans (Mell & Grance, 2011), Mell et Grance nous propose la définition suivante de l'aide à la décision :

Définition. *Le cloud computing est un modèle permettant un accès réseau omniprésent, pratique et sur demande à un pool partagé de ressources informatiques configurables (réseaux, serveurs, stockage, applications et services) pouvant être rapidement mis en service et libéré avec un effort de gestion minimal. Interaction fournisseur de services. Ce modèle de cloud comprend cinq caractéristiques essentielles, trois modèles de service et quatre modèles de déploiement.*

3.5.1 Caractéristiques du CC

Dans cette sous-section, nous exposent les caractéristiques essentielles du Cloud Computing.

- **Service-libre à la demande**

Un consommateur peut provisionner unilatéralement des capacités informatiques, telles que l'heure du serveur et le stockage réseau, selon les besoins automatiquement sans nécessiter d'interaction humaine avec chaque fournisseur de services.

- **Large accès au réseau**

Les fonctionnalités sont disponibles sur le réseau et accessibles via des mécanismes standard qui favorisent leur utilisation par des plates-formes clientes hétérogènes ou épaisses (téléphones mobiles, tablettes, ordinateurs portables et stations de travail, par exemple).

- **Mise en commun des ressources**

Les ressources informatiques du fournisseur sont regroupées pour desservir plusieurs consommateurs à l'aide d'un modèle multi-locataire, différentes ressources physiques et virtuelles étant attribuées et réaffectées de manière dynamique en fonction de la demande des consommateurs. Il existe un sentiment d'indépendance de localisation en ce que le client n'a généralement aucun contrôle ni connaissance sur la localisation exacte des ressources fournies, mais peut éventuellement spécifier une localisation à un niveau d'abstraction supérieur (par exemple, pays, état ou centre de données). Les exemples de ressources incluent le stockage, le traitement, la mémoire et la bande passante du réseau.

- **Élasticité rapide**

Les capacités peuvent être provisionnées et libérées de manière élastique, dans certains cas automatiquement, pour évoluer rapidement vers l'extérieur et vers l'intérieur en fonction de la demande. Pour le consommateur, les fonctionnalités disponibles pour le provisioning semblent souvent illimitées et peuvent être adaptées à n'importe quelle quantité et à tout moment.

- **Service mesuré**

Les systèmes en nuage contrôlent et optimisent automatiquement l'utilisation des ressources en exploitant une capacité de mesure 1 à un niveau d'abstraction approprié au type de service (par exemple, comptes de stockage, de traitement, de bande passante et d'utilisateurs actifs). L'utilisation des ressources peut être surveillée, contrôlée et rapportée, offrant une transparence à la fois au fournisseur et au consommateur du service utilisé.

3.5.2 Modèles de livraison de service

Le Cloud Computing offre différents modèles de livraison des services, ces modèles sont illustré dans la Figure 3.3. Nous verrons les exposer dans cette sous-section.

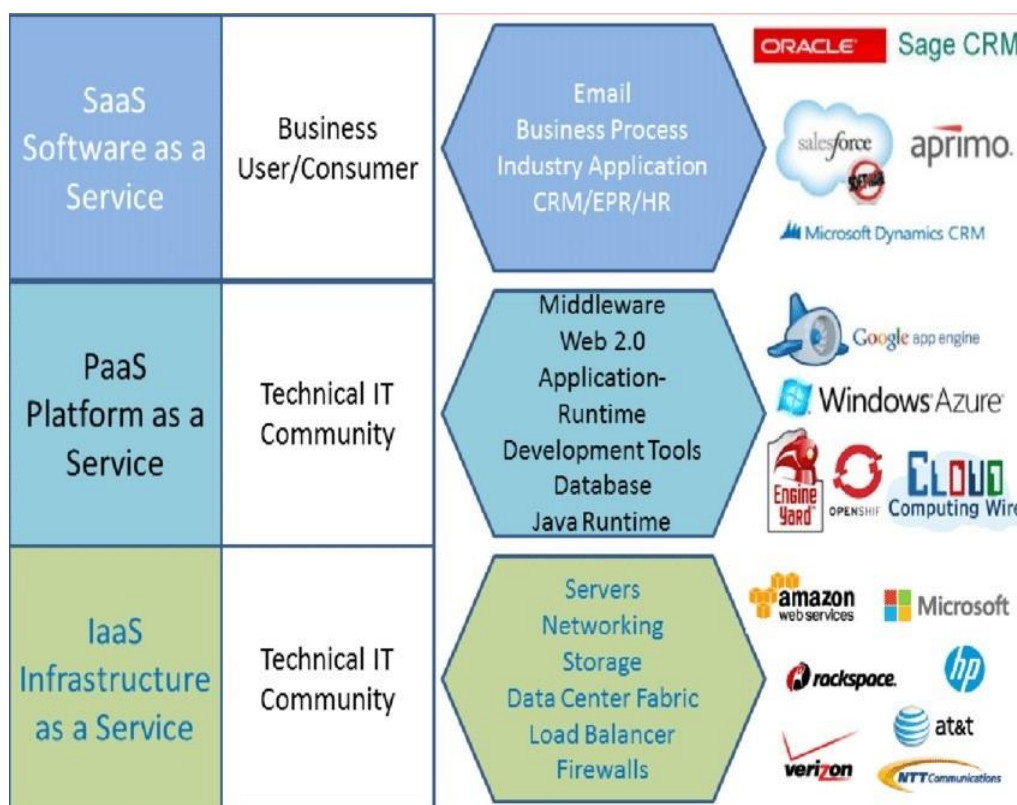


Figure 3.3 Les Modèles de Livraison de Services du Cloud (V. Kumar, Laghari, Karim, Shakir, & Brohi, 2019)

- **Le modèle Software-as-a-Service (SaaS)**

La capacité fournie au consommateur consiste à utiliser les applications du fournisseur s'exécutant sur une infrastructure en nuage. Les applications sont accessibles à partir de divers périphériques clients via une interface de client léger, telle qu'un navigateur Web (par exemple, une messagerie Web), ou une interface de programme. Le consommateur ne gère ni ne contrôle l'infrastructure cloud sous-jacente, y compris le réseau, les serveurs, les systèmes d'exploitation, le stockage ou même les fonctionnalités d'applications individuelles, à l'exception possible des paramètres de configuration d'application spécifiques à un utilisateur.

- **Le modèle Platform as a Service (PaaS)**

La capacité offerte au consommateur consiste à déployer sur l'infrastructure en nuage des applications créées ou acquises par le consommateur créé à l'aide de langages de programmation, de bibliothèques, de services et d'outils pris en charge par le fournisseur. Le consommateur ne gère ni ne contrôle l'infrastructure en nuage sous-jacente, y compris le réseau, serveurs, systèmes d'exploitation ou stockage, mais contrôle les applications déployées et éventuellement les paramètres de configuration de l'environnement d'hébergement d'applications.

- **Le modèle Infrastructure as a Service (IaaS)**

La capacité fournie au consommateur consiste à fournir un traitement, un stockage, des réseaux et d'autres ressources informatiques fondamentales permettant au client de déployer et d'exécuter des logiciels arbitraires, pouvant inclure des systèmes d'exploitation et des applications. Le consommateur ne gère ni ne contrôle

Chapitre 3. DSS pour les Bibliothèques Numériques

l'infrastructure cloud sous-jacente, mais il contrôle les systèmes d'exploitation, le stockage et les applications déployées. Et éventuellement un contrôle limité de certains composants réseau (par exemple, les pare-feu hôtes).

3.5.3 Modèles de déploiement du Cloud

Le CC déploie ses services selon des différents modèles de déploiement de services. Ces modèles permettent de définir le degré d'accès d'un utilisateur aux fournisseurs de services Cloud. Ces modèles sont divisés en quatre grandes catégories, voir Figure 3.4.

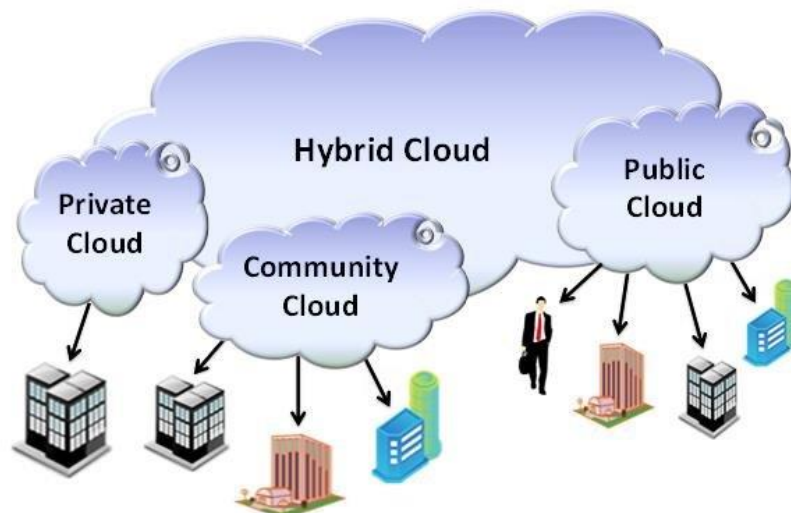


Figure 3.4 Les Moèles de Déploiement Du Cloud (Bamiah, 2015)

1. Cloud privé.

L'infrastructure en nuage est configurée pour une utilisation exclusive par une seule organisation comprenant plusieurs consommateurs (par exemple, des unités commerciales). Il peut être possédé, géré et exploité par l'organisation, un tiers, ou une combinaison de ceux-ci, et il peut exister sur place ou hors site.

2. Cloud communautaire.

L'infrastructure cloud est conçue pour une utilisation exclusive par une communauté spécifique de consommateurs d'organisations partageant des préoccupations communes (par exemple, mission, exigences de sécurité, règles et considérations de conformité). Il peut être possédé, géré et exploité par une ou plusieurs organisations de la communauté, par un tiers, ou par une combinaison de ceux-ci, et il peut exister sur place ou à l'extérieur.

3. Cloud public.

L'infrastructure de cloud est conçue pour une utilisation ouverte par le grand public. Il peut être détenu, géré et exploité par une entreprise, un établissement universitaire ou un organisme gouvernemental, ou une combinaison de ceux-ci. Il existe dans les locaux du fournisseur de cloud.

4. Cloud hybride.

L'infrastructure en nuage est composée de deux ou plusieurs infrastructures en nuage distinctes (privée, communautaire ou publique) qui restent des entités uniques, mais qui sont liées entre elles par une technologie standard ou propriétaire qui permet la portabilité des données et des applications des nuages).

3.6 Travaux connexes

Le problème de l'évaluation des bibliothèques numériques au sens large a été étudié par différents chercheurs de la discipline informatique. A cet égard, et en raison de sa nature complexe, la question de l'évaluation des bibliothèques numériques et de ses services est à la croisée de plusieurs travaux de recherche. Cependant, la littérature existante est riche en travaux qui ont tenté de relever les défis inhérents à cette problématique de recherche.

Cette section est dédiée à l'exploration, à l'étude et à l'analyse des travaux connexes ayant traité, d'une manière ou d'une autre, les questions de l'évaluation des bibliothèques numériques.

3.6.1 Les approches générales pour l'évaluation des BNs

L'évaluation des services de bibliothèques numériques sur la base des critères retenus a particulièrement retenu l'attention des gestionnaires des bibliothèques numériques (Bhadoria, Chaudhari, & Vidanagama, 2018; Kassim, 2017). Par conséquent, des approches très variées ont été proposées dans divers aspects de l'évaluation des bibliothèques numériques pour répondre aux questions liées la gestion de ces bibliothèques, à l'amélioration de leurs services et leurs collections. Cependant, ces approches restent de manière naturelle très liées à l'évaluation des bibliothèques numériques d'une manière générale. Autrement dit, elles se focalisent, généralement, sur des aspects généraux du problème de l'évaluation des bibliothèques numériques.

Ces dernières années, les chercheurs ont montré un intérêt substantiel dans la littérature. De nombreux chercheurs ont étudié la gestion des bibliothèques numériques et de nombreuses approches ont été proposées pour aider les gestionnaires des bibliothèques numériques à prendre des décisions efficaces dans les différentes tâches de gestion et d'amélioration des bibliothèques numériques. La plupart des travaux proposés tel que (Ahmad & Abawajy, 2014; Cabrerizo, Morente-Molinera, Pérez, López-Gijón, & Herrera-Viedma, 2015; Fuhr, Hansen, Mabe, Micsik, & Sølvsberg, 2001; Hu & Yu, 2019; Li & Liu, 2019; Siguenza Guzman, Saquicela, & Cattrysse, 2014; Tramullas, Sánchez-Casabón, & Garrido-Picazo, 2013; Yue, 2019) mettent l'accent sur une évaluation des services et des collections des bibliothèques numériques de façon très générale. Parmi ces approches proposées, certains DSS ont été proposés (Cabrerizo et al., 2015; Siguenza Guzman et al., 2014) ou des modèles dans DSS pour assister les gestionnaires des BNs dans leurs travaux par les informations offertes. Un autre groupe de chercheurs a proposé un DSS qui génère des recommandations sur les décisions de gestion pour aider les gestionnaire à améliorer ses services et à gérer efficacement d'autres problèmes spécifiques à la gestion des bibliothèques numériques (Chao, Zhang, & Xing, 2011; Gutwin, Paynter, Witten, Nevill-Manning, & Frank, 1999; Kang, Kim, & Ko, 2004; Kao, Chang, & Lin, 2003; Yue, 2019).

Chapitre 3. DSS pour les Bibliothèques Numériques

Le tableau. 3.1 résume les avantages et les désavantages de chacun d'un ensemble des travaux sélectionnés parmi les travaux existants.

3.6.2 Analyse comparative des approches existantes

Le Tableau. 3.1 expose une étude comparative des travaux de recherche ayant abordé, plus ou moins, les questions de l'évaluation des bibliothèques numériques. Lors de cette étude comparative, les critères suivants ont été retenus afin de faire ressortir les points faibles et les points forts de chaque approche.

- **Evaluation des Services** : ce critère désigne l'évaluation effectuée par une approche pour déterminer les niveaux des performances des services des bibliothèques numériques ;
- **Evaluation de Collection** : décrit la détermination de la richesse de la collection de la bibliothèque numérique.
- **Evaluation de Ressources** : c'est l'évaluation effectuée par une approche pour donner le niveau d'utilisation d'une ressource numérique (livre numérique, revue numérique, ou base de donnée)
- **Recommandations** : ce critère explique si une approche fournit des suggestions aux gestionnaires des bibliothèques numériques dans le but de les aider à améliorer les services de ces bibliothèques numériques.
- **Sélection de ressource** : désigne les informations fournies par une approche dans le but d'aider les gestionnaires dans le processus de sélection des ressources numériques pour acquérir, abonner, ou désabonner ...etc.
- **Evaluation des régions de collection** : illustre si une approche évalue les sous collection dans la collection de la bibliothèque numérique.

Dans le Tableau. 3.1, le signe plus (+) signifie que le travail fournit le service indiqué dans la colonne et le signe moins (-) signifie l'absence du service. Par exemple, dans la colonne intitulée "Recommandation", le signe (+) signifie que le travail fournit des suggestions (recommandations) de décisions de gestion aux gestionnaires des BNs et (-) signifie qu'il n'y a pas de suggestions fournies par ce travail. Par exemple, Kao et al. (2003) se concentrent sur l'évaluation des services et de la collection de bibliothèques numériques. Cependant, ils n'ont pas fourni des autres services aux gestionnaires (évaluation des ressources, recommandations, sélection du matériel et recommandations pour améliorer la collection de la bibliothèque numérique).

Chapitre 3. DSS pour les Bibliothèques Numériques

Reference	Evaluation des Services	Evaluation de Collection	Evaluation de Ressource	Recommandation	Sélection de ressource	Evaluation des Régions de Collection
Kao et al. (2003)	+	+	-	-	-	-
Kadir et al. (2009)	+	+	-	-	-	-
Sutton (2013)	-	+	+	-	-	-
Ahmad et al. (2014)	+	-	-	-	-	-
Siguenza et al. (2014)	+	+	-	-	-	-
Cabrerizo et al. (2015)	+	+	-	+	-	-
Graf et al (2017)	-	-	+	-	-	-
Yue, C. (2019)	+	+	-	-	-	-

Tableau 3.1 Comparaison entre les travaux existants (Hemili et al., 2019)

A partir de l'étude précédente, il apparaît clairement que chacun des approches analysées présente plusieurs points communs avec notre problème. Néanmoins, les approches qui sont proposées présentent plusieurs insuffisances pour la prise en charge des problèmes liés à la gestion des collections numériques des bibliothèques numériques académiques (voir dernières colonnes).

La plupart de ces travaux antérieurs n'ont pas permis d'évaluer les ressources numériques pour le développement de collections. En outre, de nombreuses études antérieures ont montré que le DSS peut constituer un outil efficace pour la gestion des bibliothèques numériques (Cabrerizo et al., 2015; Siguenza Guzman et al., 2014; Tramullas et al., 2013). Les travaux substantiels discutés dans le Tableau. 3.1 se concentrent sur des critères économiques tels que le prix et le budget pour effectuer leurs évaluations.

Chapitre 3. DSS pour les Bibliothèques Numériques

Avec le nouvel environnement d'un marché en croissance des ressources numériques, une tendance importante dans le développement et la gestion des collections consiste à créer des collections basées sur les besoins directs des clients (Lehman, 2014). La sélection de livres numériques, de revues / périodiques numériques et de bases de données est très complexe en raison du volume croissant de ressources numériques et de la population d'utilisateurs.

En résumé, les approches proposées ne peuvent être appliquées dans le cadre de la gestion et du développement des collections des BNAs pour que les informations fournies par ces approches sont des informations générales inutiles ou la manque des informations nécessaires dans les processus de la prise de décision.

Les principaux résultats qu'on a conclus à partir de cette étude sont :

- Le besoin d'utilisation des méthodes d'aide à la décision multicritère (MCDM), ces méthodes nous permettent d'impliquer plus de critères et de tirer plus d'information pour une évaluation plus efficace.
- La nécessité des évaluations particulières des parties de la BN pour offrir des informations très utiles, ce que permet aux gestionnaires de prendre des décisions critiques ;
- Le besoin de fournir des recommandations aux gestionnaires qui va rendre leurs travaux plus simples.
- Les DSSs ont prouvé leurs efficacités dans le support des gestionnaires des bibliothèques numériques

3.7 Conclusion

Ce chapitre a présenté des éléments issus de l'aide à la décision et différentes techniques que ce domaine regroupe, la technologie du CC et ses modèles de services et de déploiement, le DW et ses concepts de base, et une étude comparative des travaux connexes. Notre but est d'apprendre comment se comporte un gestionnaire de BN quand il s'agit de prendre des décisions critiques puis de l'aider à prendre ces décisions efficacement en lui offrent les besoins des utilisateurs sous forme des évaluations et des sélections des ressources numériques. Le système que nous voulons mettre en place correspond donc totalement à un système d'aide à la décision.

Bien que la technologie de CC et Data Warehousing aurait pu être utilisée, cela aurait nécessité de centraliser les préférences des utilisateurs. Or comme nous l'avons mis en avant dans le chapitre sur la gestion des BNs, cette centralisation peut poser des problèmes liés à la puissance de calculs. Nous avons donc choisi de rester sur une approche basée sur les besoins des utilisateurs.

Chapitre 4

Utilisation de MCDA pour l'Evaluation de Collection

Contenu

4.1	<i>Introduction</i>	67
4.2	<i>Prise de décision dans la gestion de collection</i>	68
4.3	<i>L'évaluation de collection</i>	70
4.3.1	<i>L'évaluation de collection en tant qu'outil de gestion</i>	70
4.3.2	<i>Les buts de l'évaluation de collection</i>	71
4.4	<i>Techniques d'évaluation de collection</i>	72
4.5	<i>Critères d'évaluation des collections</i>	73
4.6	<i>Approche multicritères basée sur la méthode PROMETHEE II pour l'évaluation des collections</i>	74
4.7	<i>Conclusion</i>	76

4.1 Introduction

L'objectif principal des bibliothèques numériques universitaires est de satisfaire les besoins en informations de leurs enseignants, étudiants, chercheurs et staff académiques. Dans le but de développer une collection riche qui réponde aux besoins appropriés de sa clientèle, les gestionnaires doivent prendre des décisions complexes dans les processus de développement et de gestion de la collection. Dans la gestion des collections, les gestionnaires spécialisés dans la gestion des collections ont la responsabilité de maintenir la collection utile, pertinente et à jour en enlevant les ressources numériques inutiles et en acquérant de nouvelles ressources numériques (Larson, 2012), toutes les ressources numériques doivent être évaluées pour déterminer s'il y a des ressources qui doivent être enlevées (Snyder, 2014). La sélection de livres numériques pour le désherbage ou la révision de périodiques à des fins de renouvellement ou d'annulation implique de nombreux processus décisionnels (Johnson, 2014; Polanka, 2012).

Au cours des dernières années, la croissance des collections, le budget limité et la limitation dans les ressources en personnel rendent le processus de prise de décision critique et extrêmement complexe (Disher, 2014; Johnson, 2014). L'enlèvement des ressources numériques peut être gênant pour les utilisateurs et les gestionnaires (Walters, 2013). Le climat budgétaire oblige les bibliothèques à être plus sélectives en matière de désherbage, de renouvellement, et souvent à envisager des annulations des abonnements (Saracevic, 2000; Sutton, 2013).

Plusieurs approches ont été proposées pour aider les gestionnaires à améliorer leurs collections (Ahmad & Abawajy, 2014; Cabrerizo et al., 2015; Siguenza Guzman et al., 2014; Sutton, 2013), ces approches étant axées sur l'évaluation générale des services et des collections (Ahmad & Abawajy, 2014; Saracevic, 2000; Sutton, 2013), et certaines d'entre elles ont tenté de donner des recommandations afin d'améliorer les services et les collections (Cabrerizo et al., 2015). Les approches proposées ont essayé d'aider les gestionnaires dans les activités de gestion en général et la plupart d'entre elles ne donnent qu'une évaluation des services et des collections ou de l'un d'eux. Cependant, évaluer la collection de bibliothèques numériques universitaires de façon globale aide les gestionnaires dans le processus de gestion, mais il est plus utile pour les gestionnaires de fournir une évaluation de sous collections dans la collection, et aussi une évaluation des ressources spécifiques dans la collection et de générer ou proposer des suggestions de décision (Que faire?). Ainsi, il est très utile de renseigner les gestionnaires pour prendre les bonnes décisions.

La méthode traditionnelle utilisée par les gestionnaires dans les processus décisionnels se concentre sur les nombres statistiques eux-mêmes pour prouver leurs décisions. Utiliser la ressource numérique au coût le plus bas ou basé sur d'autres fonctionnalités ou critères n'est pas un bon moyen de sélectionner la décision la plus efficace en raison de l'ignorance des autres critères. En raison du nombre de critères et des volumes de données impliqués dans l'évaluation des ressources numériques et des collections, l'analyse multicritères devient primordiale dans le processus de prise de décision dans la gestion des collections des bibliothèques numériques universitaires.

Ce chapitre porte sur les questions de prise de décision en matière de gestion de collections de bibliothèques numériques universitaires. Il montre comment nous pouvons utiliser MCDA pour aider les gestionnaires dans les processus de prise de

Chapitre 4. Utilisation de MCDA pour l'Évaluation de Collection

décision en matière de gestion des collections. Les processus de gestion de la collection tels que le désherbage des livres numériques, l'annulation ou le renouvellement des abonnements aux revues numériques impliquent un processus de décision multicritères pour sélectionner la meilleure décision pour chaque ressource numérique (livre numérique, revue numérique et base de données). Dans ce chapitre, nous avons utilisé PROMETHEE II comme méthode de MCDA dans le but d'évaluer la collection de bibliothèque numérique.

Les sections qui construisent ce chapitre sont présentées comme suit. La section 2 explore dans la littérature les processus et les critères d'évaluation des collections des bibliothèques numériques. La section 3 présente la méthode basée sur PROMETHEE II pour évaluer la collection. Enfin, la section 4 présente les conclusions et les perspectives.

4.2 Prise de décision dans la gestion de collection

La gestion des collections représente une partie très importante dans la gestion des bibliothèques numériques universitaires, car la satisfaction des besoins en informations des utilisateurs est le principal objectif des bibliothèques numériques universitaires. La gestion de la collection d'aujourd'hui est soumise à de fortes pressions de la part de la communauté des utilisateurs et de l'institut mère en raison du budget limité et de la croissance de la collection. Cela implique la nécessité d'une gestion active de la collection.

La gestion de la collection couvre les décisions relatives au désherbage, à l'annulation des abonnements aux ressources périodiques, ou aux renouvellements et à la conservation, ainsi que les activités qui sous-tendent ces décisions, telles que les études d'utilisation et l'évaluation coûts / avantages (Larson, 2012). Afin de gérer les collections, les gestionnaires assument de nombreuses responsabilités, notamment la sélection de supports dans tous les formats de désherbage (Hightower & Gantt, 2012), l'annulation, le renouvellement et la conservation, l'évaluation des collections et des services associés, l'utilisation des collections et les expériences des utilisateurs, la résolution des problèmes de ressources sélectionnés, et la gestion des allocations budgétaires.

De nombreux travaux ont essayé d'aider les gestionnaires dans leurs activités de gestion, comme présenté dans le Tableau. 4.1, les travaux proposés ont donné une évaluation générale des services et des collections de bibliothèques numériques, les sous collections n'ont pas été évalués et les ressources n'ont pas été évalués ni sélectionnés pour des activités de gestion. La sélection des ressources à désherber, désabonner ou renouveler est un processus compliqué en raison de la multiplicité des métriques impliquées dans la sélection de ces ressources, aussi l'allocation de budget pour améliorer les sous collections devient très compliquée à cause de budget limité de ces bibliothèques.

Chapitre 4. Utilisation de MCDA pour l'Évaluation de Collection

Reference	Evaluation des Services	Evaluation de Collection	Evaluation des Ressources	Evaluation des Sous Collections	Recommandations
Cabrerizo et al. dans (Cabrerizo et al., 2015)	Fourni	Fourni	Absent	Absent	Fourni
Siguenza et al. dans (Siguenza Guzman et al., 2014)	Fourni	Fourni	Absent	Absent	Absent
Sutton dans (Sutton, 2013)	Absent	Fourni	Fourni	Absent	Absent
Graf Et al. dans (Graf, Ryan, Houzanme , & Gordea, 2017)	Absent	Absent	Fourni	Absent	Absent
Ahmad et al. dans (Ahmad & Abawajy, 2014)	Fourni	Absent	Absent	Absent	Absent
Kadir et al. dans (R. Kadir, W. Dollah, F. A. Saaid, & S. Diljit, 2009)	Fourni	Fourni	Absent	Absent	Absent
Kao et al. dans (Kao et al., 2003)	Fourni	Fourni	Absent	Absent	Absent

Tableau 4.1 Comparaison des travaux qui offrent une évaluation de collection (Hemili & Laouar, 2018b)

Par conséquent, dans le présent chapitre, notre objectif est d'aider les gestionnaires à utiliser MCDA afin d'évaluer les collections, les sous collections et ressources numériques pour différentes activités de gestion tel que l'allocation de budget. En fonction de ces évaluations, nous déterminons les sous collections qui nécessitent plus d'inversement pour les questions d'améliorations. Le premier objectif de la MCDA est de fournir des outils ou des méthodes qui aident les décideurs à résoudre les problèmes

Chapitre 4. Utilisation de MCDA pour l'Évaluation de Collection

de tous les domaines de la vie lorsque de nombreuses caractéristiques sont impliquées dans ce processus décisionnel (Vincke, 1992).

Le problème dans lequel différentes métriques (critères) ont contribué au processus de prise de décision est généralement résolu par une analyse multicritère de décision (MCDA) (Roy, 1996). Dans les problèmes de MCDA, le décideur doit choisir parmi un ensemble de choix, qui sont différenciés par de multiples caractéristiques (Greco, Figueira, & Ehrgott, 2005). Dans les problèmes du monde réel de MCDA, de nombreuses méthodes de MCDA ont été utilisées; toutefois, des méthodes basées sur des relations de supériorité avec incomparabilité (par exemple, ELECTRE) ont prouvé qu'elle pouvait être efficace pour ce type de problèmes. Dans la sous-section suivante, nous présentons les critères qui seront utilisés dans la sélection des matériaux.

4.3 L'évaluation de collection

Les gestionnaires effectuent continuellement une analyse sur les collections des bibliothèques Numériques. Cette analyse fournit aux gestionnaires des informations sur divers aspects de la collection, notamment le nombre d'œuvres et de titres dans un sujet particulier; formats représentés; âge des ressources; la profondeur de la couverture; langue dans laquelle les ressources sont disponibles; niveau d'utilisation de la collection par les utilisateurs. L'objectif réel de cette analyse est de mesurer l'utilité de la collection - l'efficacité avec laquelle la collection répond à l'objectif auquel elle est destinée et, par extension, l'efficacité avec laquelle la bibliothèque dépense des fonds pour développer et maintenir cette collection.

Les gestionnaires des bibliothèques s'intéressent de plus en plus à l'identification de métriques importantes, c'est-à-dire des métriques permettant de surveiller efficacement les performances de la bibliothèque dans la réalisation de ses objectifs (Henry, Longstaff, & Van Kampen, 2008). Des métriques significatives peuvent éclairer les décisions en matière de développement et de gestion des collections et démontrer la valeur pour les bibliothèques. L'aspect difficile de mesurer et de démontrer la valeur consiste à utiliser des méthodes significatives qui ne sont pas trop onéreuses dans leur exécution.

4.3.1 L'évaluation de collection en tant qu'outil de gestion

L'évaluation de la collection fait partie de la gestion efficace des ressources. Plusieurs termes peuvent être utilisés dans les discussions sur l'évaluation des collections. Celles-ci incluent la cartographie de la collection, l'examen de la collection, l'analyse de la collection et l'évaluation de la collection. Bien qu'elles soient souvent utilisées de manière interchangeable, évaluation et estimation peuvent être distinguées en fonction de l'intention de l'analyse. Le but de l'évaluation est de déterminer dans quelle mesure la collection répond aux objectifs, aux besoins et à la mission de la bibliothèque ou de l'organisation mère. La collection (ressources accessible à distance) est évaluée dans le contexte local. Les buts et objectifs de la bibliothèque numérique doivent donc être clairement énoncés avant toute évaluation significative de la collection de la bibliothèque. Une fois les objectifs de la collection ont été assignés à des domaines ou à des communautés d'utilisateurs, la bibliothèque peut évaluer si elle a rassemblé au niveau souhaité. L'évaluation cherche à examiner ou à décrire les collections, soit dans leurs propres termes, soit par rapport à d'autres collections et mécanismes de vérification, tels que des listes. Bien que l'estimation soit un processus

Chapitre 4. Utilisation de MCDA pour l'Évaluation de Collection

plus abstrait, car elle ne tient pas compte de l'utilisation et des utilisateurs de la collection, les résultats n'ont de sens que s'ils sont pris en compte dans le contexte local. Par exemple, la mesure dans laquelle une bibliothèque a des titres sur une liste principale n'a de sens que si la bibliothèque a pour objectif de fournir ces documents (ressources).

L'évaluation permet de mieux comprendre la collection et la communauté d'utilisateurs. Les gestionnaires des bibliothèques numériques obtiennent des informations qui les aident à déterminer si une collection atteint ses objectifs, dans quelle mesure elle sert les utilisateurs, de quelle manière ou dans quels domaines (sous collections) elle est déficiente et ce qu'il reste à faire pour l'améliorer. Au fur et à mesure que les gestionnaires en apprennent davantage sur la collection et son utilité, ils sont en mesure de gérer la collection (croissance, conservation, stockage, retrait et annulation des abonnements à des périodiques et autres ressources continues) en fonction des besoins des utilisateurs, de la bibliothèque et des organisations parents.

4.3.2 Les buts de l'évaluation de collection

L'évaluation de collection peut également fournir des informations pouvant être utilisées à de nombreux buts (Bushing, 2006; Intner & Futas, 1994; Lakos, 1999).

- Elle peut en résulter un profil de sujet détaillé qui informe les nouveaux membres du personnel de la bibliothèque et les utilisateurs de la nature de la collection.
- Elle peut aider à la rédaction ou à la révision d'une politique de développement de collections et indiquer l'efficacité d'une politique existante. Cela peut aider à expliquer les décisions et les dépenses. Par exemple, la forte utilisation d'une ressource numérique peut expliquer l'attribution d'un pourcentage croissant du budget total des acquisitions au renouvellement de l'abonnement à cette ressource.
- Les informations sur les domaines (sous collections nécessitant un renforcement, un désherbage, ou une mise à jour) collectées à travers une évaluation de collection peuvent être utilisées dans le processus de planification, notamment pour justifier les demandes de budget et les référendums sur le financement. Il peut guider et éclairer les décisions et les décisions politiques dans l'ensemble de la bibliothèque numérique, y compris les affectations budgétaires et de dotation en personnel. Les projets d'analyse axés sur l'état des ressources et leur disponibilité peuvent être utilisés pour la préparation aux catastrophes, l'inventaire et la planification de l'espace.
- Les rapports de l'évaluation de collections peuvent être utilisés dans les rapports d'accréditation et à d'autres buts externes. Certaines bibliothèques universitaires participent à la planification institutionnelle de programmes de nouvelles spécialités.
- Une évaluation spécifique et détaillée de collection peut démontrer à quel point une bibliothèque numérique peut ou ne peut pas soutenir un nouveau programme ou une nouvelle spécialité.
- Les informations sur les points forts de la collection peuvent être utilisées pour recruter de nouveaux membres de la faculté et des étudiants.
- Les résultats des évaluations de collection peuvent fournir aux bibliothèques numériques des informations permettant de documenter leur capacité à soutenir de nouveaux programmes de recherche et développement.
- Les évaluations des sous collections peuvent documenter l'âge des collections par domaines ou sujet et comparer l'activité de circulation des titres détenus par

domaines, montrant ainsi les domaines (sous collections) pouvant nécessiter d'être renforcés.

- Les informations peuvent être recueillies durant l'évaluation de collection pouvant être utilisée dans des communiqués de presse, des rapports de bibliothèque et des lettres d'information, ainsi que pour des propositions de subvention.
- L'évaluation des collections permet à une bibliothèque de partager des informations avec d'autres bibliothèques avec lesquelles elle est impliquée par partenariats existants ou proposés.
- L'analyse de collection peut être utilisée pour démontrer la responsabilité en suivant les progrès accomplis dans la réalisation des objectifs de performance et en montrant comment les investissements sont utilisés efficacement.

4.4 Techniques d'évaluation de collection

Les méthodes et techniques d'évaluation des collections vont des évaluations descriptives impressionnistes aux analyses statistiques complexes (McClure, 2009). Elles peuvent être laborieuses et prendre beaucoup de temps et, comme le remarque McClure, sont parfois exaspérantes (McClure, 2009). Toutes ces techniques cherchent à fournir des informations organisées, pertinentes, spécifiques et précises sur la collection. Deux différentes topologies sont utilisées pour examiner les différentes méthodes d'analyse: les techniques sont soit quantitatives ou qualitatives. Le Tableau 4.2 représente ces topologies de laquelle différentes techniques sont organisées.

Quantitatives	Qualitatives
<ul style="list-style-type: none"> • Statistiques de circulation • Statistiques sur l'utilisation interne • "Clics" et téléchargements (par exemple, les traces de transactions, données fournies par le fournisseur de ressources numériques) • Coût à l'utilisation 	<ul style="list-style-type: none"> • Sondages d'opinion des utilisateurs (par exemple, LibQual +, basé sur le Web, courrier numérique) • Observation de l'utilisateur • Groupes de discussion • Tests d'utilisation

Tableau 4.2 Les Méthodes d'Évaluation des Collections (Nicholson, 2003)

L'évaluation quantitative compte des choses. Elle mesure les titres, les transactions de circulation, les transactions d'accès et de téléchargement avec des ressources numériques et l'argent dépensé. Certaines évaluations quantitatives comparent les mesures par rapport au temps dans une bibliothèque et avec d'autres bibliothèques. Elles tiennent compte de ratios tels que les dépenses pour les périodiques par rapport aux dépenses pour les monographies. Une bibliothèque universitaire peut analyser les dépenses totales de la collection par rapport au nombre d'étudiants, de membres de la faculté et de spécialités. Une bibliothèque publique peut prendre en compte les dépenses annuelles ou les transactions de diffusion par groupe d'utilisateurs ou par succursale. Les méthodes quantitatives démontrent la croissance et l'utilisation des collections en examinant les statistiques de circulation, l'utilisation des ressources numériques, les informations budgétaires. Une fois les lignes de base sont établies, il est possible de mesurer la taille, la croissance et l'utilisation d'une collection avec le temps.

Les systèmes d'information ont grandement facilité la collecte et l'analyse des données de collections. Les questionnaires peuvent extraire des données directement à

Chapitre 4. Utilisation de MCDA pour l'Évaluation de Collection

partir d'un système local en utilisant des rapports fournis par le système ou en écrivant des générateurs de rapports personnalisés (Stowers & Tucker, 2009).

L'analyse qualitative est plus subjective que l'analyse quantitative car elle dépend de la perception, de l'opinion et du contexte dans lequel les données sont collectées. Gorman et Crayton proposent cette définition: « La recherche qualitative est un processus d'enquête qui tire des données du contexte dans lequel les événements se produisent, dans le but de décrire ces événements, comme moyen de déterminer le processus dans lequel: les événements sont intégrés et les perspectives de ceux qui participent aux événements, utilisant l'induction pour dériver des explications possibles basées sur des phénomènes observés. " (Gorman, Clayton, Shep, & Clayton, 2005).

L'analyse qualitative d'une collection a pour objectif de déterminer les forces et les faiblesses de la collection, qui reflètent les décisions délibérées de ne pas collecter, et le degré de satisfaction de la collection par rapport aux besoins des utilisateurs. Cela dépend de l'opinion des gestionnaires locaux et des experts externes et des perceptions des utilisateurs. Même lorsque les collections sont comparées à des listes externes, ces listes sont elles-mêmes le résultat d'une opinion éclairée sur ce qui constitue une bonne collection, ce qui caractérise une collection désignée comme niveau de collecte spécifique ou ce qui constitue une collection appropriée pour un groupe d'utilisateurs spécifique.

Toutes les évaluations de collection, qu'elles soient qualitatives ou quantitatives, doivent utiliser des méthodes de recherche éprouvées. Celles-ci exigent une compréhension claire de ce qui est mesuré, de la manière de la mesurer et de l'interprétation des résultats. L'évaluation de collection commence par une question non ambiguë à laquelle il faut répondre. Un projet de recherche bien fait produit des informations à la fois fiables (susceptibles de donner les mêmes résultats une fois répétées) et valides (mesure ce qu'il est prévu de mesurer). En d'autres termes, les résultats sont répétables et les conclusions sont vraies.

4.5 Critères d'évaluation des collections

L'évaluation de collection fait intervenir plusieurs critères. Car la collection est un ensemble de ressources numériques, le processus d'évaluation des ressources numériques est la base de l'évaluation de collection. L'évaluation de la collection commence par l'évaluation des ressources numériques, une fois les ressources sont évaluées, on peut obtenir une évaluation globale de la collection et des sous collections (Waugh, Donlin, & Braunstein, 2015; Wilson, 2004), cette évaluation peut être utilisé par la suite pour accomplir les tâches de désabonnement des ressources numériques inutilisables, de renouvellement ou acquièrent des nouvelles ressources. Les gestionnaires doivent régulièrement effectuer ces tâches afin de préserver les collections pour leurs clients (utilisateurs). c'est un travail très important qui nécessite une grande attention (Manley, 1996). Dans notre travail, nous avons choisi des critères considérés comme des facteurs efficaces dans l'évaluation des ressources dans la littérature (Johnson, 2014; Polanka, 2012; Snyder, 2014). Les critères utilisés dans le processus d'évaluation sont illustrés dans le Tableau. 4.3 Et expliqués dans ce qui suit:

Criteria	U	P	A	Ag	L	Ath	C
Actions							
Material1	U1	P1	A1	Ag1	L1	Ath1	C1
Material2	U2	P2	A2	Ag2	L2	Ath2	C2

Tableau 4.3 Les Critères d'évaluation (Hemili & Laouar, 2018b)

- **Utilisation (U):** nombre de téléchargements ou de lectures
- **Pertinence (P):** le contenu est-il toujours pertinent?
- **Accessibilité (A):** est-il disponible et facilement accessible ailleurs?
- **Age (Ag):** Année de publication
- **Langue (L):** est-ce dans une langue que les utilisateurs actuels et futurs peuvent lire?
- **Auteur (Ath):** Numéro de lecture de l'auteur
- **Coût (C):** prix de la numérisation ou du renouvellement
- **Sujet (s):** Est-ce que c'est dans une langue que les utilisateurs actuels et futurs peuvent lire?

Les poids des critères utilisés dans ce travail sont déterminés par les gestionnaires, la section suivante présente l'approche proposé pour résoudre le problème d'évaluation des collections et les étapes de la méthode PROMETHEE II afin de résoudre le problème.

4.6 Approche multicritères basée sur la méthode PROMETHEE II pour l'évaluation des collections

Dans différents problèmes multicritères, et dans le but de trouver des solutions, les chercheurs ont utilisé un riche ensemble de méthodes telles que PROMETHEE (I, II et III), AHP, ELECTRE (II, III et IV), TOPSIS et SAW. Pour l'évaluation des collections pour différentes activités de gestion des collections, nous avons sélectionné PROMETHEE II comme méthode d'évaluation des ressources. PROMETHEE II est une méthode de MCDA. Cette méthode sera utilisée lorsque plusieurs critères sont utilisés dans le processus de prise de décision. Elle compare toutes les options (choix) globales impliquées dans la détermination des scores de flux net, en se basant sur ces scores on obtient une évaluation totale d'alternatives (Epepe & De Smet, 2014).

Dans la première partie de notre travail, nous présentons le problème dans une matrice de décision et nous utilisons PROMETHEE II (Brans, Vincke, & Mareschal, 1986) pour évaluer et ordonner les actions en fonction des informations fournies par les utilisateurs de bibliothèques numériques universitaires.

Le problème de l'évaluation des ressources numériques est un problème multicritère; par conséquent, nous avons utilisé une matrice pour le représenter. On appelle matrice (1) Une matrice de décision, il s'agit d'un $M * N$, un élément de cette matrice représente une évaluation d'une alternative sur un critère, identique à tous les éléments. Les poids des critères sont déterminés par les gestionnaires.

$$\text{Actions} \begin{cases} g_1 & g_2 & \dots & g_m \\ a_1 & v_{11} & v_{12} & \dots & v_{1m} \\ a_2 & v_{21} & v_{22} & \dots & v_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ a_n & v_{n1} & v_{n2} & \dots & v_{nm} \\ w_1 & w_2 & \dots & w_m \end{cases} \quad (1)$$

- **Etape 1**

Premièrement, nous construisons notre matrice de problèmes qui contient toutes les évaluations comme dans le Tableau. 4.3. Les gestionnaires fournissent les données contenant les préférences, puis les données transformées dans un tableau de préférences.

- **Etape 2**

Après la construction de la matrice de décision, nous comparons maintenant chaque deux ressource numérique (alternatives) par chaque critère afin de déterminer la valeur de différence entre les ressources. Si $g_i(a_j)$ est l'évaluation du critère g_i pour la ressource a_j , alors la différence entre deux ressources numériques et est représentée par $d_j(a_i, a_k)$.

$$d_j(a_i - a_k) = g_j(a_i) - g_j(a_k) \quad (2)$$

Après avoir obtenu les différences entre les ressources numériques, nous transformons ces différences en valeurs de préférences confinées entre 0 et 1. Lors de la transformation des différences en préférences, nous considérons qu'une ressource numérique est meilleure qu'une autre ressource numérique lorsque ses performances sont strictement plus gros que l'autre, si ses performances sont strictement meilleures. $P_j(a_i, a_k)$ est le montant de la préférence entre deux ressources par rapport au critère j . Pour calculer les quantités de préférences entre ressources numériques, nous utilisons la formule suivante:

$$\begin{cases} 0, & d_j(a_i, a_k) \leq 0 \\ 1, & d_j(a_i, a_k) > 0 \end{cases} \quad (3)$$

- **Etape 3**

La pondération des critères est très importante dans la MCAD pour donner la bonne évaluation à nos ressources numériques, car ce travail incombe aux gestionnaires qui choisissent les valeurs de pondération utilisées par la méthode MCDA.

- **Etape 4**

Une fois la valeur de préférence entre deux ressources numériques pour chaque critère est obtenue, nous calculons la valeur totale de préférence entre deux ressources numériques pour tous les critères (fonctionnalités), et nous la répétons pour chaque couple de ressources numériques. Pour obtenir la valeur de préférence totale, nous utilisons des poids déterminés par les gestionnaires et la formule suivante:

$$\pi(a_i, a_i) = \sum_{j \in C} w_j \times p_j(a_i, a_k) \quad (4)$$

Chapitre 4. Utilisation de MCDA pour l'Evaluation de Collection

La valeur totale de la préférence calculée représente la meilleure ressource entre deux ressources numériques pour tous les critères.

- **Etape 5**

Pour mettre toutes les ressources numériques dans une liste ordonnée, nous devons calculer les flux de classement (positifs et négatifs). Ces flux nous permettent de déterminer la position de toutes les ressources numériques et enfin de les hiérarchiser. Pour le calcul et on utilise:

$$\Phi^+(a_i) = \frac{1}{n-1} \sum_{a_k \in A} \pi_i(a_i, a_k) \quad (4)$$

$$\Phi^-(a_i) = \frac{1}{n-1} \sum_{a_k \in A} \pi_i(a_k, a_i) \quad (5)$$

- **Etape 6**

Enfin, en utilisant les flux positifs et négatifs en calculant les flux globaux, et en l'utilisant comme un niveau d'évaluation des ressources, nous déterminons comme suit:

$$\Phi(a_i) = \Phi^+(a_i) + \Phi^-(a_i) \quad (6)$$

- **Etape 7**

La méthode utilisée PROMETHE II nous donne une évaluation pour chaque ressource numériques (le flux globale $\Phi(a_i)$). Les ressources numériques avec la valeur la plus grande du flux total sont meilleures que les autres. Pour évaluer les sous collections, nous utilisant la formule suivant :

$$SC_j = \frac{\sum_{i=1}^{i=n} \Phi_{a_i}}{n} \quad (7)$$

SC_j représente le niveau d'évaluation de la sous collection j, et Φ_{a_i} représente les ressources de SC_j .

- **Etape 8**

Dans cette dernière étape, on calcule l'évaluation globale de la collection C, en utilisant la formule suivant :

$$C = \frac{\sum_{j=1}^{j=n} SC_j}{n} \quad (7)$$

4.7 Conclusion

Dans ce chapitre, l'utilisation de la méthode PROMETHEE II pour évaluer les ressources, les sous collection et collection de la bibliothèque numérique a été

Chapitre 4. Utilisation de MCDA pour l'Évaluation de Collection

présentée. Le but de l'étude était d'aider les gestionnaires dans les activités de gestion des collections en lui proposant une évaluation des ressources et de collection que l'aidant dans les processus de prise de décision.

Le problème d'évaluation des ressources numériques est très complexe, le processus de recherche de solutions à ce problème est très difficile et gênant pour les gestionnaires. Les anciens processus d'évaluation utilisés par les gestionnaires dans de nombreuses bibliothèques numériques universitaires sont basés sur des critères économiques et centrés sur des nombres statistiques pour prouver leurs décisions.

Les travaux de recherche existants qui visent à aider les gestionnaires présentent plusieurs méthodes d'évaluation des bibliothèques numériques universitaires, mais aucune d'entre elles ne se concentre sur des activités spécifiques. Ce travail visait à aider les gestionnaires à prendre des décisions dans le cadre d'activités spécifiques de gestion des collections en évaluant les sous collections et les ressources numériques. Différents critères ont été utilisés dans le processus d'évaluation et l'approche proposée a prouvé que MCDA peut offrir des outils efficaces pour appuyer la prise de décision dans la gestion de la collection.

Chapitre 5

Un Système d’Aide à la Décision pour Gérer le Développement des Collections

Contenu

5.1	<i>Introduction</i>	79
5.2	<i>Évaluation et sélection des ressources numériques</i>	80
5.3	<i>DSS pour soutenir le développement de la collection dirigé par les demandes des utilisateurs (CDDSS)</i>	82
5.3.1	<i>Critères considérées par le DSS</i>	83
5.3.2	<i>Analyse de fichier log pour les bibliothèques numériques académiques</i> ..	85
5.3.3	<i>Phase 1: extraction des préférences des utilisateurs à partir des requêtes de recherche</i>	87
5.3.4	<i>Phase 2: Détermination des poids des critères</i>	87
5.3.5	<i>Phase 3: Evaluation et classement des ressources numériques</i>	88
5.4	<i>L’implémentation d’un prototype du CDDSS</i>	89
5.5	<i>Data Warehouse pour CDDSS</i>	94
5.5.1.	<i>Les composantes de l’architecture proposée</i>	94
5.5.2.	<i>Schéma des données</i>	95
5.6	<i>Proposition d’une Architecture Cloud pour le CDDSS développé</i>	96
5.6.1.	<i>Logiciel de gestion Cloud</i>	96
5.6.2.	<i>Les composantes de l’architecture Cloud</i>	97
5.7	<i>Conclusion</i>	100

5.1 Introduction

Les bibliothèques numériques académiques offrent aux utilisateurs une collection organisée de ressources numériques (e-ressources). La collection de ressources numériques d'une bibliothèque numérique académique doit satisfaire les besoins d'informations de leurs utilisateurs pour éviter que les bibliothèques numériques académiques tombent dans l'oubli et mettent fin à leurs activités (Cabrerizo et al., 2015; R. Kadir et al., 2009). Par conséquent, les bibliothèques numériques académiques évoluent continuellement pour répondre aux besoins d'information pour l'enseignement et l'apprentissage, et pour identifier les problèmes à poursuivre pour un développement continu.

Ces dernières années, la croissance dans le nombre des ressources numériques disponibles et de la population des utilisateurs a rendu le processus de développement de la collection très difficile à gérer. Le grand nombre de sources de données et les grands volumes de données à analyser sont les principaux défis auxquels les gestionnaires des bibliothèques numériques sont confrontés dans le processus de développement de la collection (Koenig, 1985; Tramullas et al., 2013). Ces difficultés peuvent amener les gestionnaires des bibliothèques numériques à faire de mauvaises analyses et à les obliger à prendre des décisions en se basant sur des informations fausses ou insuffisantes, ce qui signifie des utilisateurs insatisfaits et les bibliothèques perdent leurs utilisateurs et leurs argents.

Le processus fondamental dans le développement de la collection est la sélection des ressources numériques pour performer les différentes activités de développement de la collection, comme l'achat ou l'abonnement à de nouvelles ressources. Le processus de sélection peut être très complexe, à cause des nombreux facteurs pouvant être impliqués dans cet processus et qui doivent être pris en compte, comme les besoins d'information de la communauté des utilisateurs, les prix des ressources numériques et le budget destiné à la développement de collection de la bibliothèque numérique académique (Siguenza Guzman et al., 2014).

Dans ce chapitre, nous exposons le système d'aide à la décision (DSS) basé sur la méthode PROMETHEE II et la procédure SIMOS révisée (Brans et al., 1986; José Figueira & Bernard Roy, 2002) pour évaluer et sélectionner les ressources numériques en fonction des besoins des utilisateurs, ainsi, nous présentons une architecture d'un Data Warehouse pour intégrer les données et une architecture d'un Cloud privé pour déployer le DSS et le Data Warehouse. Les étapes du DSS sont représentées dans la Figure 5.1.

• Contributions majeurs

- La proposition d'un système décisionnel basé sur l'analyse multicritère pour diriger le développement des collections.
- L'analyse du fichier log qui permet au DSS d'utiliser les besoins réels des utilisateurs pour évaluer et sélectionner des ressources numériques.
- La proposition d'une architecture d'un Data Warehouse afin de résoudre le problème d'hétérogénéité des sources des données et de formats.
- L'utilisation de la technologie Cloud Computing pour le déploiement du DSS et du Data Warehouse, afin de résoudre les problèmes liés à la puissance de calcul et au stockage.

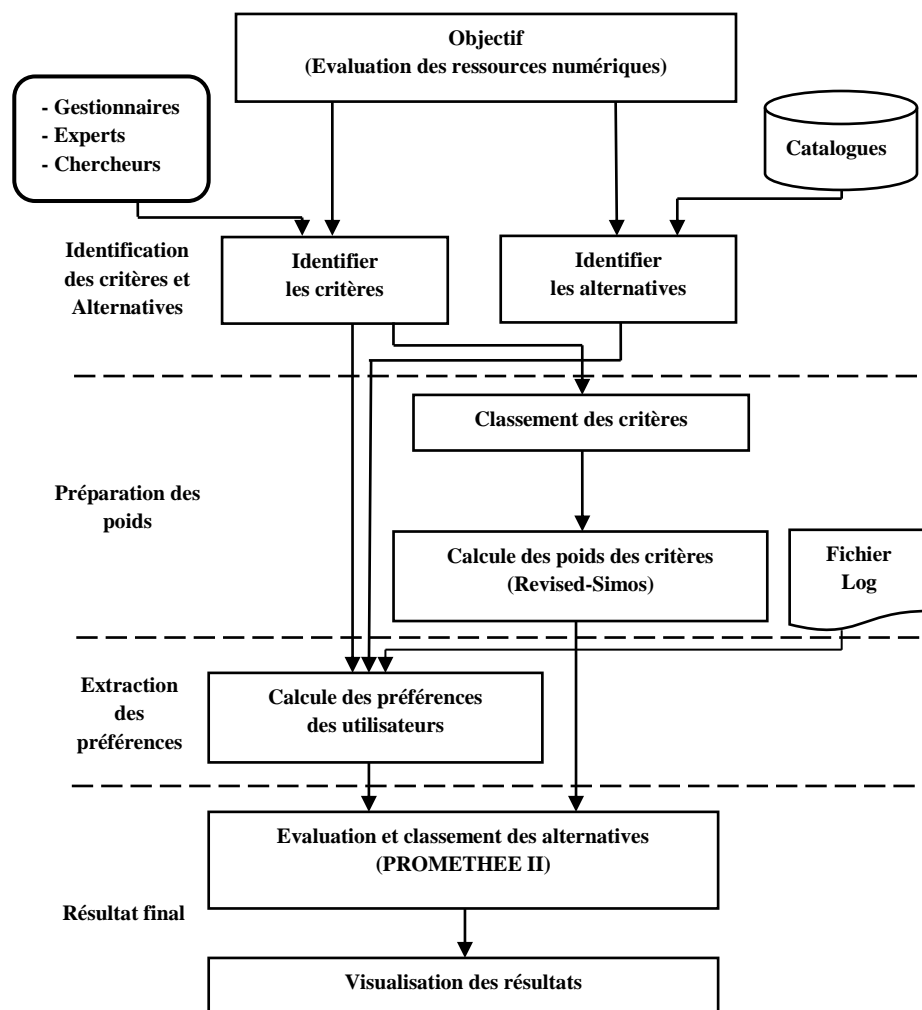


Figure 5.1 Les Etapes du DSS (Hemili et al., 2019)

5.2 Évaluation et sélection des ressources numériques

Le processus de développement de la collection est basé sur la sélection de ressources numériques, voir la Figure 5.2, les gestionnaires en tant que des décideurs disposant de processus informels pour sélectionner ces ressources numériques. Les gestionnaires jouent un rôle central dans le processus de sélection. Ils recommandent l'acquisition de ressources numériques basé sur la base des recommandations des gestionnaires en tant que facteur majeur, de la recommandation de la faculté et de la recommandation des étudiants (Flatley & Prock, 2009; Lgimol, 2015).

La sélection et l'acquisition de ressources numériques peuvent potentiellement avoir des conséquences importantes sur la viabilité future des ressources numériques en tant que produit (S. Eom & Laouar, 2017; R. Kadir et al., 2009; Magdalini, Richard, & Jennifer, 2012). Plus généralement, il est clairement nécessaire d'améliorer la sélection (Thompson & Sharp, 2009).

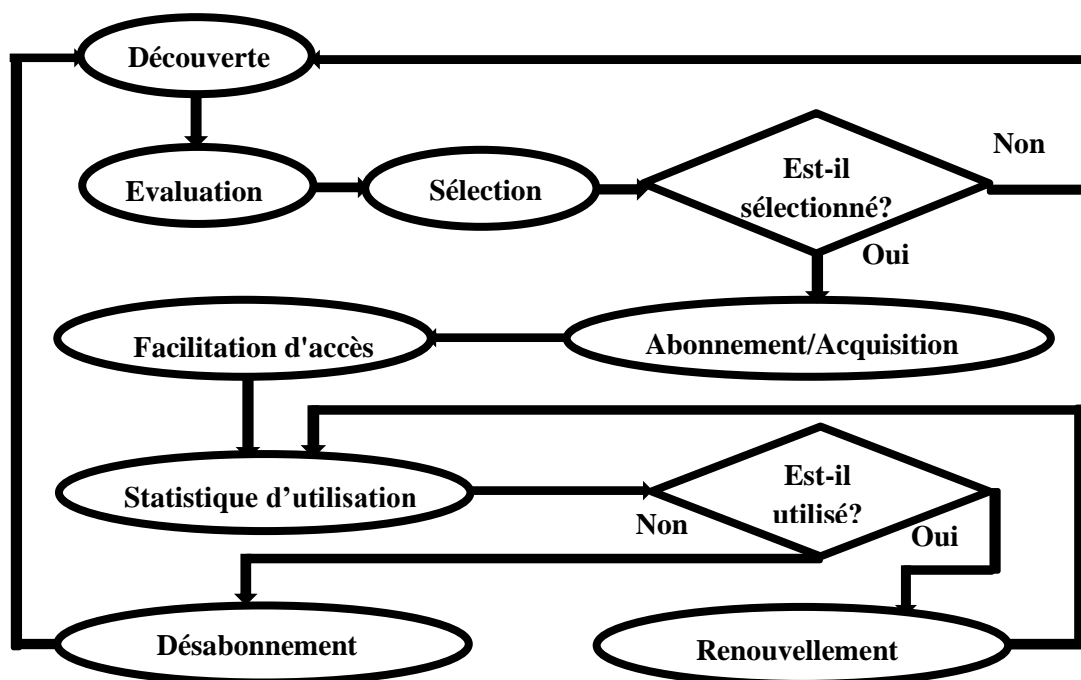


Figure 5.2 Cycle de Vie des Ressources Electroniques (Anbu, Kataria, & Ram, 2013).

En pratique, les gestionnaires ont tendance à s'appuyer à la fois sur les recommandations des académiciens de l'institut mère de la bibliothèque numérique et sur les bases de données de leurs fournisseurs pour définir ou bien déterminer les ressources numériques à acquérir. Néanmoins, l'absence d'outils de sélection fiables reste un problème majeur pour identifier les ressources numériques les plus pertinentes pour les bibliothèques universitaires (Bendib, Ridda Laouar, Hacken, & Miles, 2014; Kassim, 2017; Lonsdale & Armstrong, 2001; Walters, 2013).

Pour répondre au problème posé, le DSS basé sur l'aide à la décision multicritères (MCDM) est utilisé dans ce travail.

Le MCDM a constamment été une discipline de référence majeure au cours des dernières décennies pour soutenir les problèmes impliquant des problèmes de MCDM (S. Eom, 2016). MCDM traite généralement de nombreuses problèmes non dénommés qui impliquent plusieurs attributs, objectifs et buts (Zeleny, 1982).

Les systèmes d'aide à la décision multicritères (MCDSS) représentent les systèmes d'aide à la décision (DSS) intégrés au modèle MCDM. Le MCDSS peut être globalement classé en un MCDSS orienté données généralisé qui est basé sur des modèles de prise de décision multi-attributs (Benatia, Laouar, Bendjenna, & Eom, 2016; Jelassi, Jarke, & Stohr, 1985), et en un MCDSS orienté modèle, qui est basé sur de modèles décisionnels multi-objectifs (H. B. Eom, Lee, Snyder, & Ford, 1987-1988).

Les MCDSSs permettent aux décideurs de traiter des informations quantitatives et qualitatives essentielles aux problèmes mal structurés lors de la construction de la structure multi-objectifs. MCDSS fournit la procédure de recherche de solution interactive conçue pour faire des compromis continus entre diverses alternatives jusqu'à l'obtention de la meilleure solution disponible (ou de compromis). Tout au long du processus décisionnel, le MCDSS insiste sur les rôles cruciaux du jugement du décideur ou de la rationalité limitée, qui reflètent mieux les comportements cognitifs réels du décideur. Par conséquent, il est clair que MCDM complète DSS et vice versa.

5.3 DSS pour soutenir le développement de la collection dirigé par les demandes des utilisateurs (CDDSS)

Dans cette section, nous présentons le DSS qui génère un classement des ressources numériques pour le personnel des bibliothèques numériques académiques afin d'améliorer les collections proposées par les bibliothèques numériques académiques. La Figure 5.3 montre les composantes du système décisionnel proposé.

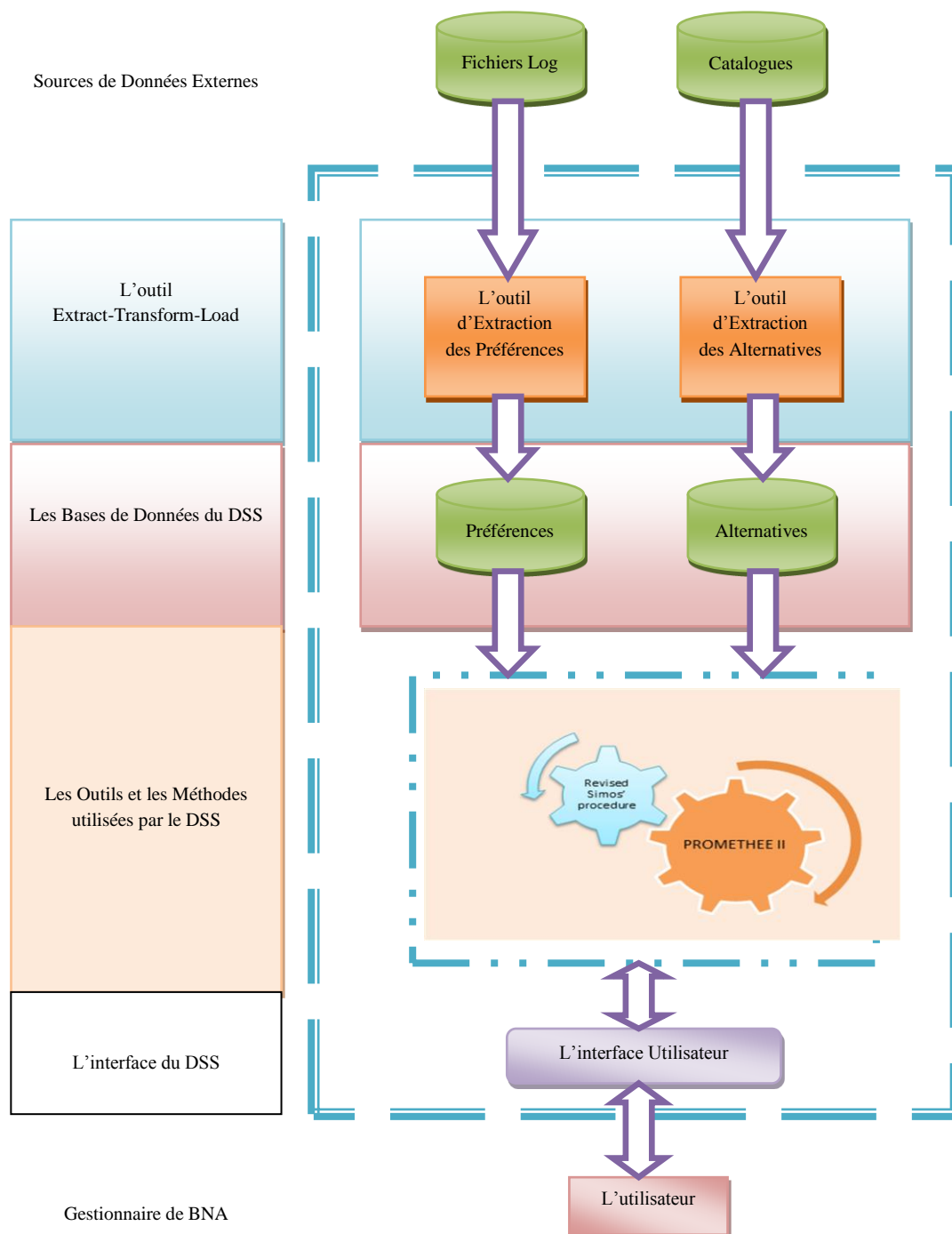


Figure 5.3 Architecture générale du système proposé (Hemili et al., 2019)

- La première composante est constituée des outils Extract-Transform-Load (ETL) ou Extraire-Transformer-Charger en français, elle joue les rôles suivants:

- Elle extrait les requêtes de recherche des utilisateurs à partir des fichiers logs ;
 - Elle extrait les alternatives (les ressources numériques) à partir des catalogues des ressources numériques ;
 - Elle analyse les requêtes pour les transformer à des préférences des utilisateurs ;
 - Les outils ETL charge les préférences et les alternatives dans les bases des données du DSS.
- La deuxième composante représente les deux bases de données du DSS,
 - Au niveau de la première base de données on trouve les préférences des utilisateurs obtenues par la première composante ;
 - Au niveau de la deuxième base de données on trouve les alternatives extrait à partir des catalogues par l'ETL.
 - La troisième composante est la base de modèle, elle est constituée de la méthode PROMETHEE II et la procédure Simos-Révisée (Revised-Simos), les rôles des modèles sont :
 - Déterminer la pondération ou les poids des critères à l'aide de la procédure Simos-Révisée ;
 - Réaliser une analyse multicritère afin d'établir le classement des alternatives à l'aide de la méthode PROMETHEE II.
 - La quatrième et la dernière composante du DSS représente l'interface utilisateur du DSS qui permet aux utilisateurs (les gestionnaires des bibliothèques numériques académiques) du DSS :
 - D'utiliser le DSS de façon très simple, voir Figure 5.3.

Dans les sous-sections suivantes, la première sous-section 5.4.1 présente les critères pris en compte ou bien considérées par le CDDSS dans le processus d'évaluation et de sélection des ressources numériques. Après, la sous-section suivante 5.4.2 présente l'analyse fichiers logs qui contient les requêtes des utilisateurs des BNAs pour l'extraction des préférences de ces utilisateurs. Les restes sous-sections exposent les phases d'évaluation et de classement des alternatives (ressources numériques), chaque sous-section présente une phase exécutée durant l'utilisation du CDDSS en détails.

5.3.1 Critères considérées par le DSS

Comme mentionné brièvement plus tôt, de nombreux gestionnaires des bibliothèques numériques fondent leurs décisions de développement de collection sur les acquisitions dirigées par la demande des clients (Lehman, 2014). Par conséquent, pour établir les critères qui seront pris en compte par le CDDSS, les chercheurs de ce travail analysent les requêtes de recherche présentes dans les fichiers journaux (logs) d'une bibliothèque numérique académique. Ils ont divisé les requêtes de recherche entrées par les utilisateurs dans le champ de recherche en plusieurs types différents tels que titres, mots-clés, noms d'auteurs et revues, noms de bases de données, etc. Parmi les types de requêtes de recherche saisies, les chercheurs ont sélectionné onze types d'entre eux et les considérées comme des critères pour la prise de décision. Ces critères

Chapitre 5. DSS pour Gérer le Développement des Collections

sont utilisés par le DSS car ils représentent les informations les plus liés aux préférences des utilisateurs de bibliothèques numériques académiques.

Les documents numériques sont une forme majeure de ressources numériques générées par un logiciel. Les informations sur les documents numériques peuvent être interprétées et traitées par un ordinateur et d'autres dispositifs numériques. Plusieurs identificateurs associés aux ressources numériques peuvent être utilisés pour traiter les ressources numériques par le CDDSS. Ils comprennent les identifiants de document numérique, les mots clés de document numérique, les dates de document numérique et les noms de journaux numériques. Les critères pris en compte par le CDDSS sont présentés dans les catégories suivantes.

A. Critères basés sur l'identifiants de documents numériques

- DOI: nombre d'occurrences de DOI de documents numériques dans les chaînes de requête de l'utilisateur.
- Titre: nombre d'occurrences du titre du document numérique dans les chaînes de requête de l'utilisateur.
- Title_Part: nombre d'occurrences de mots-clés du document numérique title_part dans les chaînes de requête de l'utilisateur.

B. Critères basés sur les mots-clés du document numériques

- Kw1: nombre d'occurrences du titre du premier mot clé du document numérique dans les chaînes de requête de l'utilisateur.
- Kw2: nombre d'occurrences du titre du deuxième mot clé du document numérique dans les chaînes de requête de l'utilisateur.
- Kw3: nombre d'occurrences du titre du troisième mot clé du document numérique dans les chaînes de requête de l'utilisateur.

C. Critères basés sur la date de publication du document numériques

- Année: L'année de publication du document numérique, le plus récent document numérique est le plus préféré.
- Mois : mois de publication.

D. Critères basés sur la revue numérique

- Nom: nombre d'occurrences de nom de journal numérique dans les chaînes de la requête.
- ISSN: nombre d'occurrences d'ISSN de documents numérique dans les chaînes de requête de l'utilisateur.

5.3.2 Analyse de fichier log pour les bibliothèques numériques académiques

Les utilisateurs interagissent avec les systèmes de recherche d'informations de bibliothèques numériques académiques en créant et en entrant des chaînes de requête dans le champ de recherche. Les systèmes de recherche d'informations dans les bibliothèques numériques collectent des données d'interaction sur les utilisateurs, voir Figure 5.4. Toutes les activités des utilisateurs sont automatiquement collectées, consignées et horodatées dans des fichiers journaux (logs). Ces activités sont associées aux utilisateurs avec des identifiants.

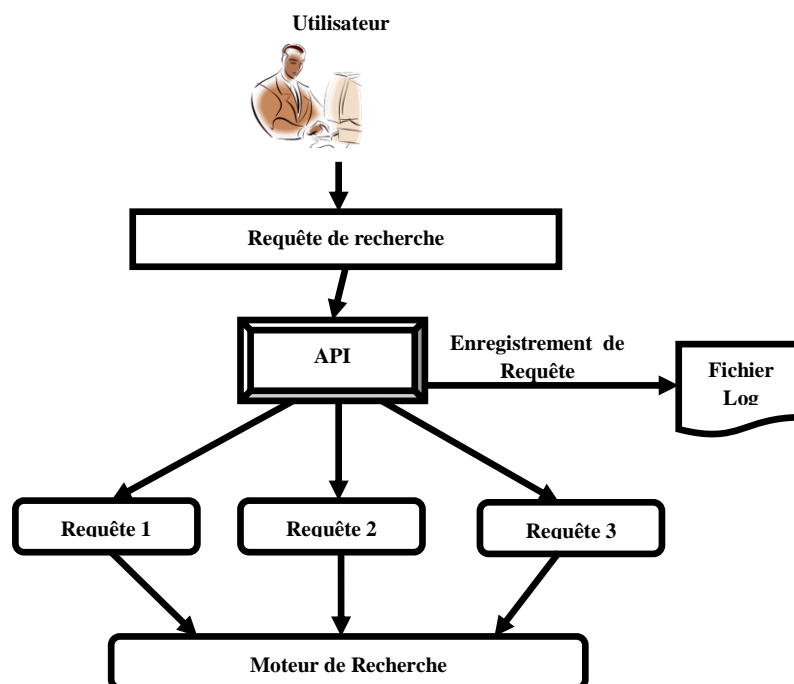


Figure 5.4 Obtenez des requêtes à partir de l'application de recherche (Priya. P, 2019)

Les activités sauvegardées peuvent inclure des chaînes de requêtes de recherche, des options de recherche, une liste des résultats renvoyés par l'IR (système de recherche) et la taille des ensembles de résultats. Les types de requêtes de recherche entrés sont les suivants:

- Recherche booléenne ;
- Recherche avec classement.

Ces types se diffèrent par :

- La proximité du terme (dans le même matériau, même page ou première page);
- Nombre maximal de documents à retourner;
- Nombre de documents à afficher sur chaque page de résultats.

Le fichier journal (log) enregistre le nombre de matériaux résultant que l'utilisateur choisit d'afficher pour chaque chaîne de requête, ainsi que la position des matériaux choisis par l'utilisateur dans la liste des résultats (H. Han, Jeong, & Wolfram, 2014; Jones, Cunningham, McNab, & Boddie, 2000).

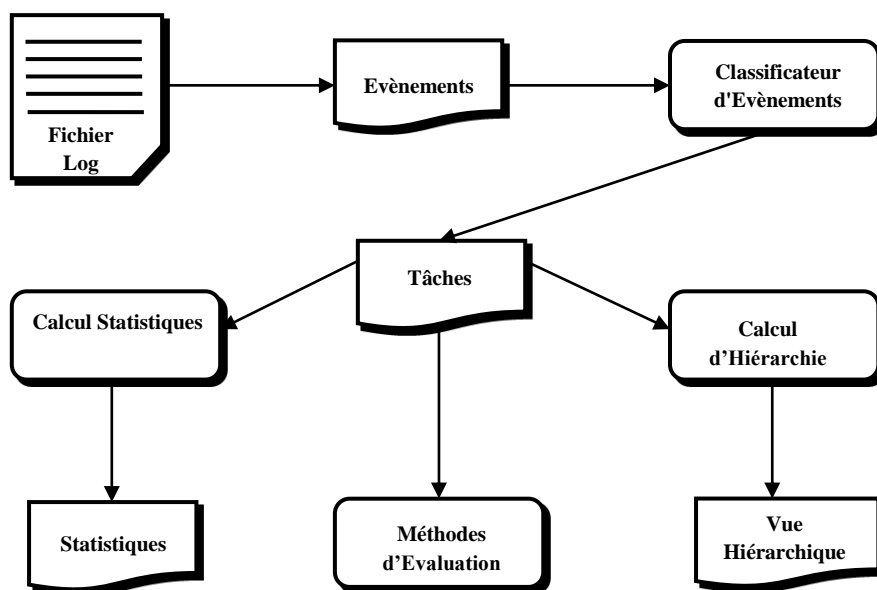


Figure 5.5 Les Etapes d'Analyse d'un Fichier Log (Mosqueira-Rey & Moret-Bonillo, 2006).

Étant donné que les journaux de transactions contiennent cette quantité d'informations, ils représentent une source d'informations très importante pour les chercheurs et les bibliothèques numériques, qui peuvent être utilisés pour identifier et comprendre les utilisateurs (Chao et al., 2011; H. Han et al., 2014; Jansen, Spink, & Saracevic, 2000; Jones et al., 2000; J. Zhang, Wolfram, & Wang, 2009).

Cependant, la Figure 5.5 illustre les étapes d'analyse des fichiers log. Dans ce travail nous sommes intéressés à déterminer les besoins d'informations des utilisateurs à partir des requêtes de recherche des utilisateurs. Pour accomplir cette tâche nous utilisons l'algorithme présenté dans la figure 5.6:

```

Ligne = lire une ligne du fichier log ;
Evènement = type d'évènement dans Ligne ;
Si l'évènement = recherche d'utilisateur alors {
    Requête = Extraire la requête de recherche de
    Ligne ;
    Si la requête n'existe pas dans la base de données
    des traces alors
    ajouter la requête dans la base de données des traces ;
    Fin ;}
Fin ;
Pour chaque utilisateur faire :
    Récupérer la liste des requêtes ;
    Pour chaque requête faire {
    
```

Figure 5.6 Algorithme d'Analyse du Fichier Log (Hemili et al., 2019)

5.3.3 Phase 1: extraction des préférences des utilisateurs à partir des requêtes de recherche

Afin de déterminer les préférences des utilisateurs, les chercheurs extraient l'ensemble des requêtes de recherche des utilisateurs à partir des fichiers journaux, puis analysent les requêtes de recherche afin de connaître leur type (titre, partie du titre, mot-clé, nom de l'auteur, nom du journal, nom de la base de données, année, prix, etc.), les chercheurs convertissent ensuite chaque requête de recherche en une préférence de la table d'évaluation. Par exemple, si la requête de recherche est un titre de livre numérique, les chercheurs ajoutent 1 dans son critère de titre. Le résultat de ce processus est un tableau d'évaluation voire Tableau. 5.1.

Criteria Actions	E-document ID				Keyword(s)		Author(s)		Date	E-journal
	DOI	ISSN	Title	T_Part	KW1	KW2	A1	A2	Year	Name
E-document1	D1	I1	T1	TP1	K11	K21	A11	A21	Y1	N1
E-document1	D2	I2	T2	TP2	K12	K22	A12	A22	Y2	N2

Tableau 5.1 Préférences des utilisateurs (Hemili et al., 2019)

Dans le Tableau. 5.1, D1 indique le nombre des utilisateurs qui ont cherchés du DOI du document numérique 1, ou réellement D1 représente le nombre des utilisateurs ont demandés le document 1, ces demandes existes dans les requêtes de recherche des utilisateurs.

5.3.4 Phase 2: Détermination des poids des critères

Le processus de détermination de la pondération des poids des critères est un processus clé dans toute analyse multicritère. La littérature suggère de nombreuses méthodes pour calculer les valeurs de pondération des critères. Parmi ces méthodes, les chercheurs ont choisi une procédure très simple donnant une valeur pondérale appropriée, appelée procédure Simos-Révisée (Rvised-Simos) proposée par J. Simos (Brans et al., 1986), et révisée par Figueira and Roy(2002), Simos a proposée de donner à un décideur un ensemble de cartes contenant les noms des critères et des cartes blanches, le DM classe ces cartes du critère le moins important au critère le plus important et insère des cartes blanches entre deux critères si le critère est très important l'autre, ils utilisent ensuite les deux formules suivantes (Shanian, Milani, Carson, & Abeyaratne, 2008) pour déterminer les valeurs numériques des poids.

Poids non normalisé

$$\hat{w}_j = 1 + \frac{(z - 1) \left[l(j) - 1 + \sum_{s=1}^{l(j)-1} e_s \right]}{v - 1 + \sum_{s=1}^{v-1} e_s} \quad (1)$$

Poids normalisé

$$w_j = \frac{\hat{w}_j}{\sum_{k=1}^m \hat{w}_k} \quad (2)$$

5.3.5 Phase 3: Evaluation et classement des ressources numériques

Dans cette section, nous avons appliqué la méthode PROMETHEE II (Brans et al., 1986) pour classer les alternatives en fonction des informations relatives aux préférences de chaque utilisateur. Les principales étapes de PROMETHEE II peuvent être décrites comme suit:

Étape 1: Élaboration de la matrice de performance des alternatives comme présentée dans le Tableau. 5.2.

a	$g_1()$	$g_2()$...	$g_j()$...	$g_m()$
a_1	$g_1(a_1)$	$g_2(a_1)$...	$g_j(a_1)$...	$g_m(a_1)$
...
a_i	$g_1(a_i)$	$g_2(a_i)$...	$g_j(a_i)$...	$g_m(a_i)$
...
a_n	$g_1(a_n)$	$g_2(a_n)$...	$g_j(a_n)$...	$g_m(a_n)$

Tableau 5.2 Les préférences d'évaluation

Cette étape s'exécute en phase 1 et le résultat est un tableau identique au Tableau. 5.1.

Étape 2: Spécification de la fonction de préférence marginale. Dans ce cas, pour le critère g_j , cette fonction traduit la différence entre les évaluations de deux alternatives:

$$d_j(a_i - a_k) = p_j(a_i) - p_j(a_k) \quad (3)$$

En un indice de préférence marginal $p_j(\alpha_i, \alpha_k)$ compris entre 0 et 1. Lorsque l'on compare les alternatives, cette fonction implique qu'une alternative est préférée à une autre sur g_j si ses performances sont strictement meilleures. $p_j(\alpha_i, \alpha_k)$ est la valeur du degré de préférence d'un critère g_j pour deux alternatives α_i et α_k . Les fonctions de préférence utilisées pour calculer ces degrés de préférence sont définies par:

$$p_i(a_i, a_k) = \begin{cases} 0, & d_j(a_i, a_k) \leq 0 \\ 1, & d_j(a_i, a_k) > 0 \end{cases} \quad (4)$$

Étape 3: Spécification ou définition de l'importance relative de chaque critère (c'est-à-dire son poids w_j). Dans cette étude, les poids sont déterminés à l'aide de la procédure Simos-Révisée décrite dans la sous-section 5.4.4.

Étape 4: Déterminez l'indice de préférence complet pour la préférence de α_i par rapport à α_k pour toutes les paires d'alternatives $\alpha_i, \alpha_k \in A$:

$$\pi(a_i, a_k) = \sum_{j \in C} w_j \times p_j(a_i, a_k) \quad (5)$$

Chapitre 5. DSS pour Gérer le Développement des Collections

Cet indice mesure la force avec laquelle α_i est préféré à α_k , compte tenu de tous les critères considérés conjointement sur une échelle de 0 à 1.

Étape 5: Agrégation des indices de préférence aux flux de classement positifs $\Phi^+(\alpha_i)$ et négatifs $\Phi^-(\alpha_i)$ qui expriment respectivement combien α_i est supérieur à toutes les autres alternatives n-1:

$$\Phi^+(a_i) = \frac{1}{n-1} \sum_{a_k \in A} \pi_i(a_i, a_k) \quad (6)$$

$$\Phi^-(a_i) = \frac{1}{n-1} \sum_{a_k \in A} \pi_i(a_k, a_i) \quad (7)$$

Ces indices peuvent être interprétés comme la force et la faiblesse globales de α_i . Enfin, le flux de classement $\Phi(\alpha_i)$ reflète la différence entre les flux positifs et négatifs de α_i :

$$\Phi(a_i) = \Phi^+(a_i) + \Phi^-(a_i) \quad (8)$$

Étape 6: PROMETHE II fournit le classement complet des alternatives par le biais de flux nets complets $\Phi(\alpha_i)$. Évidemment, le flux le plus élevé est le mieux. Notez que dans cette approche, seules les relations de préférence P et d'indifférence I sont possibles pour la comparaison de chaque paire d'alternatives.

5.4 L'implémentation d'un prototype du CDDSS

Le prototype du DSS est basé sur le langage de programmation Java et la base de données MySQL. Le logiciel Java® 2 Platform est utilisé pour créer et exécuter les applications, les applets et les servlets de la plateforme Java. MySQL est la plateforme de base de données open source la plus utilisée au monde. La dernière version de MySQL 8.0 crée et déploie la nouvelle génération d'applications Web, mobiles et Cloud / SaaS / PaaS / DBaaS. Plus d'informations peuvent être trouvées à <https://www.mysql.com/>.

Comme montre la Figure. 5.7, le système CDDSS proposé évalue et classe les ressources numériques existantes sur le marché en fonction des requêtes de recherche des utilisateurs afin de sélectionner les ressources numériques appropriées à acquérir. Avant que le DSS sélectionne les ressources numériques, les chercheurs doivent fournir des alternatives, des poids de critères et des valeurs de critères (préférences) au DSS.

En utilisant le bouton « Import Data » dans le menu illustré dans la Figure. 5.7, un nouveau menu s'ouvre, il contient deux options :

- A. La première option « Import catalog » permet au CDDSS de :
 - Charger le fichier catalogue ;
 - Eliminer les données inutiles ;

Chapitre 5. DSS pour Gérer le Développement des Collections

- Stocker les alternatives dans la base de données relationnelle (MySQL).

B. La deuxième option « Import Log File » permet au CDDSS de :

- Charger le fichier Log ;
- Extraire les requêtes ;
- Eliminer les redondances des requêtes pour chaque utilisateur ;
- Sauvegarder les requêtes dans la base de données relationnelle ;
- Appliquer l'algorithme dans la figure 5.6 pour obtenir les préférences et construire la matrice de performance des alternatives, comme indiqué dans le Tableau. 5.2.



Figure 5.7 Capture de l'interface du CDDSS (Hemili et al., 2019)

Une fois que les alternatives et les critères ont été obtenus, le DSS évalue et classe les ressources numériques à l'aide de la méthode PROMETHEE II, puis génère un rapport, qui peut être analysé par les gestionnaires des bibliothèques numériques académiques, comme on montre dans la Figure 5.8.

Ce rapport contient des informations sur les ressources numériques, pour chaque type de ressources numériques (livres numériques, revues numériques, bases de données). Le DSS fournit une liste classée des plus importants aux moins importants, les gestionnaires peuvent utiliser pour sélectionner les ressources numériques les plus importantes à acquérir par les utilisateurs.

Ebook Title	Year	ISBN	Package Name
Anticandidal Agents	2017	978-0-12-811311-0	eBook - Biochemistry, Genetics and Molecular Biology 2016
Dust Explosion Dynamics	2017	978-0-12-803771-3	eBook - Chemical Engineering 2016
Adsorption-Driers for Divided Solids	2017	978-1-78548-179-6	eBook - Chemical Engineering 2016
Studies in Natural Products Chemistry	2016	978-0-444-63749-9	eBook - Chemistry 2016
Example-Based Super Resolution	2017	978-0-12-809703-8	eBook - Engineering 2016
Expanded PTFE Applications Handbook	2017	978-1-4377-7855-7	eBook - Engineering 2016
Science and Principles of Biodegradable and Bioresorbable Medical P...	2017	978-0-08-100372-5	eBook - Materials Science 2016
Bioinspired Materials for Medical Applications	2017	978-0-08-100741-9	eBook - Materials Science 2016
The Science of Armour Materials	2017	978-0-08-101002-0	eBook - Materials Science 2016
An Invitation to Applied Mathematics	2017	978-0-12-804153-6	eBook - Mathematics 2016
Pathways into Information Literacy and Communities of Practice	2017	978-0-08-100673-3	eBook - Social Sciences 2016
Kelley and Firestein's Textbook of Rheumatology (Tenth Edition)	2017	978-0-323-31696-5	eBook - Specialty Medicine 2016
Die Psychiatrische Notfallmedizin (Dritte Ausgabe)	2017	978-3-437-22881-0	eBook - German Medical Collection 2016

Figure 5.8 Le rapport principale du DSS (Hemili et al., 2019)

La Figure. 5.8 montre les ressources numériques classées des plus importants aux moins importants.

En plus du rapport principal, le CDDSS fournit les flux de classement positifs, négatifs et complets comme dans les colonnes 2, 3 et 4 dans la Figure. 5.9, ce qui représente une information très importante pouvant être utiliser comme une preuve dans la prise de décision par les gestionnaires.

Rank	action	Phi	Phi+	Phi-
1	Computational Finance	0,5786	0,7357	0,1571
2	An Invitation to Applied	0,1643	0,5429	0,3786
3	Time and Methods in	0,1500	0,5214	0,3714
4	Smart Buildings	-0,0143	0,4214	0,4357
5	Time-Frequency Signal	-0,1357	0,3286	0,4643
6	Discover Digital	-0,1786	0,2857	0,4643
7	Discover Digital	-0,2500	0,2214	0,4714
8	4G LTE-Advanced Pro	-0,3143	0,1786	0,4929

Figure 5.9 Table d'évaluation (Hemili et al., 2019)

Cependant, pour mettre la différence entre les ressources plus claire, le CDDSS offre une autre représentation des résultats, comme le montre la Figure. 5.10. Elle montre la différence entre chaque paire de ressources numériques, et donne une meilleure vue pour les gestionnaires et rend leurs tâches beaucoup plus faciles.

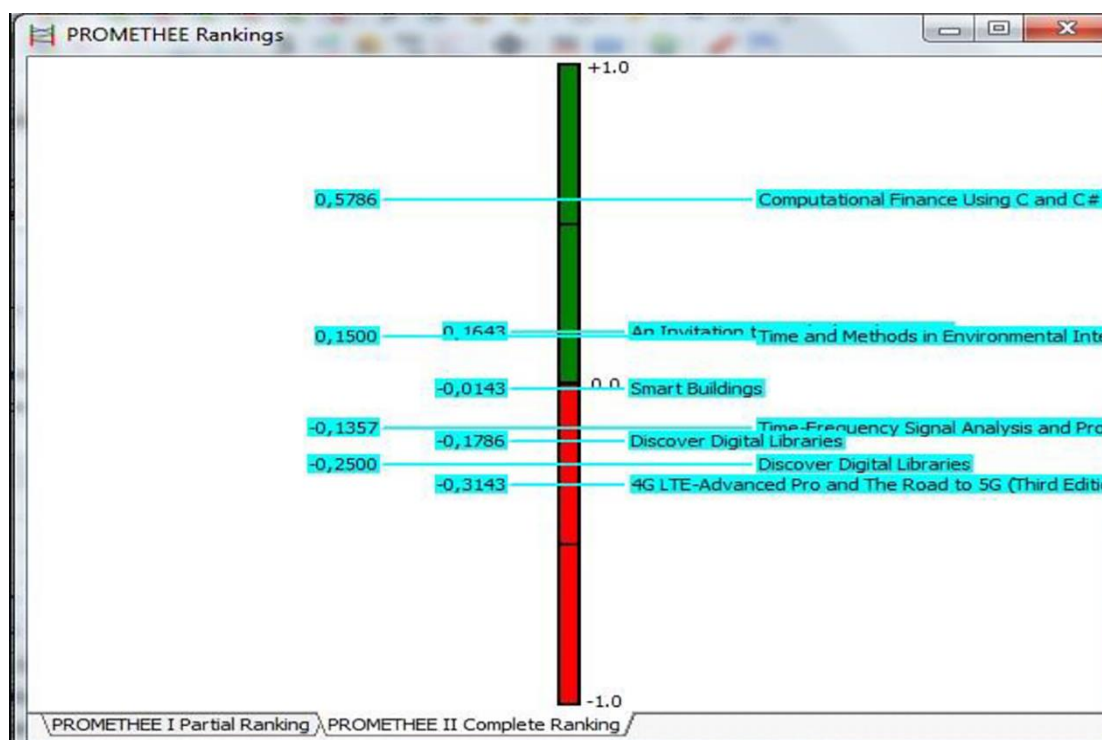


Figure 5.10 Classement totale (Hemili et al., 2019)

Comme le CDDSS évalue non seulement les livres numériques, mais également tous les types de ressources numériques telles que les journaux numériques, les bases de données et les packages dans des bases de données, le CDDSS offre de nombreuses représentations permettant aux gestionnaires de voir la différence entre les alternatives et de permettre de comparer les ressources numériques du même type (livres numériques, revues numériques, bases de données et packages dans des bases de données), ces représentations fournissent les informations appropriées sur les différences entre les ressources numériques. Les formules utilisées pour évaluer les packages sont les suivantes :

- **Evaluation des Packages**

Pour évaluer chaque package (P_j) nous utilisons la formule suivante :

$$P_j = \frac{\sum_{i=1}^{i=n} \Phi_{a_i}}{n} \quad (9)$$

P_j représente le niveau d'évaluation de la package j , et Φ_{a_i} représente l'évaluation de la ressource a_i de la package j .

- **Evaluation des Bases de Données**

Nous calculons l'évaluation de chaque base de données, en utilisant la formule suivante :

$$DB_j = \frac{\sum_{i=1}^{i=n} P_i}{n} \quad (10)$$

Chapitre 5. DSS pour Gérer le Développement des Collections

DB_j représente l'évaluation de la base de données (DB) j et P_i représente l'évaluation de la package i de la DB j .

Le système DSS offre aux gestionnaires la possibilité de les distinguer. Chaque base de données et chaque ensemble de bases de données peuvent être représentée par deux barres, voir Figure 5.11, la première représente le niveau d'évaluation des livres numériques et la seconde, le niveau d'évaluation des revues numériques. Les gestionnaires peuvent utiliser les informations de l'histogramme pour effectuer des comparaisons entre bases de données et packages, ainsi que pour choisir les ressources numériques nécessaires aux utilisateurs. Cet histogramme lui permet de déterminer le package dans lequel la base de données est la meilleure option est la plus importante pour les utilisateurs.

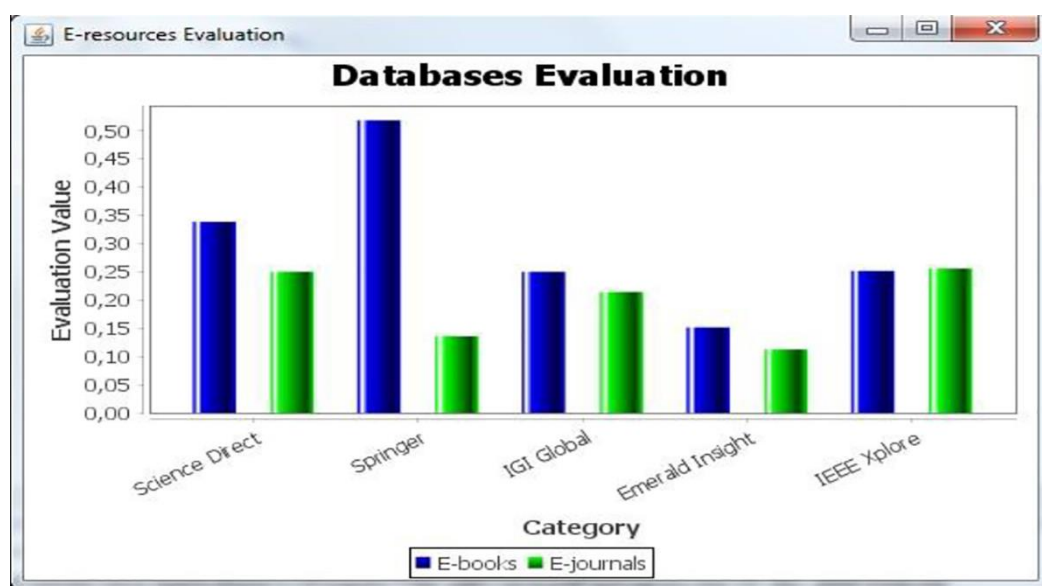


Figure 5.11 BarChart d'évaluation (Hemili et al., 2019)

Le CDDSS peut être utilisé pour évaluer les ressources numériques de collections de bibliothèques numériques académiques offrent aux gestionnaires une évaluation de leurs ressources numériques, l'évaluation peut aider les gestionnaires dans différentes tâches de gestion des collections, mentionnent les chercheurs:

- Renouvellement dans des revues numériques ou bases de données ou packages dans des bases de données ;
- Annulation d'abonnements aux ressources numériques s'ils ont une évaluation négative ;
- Remplacement des ressources numériques existantes dans une collection de bibliothèque numérique avec de meilleures ressources numériques qui n'existent pas dans la collection en supprimant les ressources numériques non utilisées et en achetant de nouvelles.

Les bibliothécaires seront en mesure d'accomplir ces tâches en prenant des décisions plus efficaces, ce qui sert l'objectif principal des bibliothèques numériques en général et des bibliothèques numériques universitaires en particulier, qui satisfait les utilisateurs en fournissant une collection riche.

En plus de satisfaire les utilisateurs des BNAs, le CDDSS a de nombreux effets sur les bibliothèques numériques académiques, entre autres, il contribue à réduire

l'inconfort des gestionnaires en les fournissant les informations nécessaires dont ils ont besoin et en améliorant la visibilité des besoins des utilisateurs pour les gestionnaires, ce qui les rend plus accessibles et les permet de travailler plus facilement. En outre, le CDDSS augmentera l'utilisation des ressources numériques par les utilisateurs et attirera de nouveaux utilisateurs qui dirigeront les bibliothèques numériques académiques pour en tirer profit et continuer à fonctionner.

La principale nouveauté de la CDDSS proposée dans ce chapitre est qu'elle sélectionne des ressources numériques en fonction des besoins en informations des utilisateurs afin d'améliorer la collection proposée par les bibliothèques numériques académiques afin d'augmenter le nombre d'utilisateurs, et de fournir aux bibliothécaires une vue globale sur le comportement et le changement du comportement dans la communauté des utilisateurs qui les permettent de mieux comprendre leurs utilisateurs et de réagir aux changements en prenant des décisions efficaces.

5.5 Data Warehouse pour CDDSS

Dans cette section, nous proposons une architecture d'un Data Warehouse, voir Figure 5.12. Elle nous permet d'intégrer des données viennent de sources de données opérationnelles hétérogènes indépendantes et de créer une vue unique sur la bibliothèque numérique.

5.5.1. Les composants de l'architecture proposée

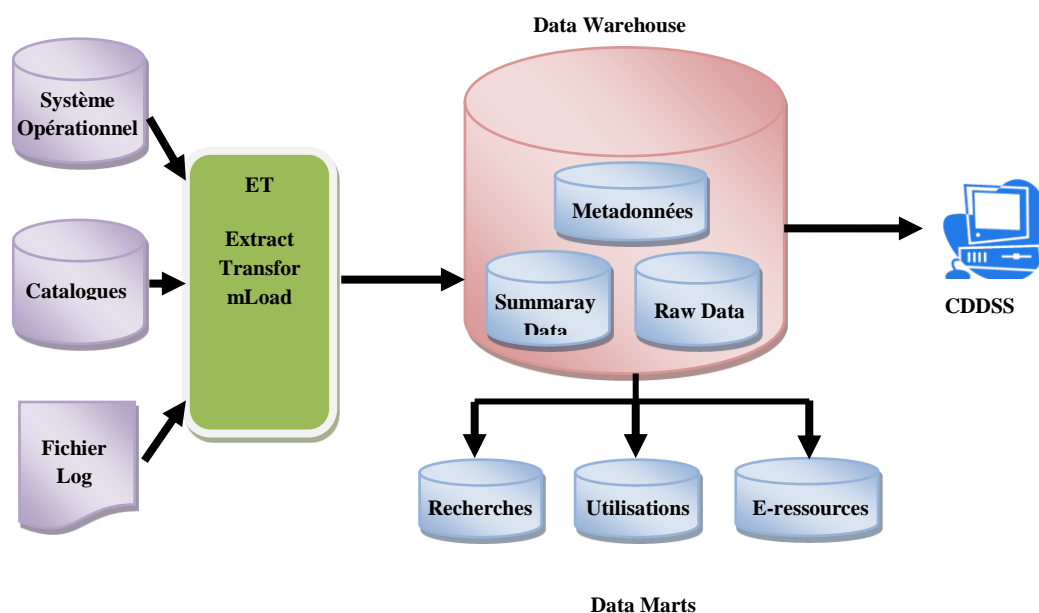


Figure 5.12 Data Warehouse pour CDDSS (Hemili et al., 2019)

Dans ce qui suit nous expliquons les composants de notre architecture :

- Data Marts : Un Data Mart est un élément important du Data Warehouse. Il peut être considéré comme le sous-ensemble d'un Data Warehouse axé sur une ligne d'activité particulière comme les recherches, les ventes, le marketing, etc., voir Figure 5.13.

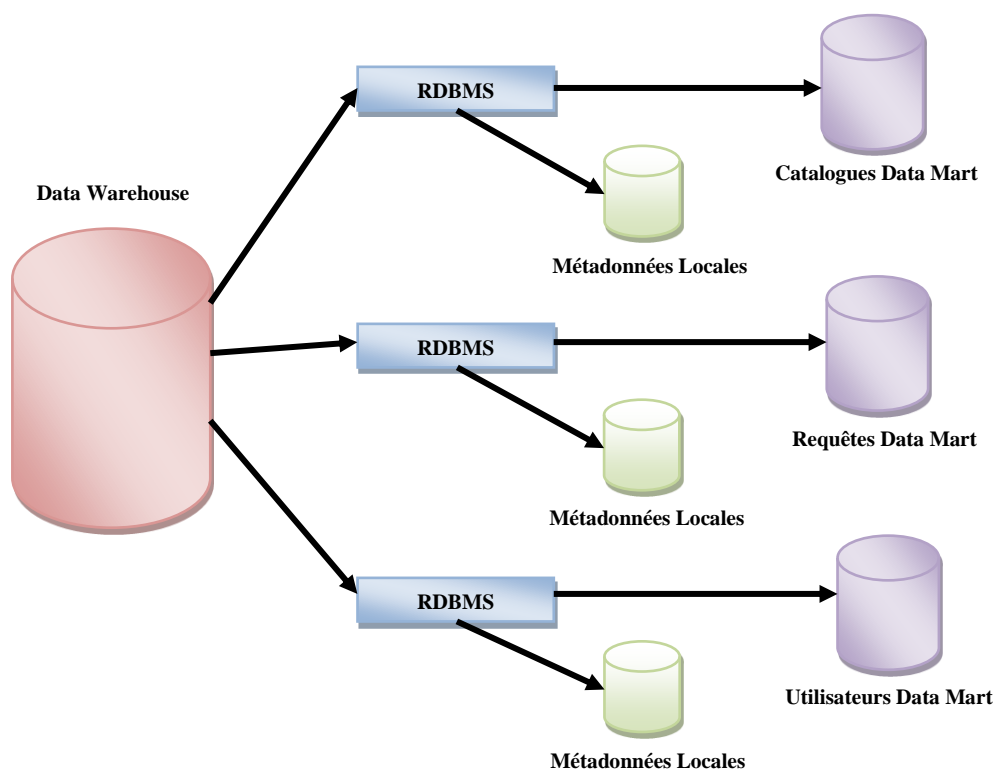


Figure 5.13 Architecture des Data Marts

- CDDSS : Collection Development Decision Support System est le système décisionnel qui exploite les données du Data Warehouse.
- ETL : Le processus ETL aide à collecter les données de différentes sources dans un seul Data Warehouse. Le processus de transformation, d'extraction et de chargement est utilisé pour construire le Data Warehouse.
- Métadonnées : Les données sur des données sont appelées métadonnées. Ils aident à obtenir les informations sur les données. Par exemple, si nous avons une bibliothèque numérique, les données sur utilisateurs et les données sur les ressources numériques peuvent être appelées des métadonnées.

5.5.2. Schéma des données

La modélisation des données de notre Data Warehouse est illustrée dans la Figure 5.14.

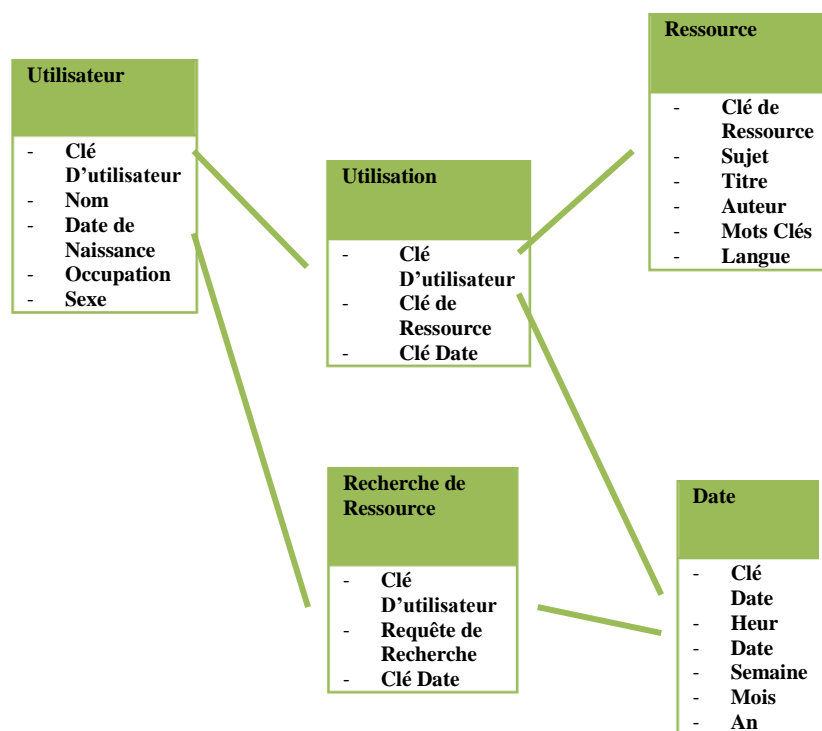


Figure 5.14 Schéma des Données dans le Data Warehouse

5.6 Proposition d'une Architecture Cloud pour le CDDSS développé

Dans le but de répondre à la problématique de la puissance de calcul on a proposé une architecture d'un système décisionnel basé Cloud. Dans la partie suivante, nous présentons les différentes composantes de cette architecture afin de gérer le développement des collections dans les bibliothèques numériques académiques. L'architecture du DSS basé Cloud proposée est illustrée par la Figure 5.15.

5.6.1. Logiciel de gestion Cloud

Dans cette sous-section nous présentons Amazon Elastic Compute Cloud (Amazon EC2) pour l'utiliser par la suite comme un logiciel de gestion de services Cloud pour expliquer le déploiement du DSS et du Data Warehouse. Les composantes principales d'Amazon EC2 sont:

- Une Amazon Machine Image (AMI) est une template qui contient une configuration logicielle (par exemple, un système d'exploitation, un serveur d'applications et des applications). À partir d'un AMI, nous pouvons lancer une instance (VM) ou plusieurs instances (VMs)
- Amazon Elastic Block Store (Amazon EBS) fournit des volumes de stockage de niveau bloc à utiliser avec les instances EC2.
- An Elastic IP address nous permet d'obtenir une adresse IPv4 statique pour notre compte AWS.

Chapitre 5. DSS pour Gérer le Développement des Collections

AWS propose plusieurs outils et services pour permettre le déploiement d'applications Java gérées par AWS et gérées par le client. Le tableau 5.3 est une référence de haut niveau pour aider à identifier l'option la plus appropriée pour un scénario spécifique.

Caractéristiques d'application	Outils d'emballage	Mécanisme de déploiement	Méthode / environnement de déploiement
Applications Java personnalisées développées dans Eclipse	Éclipse	Déploiement en un seul clic depuis Eclipse	AWS Toolkit for Eclipse
Applications Web Java déployées en tant que fichier JAR, WAR ou ZIP et nécessitant un minimum de modifications du système d'exploitation	JAR, WAR ou ZIP	Déploiement automatisé d'applications packagées à l'aide d'AWS Elastic Beanstalk	AWS Elastic Beanstalk
Toute application Java ou configuration de serveur, en particulier celles nécessitant un système d'exploitation personnalisé ou des installateurs tiers	Installateurs personnalisés existants, archive d'application (JAR, WAR, ZIP), copie manuelle de fichiers, etc.	Outils et processus de déploiement de logiciels existants ou services de déploiement automatisé, tels que AWS CodeDeploy ou AWS OpsWorks.	Instances EC2

Tableau 5.3 Services de Déploiement d'une Application Java (Amazon.com, 2019a)

5.6.2. Les composantes de l'architecture Cloud

La Figure 5.15 présente les principales composantes de l'architecture proposée : le modèle de déploiement de notre DSS dans le Cloud, la couche infrastructure, la couche plateforme et la couche software.

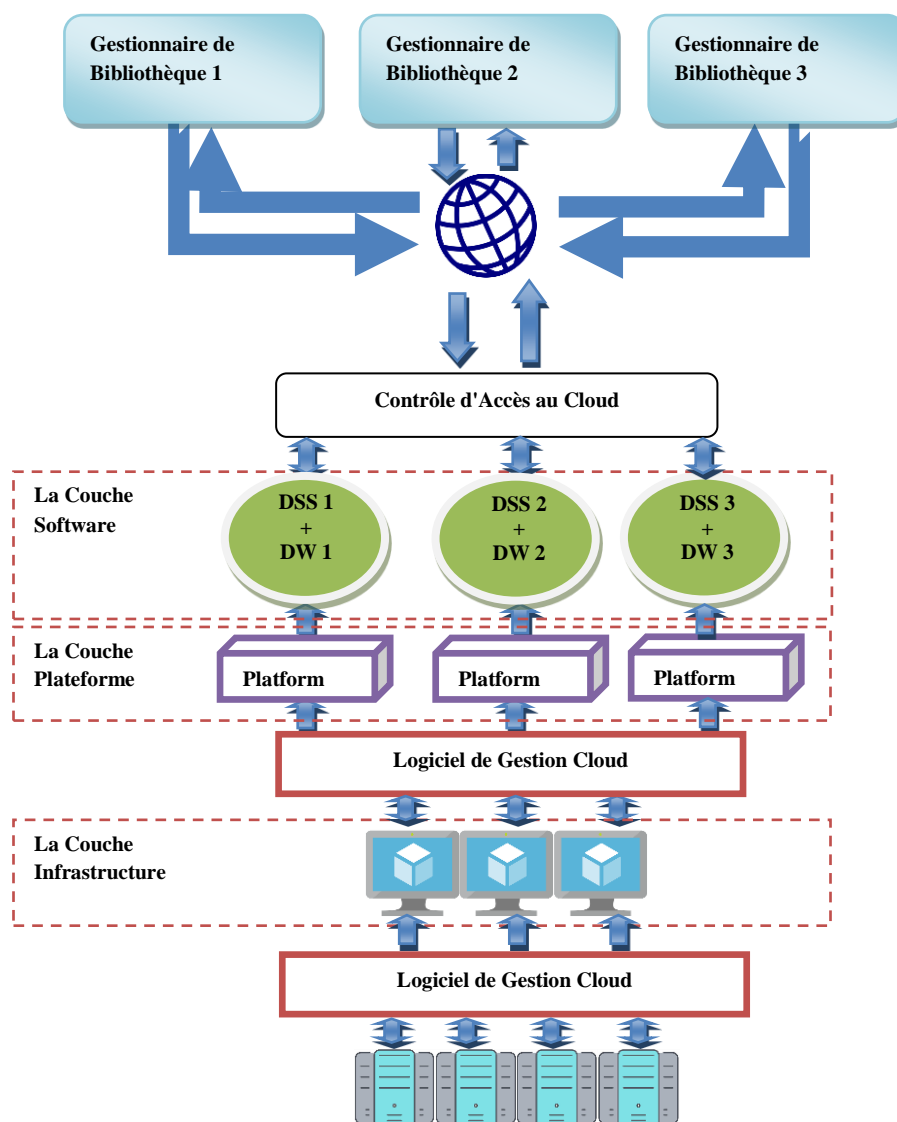


Figure 5.15 Architecture du DSS basé Cloud pour les Bibliothèques Numériques

- **Modèle de déploiement du DSS dans Cloud**

Comme nous avons expliqué précédemment, Le Cloud Computing offre différents modèles de déploiement de services. Ces modèles permettent de définir le degré d'accès d'un utilisateur aux fournisseurs de services Cloud.

Notre système d'aide à la décision proposé est destiné aux gestionnaires des collections des bibliothèques numériques, et il traite des données qui contiennent des informations privées sur les utilisateurs des bibliothèques numériques. Donc, nous avons choisi d'intégrer notre DSS dans un Cloud privé, pour protéger les informations privées des utilisateurs et protéger les données des bibliothèques numériques. En déployant notre DSS dans un Cloud privé le contrôle d'accès au Cloud autorise l'accès au DSS uniquement pour les gestionnaires.

- **La couche infrastructure**

Un fournisseur de Cloud héberge les composants d'infrastructure traditionnellement présents dans un centre de données local, y compris les serveurs, le matériel de stockage et de réseau.

La couche infrastructure fourni aux utilisateurs du Cloud des ressources virtuelles, ces ressources possèdent une grande puissance de calcul, de traitement et de stockage. La Figure 5.16 illustre les étapes de créer une VM en utilisant Amazon EC2.

Dans notre cas, cette couche a été utilisée pour créer une machine virtuelle pour chaque bibliothèque numérique. Nous allons utiliser chaque machine virtuelle pour déployer un système décisionnel d'une bibliothèque numérique et un Data Warehouse.

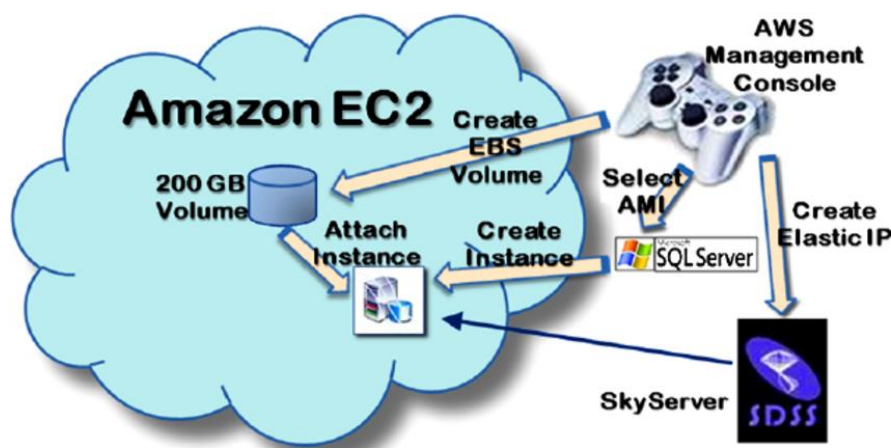


Figure 5.16 Les Etapes de Créer une VM en Utilisant Amazon EC2 (Thakar, Szalay, Church, & Terzis, 2011)

- **La couche plateforme**

Cette couche fournit une plate-forme qui nous permettant d'exécuter et de gérer notre DSS sans la complexité de la construction et de la maintenance infrastructure généralement associée au développement et au lancement d'une application. La couche plateforme repose sur la couche infrastructure. Elle utilise toutes les ressources fournies de la machine virtuelle comme la puissance de calcul.

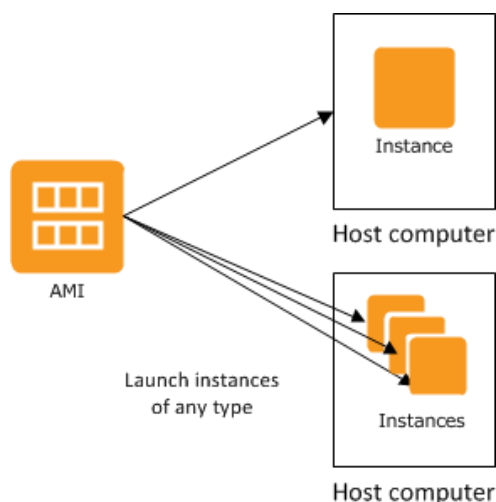


Figure 5.17 Amazon AMI permet de Créer des VMs avec plateformes (amazon.com, 2019b)

Pour déployer notre DSS dans un Cloud privé nous allons utiliser le modèle de livraison de service PaaS. Ce modèle nous permet de créer une plateforme personnalisée qui nous permet de déployer notre DSS de façon efficace.

Cette plateforme contient les outils nécessaires pour l'exécution de notre DSS tels que : un système d'exploitation qui nous permet d'utiliser la machine virtuelle, un système de gestion de base de données comme MySQL qui permet au DSS de stocker les données et de gérer ces données, une machine virtuelle Java (JVM) qui exécute des programmes compilés sous forme de bytecode Java. AMI d'Amazon permet de créer des plateformes spécialisés (Instances), voir Figure 5.17.

- **La couche software**

Cette couche contient le DSS proposé, et elle permet aux gestionnaires des bibliothèques numériques académiques d'accéder à notre système d'aide à la décision. Les gestionnaires accèdent au DSS via l'Internet, et ils l'utilisent pour accomplir les différentes tâches de développement de collection, en profitant de temps de réponse minimisé du Cloud et de prix réduit par rapport à l'achat des nouveaux serveurs.

5.7 Conclusion

Dans ce chapitre, nous avons présenté notre DSS pour la gestion de développement des collections des bibliothèques numériques académiques, et nous avons proposé une architecture Data Warehouse pour l'intégration des données, et une architecture Cloud pour le déploiement du DSS et du Data Warehouse.

Donc, Nous avons également traité le problème de choix de ressources pour aider les gestionnaires des BNAs dans les différentes tâches de gestion des collections. Par conséquent, l'objectif principal du DSS proposée est de satisfaire les besoins d'information des utilisateurs des BNAs tout en tenant compte du profit économique des BNAs.

Ainsi, nous avons répondu aux problèmes d'hétérogénéité des données et de puissances de calcul par la proposition d'une architecture Data Warehouse et une architecture Cloud.

Chapitre 5. DSS pour Gérer le Développement des Collections

Enfin, nous avons comme perspective l'implémentation de notre architecture Cloud proposée dans un Cloud privé. Pour accomplir ce but là nous allons réaliser deux comparaisons. Une comparaison entre les systèmes de gestion du Cloud comme Apache CloudStack, OpenStack, Rackspace Cloud Management et Amazon EC2 afin de choisir le système de gestion du Cloud qui nous permet de réaliser notre architecture efficacement. L'autre comparaison entre les fournisseurs de services Cloud comme Amazon, Google et Microsoft pour choisir le fournisseur le plus convenable aux bibliothèques numériques par rapport à leurs système de gestion du Cloud utilisé, le prix , la sécurité et autre critères.

Chapitre 6 Conclusion Générale

Contenu

6.1	<i>Contributions</i>	103
6.2	<i>Perspectives</i>	104

Chapitre 6. Conclusion Générale

Dans ce dernier chapitre de cette thèse, nous allons présenter les contributions apportées durant cette thèse et des perspectives dans notre domaine de recherche.

6.1 Contributions

Dans cette thèse, nous avons abordé le problème de l'évaluation et l'amélioration des collections dans bibliothèques numériques académiques (BNAs). Pour résoudre ce problème, nous avons proposé un système d'aide à la décision (DSS) pour l'évaluation des ressources numériques et des collections, et plus précisément, sélectionner les ressources numériques pour les différentes tâches de développement et de gestion des collections. Ce système vise à permettre les gestionnaires des BNAs de satisfaire les besoins d'information des utilisateurs ainsi que le profit économique des BNAs.

Un certain nombre de contributions majeurs ont été réalisés pendant la durée de cette thèse. Ces contributions peuvent être résumées comme suit:

- Nous avons proposé un système d'aide à la décision pour l'évaluation et sélection des ressources numériques (Livre numériques, Journal numériques et base de données). Ce système permet aux gestionnaires des BNAs de prendre efficacement des décisions dans les différentes activités de développement et de gestion des collections, tel que l'acquisition, l'abonnement ou le désabonnement des ressources numériques.
- Dans l'approche proposée, nous avons traité le problème de la détermination des besoins d'information des utilisateurs. Pour résoudre ce problème nous avons proposé une analyse des requêtes envoyées par les utilisateurs via le moteur de recherche de la BNA. Le résultat de cette analyse a été utilisé pour construire la table de perspectives des utilisateurs, ce tableau contient les besoins d'information des utilisateurs.
- Nous avons proposé l'utilisation de la méthode d'aide à la décision multicritère PROMETHE II, cette méthode nous permet d'évaluer les ressources numériques en comparant chaque ressource avec les autres, ces comparaisons assurent une très bonne évaluation et classement des ressources numériques. Par conséquent, les gestionnaires ont des informations importantes pour performer une bon développement ou gestion des collections des BNAs.
- L'utilisation de la méthode SIMOS révisé qui permet aux gestionnaires de déterminer les poids des critères de façons très simple, claire et efficace.
- Le DSS proposée a été mise en œuvre par un prototype logiciel. L'outil développé offre diverses fonctionnalités très utiles aux gestionnaires des BNAs. Le système fourni aux gestionnaires des pleines d'informations sur les ressources numériques de la collection. Le prototype offre des différentes représentations (Liste classé, Bar Chart ... etc) de l'information qui le permet de passer l'information clairement.
- En termes d'évaluation, le DSS proposé permet de déterminer les besoins des utilisateurs d'une manière plus efficace, et fournir ces informations aux gestionnaires de façon très claire.
- Nous avons proposé une architecture d'un Data Warehouse pour l'intégration des données.

Chapitre 6. Conclusion Générale

- Nous avons proposé une architecture Cloud pour le déploiement du DSS dans un Cloud.

6.2 Perspectives

Nous présentons ci-dessous des nouvelles directions de recherches futures dans notre domaine de recherche

- Les méthodes d'aide à la décision multicritère comme la méthode PROMETHEE II ont besoin beaucoup de calculs pour comparer un grand nombre des alternatives. Ce besoin exige une puissance de calcul élevé. Pour répondre à ce problème nous avons proposé une architecture du DSS basé Cloud, et comme des travaux futurs nous allons implémenter l'architecture proposée.
- Les gestionnaires des BNAs ont besoin de comprendre le comportement des utilisateurs (leurs orientations actuelles et futures), pour réagir rapidement avec les besoins croissant de ces utilisateurs. Data Warehousing offre l'enregistrement des données des utilisateurs de façon historique, ces données ne sera jamais mis à jour. Par conséquent, Data Warehousing nous offre les données nécessaires pour comprendre le comportement des utilisateurs. On a proposé une architecture Data Warehouse pour notre CDDSS, dans les prochains travaux on va essayer utiliser Data Warehousing pour intégrer tous les données de gestion des bibliothèques numériques dans un Data Warehouse pour améliorer la gestion des BNA.

Liste des publications et communications

Les publications et les communications qui ont été produites tous au long de cette thèse sont les suivantes :

Revue internationale

Hemili, M., Laouar, M. R., & Eom, S. (2019). A Decision Support System for Managing Demand-Driven Collection Development in University Digital Libraries. *International Journal of Information Systems and Social Change (IJSSC)*, 10(4).

Conférences internationales

1. Hemili, M., & Laouar, M. R. (2018, March). E-libraries decision support system for e-resources selection. In *Proceedings of the 8th International Conference on Information Systems and Technologies* (p. 12). ACM.

2. Hemili, M., & Laouar, M. R. (2018, October). Use of Multi-criteria decision analysis to make collection management decisions. In *2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)* (pp. 1-5). IEEE.

Synposium

1ER SYNPOSIUM DES DOCTORANTS EN INFORMATIQUE TUNISO-ALGERIENS 2017 (SDTA 2017), Tunisie, Hammamet du 24 au 26 avril 2017

Workshops

1. HEMILI, Mohamed. Getting started with big data tools. Doctoral students in computer science : workshops and practices, April 15- 19, 2018.

2. HEMILI, Mohamed. How computer science is revolutionizing medicine? a Cambridge university experience. Doctoral students in computer science : workshops and practices, Tebessa, April 15- 19, 2018.

3. HEMILI, Mohamed. Deep learning : principles, techniques and tools. Doctoral students in computer science : workshops and practices, Tebessa, April 15- 19, 2018.

4. HEMILI, Mohamed. Scientific paper : how to succeed in redaction and publication? Doctoral students in computer science : workshops and practices, Tebessa, April 15- 19, 2018

5. HEMILI, Mohamed. The Islamic golden age and its impact in science and technologies. Doctoral students in computer science : workshops and practices, Tebessa, April 15- 19, 2018

6. HEMILI, Mohamed. The Smart -environments; from theory to practice. Doctoral students in computer science : workshops and practices, Tebessa, April 15- 19, 2018

Bibliographie

- Abdolmohammadi, M., & Usoff, C. (2001). A longitudinal study of applicable decision aids for detailed tasks in a financial audit. *Intelligent Systems in Accounting, Finance & Management*, 10(3), 139-154.
- Ahmad, M., & Abawajy, J. H. (2014). Digital library service quality assessment model. *Procedia-social and behavioral sciences*, 129, 571-580.
- Al-Natsheh, H. T., Martinet, L., Muhlenbach, F., Rico, F., & Zighed, D. A. (2018). Metadata Enrichment of Multi-disciplinary Digital Library: A Semantic-Based Approach. Paper presented at the International Conference on Theory and Practice of Digital Libraries.
- Amazon.com. (2019a). Deploying Java Web Applications. Retrieved from <https://aws.amazon.com/answers/web-applications/aws-web-app-deployment-java/>
- amazon.com. (2019b). Instances and AMIs. Retrieved from <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-instances-and-amis.html>
- Anbu, K., Kataria, S., & Ram, S. (2013). Dynamics of managing electronic resources: Electronic resource management system (ERMS) initiatives.
- Arms, W. Y. (2001). *Digital libraries*: MIT press.
- Armstrong, C., & Lonsdale, R. (2005). Challenges in managing e-books collections in UK academic libraries. *Library Collections, Acquisitions, and Technical Services*, 29(1), 33-50.
- Ashcroft, L. (2011). Ebooks in libraries: an overview of the current situation. *Library Management*, 32(6/7), 398-407.
- Association, A. L. (2019). Universal Gift Form and Instructions. *College & Research Libraries News*, 36(3), 95-96.
- Bamiah, M. A. (2015). Trusted Cloud Computing Frame Work in Critical Industrial Application. *Universiti Teknologi Malaysia*.
- Benatia, I., Laouar, M. R., Bendjenna, H., & Eom, S. B. (2016). Implementing a cloud-based decision support system in a private cloud: The infrastructure and the deployment process. *International Journal of Decision Support System Technology (IJDSST)*, 8(1), 25-42.
- Bendib, I., Ridda Laouar, M., Hacken, R., & Miles, M. (2014). Semantic ontologies for multimedia indexing (SOMI) Application in the e-library domain. *Library Hi Tech*, 32(2), 206-218.
- Bhadoria, R. S., Chaudhari, N. S., & Vidanagama, V. T. N. (2018). Analyzing the role of interfaces in enterprise service bus: a middleware epitome for service-oriented systems. *Computer Standards & Interfaces*, 55, 146-155.
- Board, A. L., & Tafuri, N. (2017). Prices of US and Foreign Published Materials 2017.
- Bouyssou, D. (1993). Décision multicritère ou aide multicritère. Newsletter of the European Working Group—Multicriteria Aid for Decisions, 1-2.
- Bowker. (2012, June 5, 2012). Publishing Market Shows Steady Title Growth in 2011 Fueled Largely by Self-Publishing Sector. Retrieved from <http://www.bowker.com/tools-resources/Bowker-Data.html>

Bibliographie

- Brans, J.-P., & Vincke, P. (1985). Note—A Preference Ranking Organisation Method: (The PROMETHEE Method for Multiple Criteria Decision-Making). *Management science*, 31(6), 647-656.
- Brans, J.-P., Vincke, P., & Mareschal, B. (1986). How to select and how to rank projects: The PROMETHEE method. *European journal of operational research*, 24(2), 228-238.
- Bushing, M. C. (2006). *Collection mapping-An Evolving Tool for Better Resources and Better Access*. Signum.
- Cabrerizo, F. J., Morente-Molinera, J. A., Pérez, I. J., López-Gijón, J., & Herrera-Viedma, E. (2015). A decision support system to develop a quality management in academic digital libraries. *Information Sciences*, 323, 48-58.
- Carrigan, D. P. (1996). Collection development—evaluation. *The Journal of academic librarianship*, 22(4), 273-278.
- Carroll, D., & Cummings, J. (2010). Data driven collection assessment using a serial decision database. *Serials Review*, 36(4), 227-239.
- Chao, L., Zhang, Y., & Xing, C. (2011). Improving query suggestion for digital libraries. Paper presented at the Computer Software and Applications Conference Workshops (COMPSACW), 2011 IEEE 35th Annual.
- Charvériat, L. (2008). A la recherche de traces informatiques... Aspects pratiques et juridiques. Retrieved from <https://www.journaldunet.com/solutions/expert/24676/a-la-recherche-de-traces-informatiques--aspects-pratiques-et-juridiques.shtml>
- Cleveland, G. (1998). Digital libraries: definitions, issues and challenges: IFLA, Universal dataflow and telecommunications core programme.
- Curtis, M. B., & Joshi, K. (1999). Developing a Data Warehouse: Some Guidelines and Suggestions. *Review of Business Information Systems (RBIS)*, 3(2), 13-26.
- Dadzie, P. S. (2005). Electronic resources: access and usage at Ashesi University College. *Campus-Wide Information Systems*, 22(5), 290-297.
- Database.com, P. (2013). Top Physics and Science Books of 2012.
- Delgado, T., González, G., Miranda, G., Navarro, D. G., & Graverán, A. (2013). Context-aware spatial decision support systems (ca-sdss): Articulating decision support systems, business intelligence and recommender systems considering the geospatial component. Paper presented at the Fourth International Workshop on Knowledge Discovery, Knowledge Management and Decision Support.
- Deutch, D., & Milo, T. (2012). Type inference and type checking for queries over execution traces. *The VLDB Journal—The International Journal on Very Large Data Bases*, 21(1), 51-68.
- Dewan, P. (2012). Are books becoming extinct in academic libraries? *New library world*, 113(1/2), 27-37.
- Disher, W. (2014). *Crash course in collection development: ABC-CLIO*.
- DLF. (1998). A working definition of digital library. Retrieved from <https://old.diglib.org/about/dldefinition.htm>
- Egan, N., Yearwood, S. L., & Kendrick, C. L. (2016). Patron-driven acquisitions at the City University of New York: A case study. *Technical Services Quarterly*, 33(2), 131-144.
- elsevier.com. (2019). Search Results.
- ENSSIB. (2012). Bibliothèques numériques. Retrieved from <https://www.enssib.fr/le-dictionnaire/bibliotheques-numeriques>

Bibliographie

- Eom, H. B., Lee, S. M., Snyder, C. A., & Ford, N. F. (1987-1988). A Multiple Criteria Decision Support System for Global Financial Planning. *Journal of Management Information Systems*, 4(3), 94-113.
- Eom, S. (2016). Longitudinal Author Cocitation Mapping: The Changing Structure of Decision Support Systems Research (1969–2012). *Foundations and Trends® in Information Systems*, 1(4), 277-384.
- Eom, S., & Laouar, M. R. (2017). Effects of Interaction on E-Learning Satisfaction and Outcome: A Review of Empirical Research and Future Research Direction. *International Journal of Information Systems and Social Change (IJISSC)*, 8(3), 58-71.
- Eppe, S., & De Smet, Y. (2014). Approximating Promethee II's net flow scores by piecewise linear value functions. *European journal of operational research*, 233(3), 651-659.
- Eschenfelder, K. R. (2008). Every library's nightmare? Digital rights management, use restrictions, and licensed scholarly digital resources. *College & Research Libraries*, 69(3), 205-226.
- Figueira, J., & Roy, B. (2002). Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European journal of operational research*, 139(2), 317-326.
- Figueira, J., & Roy, B. (2002). Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, 139(2), 317-326.
- Flatley, R., & Prock, K. (2009). E-resource collection development: A survey of current practices in academic libraries. *Library Philosophy and Practice (e-journal)*.
- Flynn, B. L. (1993). Developing Media Collections to Serve Multicultural/Multiethnic Communities. *The Acquisitions Librarian*, 5(9-10), 29-40.
- Fomba, S. (2018). Décision multicritère: un système de recommandation pour le choix de l'opérateur d'agrégation. Toulouse 1.
- Fox, E. A. (2016). Introduction to digital libraries. Paper presented at the Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries.
- Fuhr, N., Hansen, P., Mabe, M., Micsik, A., & Sølvsberg, I. (2001). Digital libraries: A generic classification and evaluation scheme. Paper presented at the International Conference on Theory and Practice of Digital Libraries.
- Gorman, G. E., Clayton, P. R., Shep, S. J., & Clayton, A. (2005). *Qualitative research for the information professional: A practical handbook*: Facet Publishing.
- Gorry, G. A., & Scott Morton, M. S. (1971). *A framework for management information systems*.
- Graf, R., Ryan, H. M., Houzanme, T., & Gordea, S. (2017). A decision support system to facilitate file format selection for digital preservation. *Libellarium: journal for the research of writing, books, and cultural heritage institutions*, 9(2).
- Greco, S., Figueira, J., & Ehrgott, M. (2005). *Multiple criteria decision analysis*. Springer's International series.
- Gutwin, C., Paynter, G., Witten, I., Nevill-Manning, C., & Frank, E. (1999). Improving browsing in digital libraries with keyphrase indexes. *Decision Support Systems*, 27(1-2), 81-104.
- Han, H., Jeong, W., & Wolfram, D. (2014). Log analysis of academic digital library: user query patterns. *iConference 2014 Proceedings*.
- Han, J., Kamber, M., & Pei, J. (2011). *Data mining concepts and techniques third edition*. The Morgan Kaufmann Series in Data Management Systems, 83-124.

Bibliographie

- Hemili, M., & Laouar, M. R. (2018a). e-libraries decision support system for e-resources selection. Paper presented at the Proceedings of the 8th International Conference on Information Systems and Technologies.
- Hemili, M., & Laouar, M. R. (2018b). Use of Multi-criteria decision analysis to make collection management decisions. Paper presented at the 2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS).
- Hemili, M., Laouar, M. R., & Eom, S. B. (2019). A Decision Support System for Managing Demand-Driven Collection Development in University Digital Libraries. *International Journal of Information Systems and Social Change (IJSSC)*, 10(4), 57-74.
- Henry, E., Longstaff, R., & Van Kampen, D. (2008). Collection analysis outcomes in an academic library. *Collection Building*, 27(3), 113-117.
- Hightower, B. E., & Gantt, J. T. (2012). Weeding nursing e-books in an academic library. *Library Collections, Acquisitions, and Technical Services*, 36(1-2), 53-57.
- Himma, K. E., & Just, P. (2007). Electronic books in the USA—their numbers and development and a comparison to Germany. *Library Hi Tech*.
- Hobohm, H.-C. (2011). Knowledge management: Libraries and librarians taking up the challenge (Vol. 108): Walter de Gruyter.
- Hu, J., & Yu, B. (2019). Analysis on User Behavior and Needs of University Libraries. *ieeexplore.ieee.org*. (2019). Search Results.
- igi-global.com. (2019). Search Results.
- Inmon, W. H. (2005). *Building the data warehouse*: John Wiley & Sons.
- Intner, S. S., & Futas, E. (1994). Evaluating public library collections: Why do it, and how to use the results. *American Libraries*, 25(5), 410-412.
- Jansen, B. J., Booth, D. L., & Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. *Information Processing & Management*, 44(3), 1251-1266.
- Jansen, B. J., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing & Management*, 36(2), 207-227.
- Jareonruen, Y., & Tuamsuk, K. (2019). Lifecycle and Requirements for Digital Collection Management of Thai Theses and Dissertations. *Journal of Information Science Theory and Practice*, 7(3), 52-64.
- Jelassi, M. T., Jarke, M. A., & Stohr, E. A. (1985). Designing a Generalized Multiple Criteria Decision Support System. *Journal of Management Information Systems*, 1(4), 24-43.
- Johnson, P. (2014). *Fundamentals of collection development and management*: American Library Association.
- Johnson, P. (2018). *Fundamentals of collection development and management*: American Library Association.
- Jones, S., Cunningham, S. J., McNab, R., & Boddie, S. (2000). A transaction log analysis of a digital library. *International Journal on Digital Libraries*, 3(2), 152-169.
- Kadir, R., Dollah, W., Saaid, F. A., & Diljit, S. (2009). A user-based measure in evaluating academic digital library. Paper presented at the International Conference on Academic Libraries.
- Kadir, R. A., Dollah, W. A., Saaid, F. A., & Diljit, S. (2009). A user-based measure in evaluating academic digital library.

Bibliographie

- Kang, J.-H., Kim, C.-S., & Ko, E.-J. (2004). An XQuery engine for digital library systems that support XML data. Paper presented at the Applications and the Internet Workshops, 2004. SAINT 2004 Workshops. 2004 International Symposium on.
- Kao, S.-C., Chang, H.-C., & Lin, C.-H. (2003). Decision support for the academic library acquisition budget allocation via circulation database mining. *Information Processing & Management*, 39(1), 133-147.
- Kassim, N. A. (2017). Evaluating users' satisfaction on academic library performance. *Malaysian Journal of Library & Information Science*, 14(2), 101-115.
- Kavulya, J. (2015). The Emerging role of Digital Libraries in E-learning: A Review of challenges and opportunities.
- Keen, P. G. (1978). *Decision support systems; an organizational perspective*. Retrieved from
- Kimball, R., & Ross, M. (2011). *The data warehouse toolkit: the complete guide to dimensional modeling*: John Wiley & Sons.
- Kimball, R., & Ross, M. (2015). *The kimball group reader: Relentlessly practical tools for data warehousing and business intelligence remastered collection*: John Wiley & Sons.
- Koenig, M. (1985). Data relationships: bibliographic information retrieval systems and database management systems. *Information technology and libraries*, 4(3), 247-272.
- Kristi Ekonen, P. P., Pentti Vattulainen. (2006). *Handbook on the International Exchange of Publications (5th ed ed.)*: Munich: G. K. Saur.
- Kumar, P. A. (2017). Impact of information technology on the collection development in university libraries of assam a study.
- Kumar, V., Laghari, A. A., Karim, S., Shakir, M., & Brohi, A. A. (2019). Comparison of fog computing & cloud computing. *Int. J. Math. Sci. Comput*, 1, 31-41.
- Kumbhar, S. S., & Harake, S. B. (2015). *Digital Library Management: Challenges and Opportunities*.
- Lai, M.-C., Wang, W.-K., Huang, H.-C., & Kao, M.-C. (2011). Linking the benchmarking tool to a knowledge-based system for performance improvement. *Expert Systems with Applications*, 38(8), 10579-10586.
- Lakos, A. (1999). The missing ingredient: culture of assessment in libraries. *Performance Measurement and Metrics: The International Journal for Library and Information Services*, 1(1), 5.
- Larson, J. (2012). *CREW: A weeding manual for modern libraries*: Texas State Library and Archives Commission.
- Lathoud, B. (2002). *Gestion formelle de la trace informatique: un moyen de renforcer la politique de sécurité d'un système d'information*. Université de Lausanne, Faculté des hautes études commerciales.
- Lehman, K. A. (2014). Collection development and management: An overview of the literature, 2011-12. *Library Resources & Technical Services*, 58(3), 169-177.
- Lgimol, B. (2015). Selection and acquisition of e-resources in academic libraries. *International Journal of Digital Library Services*.
- Li, Y., & Liu, C. (2019). Information Resource, Interface, and Tasks as User Interaction Components for Digital Library Evaluation. *Information Processing & Management*, 56(3), 704-720.
- Lonsdale, R., & Armstrong, C. (2001). Electronic books: Challenges for academic libraries. *Library Hi Tech*, 19(4), 332-339.

Bibliographie

- Magdalini, V., Richard, H., & Jennifer, R. (2012). Choosing e-books: a perspective from academic libraries. Online information review. Online Information Review.
- Manley, W. (1996). The Manley Arts: If I called this column 'weeding,' you wouldn't read it. *Booklist*, 92(13), 1108.
- Marangozova-Martin, V., & Pagano, G. (2013). Gestion de traces d'exécution pour le systèmes embarqués: contenu et stockage.
- McClure, J. Z. (2009). Collection assessment through WorldCat. *Collection Management*, 34(2), 79-93.
- Meier, J. J. (2016). The future of academic libraries: conversations with today's leaders about tomorrow. *portal: Libraries and the Academy*, 16(2), 263-288.
- Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.
- Mohamed, H., Mohamed, R., Laouar, & Sean, E., B. (2019). A Decision Support System for Managing Demand-Driven Collection Development in University Digital Libraries. *International Journal of Information Systems and Social Change (IJSSC)*, 10(4).
- Mosqueira-Rey, E., & Moret-Bonillo, V. (2006). A tool for the usability analysis of computerised systems. Paper presented at the Proceedings of the 10th WSEAS international conference on Computers.
- Nicholson, S. (2003). Bibliomining for automated collection development in a digital library setting: Using data mining to discover Web-based scholarly research works. *Journal of the American Society for information science and technology*, 54(12), 1081-1090.
- NISO. (2005). Information services and use: metrics & statistics for libraries and information providers: data dictionary. Retrieved from <https://www.niso.org/publications/z397-2013>
- Olson, D. L. (1996). *Smart Decision Aids for Selection Problems* (pp. 34-48): Springer.
- Ozmutlu, S., Spink, A., & Ozmutlu, H. C. (2003). Multimedia web searching trends: 1997–2001. *Information Processing & Management*, 39(4), 611-621.
- Phillips, M., & Kraemer, A. (2017). Exploring Usage Data Across Multiple Collection Types.
- Polanka, S. (2012). *No shelf required 2: use and management of electronic books* (Vol. 2): American Library Association.
- Pomerol, J.-C. (1992). Aide à la décision et IA. L'intelligence Artificielle une discipline et un carrefour.
- Ponniah, P. (2011). *Data warehousing fundamentals for IT professionals*: John Wiley & Sons.
- Powell, A. (1994). Management models and measurement in the virtual library. *Special libraries*, 85(4), 260-263.
- Priya, P. P. V. (2019). Semantic-based Recommendation Tool for Library Management System using Domain Specific Ontology. *International Journal of Recent Technology and Engineering*, 8(3).
- Roy, B. (1985). *Méthodologie multicritère d'aide à la décision*: Economica.
- Roy, B. (1996). *Multicriteria methodology for decision aiding*, volume 12 of nonconvex optimization and its applications: Kluwer Academic Publishers, Dordrecht.
- Rutledge, J., & Swindler, L. (1987). The selection decision: defining criteria and establishing priorities. *College & Research Libraries*, 48(2), 123-131.
- Ryan, J., McClure, C. R., & Bertot, J. C. (2001). Choosing measures to evaluate networked information resources and services: Selected issues. CR McClure &

Bibliographie

- JC Bertot, Evaluating networked information services: Techniques, policy, and issues, 111-136.
- Saaty, T. L. (1988). What is the analytic hierarchy process? Mathematical models for decision support (pp. 109-121): Springer.
- Saracevic, T. (2000). Digital library evaluation: Toward an evolution of concepts.
- Saurabh, A. K., & Nagpal, B. (2011). A survey on current security strategies in data warehouses. *International Journal of Engineering Science and Technology*, 3(4).
- Schärlig, A. (1985). Décider sur plusieurs critères: panorama de l'aide à la décision multicritère (Vol. 1): PPUR presses polytechniques.
- Schwartz, C. (2000). Digital libraries: An overview. *The Journal of academic librarianship*, 26(6), 385-393.
- sciencedirect.com. (2019). Search Results.
- Sen, A., & Sinha, A. P. (2005). A comparison of data warehousing methodologies. *Communications of the ACM*, 48(3), 79-84.
- Shanian, A., Milani, A., Carson, C., & Abeyaratne, R. (2008). A new application of ELECTRE III and revised Simos' procedure for group material selection under weighting uncertainty. *Knowledge-Based Systems*, 21(7), 709-720.
- Shenton, H. (2003). Life cycle collection management. *Liber Quarterly*, 13(3/4).
- Shreeves, E. (2018). *The New Dynamics and Economics of Cooperative Collection Development*: Routledge.
- Siguenza Guzman, L., Saquicela, V., & Cattrysse, D. (2014). Design of an integrated decision support system for library holistic evaluation. Paper presented at the Proceedings of IATUL Conferences.
- Simon, H., & MARCH, J. (1976). *Administrative behavior organization*: New York: Free Press.
- Simon, H. A. (1987). Making management decisions: The role of intuition and emotion. *Academy of Management Perspectives*, 1(1), 57-64.
- Singh, H. S. (1997). *Data warehousing: concepts, technologies, implementations, and management*: Prentice-Hall, Inc.
- Slater, R. (2010). Why aren't e-books gaining more ground in academic libraries? E-book use and perceptions: A review of published literature and research. *Journal of Web Librarianship*, 4(4), 305-331.
- Snyder, C. E. (2014). Data-driven deselection: Multiple point data using a decision support tool in an academic library. *Collection Management*, 39(1), 17-31.
- Sreenivasulu, V. (2000). The role of a digital librarian in the management of digital information systems (DIS). *The electronic library*, 18(1), 12-20.
- Stowers, E., & Tucker, C. (2009). Using link resolver reports for collection management. *Serials Review*, 35(1), 28-34.
- Sutton, S. (2013). A model for electronic resources value assessment. *The Serials Librarian*, 64(1-4), 245-253.
- Tenopir, C., & King, D. W. (2000). Towards electronic journals: realities for scientists, librarians, and publishers: Special Libraries Association Washington, DC.
- Thakar, A., Szalay, A., Church, K., & Terzis, A. (2011). Large science databases—are cloud services ready for them? *Scientific Programming*, 19(2-3), 147-159.
- Thohira, M., Chambers, M. B., & Sprague, N. (2010). Full-text databases: A case study revisited a decade later. *Serials Review*, 36(3), 152-160.
- Thompson, S., & Sharp, S. (2009). E-books in academic libraries: Lessons learned and new challenges. *Serials*, 22(2).

Bibliographie

- Tjondronegoro, D., Spink, A., & Jansen, B. J. (2009). A study and comparison of multimedia Web searching: 1997–2006. *Journal of the American Society for information science and technology*, 60(9), 1756-1768.
- Townley, C. T. (2001). Knowledge management and academic libraries. *College & Research Libraries*, 62(1), 44-55.
- Tramullas, J., Sánchez-Casabón, A.-I., & Garrido-Picazo, P. (2013). An evaluation based on the digital library user: an experience with greenstone software. *Procedia-social and behavioral sciences*, 73, 167-174.
- Treviño, R. Z., & Treviño, R. Z. (2006). *The Pura Belpré Awards: Celebrating Latino Authors and Illustrators*: American Library Association.
- Tzeng, G.-H., & Huang, J.-J. (2011). *Multiple attribute decision making: methods and applications*: Chapman and Hall/CRC.
- Vincke, P. (1992). *Multicriteria decision-aid*: John Wiley & Sons.
- Wah, T. Y., Peng, N. H., & Hok, C. S. (2007). Building data warehouse. Paper presented at the Proceedings of the 24th South East Asia Regional Computer Conference.
- Waller, V. (2009). How do virtual visitors get to the library? *The electronic library*, 27(5), 815-830.
- Walters, W. H. (2013). E-books in academic libraries: Challenges for acquisition and collection management. *portal: Libraries and the Academy*, 13(2), 187-211.
- Wang, P., Berry, M. W., & Yang, Y. (2003). Mining longitudinal Web queries: Trends and patterns. *Journal of the American Society for information science and technology*, 54(8), 743-758.
- Waugh, M., Donlin, M., & Braunstein, S. (2015). Next-generation collection management: A case study of quality control and weeding e-books in an academic library. *Collection Management*, 40(1), 17-26.
- Weicher, M., & Zhang, T. X. (2012). Unbundling the “big deal” with pay-per-view of e-journal articles. *The Serials Librarian*, 63(1), 28-37.
- Wikipedia. (2019). Books Published per Country per year. Retrieved from https://en.wikipedia.org/wiki/Books_published_per_country_per_year
- Williams, L. B. (1991). Subject Knowledge for Subject Specialists: What the Novice Bibliographer Needs to Knowledge. *Collection Management*, 14(3-4), 31-47.
- Wilson, A. P. (2004). Weeding the e-book collection. *Public Libraries*, 43(3), 158-159.
- Witten, I. H., Bainbridge, D., & Nichols, D. M. (2009). *How to build a digital library*: Morgan Kaufmann.
- Wolfram, D. (2008). Search characteristics in different types of Web-based IR environments: Are they the same? *Information Processing & Management*, 44(3), 1279-1292.
- Wrembel, R. (2006). *Data Warehouses and OLAP: Concepts, Architectures and Solutions*: Igi Global.
- Yue, C. (2019). A symbol-based fuzzy decision-making approach to evaluate the user satisfaction on services in academic digital libraries. *Iranian Journal of Fuzzy Systems*, 16(2), 73-86. doi:10.22111/ijfs.2019.4543
- Zeleny, M. (1982). *Multiple Criteria Decision Making*. New York: McGraw-Hill.
- Zhang, J., Wolfram, D., & Wang, P. (2009). Analysis of query keywords of sports-related queries using visualization and clustering. *Journal of the Association for Information Science and Technology*, 60(8), 1550-1571.
- Zhang, Y. (2010). Developing a holistic model for digital library evaluation. *Journal of the American Society for information science and technology*, 61(1), 88-110.

Bibliographie

Zielke, A., Sitter, H., Rampp, T., Bohrer, T., & Rothmund, M. (2001). Clinical decision-making, ultrasonography, and scores for evaluation of suspected acute appendicitis. *World journal of surgery*, 25(5), 578-584.