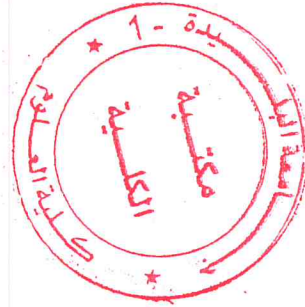


République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université Saad Dahleb Blida
Faculté des Sciences
Département de l'informatique



Mémoire de fin d'étude

En vue de l'obtention du diplôme de Master en Informatique
Option: Génie des Systèmes Informatique

Thème:

**Conception et réalisation d'un Datawarehouse pour la
realisation d'un système décisionnel**

Réalisé par:

- SAFAR ZITOUN Berkane
- MEKID Lyes

Promotrice: Mme ZAHRA Fatma Zohra
Encadreur: Mr BENAÏSSA Kheireddine

Promotion:2015/2016

MA-004-362-2

Remerciements

Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce Modeste travail.

Nous tenons à saisir cette occasion et adresser nos remerciements et nos reconnaissances à notre promotrice Mme Zahra pour ces conseils et pour sa patience avec nous.

Nous tenons aussi à remercier nos parents, pour leur encouragements, soutien et patience tout au long de notre cursus.

Nous remercions aussi les membres du jury qui nous ont fait l'honneur d'accepter de juger notre travail.

Résumé:

Le secteur de l'informatique décisionnelle en Algérie évolue dans un contexte de changement et de concurrence permanents. "El Maarifa" est un institut qui souhaite développer une solution de Business Intelligence pour ses besoins.

Les informations récupérées des différents documents et base de données reposent sur différents schémas, plusieurs formats de données et donc une hétérogénéité est identifiée, ce qui nécessite un temps considérable pour réaliser des analyses et les tableaux de bords à remettre aux décideurs. Dans cette optique, l'institut veut se doter d'un système d'information décisionnel qui permet aux décideurs d'avoir des informations fiables et utiles pouvant les aider pour prendre des décisions et réagir rapidement tout en corrigeant et intégrant leur patrimoine informationnel.

La réalisation de ce projet commencera d'abord par l'identification des besoins métiers, passant ensuite par l'analyse puis la conception d'un Datawarehouse et des outils ETL nécessaires. Ensuite, viendra la réalisation de cette plateforme en mettant à la disposition des décideurs des tableaux de bord décisionnels.

Mots Clés: BI, OBIEE, Informatique décisionnelle, Datawarehouse, ETL, Tableau de bord, Reporting, KPI.

Abstract:

The BI industry in Algeria operates in a context of permanent change and competition. "El Maarifa" is an IT service company that wants to develop a Business Intelligence solution for its needs.

The information collected from various documents and databases based on different patterns, multiple data formats and therefore heterogeneity is identified, which requires considerable time to perform analysis and dashboards to give to decision makers. In this context, "El Maarifa" Institute wants to have a decision support system that allows decision makers to have reliable and useful information that can help them to make decisions and react quickly while correcting and integrating their information assets.

The realization of this project will start first by identifying business needs, then through the analysis and design of a data warehouse and ETL tools necessary. Then will come the realization of this platform by providing decision makers in executive dashboards.

Keywords: BI, OBIEE, Business Intelligence, Data Warehousing, ETL, Dashboard, Reporting, KPI.

Contents

Contents	4
List of Figures	7
1 Le Système Décisionnel	3
1.1 Les systèmes décisionnels :	4
1.1.1 Définitions et objectifs :	4
1.1.2 Le processus décisionnel :	4
1.2 Architecture d'un système décisionnel :	5
1.3 Système opérationnelles vs décisionnelles :	6
1.4 Conclusion :	7
2 Data Warehousing	9
2.1 Définition:	10
2.2 Caractéristiques d'un Data Warehouse :	10
2.3 Objectives d'un Data Warehouse:	10
2.4 Alimentation des Data Warehouse "Extraction, Transformation, Load" :	10
2.5 Les magasins de données (Datamart):	12
2.5.1 Comparaison entre les finalités d'un datawarehouse et d'un datamart :	12
2.6 La modélisation multidimensionnelle :	13
2.6.1 Définitions :	13
2.6.2 Table de fait :	13
2.6.3 Les Dimensions :	14
2.6.4 Les Mesures :	14
2.7 Techniques de modélisation multidimensionnelle :	15
2.7.1 Schéma en étoile :	15
2.7.2 Schéma en flocon de neige :	15
2.7.3 Modèle en constellation :	16
2.8 Le concept OLAP :	16
2.8.1 Définitions :	16
2.8.2 OLAP vs OLTP :	17
2.8.3 Les différents types OLAP :	17

2.8.4	Le cube OLAP et ses caractéristiques:	18
2.8.5	Opérations sur les cubes :	19
2.9	Conclusion	21
3	Les Outils de Restitution de données	23
3.1	Tableau de bord	24
3.1.1	Objectifs d'un tableau de bord	24
3.1.2	Les indicateurs Clé de performance (KPI)	24
3.2	Reporting	25
3.2.0.1	Objectifs de Reporting	25
3.3	Data Mining	25
3.3.1	Objectifs de Data Mining	26
3.4	Conclusion	26
4	Étude de l'existant	27
4.1	Présentation de l'organisme d'accueil	28
4.1.1	Principales missions de l'institut El Maarifa :	28
4.1.2	Organigramme de l'institut :	28
4.1.3	Les directions principales de l'institut :	28
4.2	Présentation du système opérationnel existant	29
4.2.1	Etude des sources de données	29
4.3	Critique de l'existant	30
4.4	Conclusion	31
5	Analyse des besoins	33
5.1	Collecte d'informations	34
5.1.1	La documentation	34
5.1.2	L'analyse des processus fonctionnels et techniques existants	34
5.1.3	Les réunions	34
5.2	L'analyse des besoins analytiques:	34
5.2.1	Identification des tableaux de bord :	34
5.2.2	Identification des rapports	34
5.3	L'analyse des besoins techniques	35
5.4	Solution proposée	35
5.5	Conclusion	37
6	Conception du système décisionnel	39
6.1	Conception de la base de données	40
6.1.1	Choix de langage de modélisation	40
6.1.2	Définition de la méthode UML	40
6.1.3	Modélisation	40
6.1.3.1	Définition de diagramme de cas d'utilisation	41
6.1.3.2	Définition de diagramme de classe:	43
6.2	Modélisation de l'entrepôt de données	45

6.2.1	Méthode de modélisation	45
6.2.2	Le schéma en flocon de neige du magasin "Suivi- Intra"	46
6.2.3	Le schéma en flocon de neige du magasin "Suivi-Inter"	47
6.2.4	Les dimensions du système	48
6.2.5	Niveaux et hiérarchies des dimensions	52
6.2.6	La zone d'alimentation	53
6.2.7	Conception des cubes dimensionnels	55
6.2.7.1	Niveaux et hiérarchies des dimensions	56
6.2.7.2	Les cubes dimensionnels	57
6.2.8	Le système du tableau de bord	57
6.2.8.1	Méthode de construction du TDB :	57
6.2.8.2	Les indicateurs clés de performance (KPI) :	59
6.3	Conclusion	59
7	Réalisation du système décisionnel	61
7.1	Architecture technique du projet	62
7.1.1	L'environnement technique de la solution:	62
7.1.2	Description des outils utilisés:	62
7.1.3	La couche Physique:	64
7.1.4	La couche Business Model and Mapping:	66
7.1.5	Couche Présentation:	67
7.2	Construction de l'application utilisateur	67
7.3	Conclusion	70

List of Figures

1.1	Architecture d'un système décisionnel [Web2]	6
1.2	: Différences entre les systèmes opérationnelles et les données décisionnelles	7
2.1	Le Processus d'alimentations du Data Warehouse – ETL	12
2.2	Datawarehouse divisé en Datamarts	12
2.3	Comparaison entre les finalités des DataWarehouse et des DataMarts	13
2.4	Exemple d'une table de fait	14
2.5	Exemple d'une table de dimension	14
2.6	: Exemple des mesures	14
2.7	Structure d'un schéma en étoile.	15
2.8	: Structure d'un schéma en flocon de neige	16
2.9	Structure d'un schéma en constellation	16
2.10	Récapitulatif des différences entre OLTP et OLAP	17
2.11	Architecture du ROLAP [Nakache, 1998]	18
2.12	Architecture du MOLAP [Nakache, 1998]	18
2.13	Exemple d'un cube de données	19
2.14	: Exemple du Slicing [Epinasse ,2010]	20
2.15	Exemple du Dicing [Epinasse ,2010]	20
2.16	Exemple du Drill-down [Epinasse ,2010]	21
2.17	Exemple du roll-up [ESPINASSE, 2010]	21
3.1	Le " triangle " de l'indicateur : stratégie traduite en objectif, processus d'action et acteur collectif [Lorino, 2001]	25
4.1	Organigramme de l'institut El Maarifa	28
5.1	Architecture de la solution proposée	36
5.2	Exemple des fichiers sources Excel	36
6.1	Diagramme de cas d'utilisation général	42
6.2	Les spécifications de l'administrateur	42
6.3	Les spécifications des utilisateurs	43
6.4	Diagramme de classe	44

6.5	Méthodes de modélisation d'un entrepôt	46
6.6	Schéma en flocon de neige du magasin "Suivi-Intra"	47
6.7	Schéma en flocon de neige du magasin "Suivi-Inter"	48
6.8	Tableau descriptif de la dimension "Formation"	49
6.9	Tableau descriptif de la dimension "Client"	49
6.10	Tableau descriptif de la dimension "Demande"	50
6.11	Tableau descriptif de la dimension "Etat-Demande"	51
6.12	Tableau descriptif de la dimension "Temps"	51
6.13	: Tableaux descriptif de la dimension "Adresse"	52
6.14	Le processus général du chargement du DW	55
6.15	Liste des niveaux et des hiérarchies des dimensions du data warehouse.	56
6.16	Les 10 étapes de la méthode GIMSI, [Fernandez, 2006]	58
6.17	Les indicateurs du tableau de bord	59
7.1	Outils utilisés pour la réalisation de notre système	62
7.2	Architecture d'OBIEE	64
7.3	Représentation des trois couches réalisées	65
7.4	Schéma relationnel de la base de données BDD-IM	66
7.5	Schéma en flocon de neige de suivi-intra sur la couche logique	67
7.6	Exemple de traitement d'analyse sous OBIEE.	68
7.7	Exemple d'un tableau de bord	69
7.8	Exemple d'un rapport	69

Introduction Générale

Contexte de l'étude

Pendant longtemps, le succès d'une entreprise se mesurait exclusivement par des critères tels que la bonne gestion des finances, le contrôle des ressources humaines, la gestion des ressources physiques ... etc. Aujourd'hui, le succès est mesuré en fonction du contrôle que possède l'entreprise sur l'information pertinente en son domaine. Il est primordial pour un agent de l'entreprise de pouvoir accéder aux informations, de les analyser et de les trier afin de mieux les exploiter. Le traitement de l'information est donc aujourd'hui un facteur déterminant pour le succès de l'entreprise.

Pour se faire, les dirigeants de l'entreprise doivent prendre notamment les décisions les plus opportunes. Ces décisions, qui influenceront grandement sur la stratégie de l'entreprise et donc sur son devenir, ne doivent pas être prises ni à la légère, ni de manière trop hâtive, compte tenu de leurs conséquences sur la survie de l'entreprise. Il s'agit de prendre des décisions fondées, basées sur des informations claires, fiables et pertinentes. Le problème est de savoir donc comment identifier et présenter ces informations à qui de droit, sachant par ailleurs que les entreprises croulent d'une part sous une masse considérable de données et que d'autre part les systèmes opérationnels " transactionnels " s'avèrent limités, voire inaptes à fournir de telles informations et constituer par la même un support appréciable à la prise de décision.

C'est dans ce contexte que les " systèmes décisionnels " ont vu le jour. Ils offrent aux décideurs des informations de qualité sur lesquelles ils pourront s'appuyer pour arrêter leurs choix décisionnels. Pour se faire, ces systèmes utilisent un large éventail de technologies et de méthodes, dont les " entrepôts de données " représentent l'élément principal et incontournable pour la mise en place d'un bon système décisionnel.

L'Institut "El Maarifa" génère des données complexes et volumineuses. Ces données représentent une source précieuse d'informations, qui serait à même d'améliorer de façon significative le processus de prise de décision. Cependant, ces données ne sont pas exploitées de manière satisfaisante, hypothéquant ainsi le processus de prise de décision à tous les niveaux de l'institut.

Le présent projet tend à mettre en place un système, en mesure de consolider les données issues des systèmes transactionnels, et d'offrir des informations de qualité pour les décideurs. Il s'agit en fait de mettre à la disposition des décideurs des données à même de les éclairer et leur faciliter une prise de décision prompte en connaissance de cause. Un tel système requiert la mise en place d'un système décisionnel fiable contenant les informations nécessaires à l'accomplissement des processus décisionnels.

Problématique

L'Institut "El Maarifa" a pour mission de fournir de la formation de haut niveau dans les domaines du management, des langues et du développement humains, respectant les standards internationaux et garantissant à l'apprenant, l'acquisition de compétences avérées lui ouvrant de nouvelles perspectives pour sa carrière. Appelé à interagir avec ses clients sur différentes phases de la formation, l'institut veut se doter d'un système décisionnel, dans un souci de suivi de la clientèle et de gestion de la formation. Dans un pareil contexte, la plus simple des opérations d'analyse devient une tâche ardue. En effet, les décideurs de l'institut se trouvent dans l'incapacité de faire des analyses fiables, efficaces et à des moments opportuns sans engager des moyens considérables sur des périodes plus ou moins longues. Ainsi, les principales difficultés rencontrées peuvent être résumées en:

- **Des données non structurées** : ne sont pas organisées dans une perspective décisionnelle, et donc ne sont pas exploitées de la meilleure façon, ce qui rend les requêtes statistiques lentes.
- **Difficultés dans l'élaboration des rapports d'activité** : L'élaboration des rapports d'activité fait intervenir, généralement, plusieurs intermédiaires. Les retards enregistrés, parfois, font que le rapport d'activité est élaboré sur la base d'une consolidation antérieure, en sachant pertinemment que les données ne sont pas à jour
- **Lenteur de la procédure de Reporting** : La politique de Reporting actuelle, qui du reste est quasi manuelle, connaît des lenteurs qui n'arrangent pas les décideurs. Ceux-ci ont besoin d'informations fiables et dans des délais raisonnables.
- **Insuffisance du module " Statistique "** : Afin de produire et offrir un moyen de suivi des activités de la distribution, un module " Statistique " a été développé et intégré dans le système d'information existant . Ce dernier fournit des états statistiques permettant, aux décideurs l'analyse et la prise de décision. Cependant, ce module connaît quelques problèmes dû au fait qu'il interroge directement la base de données en production. En effet le lancement de la production de n'importe quel rapport du module pénalise le système.

Objectifs :

Afin de palier aux problèmes suscités, l'Institut "El Maarifa" s'est engagé à créer un système d'aide à la décision. Ce système doit assurer un processus simplifiant toute la chaîne depuis le traitement des données,

jusqu'à la restitution des résultats d'analyse. Pour cela, des objectifs organisationnels et fonctionnels ont été tracés à savoir :

- Rendre disponible à tout moment des informations nécessaires à la prise de décision.
- Automatiser le processus de préparation des données pour les deux types de formations (qualifiantes et diplômantes).
- Offrir aux décideurs et aux analystes la possibilité de faire des analyses appropriées.
- Minimiser la durée d'élaboration des rapports au niveau de l'entreprise.
- Visualiser les données via un tableau de bord.
- Offrir des informations fiables, cohérentes et pertinentes.

Organisation du mémoire :

Après cette introduction générale dans laquelle nous avons présenté le contexte général du projet, ainsi que la problématique et les objectifs visés, nous avons retenu pour ce mémoire, une organisation en deux parties : Pratique et Théorique.

La partie théorique sera consacrée à une synthèse bibliographique, où nous présenterons quelques généralités sur les systèmes décisionnels, les entrepôts de données, la modélisation multidimensionnelle ainsi que le concept OLAP. tandis que la partie pratique sera divisée en 4 chapitres:

- **Chapitre I :** Dans ce chapitre, nous allons faire dans une première partie la présentation de l'organisme d'accueil, son environnement, sa structure ainsi que ses activités. Nous passerons par la suite à l'analyse de l'existant du système existant pour finir avec des critiques de l'existant.
- **Chapitre II :** Dans ce chapitre, nous allons voir les techniques de collecte d'information utilisées, la démarche suivie afin d'identifier les besoins pour finir avec une proposition de la solution.
- **Chapitre III :** : ce chapitre portera sur la conception de la solution proposée.
- **Chapitre IV :** : dans ce chapitre, nous présenterons l'implémentation de notre système réalisé, puis nous effectuerons quelques tests d'évaluation de notre module.

Une conclusion générale est proposée afin de synthétiser le travail réalisé et de citer les perspectives du projet.

Chapter 1

Le Système Décisionnel

La notion de Business Intelligence (BI) englobe les solutions informatiques dont le but est de consolider les informations disponibles au sein des bases de données de l'entreprise. En effet, les entreprises, et plus globalement les organisations, se trouvent confrontées à des environnements de plus en plus complexes et compétitifs dans lesquels le pilotage implique des choix qui doivent être faits dans des temps très courts, tout en prenant en compte un volume d'informations toujours plus important. Dans ce chapitre nous présentons les systèmes décisionnels. Nous allons détailler l'architecture générale d'un système décisionnel. Nous allons passer par la suite à la modélisation dimensionnelle et l'exploitation des données dans les modèles dimensionnels.

1.1 Les systèmes décisionnels :

1.1.1 Définitions et objectifs :

Les systèmes décisionnels s'inscrivent dans la famille des systèmes d'informations, ils sont souvent appelés système d'information décisionnel (SID). Parmi les nombreuses définitions des systèmes décisionnels, nous retenons celle de Devisy [A.Devisy, 2002] qui définit le SID comme : Un système permettant aux décideurs de l'entreprise de disposer d'informations pertinentes et d'outils d'analyse puissants pour aider à prendre les bonnes décisions au bon moment. Un système décisionnel a pour objectifs :

- Avoir une vision globale, stratégique, homogène et cohérente des informations manipulées par l'entreprise.
- Accéder simplement et intuitivement aux données pertinentes.
- Afficher les résultats sous différentes formes (tables, courbes, graphes...etc.)
- Faciliter l'exploitation des données pour l'aide à la prise de décision.

1.1.2 Le processus décisionnel :

Un projet décisionnel est composé de 4 phases clés [Web1] :

- **La phase de collecte :**
L'objectif de cette phase étant essentiellement d'identifier les sources de données qui peuvent être sous différents formats (fichiers plats ou bases de données) qui sont en général hétérogènes.
- **La phase d'intégration :**
C'est à ce niveau qu'apparaît la première couche logicielle de la chaîne décisionnelle à savoir l'ETL. Elle permet l'extraction de données provenant de divers systèmes, la transformation de ces données et de leur chargement dans une zone intermédiaire appelée Staging Area. Elle est destinée à stocker les données issues des systèmes opérationnels avant leur intégration dans l'entrepôt de données.
- **La phase d'organisation :**
La troisième phase permet de stocker les données dans un entrepôt appelé DataWarehouse. Cette structure de données généralement dénormalisée pour pouvoir optimiser les temps de réponses lorsque l'on fait des analyses de type OLAP qui se réfère à une base de données multidimensionnelle.
- **La phase de restitution :**
La dernière phase concerne l'accès et la diffusion des résultats, on

distingue à ce niveau plusieurs types d'outils :

- Les outils de Reporting permettent de générer des rapports périodiques, pré formatés et paramétrables par les opérationnels. Ils facilitent la création de rapports par les utilisateurs eux-mêmes, en interrogeant le DW grâce à des analyses croisées, à travers une couche d'abstraction orientée métier. Ils permettent également la production de tableaux de bord avec des indicateurs de haut niveau pour les managers, synthétisant différents critères de performance.

- Les outils d'analyse OLAP permettent de structurer les données sous forme de cubes multidimensionnels et de naviguer dans les différentes dimensions. Ce qui permet d'obtenir immédiatement plusieurs représentations d'un même résultat.

- Les outils de Datamining offrent une analyse plus poussée permettant de découvrir des connaissances implicites dans les données comme la détection de corrélations et des tendances entre les différentes variables étudiés, les segmentations ou encore des prévisions. Le Datamining est basé sur des algorithmes statistiques et mathématiques, et sur des hypothèses métier.

1.2 Architecture d'un système décisionnel :

L'architecture type d'un système décisionnel peut être représentée en quatre niveaux où chaque niveau représente une phase de décision :

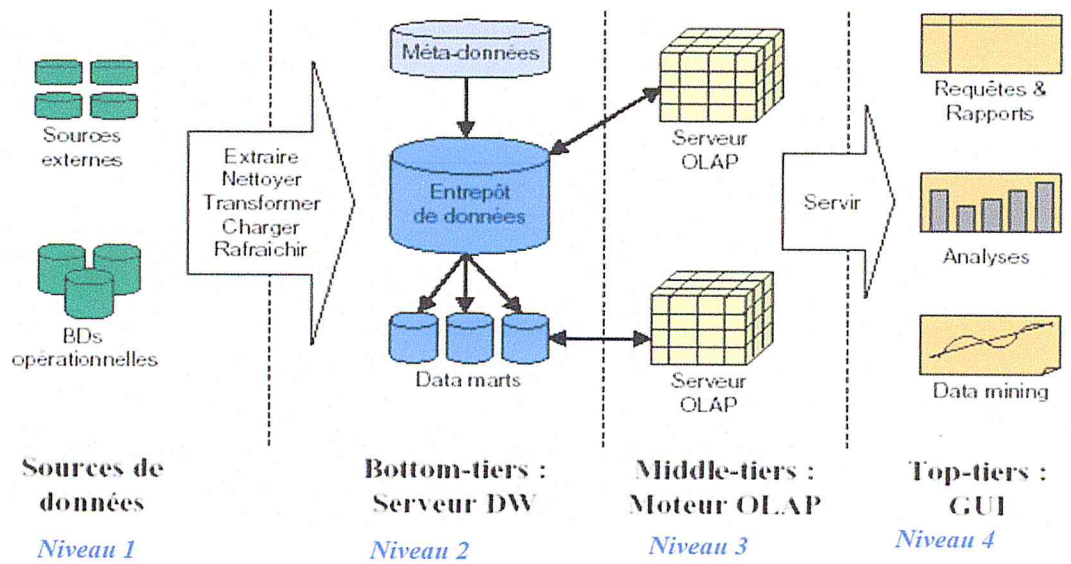


Figure 1.1: Architecture d'un système décisionnel [Web2]

1.3 Système opérationnelles vs décisionnelles :

Le tableau 1, est un comparatif des systèmes opérationnelles et décisionnelles selon les données et l'usage.

	Système opérationnelles	Système décisionnelles
Les Données	Orientées application, détaillées, précises au moment de l'accès.	Orientées activité (thème, sujet), condensées, représentent des données historiques.
	Données atomiques.	Données générales et détaillées.
	Petite volumétrie des données	Gros volumes de données.
	Lecture, écriture et modification des données.	Données en lecture seule.
	Fragmentés, Hétérogènes.	Centralisés.
L'Usage	Assure l'activité au quotidien.	Permet l'analyse et la prise de décision.
	Grand nombre d'utilisateurs.	Petit nombre d'utilisateurs.
	Projets comportant peu de risques.	Projets très risqués.
	Accès extrêmement rapides.	Rapidité suggérée.

Figure 1.2: : Différences entre les systèmes opérationnelles et les données décisionnelles

1.4 Conclusion :

Le système décisionnel est devenu un sujet incontournable et une composante essentielle au sein d'une entreprise afin d'optimiser le pilotage de son activités. Nous avons abordé dans cette première partie la notion de système décisionnel, ses objectifs, son architecture et de ses phases, notamment le Data Warehouse. Ce dernier est considéré comme étant le cœur de l'informatique décisionnelle, d'où l'intérêt de poursuivre la suite en détaillant le processus de conception et l'architecture du DW.

Chapter 2

Data Warehousing

Les besoins décisionnels exprimés par les entreprises n'étaient pas satisfaits par les systèmes traditionnels de bases de données, ce qui a donné naissance aux systèmes décisionnels basés sur les entrepôts de données. Dans cette partie, nous allons définir les concepts du Data Warehouse, ses caractéristiques, ses objectifs, ses différentes architectures, ses approches de construction et enfin le processus d'alimentation ETL.

2.1 Définition:

Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et évolutives dans le temps, organisées pour le support d'un processus d'aide à la décision. [Inmon, 2002].

2.2 Caractéristiques d'un Data Warehouse :

Dans un entrepôt, les données sont :

- **Orienté sujets** : les données collectées doivent être orientées "métier" et donc triées par thème.
- **Intégrées** : les données doivent être mises en forme et unifiées afin d'avoir un état cohérent.
- **Non volatiles** : afin de conserver la traçabilité des informations et des décisions prises, les informations stockées au sein de l'entrepôt de données ne peuvent être supprimées.
- **Historiées** : l'historisation est nécessaire pour suivre dans le temps l'évolution des différentes valeurs des indicateurs à analyser. Ainsi, un référentiel temps doit être associé aux données afin de permettre l'identification dans la durée de valeurs précises.

2.3 Objectives d'un Data Warehouse:

Parmi les objectifs des data warehouses, nous citons [Web3] :

- Développement d'applications décisionnelles et de pilotage de l'entreprise et de ses processus.
- Offrir une vision globale et orientée métiers de toutes les données que manipule l'entreprise.
- Jouer un rôle de référentiel pour l'entreprise puisqu'il permet de fédérer des données souvent éparpillées dans différentes bases de données.
- Faire face aux changements du marché et de l'entreprise.
- Offrir une information compréhensible, utile et rapide.

2.4 Alimentation des Data Warehouse "Extraction, Transformation, Load" :

L'ETL est le processus qui permet de charger un data warehouse à partir de données externes, généralement issues de bases transactionnelles.

2.1 Définition:

Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et évolutives dans le temps, organisées pour le support d'un processus d'aide à la décision. [Inmon, 2002].

2.2 Caractéristiques d'un Data Warehouse :

Dans un entrepôt, les données sont :

- **Orienté sujets** : les données collectées doivent être orientées "métier" et donc triées par thème.
- **Intégrées** : les données doivent être mises en forme et unifiées afin d'avoir un état cohérent.
- **Non volatiles** : afin de conserver la traçabilité des informations et des décisions prises, les informations stockées au sein de l'entrepôt de données ne peuvent être supprimées.
- **Historiées** : l'historisation est nécessaire pour suivre dans le temps l'évolution des différentes valeurs des indicateurs à analyser. Ainsi, un référentiel temps doit être associé aux données afin de permettre l'identification dans la durée de valeurs précises.

2.3 Objectives d'un Data Warehouse:

Parmi les objectifs des data warehouses, nous citons [Web3] :

- Développement d'applications décisionnelles et de pilotage de l'entreprise et de ses processus.
- Offrir une vision globale et orientée métiers de toutes les données que manipule l'entreprise.
- Jouer un rôle de référentiel pour l'entreprise puisqu'il permet de fédérer des données souvent éparpillées dans différentes bases de données.
- Faire face aux changements du marché et de l'entreprise.
- Offrir une information compréhensible, utile et rapide.

2.4 Alimentation des Data Warehouse "Extraction, Transformation, Load" :

L'ETL est le processus qui permet de charger un data warehouse à partir de données externes, généralement issues de bases transactionnelles.

Son rôle est de récupérer ces données et de les traiter pour qu'elles correspondent aux besoins du modèle dimensionnel.

En général les données sources doivent être nettoyées et aménagées pour être exploitables par les outils décisionnels. La précision et la vitesse de la plateforme décisionnelle toute entière dépendent des processus ETL, notamment :

- **Extraction des données :** L'extraction des données des systèmes sources est sans doute la tâche la plus vaste du projet Entrepôt de données. [Kimball et Al, 2007].
L'extraction des données est la première étape dans les systèmes ETL. C'est la lecture des données à partir des systèmes sources généralement hétérogènes. Les contraintes de cette étape sont :
 - Elle doit être réalisée rapidement pour ne pas surcharger le système source.
 - Elle implique une gestion des mises à jour des dimensions changeantes.
 - Elle doit exploiter le minimum possible des ressources du système source.
- **Transformation des données :** La transformation est la seconde phase de processus. C'est une suite d'opérations qui a pour but de rendre les données cibles homogènes pour qu'elles puissent être traitées de façon cohérente.
- **Chargement des données :** Cette phase consiste à charger les données dans le Data Warehouse. Ce processus varie selon les critères et les différents besoins de l'entreprise, afin de rendre le processus de chargement efficace.

Nous récapitulons les étapes de l'ETL dans la figure suivante :

2.5 Les magasins de données (Daramart):

” Le Datamart est un sous-ensemble du Data Warehouse, constitué de tables au niveau détail et à des niveaux plus agrégés, permettant de restituer tout le spectre d'une activité métier. L'ensemble des Datamarts de l'entreprise constitue le Data Warehouse ” [Kimball, 2004].

2.5.1 Comparaison entre les finalités d'un datawarehouse et d'un datamart :

Le tableau suivant montre une comparaison entre un Data Warehouse et un Data Mart :

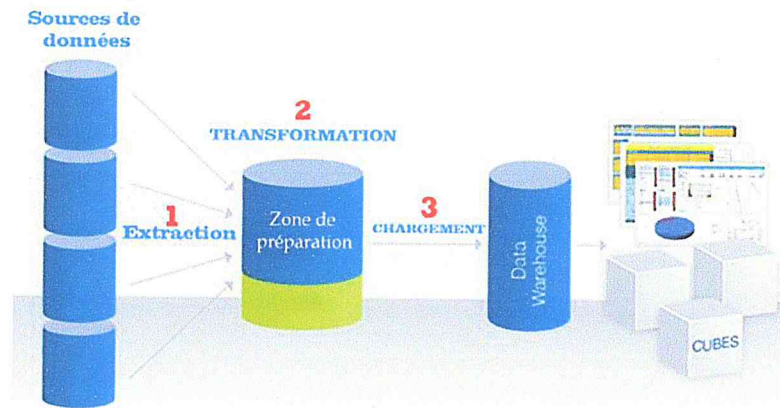


Figure 2.1: Le Processus d'alimentations du Data Warehouse – ETL

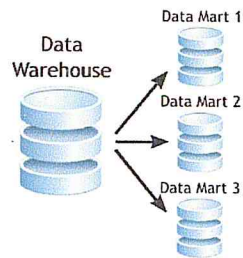


Figure 2.2: Datawarehouse divisé en Datamarts

2.6 La modélisation multidimensionnelle :

2.6.1 Définitions :

”La modélisation multidimensionnelle est une méthode de conception logique qui vise à présenter les données sous forme standardisée, intuitive et qui permet des accès hautement performants. Elle adhère totalement à la dimensionnalité ainsi qu’à une discipline qui exploite le modèle relationnel tout en le limitant sérieusement.” [Kimball et Al, 2005].

2.6.2 Table de fait :

”Une table de fait est une table qui contient les données observables et mesurables (les faits) que l’on possède sur un sujet et que l’on veut étudier, selon divers axes d’analyse. Un fait est une observation de marché. Dans le monde d’entreprise de nombreux faits sont numériques ; bien que certains puissent être valorisés par des informations textuelles.” [Kimball et Al, 2005].

Caractéristiques	DataWarehouse	DataMart
Cible entreprise	Toute l'entreprise	Département
Implication du service informatique	Elevé	Faible ou moyen
Champ applicatif	Multi sujets, neutre	Quelque sujet, spécifique
Sources de données	multiples	Quelques unes
Stockage	Bases de données	Plusieurs bases distribuées
Tailles	Centaine de GO et plus	Une à 2 dizaines de GO

Figure 2.3: Comparaison entre les finalités des DataWarehouse et des Data-Marts

FAITS
N° Commande
N° Client
Code produit
Code vendeur
Quantité
Prix total

Figure 2.4: Exemple d'une table de fait

2.6.3 Les Dimensions :

"Les tables de dimension sont les tables qui accompagnent une table de faits, elles contiennent les descriptions textuelles de l'activité. Une table de dimension est constituée de nombreuses colonnes qui décrivent une ligne. C'est grâce à cette table que l'entrepôt de données est compréhensible et utilisable; elles permettent des analyses en tranches et en dés. Une table de dimension est une table qui contient la description textuelle du métier." [Kimball et Al, 2005].

CLIENTS
N° Client
Nom client
Ville
Pays

Figure 2.5: Exemple d'une table de dimension

2.6.4 Les Mesures :

Ce sont les éléments de données sur lesquels portent les analyses en fonctions des différentes dimensions, ces valeurs sont le résultat d'opérations d'agrégation sur les données.

FAITS
N° Commande
N° Client
Code produit
Code vendeur
Quantité
Prix total

Mesures {

Figure 2.6: : Exemple des mesures

2.7 Techniques de modélisation multidimensionnelle :

Après avoir vu les concepts de table des dimensions et table des faits, nous allons étudier les différentes techniques de modélisation dimensionnelle de celles-ci.

2.7.1 Schéma en étoile :

Ce type de modèle est initié par Ralph Kimball. Il peut posséder un nombre quelconque de tables de dimension, qui n'ont aucune relation entre elles, reliées avec une unique table centrale qui est la table de fait. La figure ci-dessous illustre un schéma en étoile comportant une unique table de faits et quatre tables de dimension.

2.7.2 Schéma en flocon de neige :

Identique au modèle en étoile, sauf que ses branches sont éclatées en hiérarchies. Cette modélisation est généralement justifiée par l'économie d'espace de stockage, cependant elle peut s'avérer moins compréhensible pour l'utilisateur final, et très coûteuse en termes de performances.

2.7. TECHNIQUES DE MODÉLISATION MULTIDIMENSIONNELLE : 17



Figure 2.7: Structure d'un schéma en étoile.

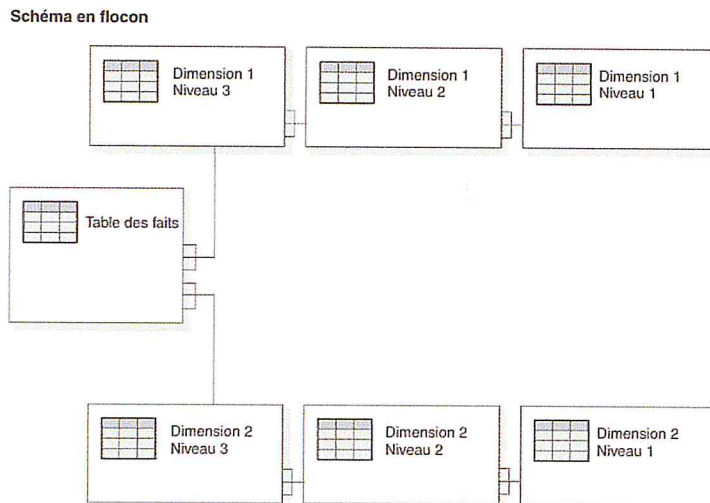


Figure 2.8: : Structure d'un schéma en flocon de neige

2.7.3 Modèle en constellation :

Ce modèle est un ensemble de schémas en étoiles et/ou en flocon dans lesquels les tables de faits se partagent certaines tables de dimensions. En d'autre terme, le schéma en constellation comporte des dimensions en commun entre plusieurs faits.

La figure 9 illustre un schéma en constellation comportant deux tables de faits et six tables de dimension dont deux dimensions sont partagées entre les deux tables de faits.

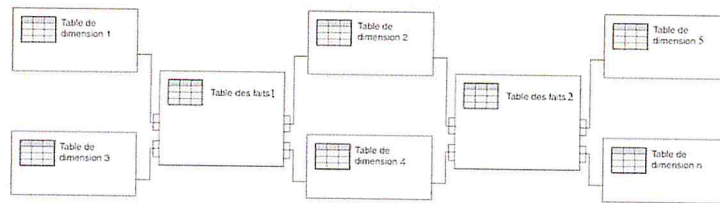


Figure 2.9: Structure d'un schéma en constellation

2.8 Le concept OLAP :

2.8.1 Définitions :

Le terme OLAP (On-Line Analytical Processing) désigne une classe de technologies conçue pour l'accès aux données et pour une analyse instantanée de ces dernières, dans le but de répondre aux besoins de Reporting et d'analyse. Ralph Kimball définit le concept OLAP comme "Activité globale de requêtage et de présentation de données textuelles et numériques contenues dans l'entrepôt de données; Style d'interrogation spécifiquement dimensionnel" [Kimball, 2005].

2.8.2 OLAP vs OLTP :

Le OLTP (On Line Transaction Processing) : " concerne l'ensemble des applications interactives qui facilitent et gèrent les transactions opérationnelles de l'entreprise ".

Le Tableau suivant dresse un comparatif entre les deux concepts :

2.8.3 Les différents types OLAP :

Au niveau logique, il existe plusieurs possibilités pour la modélisation multidimensionnelle :

- **ROLAP** : Dans le Relational OLAP, les données sont stockées dans une base de données relationnelle. Un moteur OLAP permet de simuler le fonctionnement d'un hypercube. Cela permet une facilité dans la mise à jour des données.
- **MOLAP** : Le Multidimensionnel OLAP consiste à utiliser un système multidimensionnel pur, qui gère des structures multidimensionnelles natives. Ils utilisent des tableaux à n dimensions. L'accès aux données se fait directement dans le cube. Cela permet une rapidité d'accès à l'information mais augmente le temps de mise à jour.

Caractéristiques	OLTP	OLAP
Utilisation	SGBD (base de production)	Data Warehouse
Source de données	Données opérationnelles	Les données OLAP viennent des différentes bases de données OLTP
Objectif	Contrôler et exécuter les tâches métier fondamentales et fréquentes	Aider à planifier, prendre des décisions d'ordre stratégique
Vitesse d'exécution	Typiquement très rapide	Dépend de la quantité de données impliquées, la vitesse peut se voir améliorer par la création d'index
Opération typique	Mise à jour	Analyse
Quantité d'informations	Faible	Importante
Taille BDD	Relativement faible (max quelques GB) si les données historisées sont archivées	Importante (pouvant atteindre plusieurs TB) due à l'existence des structures d'agrégations et des données historisées.
Ancienneté des données	Récente	Historique

Figure 2.10: Récapitulatif des différences entre OLTP et OLAP

- **HOLAP** : HOLAP (Hybrid OLAP) est un hybride entre ROLAP et MOLAP. Les parties tables de faits et tables de dimensions sont stockées dans une base relationnelle standard. Tandis que le reste des données (les calculs) sont stockées dans une base multidimensionnelle.

2.8.4 Le cube OLAP et ses caractéristiques:

Un cube OLAP est une représentation abstraite d'informations multidimensionnelles exclusivement numérique utilisé par l'approche OLAP. Cette structure est prévue à des fins d'analyses interactives par une ou plusieurs personnes (souvent ni informaticiens ni statisticiens) du métier que ces données sont censées représenter.

Les cubes OLAP ont les caractéristiques suivantes : [Web5]

- Obtenir des informations déjà agrégées selon les besoins de l'utilisateur.

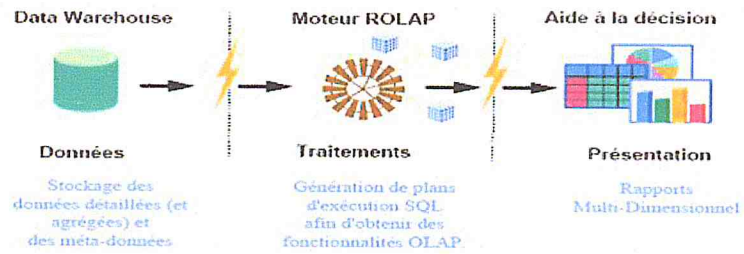


Figure 2.11: Architecture du ROLAP [Nakache, 1998]

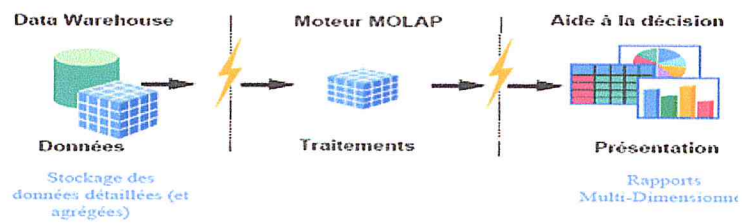


Figure 2.12: Architecture du MOLAP [Nakache, 1998]

- Simplicité et rapidité d'accès.
- Capacité à manipuler les données agrégées selon différentes dimensions.
- Un cube utilise les fonctions classiques d'agrégation : min, max, count, sum, avg, mais peut utiliser des fonctions d'agrégation spécifiques.

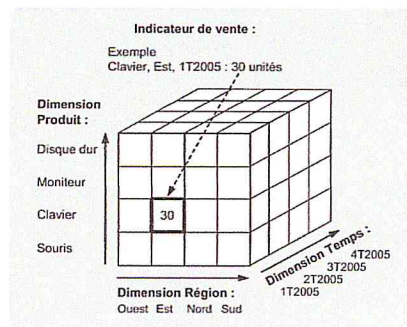


Figure 2.13: Exemple d'un cube de données

2.8.5. Opérations sur les cubes :

Une fois que le serveur OLAP a construit le cube multidimensionnel, plusieurs opérations sont possibles sur ce dernier, offrant ainsi la possibilité de naviguer dans les données qui le constituent. Ces opérations de navigation "Data Surfing" doivent être, d'une part, assez complexes pour adresser l'ensemble des données et, d'autre part, assez simples afin de permettre à l'utilisateur de circuler de manière libre et intuitive dans le modèle dimensionnel. Les opérations possibles dans le cube OLAP sont décrites en ce qui suit.

- **Slice and Dice:**

- **Le Slicing :** en général, cette opération consiste à filtrer une dimension selon une valeur ou une plage de valeurs afin de se concentrer sur une partie de données.

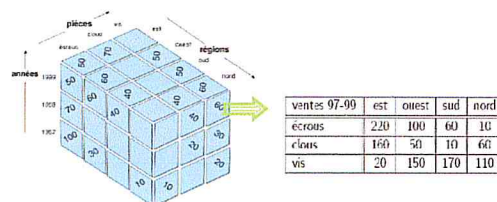


Figure 2.14: : Exemple du Slicing [Epinasse ,2010]

- **Le Dicing :** quant à lui, peut être vu comme étant une extraction d'un sous-cube.

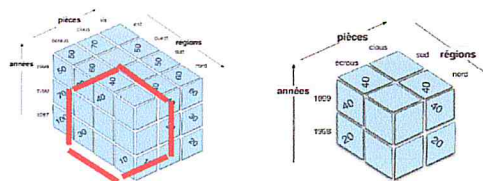


Figure 2.15: Exemple du Dicing [Epinasse ,2010]

- **Drill-down and Roll-up:**

Ce sont les méthodes les plus répandues pour une navigation dans un Data Warehouse. Elles consistent à représenter les données du cube à un niveau de granularité inférieur dans le cas du "Drill-down", ou un niveau supérieur, c'est le "Roll-up". En somme,

ces deux opérations permettent de contrôler le niveau de détail des données du cube. Ces opérations ne sont pas aussi faciles à implémenter car elles sont basées sur la notion d'une bonne hiérarchisation des attributs d'une dimension et la différenciation entre tous les niveaux de hiérarchie, disponibles dans les différentes dimensions. [Epinasse ,2010].

- Drill-down :

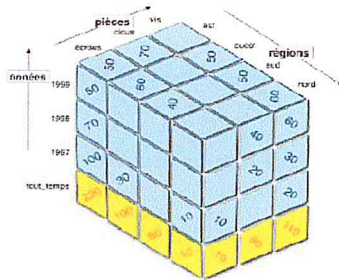


Figure 2.16: Exemple du Drill-down [Epinasse ,2010]

- Roll-up :

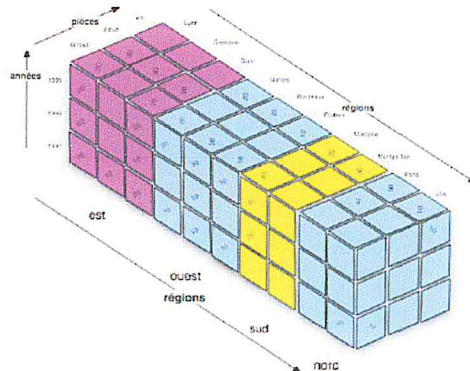


Figure 2.17: Exemple du roll-up [ESPINASSE, 2010]

2.9 Conclusion

A travers cette partie, nous avons abordé le concept du Data Warehouse qui est apparu en réponse à des besoins décisionnels. La construction d'un Data Warehouse passe par plusieurs étapes, à savoir le processus d'alimentation du DW et la modélisation dimensionnelle de ce dernier. La problématique maintenant réside autour de l'analyse des résultats de l'entreposage des données et leurs présentations adéquates aux attentes des utilisateurs finaux afin de leur donner une vision limpide sur les systèmes contrôlés. Pour remédier à cela une autre famille d'outils a

était créée, ce sont les outils de restitution que nous allons évoquer dans la prochaine partie de ce chapitre.

Chapter 3

Les Outils de Restitution de données

Une fois que les données sont transformées, organisées et chargées dans un DW bien conçu, la couche présentation vient apporter d'autres fonctionnalités complémentaires et qui permettent à l'utilisateur final d'accéder aux informations souhaitées. Dans cette partie, nous allons parcourir les outils de restitution en se focalisant sur le tableau de bord.

3.1 Tableau de bord

Un tableau de bord est un instrument de mesure de la performance facilitant le pilotage pro-actif d'une ou plusieurs activités dans le cadre d'une démarche de progrès. Le tableau de bord contribue à réduire l'incertitude et facilite la prise de risque inhérente à toutes décisions. Le tableau de bord est un instrument d'aide à la décision. [Fernandez, 2011] Nous pouvons le définir comme un outil de pilotage qui indique l'état d'avancement d'une ou plusieurs activités de l'entreprise afin d'assister les décideurs et de leur permettre de mettre en place des actions correctives.

3.1.1 Objectifs d'un tableau de bord

- Un tableau de bord permet de suivre l'évolution des résultats de l'activité d'une entreprise à tout instant et d'anticiper sa situation future.
- Il permet également de mesurer l'écart entre les résultats atteints et les objectifs fixés par les décideurs.
- Il améliore le niveau de la communication entre les différents niveaux hiérarchiques.
- C'est un outil de veille qui permet de déceler les nouveaux risques et opportunités pour l'entreprise.

3.1.2 Les indicateurs Clé de performance (KPI)

Un indicateur de performance " KPI " est une mesure ou un ensemble de mesures braquées sur un aspect critique de la performance globale de l'organisation. [Fernandez, 2006]

Selon Lorino, un KPI représente une information devant aider un acteur, individuel ou plus généralement collectif, à conduire le cours d'une action vers l'atteinte d'un objectif ou devant lui permettre d'en évaluer le résultat. [Lorino, 2001]

On en déduit que les KPIs interagissent avec trois composants à savoir : les objectifs de la stratégie, les acteurs qui sont les consommateurs de l'information (indicateurs), et les actions mises en place par ses acteurs pour réaliser ses objectifs comme le montre la figure suivante.

3.2 Reporting

Le terme "Reporting" désigne une famille d'outils de Business intelligence destinés à assurer la réalisation, la publication et la diffusion de rapports d'activité selon un format prédéterminé. Ils sont essentiellement destinés à faciliter la communication de résultats chiffrés ou d'un

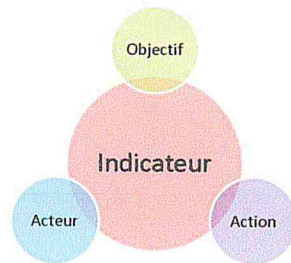


Figure 3.1: Le " triangle " de l'indicateur : stratégie traduite en objectif, processus d'action et acteur collectif [Lorino, 2001]

suivi d'avancement. [Fernandez, 2012]

L'outil offre bien entendu des fonctions spécifiques pour l'élaboration du modèle du rapport, des modules de calcul et de présentation (graphiques) afin de concevoir des comptes rendus particulièrement seyants et pertinents.

3.2.0.1 Objectifs de Reporting

Le Reporting a comme premier objectif de permettre aux utilisateurs finaux de voir les informations.

Nous pouvons résumer le rôle des outils de Reporting en trois points :

- La collecte et la présentation des données prêtes à être analysées, y compris les données historiques qui peuvent être suivis dans le temps.
- Autonomiser les utilisateurs finaux avec les connaissances nécessaires pour devenir des experts dans leur domaine d'activité.
- Avoir les chiffres sous-jacents pour sauvegarder les actions et expliquer les décisions.

3.3 Data Mining

Le Data Mining, appelé fouille de données en français, est l'analyse d'un ensemble (généralement vaste) de données des observations qui vise à trouver des relations insoupçonnées et de les résumer d'une manière à ce qu'elles soient plus compréhensible et utile pour l'utilisateur [Hand et Al, 2001]. Autrement dit, il consiste à extraire des informations pertinentes et des connaissances à partir d'un entrepôt de données en se basant sur plusieurs algorithmes et méthodes d'exploration de données.

3.3.1 Objectifs de Data Mining

Les objectifs du data mining sont par définition déterminés en fonction des données existantes et des problématiques de l'entreprise mais tout projet de fouille de données en général a pour objectifs l'un des suivants : [Web5]

- Identifier les clients fidèles, les clients rentables, ceux à fort potentiel et inversement.
- Accroître la fidélité en étant capable de proposer la bonne offre, au bon client, au bon moment, par le bon canal.
- Calculer le coût d'acquisition d'un nouveau client.
- Réduire les risques d'impayés par une meilleure identification du risque tout au long du cycle de vie du client.

3.4 Conclusion

Nous avons clôturé la partie de synthèse bibliographique par la présentation des outils qui se situent au bout de la chaîne décisionnelle, à savoir : le Data Mining, les rapports et les tableaux de bord. Le rôle majeur de ces outils étant de présenter les informations contenues dans le DW d'une manière adéquate aux attentes décisionnelles des utilisateurs finaux. A travers cette recherche bibliographique, qui a été hautement bénéfique pour nous, nous avons pu augmenter notre maîtrise en termes de méthodes et technologies du Business Intelligence. Par conséquent, nous allons essayer d'appliquer ces connaissances durant notre projet.

Part II
Partie Pratique

Chapter 4

Étude de l'existant

Après avoir fait une synthèse bibliographique sur les systèmes d'information décisionnels, il est temps de présenter l'entreprise qui nous a accueillis pour réaliser notre stage dans le cadre du projet de fin d'études. Nous commençons ce chapitre par la présentation de l'organisme d'accueil l'institut " El Maarifa " en détaillant ses structures, ses services et ses activités.

4.1 Présentation de l'organisme d'accueil

4.1.1 Principales missions de l'institut El Maarifa :

L'institut El Maarifa a comme principales missions:

- Fournir de la formation de haut niveau dans les domaines du management, des langues et du développement humains, respectant les standards internationaux et garantissant à l'apprenant l'acquisition de compétences avérées lui ouvrant de nouvelles perspectives à sa carrière.
- Offrir des formations de qualité supérieure et qui répondront aux besoins et exigences de leurs clients sur tous le territoire national, et cela à travers le respect de leurs valeurs d'efficacité, d'engagement et de qualité.

4.1.2 Organigramme de l'institut :

Le schéma ci-dessous illustre la structure générale de l'IM:

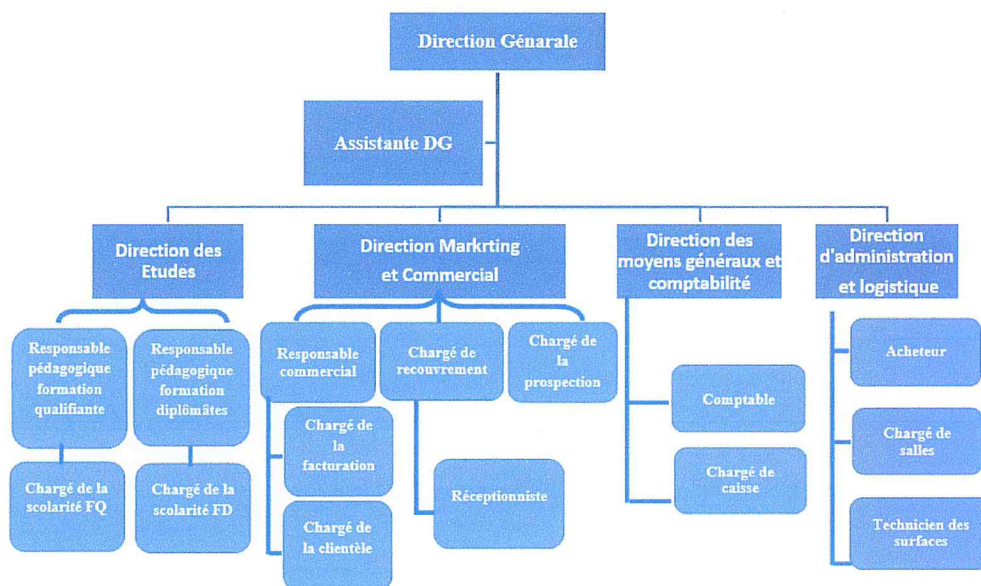


Figure 4.1: Organigramme de l'institut El Maarifa

4.1.3 Les directions principales de l'institut :

Direction des Etudes:

La direction des études est une direction d'administration des services

centraux de l'établissement. Elle est chargée d'assurer les missions dédiées à la mise en œuvre opérationnelle et au pilotage de la formation au sein de l'établissement, dans une logique de service.

Direction de Marketing et Commercial:

La direction du Marketing et Commerciale a pour mission principale de mettre en œuvre un plan marketing et commerciale de consolider l'image de marque de l'établissement. Parmi ces principales missions nous citons:

- Consulter fréquemment les collaborateurs et s'informer des tendances du marché.
- Assure le suivi des transactions commerciales.
- Détermine les actions à mener.

Direction des Moyens Généraux et de la Comptabilité:

La principale fonction de cette direction est le contrôle de l'équilibre financier de l'établissement. Elle permet également de mettre en œuvre la politique de gestion financière et comptable de l'institut et de procéder à des analyses financières périodiques.

Direction de l'administration et logistique:

La Direction d'Administration et Logistique assure d'une part une mission de soutien à la pédagogie (organisation des séquences pédagogiques et examens, équipement des locaux pédagogiques et salles), et d'autre part, elle veille à la bonne exécution des marchés qui lui sont confiés.

Nous précisons que notre stage s'est effectué au niveau de la "Direction des Etudes des Formations Qualifiantes".

4.2 Présentation du système opérationnel existant

Actuellement, l'établissement El Maarifa n'utilise aucun système d'aide à la décision. Tout processus de prise de décision se base sur des rapports. Ces rapports sous forme de fichiers plats (Excel ou autre) sont élaborés pour répondre aux besoins décisionnels souhaités et pour pouvoir analyser la situation des activités commerciales et marketing de l'IM par les responsables. Ces rapports sont issus des travaux d'analyse que les marketeurs et commerciaux réalisent souvent manuellement. Les tableaux de bord demandés sont aussi élaborés sous Microsoft Excel.

4.2.1 Etude des sources de données

Dans le but de bien comprendre les données du groupe en vue de leur intégration, nous avons bien manipulé leur principal fichier source afin

de comprendre la liaison entre les différentes tables et interroger ces derniers de façon à déterminer les principales relations entre ses tables. Le système d'information mis en place au niveau de l'établissement prend en charge la saisie et le stockage des différentes informations concernant les demandes reçues, les formations proposées et l'ensemble des clients afin d'analyser fréquemment ces données pour une gestion et pilotage plus efficace.

Le fichier Excel de suivi des offres est structuré comme suit :

- **Feuille 1 du fichier Excel:** Suivi des formations de type "Intra-entreprises"

Dans cette feuille, chaque ligne de la matrice désigne une demande d'une entreprise pour une formation.

- **Feuille 2 du fichier Excel:** Suivi des formations de type "Inter-entreprises"

Dans cette feuille, chaque ligne de la matrice désigne l'intitulé de la formation avec un éventuel code, le nom du client concerné, le total inscrits, et le mois de la formation.

- **Feuille 3 du fichier Excel:** Chiffres prévisionnels

Cette feuille est un ensemble de tableaux, chaque tableau décrit un client avec l'ensemble des formations effectuées et le montant global des formations.

- **Feuille 4 du fichier Excel:** Agenda

Cette feuille comme son nom l'indique, est un agenda des formations offertes par mois, avec la durée de chaque formation, sa date début et fin, et son prix unitaire.

4.3 Critique de l'existant

Le diagnostic est nécessaire avant tout développement d'une application, et toute prise de décision au sein d'une organisation. Ce diagnostic permettra de dépister les différentes anomalies constatées et dégager les insuffisances dans le fonctionnement du système existant. Après l'étude de l'existant de la direction des Etudes de l'établissement El Maarifa; nous avons pu identifier les insuffisances qui caractérisent le système actuel, ainsi, nous citons :

- Un manque de données sur les formations diplomates;
- Les valeurs de certains attributs ne sont pas toujours renseignées;

- Difficultés dans l'élaboration des rapports d'activité à cause d'insuffisance de données (non saisie de certaines données telles que les renseignements concernant les particuliers).

4.4 Conclusion

A travers la présentation de l'établissement, nous avons voulu donner une idée claire sur sa structure organisationnelle et son métier. Ensuite, nous avons décrit l'existant informatique au sein de ce dernier afin d'avoir une vision sur son état décisionnel. Dans le chapitre suivant nous allons identifier les besoins des utilisateurs finaux de notre système

Chapter 5

Analyse des besoins

L'analyse des besoins est la phase pendant laquelle on identifie les attentes du client afin de le satisfaire. Dans un projet décisionnel, cette étape est importante pour assurer la réussite du projet. D'où l'intérêt de ce chapitre dans lequel nous allons présenter notre démarche de collecte des besoins et la liste des différentes analyses décisionnelles définies par les utilisateurs de notre système décisionnel.

5.1 Collecte d'informations

Afin de recenser les besoins analytiques des utilisateurs de l'Institut El Maarifa, nous avons opté pour les trois techniques : la documentation, l'analyse des processus fonctionnels et techniques existants et les réunions car elles assurent des informations fiables et valides.

5.1.1 La documentation

Elle nous a permis de percevoir une idée claire sur la gestion des formations au sein de l'établissement des techniques modernes, ces différents services et processus internes, notamment des documents qui décrivent l'organisation de l'établissement, et des documents de tableaux de bords actuels qui décrivent les différents tableaux de bord et les indicateurs mis en place.

5.1.2 L'analyse des processus fonctionnels et techniques existants

Cette analyse des processus de système opérationnel, nous a permis de comprendre de plus en plus le fonctionnement de système actuel : ces entrées/sorties, ces capacités, ces limites, et les besoins fonctionnels qu'il couvre, ainsi que le parc informatique derrière : les sources de données, les applications de saisie, et les besoins techniques qu'il assure.

5.1.3 Les réunions

Afin de recueillir des informations détaillées et actualisées sur le fonctionnement de la gestion des formations de l'IM, et de mieux éclaircir les besoins espérés, nous avons prévu des réunions avec certains cadres et employés de la direction des études.

5.2 L'analyse des besoins analytiques:

5.2.1 Identification des tableaux de bord :

Au cours des démarches d'analyse et de collecte d'information, nous avons pu recenser les tableaux de bords exigés par les utilisateurs. Les TDB recensés contiennent des indicateurs par rapport à des axes bien déterminés: Le temps, les formations, le suivi, et des autres spécifications liées aux clients notamment: le type de client, région, domaine d'activité... etc.

5.2.2 Identification des rapports

Les utilisateurs entretenus nous ont exprimé les besoins suivants :

- Un rapport des segments clients, leurs secteurs d'activités et leurs agences;
- Un rapport sur les formations les plus exigées sur l'ensemble des clients;
- Un rapport sur le nombre de formation proposée, les formations clôturées par type de client (entreprise, professionnel et particulier);
- Un rapport sur le nombre d'inscrits dans chaque formation;
- Un rapport des clients acquis et perdus et leurs causes de perte.
- Bilan des formations.

5.3 L'analyse des besoins techniques

Afin de répondre à ces exigences, certaines spécifications techniques que le futur système doit respecter ont été négociées avec l'établissement.

- Le système doit permettre l'accès facile aux données, à tout moment.
- Le système doit se connecter à la base de données existante.
- Le système doit répondre aux différentes requêtes dans le plus petit temps de réponse possible.
- Le système doit être flexible et paramétrable.
- Le système doit garder l'historique des données, et générer les archives.

5.4 Solution proposée

Au bout de notre étude de l'existant et l'expression des besoins, nous avons pu proposer une solution qui va répondre aux différentes exigences recensées. Notre solution repose sur nos constats essentiellement ceux qui concernent les sources de données (des fichiers Excel), transformés selon les besoins afin de les charger dans notre entrepôt de données. Et enfin d'offrir aux utilisateurs l'accès aux différentes informations sous plusieurs façons de présentations afin d'interagir chacun selon son rôle. La solution proposée doit répondre aux différents besoins recensés, donc elle doit être la plus adéquate à leurs exigences ; c'est la solution décisionnelle avec ces différents instruments et offres que nous allons présenter par la suite.

La solution proposée est illustrée par le schéma ci-dessous :

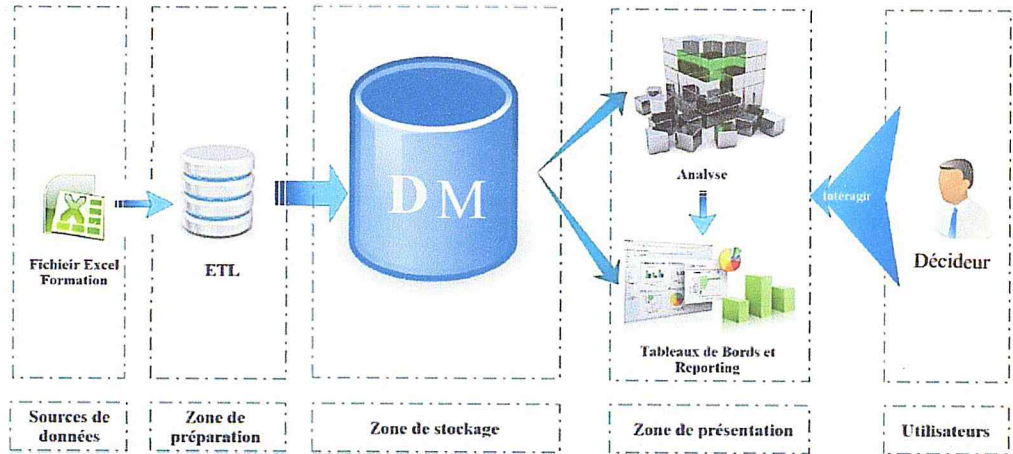


Figure 5.1: Architecture de la solution proposée

Ce système est composé de plusieurs zones :

- **La zone des sources de données:**

Cette zone représente là où notre système est alimenté par les données opérationnelles brutes. Elle contient la source de données stockée sous forme :

Des Fichiers Excel: ce sont des fichiers standardisés sous formats Excel, ils sont adaptés pour chaque type de traitement et services, afin de stocker les données nécessaires des formations et des clients. La figure suivante est une prise d'écran d'un de nos fichiers sources:

N°	Client	Type	Date de reception	Date de réponse prévu	Date de réponse réelle	Suivi						Montant de l'offre	Réponse				
						Etape 1	Etape 2	Etape 3	Etape 4	Etape 5	Etape 6		Etat	Raison	Part du n		
Janvier																	
1	Client 1	Demande	5-janv.	5-janv.	5-janv.	Of	5-janv.	Relance	24-janv.	Relance	7-févr.	Relance	7-mars	226 840,00			
2	Client 2	FD	11-janv.	12-janv.	12-janv.	Of	12-janv.							1 254,00	Retenu	On est les	
3	Client 3		30-déc.	3-janv.	11-janv.	Of	11-janv.	Relance	24-janv.	Relance				11 034,00	Non retenu	Pas de rés	
4	Client 4	Consultation	30-déc.	3-févr.	11-janv.	Of	11-janv.	Relance	24-janv.	Relance	14-févr.			59 920,00	Retenu	CV	
5	Client 5	Demande	12-janv.	14-janv.	13-janv.	Réponse	13-janv.									Non répondu	on n'assu
6	Client 6	Demande	17-janv.	19-janv.	01-févr.	Of	1-févr.	Relance	29-févr.	Relance	7-mars			6 420,00			
7	Client 7	Demande	15-déc.	17-déc.	30-déc.	Of	30-déc.	Of	18-janv.	Relance	7-févr.			182 466,00	Retenu		
8	Client 8	Demande			26-janv.	Of	26-janv.	Relance	7-févr.					130 590,00	Non retenu	Ils ont tra	
9	Client 9	FD	17-janv.	29-janv.	28-janv.	Of	28-janv.	Relance	31-janv.					12 840,00	Non retenu	pour des r	
Février																	
10	Client 10	FD	1-févr.	3-févr.	2-févr.	Of	2-févr.	Relance	7-févr.					1 070,00	Non retenu	Formation déjà cor	
11	Client 11	Demande	2-févr.	4-févr.	10-févr.	Of	10-févr.	Relance	11-févr.					31 560,00	Retenu		
12	Client 12	Demande	2-févr.	4-févr.	16-févr.	Of	16-févr.	Relance	22-févr.					133 392,00	Retenu		

Figure 5.2: Exemple des fichiers sources Excel

- **La zone de préparation des données**

C'est la zone critique de notre système. Elle est sectionnée en deux parties : la première c'est la partie qui s'occupe de la charge, la consolidation et la centralisation des données des fichiers Excel à partir de la première zone des sources de données, afin de les dresser pour la deuxième partie qui est la partie des traitements ETL. Durant cette partie, des opérations d'extraction, de traitement et de chargement sont exécutés, afin de préparer les données selon le besoin pour l'entreposage dans la zone suivante.

Afin de disposer d'une source de données homogène, les données provenant de données des fichiers Excel des différentes formations et demandes seront chargés à l'état brut dans une zone de préparation grâce au processus ETL. Cette zone de préparation (Staging Area) permet d'extraire les données source sans aucune transformation, et ce pour ne pas encombrer le système opérationnel.

Remarques:

- Notre zone de préparation prend la forme des tables relationnelles dans une base de données.
- Le Staging Area ne représente qu'une zone de passage de données.

- **La zone de stockage des données**

Elle représente le magasin de données où les données traitées sont centralisées, consolidées, historiées, normées, intégrées et prêtes pour les différentes analyses afin de répondre aux différents besoins des utilisateurs à travers la dernière zone.

- **La zone de Présentation**

Cette zone interagit directement avec l'ensemble des utilisateurs, elle représente la zone de restitution des données en instrument de présentation abstraite : des indicateurs dans des tableaux de bords et des rapports, ainsi que les différentes analyses. Tout ça, va permettre à l'ensemble des utilisateurs d'effectuer des visualisations des données de façon facile et sélective et selon des différents points de vue.

5.5 Conclusion

L'analyse des besoins nous a permis de définir l'ensemble des fonctionnalités requises dans notre nouveau système décisionnel. Nous avons

proposé un système décisionnel basé sur un entrepôt de données dont nous allons définir le schéma conceptuel dans le prochain chapitre. Finalement, nous avons présenté notre solution générale, où nous avons recensé les principaux aspects et les parties globales de notre solution, afin de les détailler et les expliquer dans les prochains chapitres.

Chapter 6

Conception du système décisionnel

Après avoir étudié et analysé les besoins des utilisateurs pour lesquels notre système décisionnel doit répondre, la conception de ce dernier est nécessaire avant sa réalisation. Cette partie est consacrée pour la conception de l'entrepôt de données, que nous avons effectué en trois étapes :

1. **La conception de la base de données** : consiste à définir les tables constituant la base de données ainsi qu'un diagramme de classe.
2. **La modélisation de l'entrepôt** : consiste à définir le schéma conceptuel de l'entrepôt (les tables de faits, les tables de dimensions,...).
3. **Le système du tableau de bord et reporting.**

6.1 Conception de la base de données

Etant donné que l'institut " El Maarifa " ne dispose pas d'une base de données, et qu'il utilise encore un système non structuré pour gérer ses activités, ce qui laisse son rendement moins efficace, élaborer un rapport de gestion des formations assurées dans ces conditions coûte énormément en temps et en énergie.

Système non-structuré et, par conséquent, lenteur, lourdeur de ce dernier, tel est le problème majeur auquel est confronté la réalisation d'une base de données (BDD IM) qui permettra d'optimiser la gestion des informations et d'assurer une meilleure analyse pour alléger les difficultés que rencontrent les responsables de gestion de formation qualifiante en termes d'organisation des informations relatives aux clients.

Cette base de données jouera le rôle d'un "Staging Area", c'est à dire un ensemble de tables qui représentent une copie conforme de la source de données dont le but principal est le passage et la transformation dans une architecture décisionnelle.

Avant d'entamer la conception de cette base, nous précisons que le fichier source tel qu'il est fourni ne peut pas être exploité, ce qui nous a mené à proposer une nouvelle conception de ce dernier qui servira par la suite lors du chargement de ses données.

6.1.1 Choix de langage de modélisation

Plusieurs langages de modélisation ont fait leur apparition, nous avons choisi la méthode UML, parce qu'elle est devenue un standard de modélisation des systèmes informatiques. [O.Guibert]

6.1.2 Définition de la méthode UML

UML (Unified Modeling Language ou " langage de modélisation unifié ") est un langage de modélisation graphique à base de pictogrammes. Il est apparu dans le monde du génie logiciel, dans le cadre de la "conception orientée objet ". Couramment utilisé dans les projets logiciels, il peut être appliqué à toutes sortes de systèmes ne se limitant pas au domaine informatique.[Web6]

6.1.3 Modélisation

La richesse d'UML réside dans le nombre de diagrammes qu'il permet construire, mais il n'est pas nécessaire d'utiliser tous les diagrammes pour modéliser un système. En général deux ou trois diagrammes suffisent, le plus important est le diagramme de classes. .[Web6]

6.1.3.1 Définition de diagramme de cas d'utilisation

Les diagrammes de cas d'utilisation sont des diagrammes UML utilisés pour donner une vue du système dans son environnement extérieur. Ils représentent la structure des grandes fonctionnalités nécessaires aux utilisateurs du système appelés "acteurs". Ils permettent aussi de définir les limites du système.

Un cas d'utilisation (Use Case) représente graphiquement un ensemble de séquences d'actions et de fonctionnalités qui sont réalisées par le système et qui produisent un résultat observable, intéressant pour un acteur particulier et son interaction avec le système. [Web7]

- **Identification des utilisateurs finaux du nouveau système :**

A travers plusieurs réunions avec le responsable de la direction des Etudes, nous avons pu repérer les utilisateurs potentiels du nouveau système. Ces utilisateurs sont exclusivement des employés du service commercial et formation qualifiante.

Les utilisateurs du nouveau système peuvent être rassemblés en trois principaux groupes qui sont:

- **Administrateur du système décisionnel :** celui qui s'occupe de la gestion et de l'administration du système décisionnel au niveau de l'institut " El Maarifa " .
- **Dirigeants :** représenté par le directeur du service de formation qualifiante et celui du service commerciale et marketing.
- **Utilisateurs métiers :** représenté par certains agents commerciaux de la direction commerciale.

- **Diagramme de cas d'utilisation général:**

Nous présenterons dans la figure suivante les principaux besoins fonctionnels de notre système:

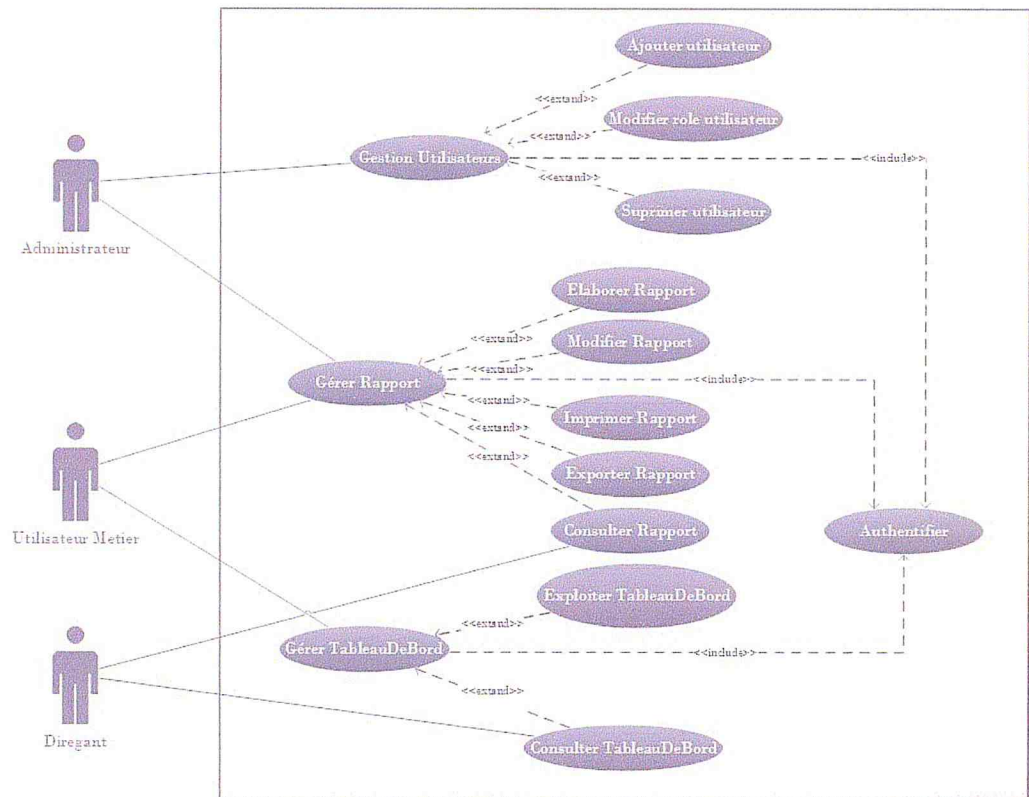


Figure 6.1: Diagramme de cas d'utilisation général

- **Description des cas d'utilisation :**

Le tableau ci-dessous contient les principales fonctionnalités de l'administrateur :

Use Case	Description
Gestion des utilisateurs	Ajouter utilisateur: Permet de rattacher un nouvel utilisateur
	Modifier utilisateur: Permet de mettre à jour les droits d'accès d'un utilisateur.
	Supprimer utilisateur: Permet de faire disparaître un utilisateur

Figure 6.2: Les spécifications de l'administrateur

Le tableau ci-dessous concerne les fonctionnalités des utilisateurs :

Use case	Description
Gestion des Rapports	Elaborer Rapport: Permet de créer un nouveau rapport.
	Modifier Rapport: Permet de mettre à jour un rapport déjà créé.
	Imprimer Rapport: En consultant un rapport, le décideur peut éventuellement l'imprimer.
	Consulter Rapport: Afficher le contenu d'un rapport déjà créé.
	Exporter Rapport: En consultant un rapport, le décideur peut également l'exporter en Excel, Pdf ou autre.
Gestion des Tableau de bord	Consulter TDB: Consulter certains indicateurs proposés par le tableau de bord.
	Exploiter TDB: Permet une synthèse rapide des actions en cours et fournit un outil d'aide au système.

Figure 6.3: Les spécifications des utilisateurs

6.1.3.2 Définition de diagramme de classe:

Un diagramme de classes UML décrit les structures des objets et des informations utilisés, à la fois en interne et en communication avec ses utilisateurs. Ses classes et relations peuvent être implémentées de différentes manières.[Web9]

Un diagramme de classe proprement réalisé permet de structurer le travail de développement de manière très efficace et permet de construire le système de manière correcte. Nous présenterons dans la figure 22 notre diagramme de classes.

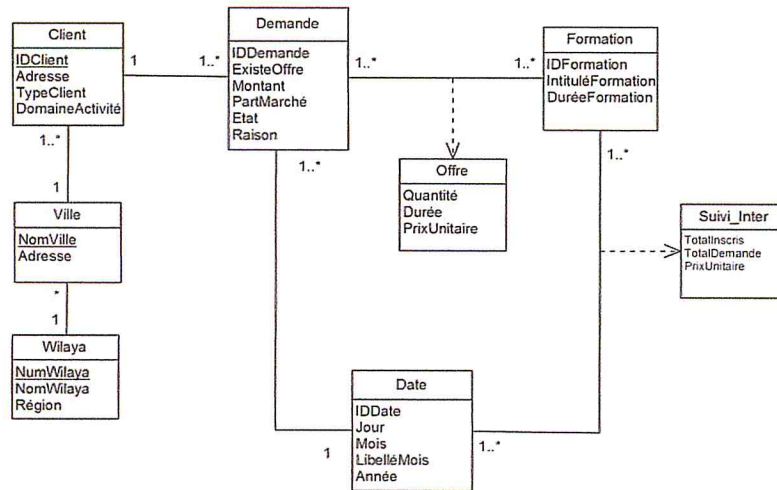


Figure 6.4: Diagramme de classe

- **Modèle relationnel de la base de données:**

Après avoir modélisé le schéma de la base de données en utilisant la modélisation UML, nous passons au modèle relationnel afin de simplifier la structure de données et de fournir un moyen plus proche de l'implémentation dans le SGBD.

Pour ce faire, nous avons suivis les règles du passage et nous avons pu produire le modèle relationnel suivant : Client (IDclient, Adress, TypeClient, DomaineActivité, Ville)

Ville (NomVille, Adress*, Numwilaya*)

Willaya (NumWilaya, NomWilaya)

Demande (IDDemande, , IDclient*, DateRecep*, ExistOff, Montant, PartMarché, Etat, Raison)

Offre (IDClient*, IDFormation*, Prixunitaire, Quantité)

Formation (IDFormation, IntituléFormation, Durée)

SuiviInter (IDFormation*, Date*, Totalinscrist, TotalDemande, Prix)

Date (IDDate, Jour, Mois , LibelléMois , Année)

6.2 Modélisation de l'entrepôt de données

Cette deuxième partie va concerner la modélisation du système décisionnel qui va être implémenté à travers un mini entrepôt. Pour la modélisation de ce dernier, nous avons suivi l'architecture proposée par Kimball qui consiste à créer les magasins de données selon les indicateurs, ces magasins vont être conçus selon l'approche ROLAP pour les nombreux avantages qu'il offre.

6.2.1 Méthode de modélisation

Nous décrivons dans le tableau ci-dessous une méthodologie par étapes pour construire l'entrepôt de données, cette méthode a été initialement proposée par Kimball et s'appelle méthodologie à neuf étapes dans la modélisation d'un entrepôt de données :

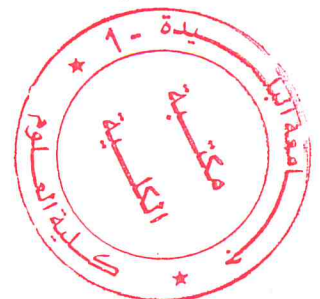
Etape	Activité	Description de l'activité
1	Choisir la procédure	Faire référence au sujet d'un magasin de données particulier
2	Choisir le grain	Décider de ce que représente un enregistrement d'une table de fait.
3	Identifier les dimensions et s'y conformer	Déterminer le contexte dans le quel nous pourrons poser des questions à propos des faits établis dans la table de fait.
4	Choisir les mesures	Déterminer les faits utilisables.
5	Emmagasiner les calculs préliminaires dans la table des faits	Réexaminer les faits choisis pour déterminer si des opportunités apparaissent d'exploiter les calculs préliminaires.
6	Finaliser les tables de dimensions	Ajouter les descriptions textuelles possibles qui seront intuitives et compréhensibles.
7	Choisir la durée de la BDD	La durée mesure le saut dans le passé qu'une table de fait permet d'effectuer.
8	Suivre les dimensions a modification lente	Appliquer une clé généralisée aux dimensions importantes afin de distinguer les multiples instantanées.
9	Décider des priorités de requêtes et des modes de requêtes	Nous prenons en considération les soucis liés au design physique.

Figure 6.5: Méthodes de modélisation d'un entrepôt

Pour proposer un schéma conceptuel de l'entrepôt du système décisionnel, nous nous sommes basés sur les besoins exprimés par les utilisateurs de l'institut et ceci selon l'analyse présentée dans le chapitre précédent.

6.2.2 Le schéma en flocon de neige du magasin "Suivi-Intra"

Le schéma suivant présentera le modèle en flocon de neige du magasin concernant la réalisation des formations de type intra-entreprise au sein de l'institut. Les mesures définies sont: Le nombre d'inscrits ainsi que le montant des formations. Les dimensions de ce magasin sont: Client, Demande, Formation, et la dimension temporelle Date.



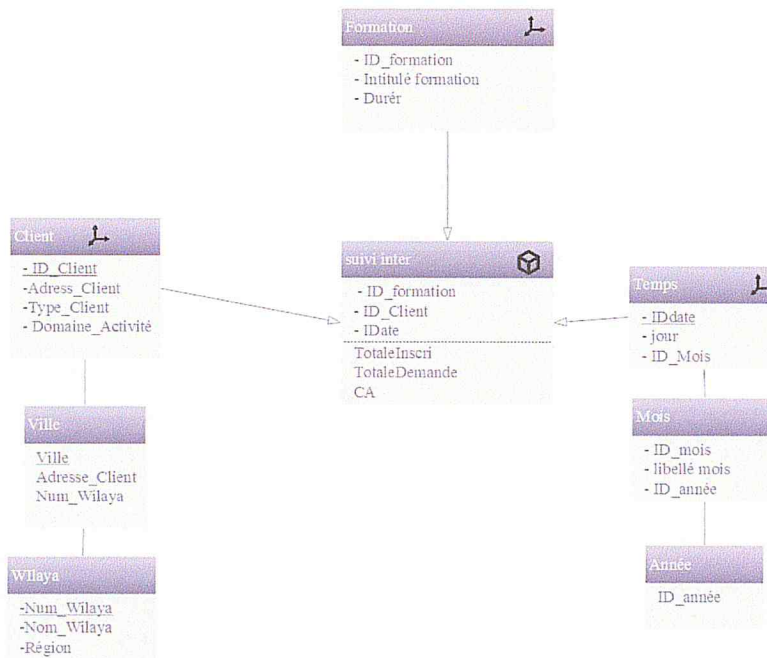


Figure 6.6: Schéma en flocon de neige du magasin "Suivi-Intra"

6.2.3 Le schéma en flocon de neige du magasin "Suivi-Inter"

Le schéma ci-dessous représente le magasin "Suivi des formations inter-entreprise". Les dimensions de ce magasin sont: Client, Formation, et la dimension temporelle Date. Les mesures sont : chiffre d'affaire, total inscrits et total demandes.

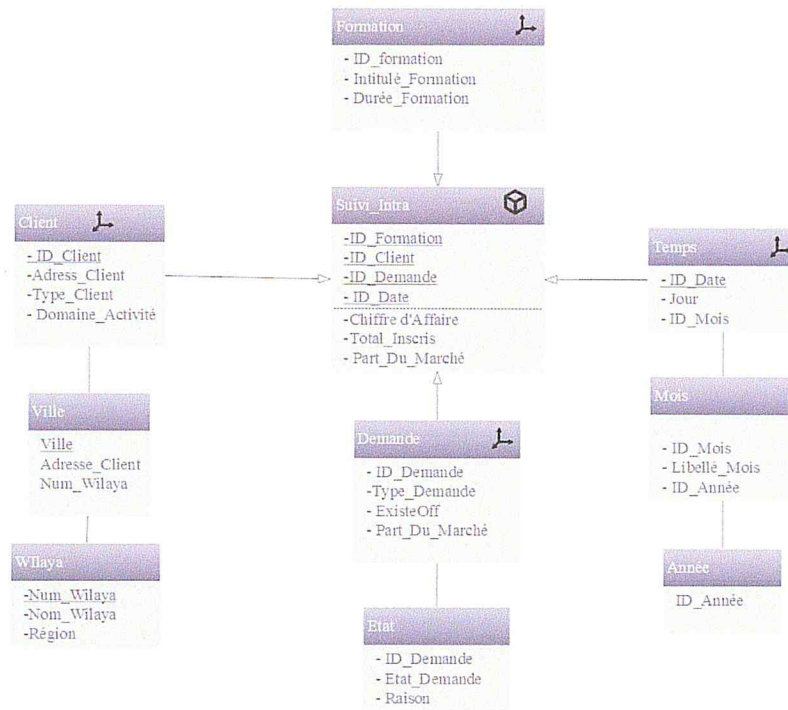


Figure 6.7: Schéma en flocon de neige du magasin "Suivi-Inter"

6.2.4 Les dimensions du système

Nous présentons dans ce qui suit en détail la signification de chaque dimension ainsi que ses attributs et leurs désignations.

- **Dimension Formation:**

Dimension Formation	C'est la dimension qui désigne l'ensemble des formations offertes par l'institut	
Nom d'attribut	Désignation	Ex. Valeurs
Intitulé_Formation	Clé de substitution de la dimension	
Durée_Formation	La durée de la formation	5j, 10j, 1m,.....
Prix_Unitaire	Le prix unitaire de la formation	8500 DA, 11000 DA,.....etc

Figure 6.8: Tableau descriptif de la dimension "Formation"

- **Dimension Client:**

Dimension Client	C'est la dimension qui désigne le client intéressé par une formation	
Nom d'attribut	Désignation	Ex. Valeurs
ID_Client	Clé de substitution de la dimension	Carama Assurance, SEAAI, Ministère de la communication, ...etc.
Type_Client	Le type client	Etatique, Privé.
Domaine_d'activité	Le domaine d'activité	Hydraulique, Pétrolière, Financière,....etc.
Adresse_Client	L'adresse du client	13, Rue khoudjet el Djeld, Les sources, Bir Mourad Rais, Alger

Figure 6.9: Tableau descriptif de la dimension "Client"

- **Dimension Demande:**

Dimension Demande	C'est la dimension qui désigne l'ensemble des demandes que l'institut a reçu.	
Nom d'attribut	Désignation	Ex. Valeurs
ID_Demande	Clé de substitution de la dimension	1,2,3.....
Type_Demande	Demande avec ou sans Restauration	0,1.
Existe_Offre	S'il existe une offre ou pas	0,1 (oui/non)
Unité_Offre	L'unité de l'offre	Groupe, Personne

Figure 6.10: Tableau descriptif de la dimension "Demande"

- **Dimension Etat-Demande:**

Dimension_Etat_Demande	C'est la dimension qui désigne l'état des demandes c.à.d. demande retenue ou non retenue	
Nom d'attribut	Désignation	Ex. Valeurs
Etat_Demande	Clé de substitution de la dimension	0,1
Raison	Les raisons pour lesquelles les demandes ne sont pas effectuées.	Formation déjà commencée, pour des raisons personnelles...

Figure 6.11: Tableau descriptif de la dimension "Etat-Demande"

- **Dimension Temps:**

Dimension_Temps	C'est la dimension qui désigne le temps	
Nom d'attribut	Désignation	Ex. Valeurs
ID_date	Clé de substitution de la dimension	05/01/2014
Jour	Le jour	05
Mois	Le mois	01
Libellé mois	Le mois en lettre	Janvier
Année	L'année	2014

Figure 6.12: Tableau descriptif de la dimension "Temps"

- **Dimension Adresse**

Dimension_Adresse	C'est la dimension qui désigne les adresses des clients	
Nom d'attribut	Désignation	Ex. Valeurs
ID_Adresse	Clé de substitution de la dimension	13, Rue khoudjet el Djeld, Les sources, Bir Mourad Rais, Alger
Ville	Nom de la Ville	Alger
Wilaya	Nom de la Wilaya	Alger

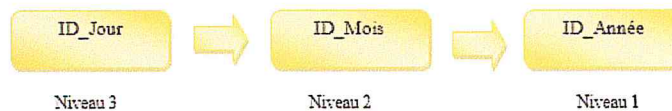
Figure 6.13: : Tableaux descriptif de la dimension "Adresse"

6.2.5 Niveaux et hiérarchies des dimensions

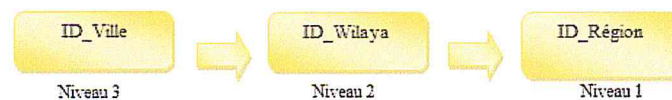
Une dimension est un ensemble de valeurs décomposables. Les valeurs d'une dimension sont généralement organisées à l'intérieur d'une hiérarchie. L'accès au niveau supérieur dans une hiérarchie est appelé "Rolling up", et au niveau inférieur "Drill down".

Dans la partie suivante, nous allons identifier les niveaux des attributs de chaque dimension avec sa hiérarchie correspondante.

- **Dimension Temps:**



- **Dimension Adresse:**



6.2.6 La zone d'alimentation

Une fois que notre entrepôt est conçu, on l'alimente à partir de la base de données (BDD-IM) que nous avons réalisé.

L'alimentation d'un DW est un processus qui s'effectue en plusieurs étapes :

- Sélection des données sources
- Extraction des données
- Transformation
- Chargement

1. Sélection de données:

La source de données de notre entrepôt est la base de données (BDD-IM) que nous avons construit à partir des données du fichier Excel.

2. Extraction des données:

Elle consiste à collecter les données utiles dans le système de production pour rafraîchir la base décisionnelle.

Au premier temps, et à partir du système opérationnel nous avons chargé les données utiles dans une base intermédiaire (base de données Staging), Cette Staging Area contient des requêtes qui vont être exécutées périodiquement à travers un petit script qu'on a implémenté, ces requêtes vont transférer les données situés dans les fichiers sources (Excel) vers notre Base de données.

3. Transformation des données:

Avant d'effectuer les traitements nécessaires sur les données extraites, on est obligé de faire des vérifications sur la qualité de ces dernières.

Les données extraites subissent des transformations avant d'être chargées dans l'entrepôt. Nous distinguerons les opérations de regroupement de données, changement des libellés des champs, création de clés (clés artificielles pour les tables de fait par exemple), calcul des données secondaires ...

4. Chargement de données:

C'est la dernière phase de l'ETL, elle passe par trois étapes :

- **Chargement de la dimension temps :**
Ce chargement s'effectue une seule fois dans le cycle de vie d'un entrepôt de données.
- **Chargement des tables de dimension :**
Ce chargement se fait à partir de la base de données, il nécessite un chargement initial et un chargement incrémental afin d'insérer

les nouvelles données. De ce fait, nous devons établir une politique de chargement permettant de garder les dimensions chargées précédemment car cela permet aux décideurs de faire des prévisions.

- **Chargement des tables des faits :**

Le chargement des tables de fait correspond au chargement des différentes mesures. Il faut s'assurer à ce stade de l'intégrité référentielle, c'est à dire pour chaque clé étrangère dans la table de fait, il doit exister une correspondance dans la table dimensionnelle. Au final, l'identificateur de chaque table des faits sera composé des clés des dimensions participantes au fait.

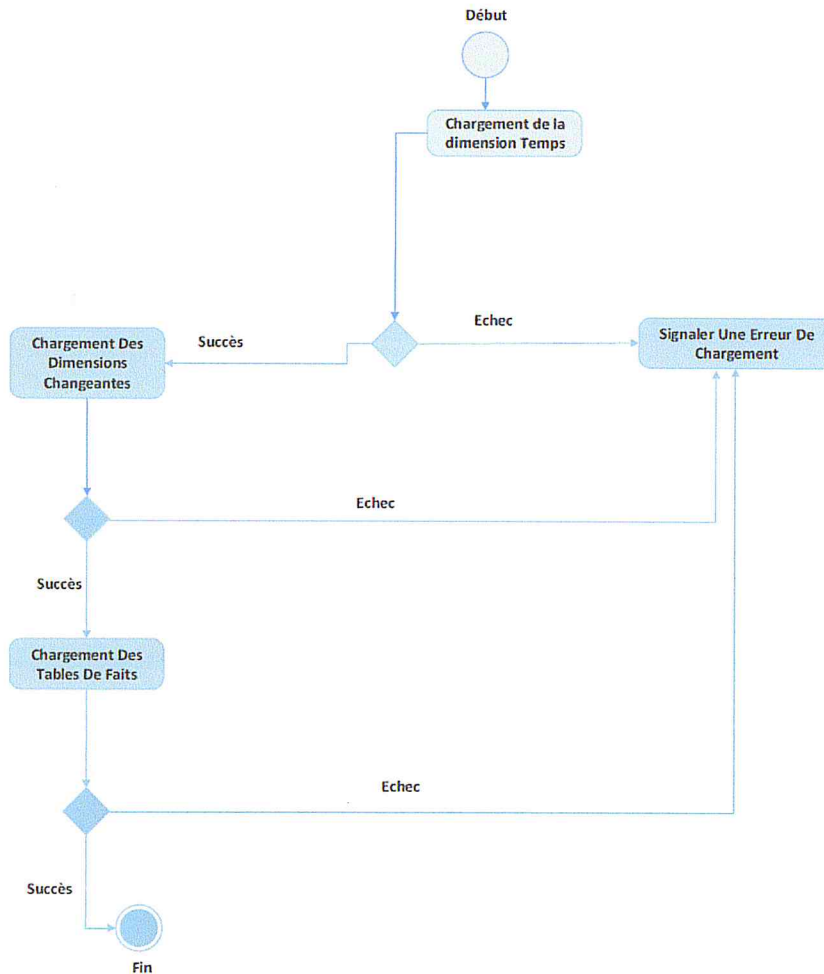


Figure 6.14: Le processus général du chargement du DW

6.2.7 Conception des cubes dimensionnels

Les utilisateurs de notre DW doivent pouvoir accéder aux données dont ils ont besoin, les extraire, les évaluer et les traiter de façon intuitive et rapide, donc il est nécessaire de mettre en place des cubes de données afin de leur permettre de faire des analyses multidimensionnelles complexes sur les données et d'exploiter de manière efficace les informations contenues dans le DW. La conception de ces cubes est basée principalement, sur les besoins recensés auprès des utilisateurs et décideurs de l'entreprise.

6.2.7.1 Niveaux et hiérarchies des dimensions

La hiérarchie d'une dimension est le classement de ses attributs selon une série de relations 1-n suivant plusieurs niveaux ou le niveau " All " est le niveau le plus haut d'une hiérarchie, il représente le niveau le plus agrégé.

La définition des hiérarchies et des niveaux d'une dimension est une étape très importante dans la conception des cubes. En effet, c'est grâce à ces hiérarchies que l'utilisateur peut naviguer et explorer les informations du cube avec le niveau de détails souhaité. Une hiérarchie se compose de plusieurs niveaux, en conséquent nous allons définir, les différents niveaux et leurs attributs, ensuite nous passons aux hiérarchies. Le tableau suivant résume l'ensemble des hiérarchies concernant les dimensions de notre DW:

Dimensions	Attributs	Niveaux	Hiérarchie
Temps	ID_Jour	N1	N1 -> N2 -> N3->ALL
	ID_Mois	N2	
	ID_Année	N3	
Address	ID_Ville	N1	N1 -> N2 -> N3->ALL
	ID_Wilaya	N2	
	ID_Region	N3	
Client	Type_Client	N1	N1-> ALL
Formation	Type_Formation	N1	N1 -> ALL

Figure 6.15: Liste des niveaux et des hiérarchies des dimensions du data warehouse.

6.2.7.2 Les cubes dimensionnels

Cubes	Mesures	Dimesions
Suivi_Inter	TotalDemande	Client
	TotalInscrit	Temps
	CA	Formation
Suivi_Intra	TotaleInscrit	Client
	Part_du_marché	Temps
	Chiffre_d'affaire	Demande
		Formation

6.2.8 Le système du tableau de bord

Les utilisateurs finaux du système décisionnel seront dotés d'un tableau de bord qui leur permettra de faire leurs analyses déjà spécifiées dans le chapitre "Identification des besoins". En effet, il existe plusieurs manières d'implémentation d'un TDB, nous pouvons développer une interface web qui permet d'afficher et de visualiser ses indicateurs, nous pouvons également faire appel à d'autres technologies de navigation tel que les portails intranet.

Nous avons dans notre cas opté pour l'implémentation du TDB dans le portail de l'entreprise.

6.2.8.1 Méthode de construction du TDB :

La conception d'un tableau de bord est une façon de présenter des indicateurs qui aident une entreprise à améliorer sa performance.

Afin de concevoir notre tableau de bord nous avons utilisé la méthode GIMSI. Cette méthode s'inscrit dans une démarche coopérative et propose une approche conceptuellement fondée sur les besoins précis par les décideurs.

- **La méthode GIMSI :**

C'est une méthode complète de mise en œuvre du projet de pilotage de la performance (Business Intelligence) centrée sur l'homme, décideur en situation.

C'est une méthode coopérative de conception de systèmes d'aide à la décision et plus précisément d'assistance au pilotage par tableaux de bord. Elle couvre tous les aspects du projet décisionnel, depuis l'élaboration de la stratégie jusqu'au choix et la mise en œuvre des progiciels. Elle traite notamment les aspects purement technologiques du projet et les questions managériales ou politico-sociologiques propres à l'entreprise [Fernandez, 2013].

- **Les étapes de la méthode GIMSI :**

La démarche est structurée en 10 étapes, chacune traitant une préoccupation particulière du projet et chacune marquant un seuil identifiable dans l'avancement du système [Fernandez, 2008].

Le tableau suivant décrit chacune de ces étapes ainsi que leurs objectifs :

Phase	Etape	Objectif
Identification <i>Quel est le contexte?</i>	1: Environnement de l'entreprise	Analyse de l'environnement économique et de la stratégie de l'entreprise afin de définir le périmètre et la portée du projet
	2: Identification de l'entreprise	Analyse des structures de l'entreprise pour identifier les processus, activités et acteurs concernés.
Conception <i>Que faut-il faire ?</i>	3: Définition des objectifs	Sélection des objectifs tactiques de chaque équipe.
	4: Construction du tableau de bord	Définition du tableau de bord de chaque équipe
	5: Choix des indicateurs	Choix des indicateurs en fonction des objectifs choisis
	6: La collecte d'information	Identification des informations nécessaires à la construction des indicateurs
	7: Le système de tableau de bord	Construction du système de tableaux de bord, contrôle de la cohérence globale
Mise en œuvre <i>Comment le faire ?</i>	8: Choix du progiciel	Elaboration de la grille de sélection pour le choix des progiciels adéquats
	9: Intégration et déploiement de la solution	Implémentation des progiciels, déploiement à l'entreprise
Suivi permanent	10: L'audit du système	Suivi permanent du système

Figure 6.16: Les 10 étapes de la méthode GIMSI, [Fernandez, 2006]

Les tableaux de bord que nous avons réalisés sont de deux types :

1. **Tableau de bord statiques** : les indicateurs et les données affichés sont figés.
2. **Tableau de bord paramétrable** : le décideur a la possibilité de choisir les données à afficher dans son tableau de bord selon le besoin, et il peut descendre dans les détails.

6.2.8.2 Les indicateurs clés de performance (KPI) :

La mise en place d'un tableau de bord nécessite une réflexion approfondie sur les indicateurs à surveiller afin de satisfaire les attentes

des utilisateurs, car l'identification d'un KPI permet, lors de la conception d'un tableau de bord, de déterminer l'objet à mesurer et de préciser la collecte d'indices représentatifs à effectuer.

Le tableau suivant résume les indicateurs à inclure dans ces tableaux de bord ainsi que leurs formules de calcul:

Indicateur	Formule de calcul
Chiffre d'affaire réalisé	Somme (Total Montant)
Client le plus rentable	Max (CA) / Client
Nombre de demande annulée	Somme (Demande non retenues)
Part de marché	Somme (Part de marché) / Somme (Offre) proposée
Taux de satisfaction	Somme (Demandes retenues) - Somme (Demande non retenues)
Formation plus demandée	Max (Inscris) / Formation

Figure 6.17: Les indicateurs du tableau de bord

6.3 Conclusion

Dans ce chapitre, nous avons, dans un premier lieu, détaillé la conception de la base de données (ou le staging area). Par la suite, nous avons présenté une étude de notre système en utilisant la modélisation UML. Nous avons également défini les différentes dimensions du système ainsi que les schémas des magasins. En dernier, nous avons présenté le système du tableau de bord en choisissant sa méthode de construction (GIMSI).

Chapter 7

Réalisation du système décisionnel

Dans ce chapitre, nous entamerons la dernière partie de notre projet qui consiste en la réalisation et la mise en place de notre solution. Pour cela une étude sur la technologie utilisée a été effectuée. Par la suite, nous implémenterons quelques captures d'écrans de l'application.

7.1 Architecture technique du projet

7.1.1 L'environnement technique de la solution:

Machines utilisés :

- **Machine Client** : Pour qu'un utilisateur puisse accéder aux tableaux de bord et visualiser les rapports, il doit disposer d'un navigateur web
- **Machine Serveur** : dans les machines serveur l'environnement suivant est installé : Oracle Business Intelligence Enterprise Edition 11g (OBIEE)

7.1.2 Description des outils utilisés:

L'architecture technique globale de notre système est schématisée dans la figure ci-dessous :

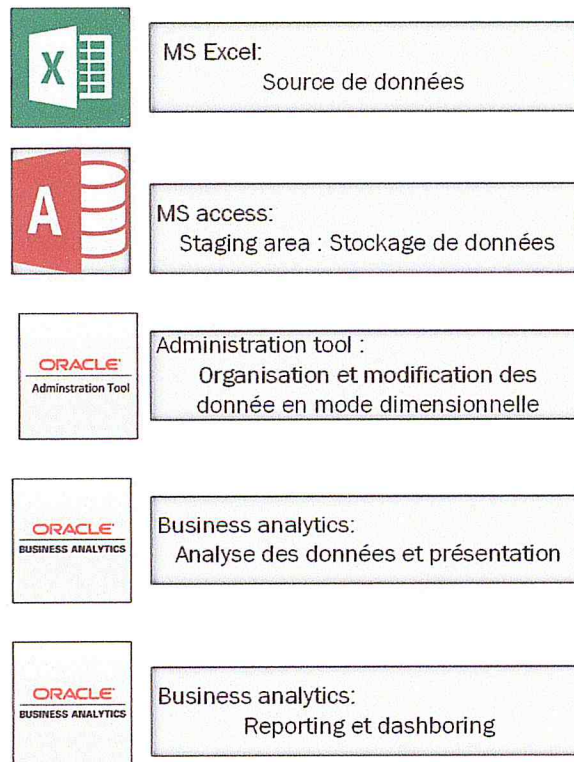


Figure 7.1: Outils utilisés pour la réalisation de notre système

Ces outils seront abordés plus en détail dans le prochain paragraphe.

- **Microsoft Access :**

Microsoft Access (officiellement Microsoft Office Access) est une base de données relationnelle éditée par Microsoft. Elle est composée de plusieurs programmes : le moteur de base de données Microsoft Jet, un éditeur graphique, une interface de type Query pour manipuler les bases de données, et le langage de programmation Visual Basic for Applications.

Nous avons utilisé la dernière version d'Access qui est Microsoft Access 2016 parce qu'elle intègre de nouvelles fonctionnalités dont de nouveaux thèmes, la modernisation des cinq modèles les plus populaires et l'exportation d'informations de sources de données liées vers Excel. [Web10]

- **Oracle Business Intelligence Enterprise Edition 11g :**

Oracle Business Intelligence Suite Enterprise Edition (OBIEE) est une plateforme unifiée et complète, répondant de besoins décisionnels. L'architecture avancée et unifiée, hautement évolutive de OBIEE procure les avantages suivants :

- Une restitution de données s'adressant à tous les utilisateurs de l'entreprise. Chaque métier peut consulter les informations nécessaires à l'atteinte de ses objectifs, sous une forme optimisée par rapport à ses besoins.
- L'intégration avec MS Office rend l'utilisation des informations encore plus performante.
- Une console d'administration centralisée permet la mise en place d'un référentiel unique pour l'ensemble des métiers de l'organisation (objets physiques, modèles logiques, dimensions, indicateurs, une approche du métier adaptable à chaque utilisateur).
- Le portail 100/100 web permet aux utilisateurs d'accéder à un ensemble de fonctionnalités natives de l'outil : tableaux de bords, analyse ad hoc, rapports de masse, alertes proactives, mode déconnecté, exports et publication vers les outils bureautique (Office et Adobe).

1. **Architecture de Oracle BI:**

2. **Oracle BI Presentation Services :** Fournit une interface permettant de présenter les données décisionnelles aux clients Web. Il s'agit d'un processus autonome qui communique avec Oracle BI Server via ODBC (Open Database Connectivity) sur TCP/IP. Il se compose d'un ensemble de clients de requête qui inclut Analysis Editor et Interactive Dashboards.

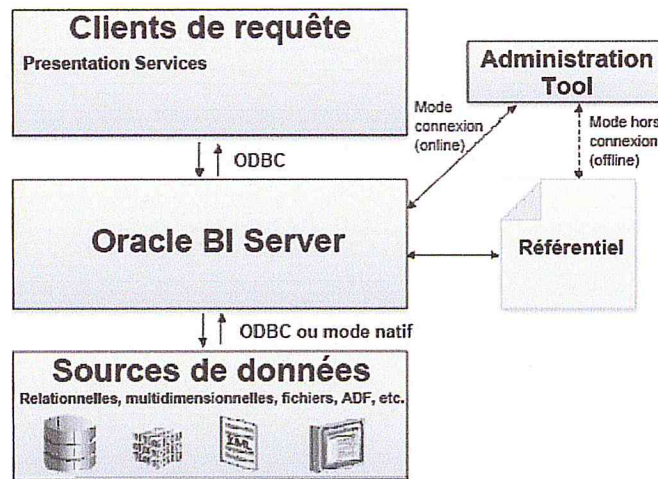


Figure 7.2: Architecture d'OBIEE

3. **Oracle BI Server:** C'est le serveur central de l'architecture Oracle Business Intelligence. Il assure un traitement efficace permettant l'accès aux sources de données physiques et la structuration intelligente des informations.
4. **Sources de données:** Elles contiennent les données que les utilisateurs veulent analyser. Elles sont accessibles via Oracle BI Server.
5. **Référentiel Oracle BI :** Il sert à stocker les métadonnées utilisées par Oracle BI Server.
6. **Oracle BI Administration Tool :** Cette interface présente le référentiel Oracle BI en trois volets distincts appelés couches :
 - Physique;
 - Business Model and Mapping (BMM);
 - Presentation.

La figure suivante représente l'interface Administration Tool de notre système :

7.1.3 La couche Physique:

Elle contient les objets représentant la source de données physiques à laquelle Oracle BI Server envoie les requêtes. Dans notre cas, la source est la base de données (BDD-IM) réalisée avec MS Access. La base de données est l'objet de plus haut niveau de la couche Physical. Chaque objet de base de données de la couche Physique contient : un pool de

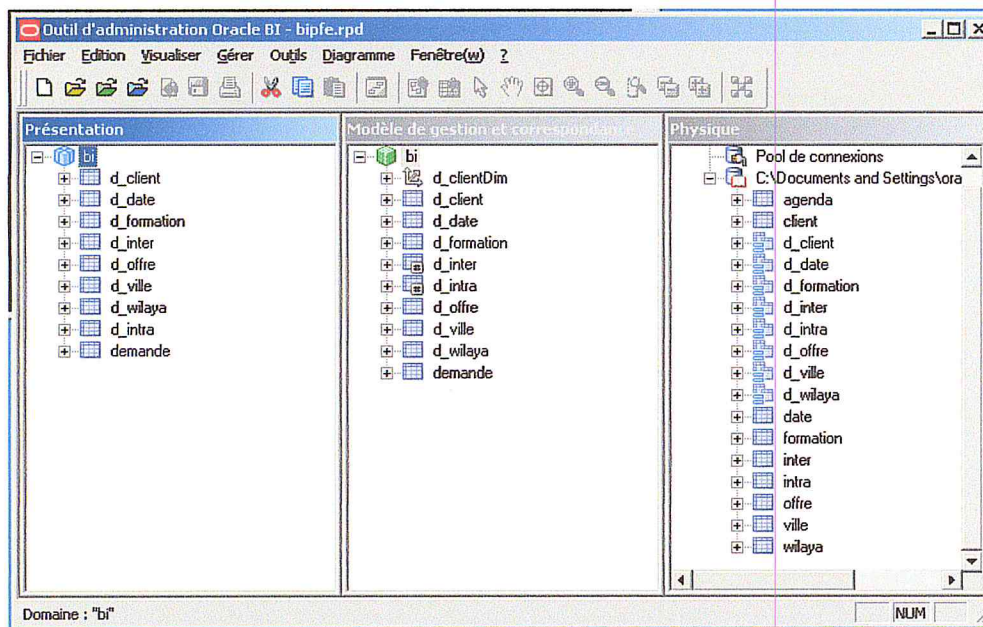


Figure 7.3: Représentation des trois couches réalisées

connexions, de dossiers, de tables, de colonnes et de clés.

La figure suivante représente la base de données source BDD-IM alimentée dans la couche physique :

7.1.4 La couche Business Model and Mapping:

C'est la couche dans laquelle les schémas physiques sont simplifiés et réorganisés pour constituer le fondement de la vue des données qui est proposée aux utilisateurs. Dans cette couche, les données sont séparées en faits et en dimensions, pour plus de simplicité. Les dimensions logiques introduisent des hiérarchies formelles dans un modèle de gestion. Elles établissent des niveaux pour le regroupement des données et les calculs, et fournissent des chemins pour l'exploration descendante des données dans les outils destinés à l'utilisateur final.

Ci-dessous, une figure de la couche logique:

7.1.5 Couche Présentation:

Elle contient les objets de présentation qui offrent aux utilisateurs une vue personnalisée d'un modèle de gestion. Elle simplifie et organise le modèle de gestion pour qu'il soit facile à comprendre et à interroger.

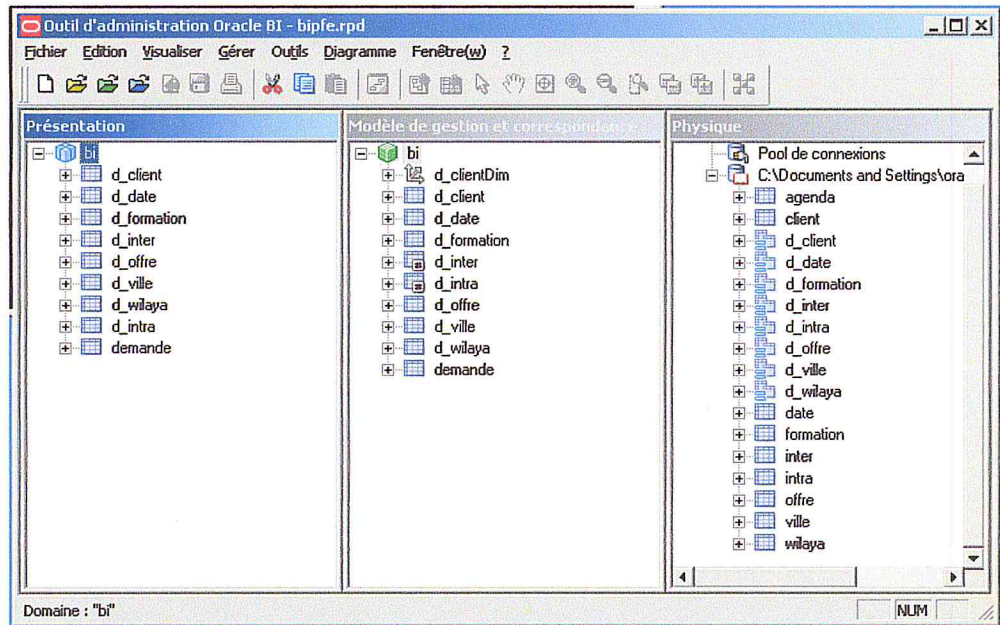


Figure 7.4: Schéma relationnel de la base de données BDD-IM

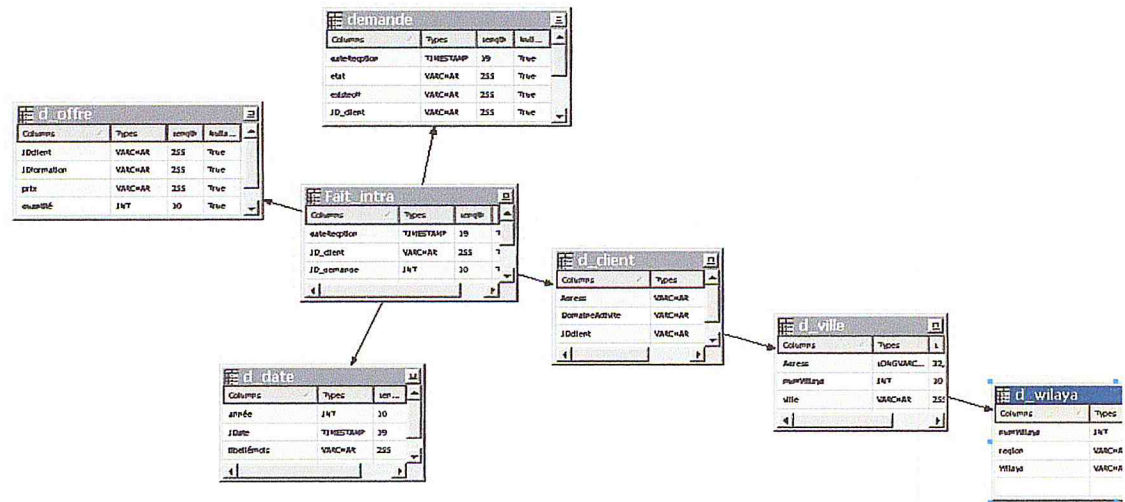


Figure 7.5: Schéma en flocon de neige de suivi-intra sur la couche logique

7.2 Construction de l'application utilisateur

Pour exposer les informations enregistrées dans l'entrepôt de données à l'utilisateur final, nous avons utilisé : Les tableaux de bord et le reporting.

- **Les TDB** : la dernière étape de notre projet est la construction d'un tableau de bord, récapitulatif des différents indicateurs que nous avons présentés dans le chapitre "conception" . Pour cela, nous avons utilisé Oracle Business Analytics. C'est un outil développé par Oracle qui offre la possibilité de créer, modifier et enregistrer des tableaux de bord.
- **Le reporting** : en utilisant l'outil Business Analytics développé par Oracle, nous avons créé des rapports par cet outil et qui présentent l'ensemble des activités dont le suivi des formations les plus exigées sur l'ensemble des clients, les clients acquis et perdus et leurs causes de perte, et d'autres rapports sous forme tabulaires ou graphiques. Nous présenterons ci-dessous quelques captures d'écrans de tableaux de bord et rapport:

Exemple de traitement d'une analyse:

La diapositive ci-dessous illustre un exemple simple de traitement d'analyse Oracle BI.

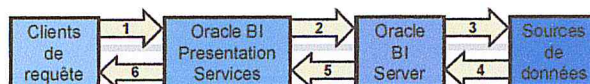


Figure 7.6: Exemple de traitement d'analyse sous OBIEE.

1. Un utilisateur consulte un tableau de bord ou lance une analyse;
2. Oracle BI Presentation Services demande à Oracle BI Server d'extraire les données demandées;
3. Oracle BI Server utilise le fichier de référentiel pour optimiser l'interrogation des sources de données.
4. Oracle BI Server reçoit les réponses des sources de données et effectue le traitement nécessaire;
5. Oracle BI Server transmet les données à Oracle BI Presentation Services;
6. Oracle BI Presentation Services formate les données et les envoie au client de requête.

La couche présentation nous a permis de :

- Construire un tableau de bord avec les indicateurs choisis au moment de la conception :

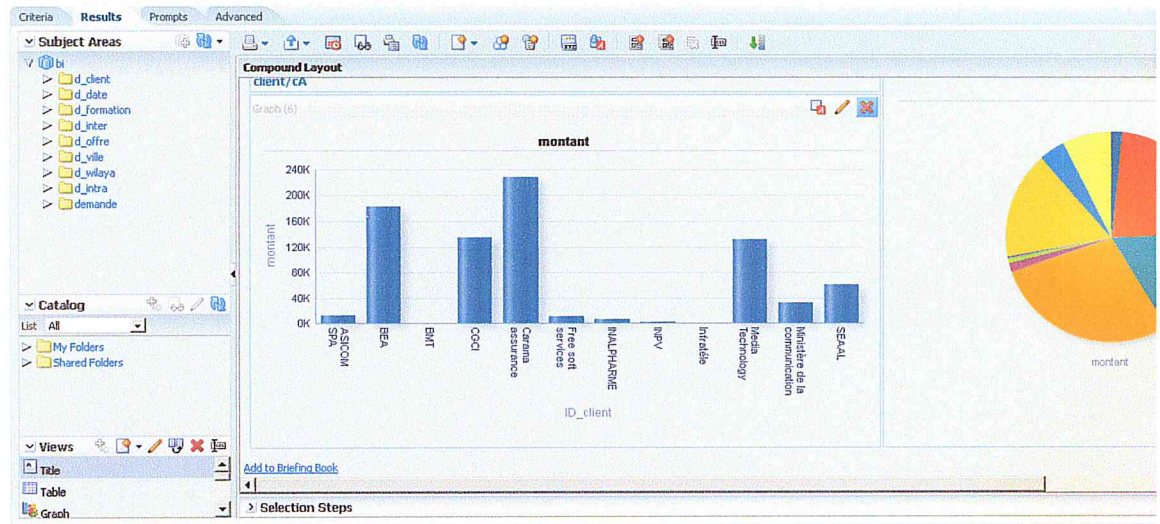


Figure 7.7: Exemple d'un tableau de bord

- Construire des rapports paramétrés à l'aide du moteur "Business Analytics" qui permet de visualiser les résultats sous plusieurs formes graphiques en offrant la possibilité d'exporter le rapport en plusieurs formats (XLS, JPEG, PDF). La figure ci-dessous montre un exemple d'un rapport exporté sous format PDF.

7.3 Conclusion

Dans ce chapitre, nous avons décrit la réalisation et le déploiement de notre système décisionnel. Dont nous avons présenté le matériel, les outils ainsi que l'architecture technique utilisée. Par la suite, nous avons donné des exemples d'utilisation des outils dans les différentes couches du système.

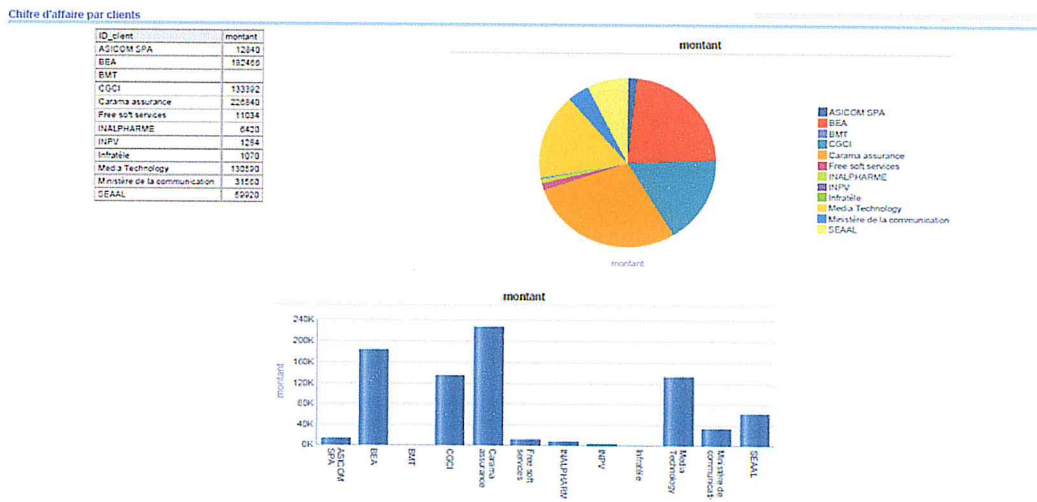


Figure 7.8: Exemple d'un rapport

Conclusion Générale et Perspectives

Le système décisionnel offre une réelle aide pour les entreprises dans leur processus d'aide à la décision. Il permet la transformation des données brutes extraites en informations afin de comprendre de la situation actuelle de l'entreprise et d'anticiper sa situation future pour un pilotage plus éclairé.

Dans ce contexte, L'institut " El Maarifa " nous a confié de concevoir et mettre en place un système d'information d'aide à la décision permettant de fournir des rapports, des tableaux de bord et des indicateurs clés de performance concernant ses processus de "Formation Qualifiante Interentreprises", et "Formation Qualifiante Intra-entreprises". Ce système lui permettra de piloter efficacement ses activités, en analysant soigneusement les données pour faire face aux multiples problèmes qu'elle rencontre. Ainsi, il permettra à ses responsables et dirigeants de prendre les décisions adéquates.

Pour réaliser ce système, nous avons étudié l'aspect théorique de notre projet afin d'approfondir nos connaissances dans le domaine de décisionnel. Par suite, nous avons conçu l'entrepôt de données en suivant la méthode de "Kimball" qui se focalise sur l'étude détaillées des besoins des utilisateurs à travers des réunions, la documentation et l'analyse des processus fonctionnels et techniques existants.

A partir des besoins collectés, nous avons conçu la zone d'entreposage en suivant la modélisation dimensionnelle en flocon de neige. Pour alimenter cette zone, nous devrions concevoir la zone d'alimentation, à partir de la base de données réalisé en premier lieu à partir des fichiers sources Excel et ce par le biais d'outil ETL. Afin de permettre aux utilisateurs d'exploiter les données stockées dans l'entrepôt, nous avons conçu un tableau de bord regroupant les indicateurs selon les objectifs des décideurs. Pour cela, nous avons suivi la méthode de conception des TDB "GIMSI".

Enfin nous avons présenté les différentes étapes de la réalisation, l'architecture de déploiement de notre solution ainsi que les différentes technologies utilisées.

Pour finir, et avant de citer les perspectives du projet, nous pouvons dire que ce stage au niveau de l'institut "El Maarifa", nous a permis d'acquérir une bonne expérience professionnelle et d'évoluer dans un domaine très intéressant à savoir le domaine des systèmes décisionnels. Afin de compléter notre projet, comme perspectives, nous proposons de :

- Suivre le déploiement afin d'apporter les corrections nécessaires sur la solution actuelle;
- Apporter des modifications nécessaires au niveau de la base de données afin de répondre à l'ensemble des besoins des utilisateurs;
- Intégrer les données des formations diplômâtes dans notre système;
- Réaliser une carte localisant les clients.

Bibliographie

[Devisy, 2002] : A.Devisy, Livre blanc e-Business intelligence, l'effet internet sur le décisionnel . Aubay, 2002.

[Fernandez, 2006] : A.Fernandez, Les nouveaux tableaux de bord des managers, Eyrolles 2006.

[Fernandez, 2008] : A.Fernandez, Les nouveaux tableaux de bord des managers pour le projet décisionnel en totalité. Edition d'organisation, (2008).

[Fernandez, 2011] : A.Fernandez, L'essentiel du tableau de bord : Méthode complète et mise en pratique avec Microsoft Excel. Eyrolles, 2011.

[Fernandez, 2012] : A.Fernandez, La création et la publication de rapport d'activité. Disponible : <http://www.piloter.org/business-intelligence/reporting.htm> , Mai 2012.

[Fernandez, 2013] : A.Fernandez, Les nouveaux tableaux de bord des managers : Le projet Business Intelligence clés en main, Édition : 6eme édition. Eyrolles, 2013.

[Guibert,] : O.Guibert, Analyse et Conception des Systèmes d'information- Méthode Objet-Le langage de modélisation objet UML. Département Informatique de l'Institut Universitaire de Technologie de l'Université de Bordeaux 1. Disponible sur: " <https://www.labri.fr/perso/guibert/DocumentsEnseignement/UML.pdf> "

[Hand et Al, 2001] : D.J.Hand, H.Mannila, and P.Smyth, Principles of Data Mining. MIT Press, 2001.

[Kimball, 2004] : R.Kimball, The Data Warehouse ETL Toolkit (3eme Edition), Wiley publishing, 2004.

[**Kimball et Al, 2005**] : R.Kimball, L.Reeves, M.Ross, W.Thornthwaite, Le data warehouse : Guide de conduit de projet. Paris: Eyrolles, 2005.

[**Kimball, 2005**] : R.Kimball, Le Data Warehouse, Guide de conduite de projet. Edition Eyrolles. (2005).

[**Kimball et Al, 2007**] : R.Kimball, M.Ross, W.Thornthwaite, Le Data Warehouse, Guide de conduite de projet (3eme Edition), Eyrolles, 2007.

[**Lorino, 2001**] : P.Lorino, Méthodes et pratiques de la performance, Editions d'Organisation, Paris, 1997- 2001.

[**Nakache, 1998**] : D.Nakache, Data Warehouse et Data Mining. Conservatoire National des Arts et Métiers de Lille; Version 1.1; 15 Juin 1998.

[**Web1**] : L'informatique Décisionnelle (BI), disponible sur : " <http://perso.univ-lyon1.fr/haytham.elghazel/BI/presentation.html>"

[**Web2**] : Qu'est-ce que l'informatique décisionnelle ?, disponible sur : " <http://business-intelligence.developpez.com/tutoriels/quest-ce-que-la-bi>"

[**Web3**] : Le Data Warehouse et les Systèmes Multidimensionnels, disponible sur : " <https://deptmedia.cnam.fr/new/spip.php?pdoc752>"

[**Web4**] : Informatique décisionnelle (BI) : OLAP et la modélisation dimensionnelle, disponible sur : " <http://www-igm.univ-mlv.fr/dr/XPOSE2009/informatique>

[**Web5**] : Les objectifs du data mining sont par définition déterminés en fonction des données existantes et des problématiques de l'entreprise mais tout projet de fouille de données en général a pour objectifs l'un des suivants :

[**Web6**] : UML, Langage de modélisation objet unifié, disponible sur: " <http://manu.kegtux.org/UML/Cours/CoursUML1.pdf> "

[**Web7**] : Introduction à UML, disponible sur : " <http://www.commentcamarche.net/cont/introduction-à-uml>"

[**Web8**] : T.Chaari, Atelier UML, disponible sur : " <http://www.redcad.org/members>"

[Web9] : UML Class Diagrams: Reference, disponible sur : "
<https://msdn.microsoft.com/fr-fr/library/dd409437.aspx>"

[Web10] : Microsoft Access, disponible sur : "<http://fr.wikipedia.org/wiki/MicrosoftAccess>"