

**Ministère de l'enseignement supérieur et de la recherche scientifique**

**Université Saad DAHLAB de Blida**



**Faculté des Sciences**

**Département d'Informatique**

Mémoire Présenté par :

- **Mr. ELMOKREFI** ishak
- **Mr. AKSIL** mohamed raouf

En vue d'obtenir le diplôme de **Master**

**Domaine** : Mathématique et Informatique

**Filière** : Informatique

**Option** : Ingénierie de logiciels

**Sujet** : Conception et réalisation d'un système de reporting  
décisionnel pour DREM

Soutenu le : 29/09/2011,

devant le jury composé de :

Mme. Bensettiti

Président

Mlle. AMEUR

Rapporteur

Mme. Boucetta

Examineur

Mme. Touahri

Examineur

Mme. Boumehdi

Examineur

MA-004-72-1

**Ministère de l'enseignement supérieur et de la recherche scientifique**

**Université Saad DAHLAB de Blida**



**Faculté des Sciences**

**Département d'Informatique**

Mémoire Présenté par :

- **Mr. ELMOKREFI ishak**
- **Mr. AKSIL mohamed raouf**

En vue d'obtenir le diplôme de **Master**

**Domaine : Mathématique et Informatique**

**Filière : Informatique**

**Option : Ingénierie de logiciels**

**Sujet : Conception et réalisation d'un système de reporting  
décisionnel pour DREM**

Soutenu le : 29/09/2011

devant le jury composé de :

Mme. Bensettiti

Président

Mlle. AMEUR

Rapporteur

Mme. Boucetta

Examineur

Mme. Touahri

Examineur

Mme. Boumehdi

Examineur

“إن استطعت فكن عالما ... فإن لم تستطع فكن متعلما ... فإن لم تستطع  
فأحبهم .... فإن لم تستطع فلا تبغضهم “

عمر بن عبد العزيز

## Remerciements

*Nous tenons tout d'abord à remercier Dieu Tout Puissant, tout Clément qui nous a aidés à élaborer ce travail et aboutir à ce niveau d'études.*

*Nous exprimons notre profonde gratitude à tous ceux et celles qui, de près ou de loin, nous ont aidé dans la réalisation de ce travail comme ceux et celles qui ont bien voulu accepter de le juger.*

*Que Melle Ameer de l'Université de Blida veuille bien trouver le témoignage de toute notre reconnaissance pour avoir dirigé cette thèse et de nous avoir apporté conseils et orientations tout au long de sa réalisation.*

*Nous remercions très sincèrement tous les cadres de la DREM pour leur aide et leur accueil chaleureux et ce au cours de notre stage pratique, plus particulièrement, Chef de service d'information Mr.Rabehallah Omar, chef d'agence ALEM d'elaffroun Mr.Guebli Youcef et tous les agents de cette agence (surtout Mr.bentaiba), Chef de Service Statistiques ainsi que Mme Boudissa, Chef, pour leur patience et tous leurs efforts.*

*Nous ne pouvons oublier nos parents ainsi que nos frères et sœurs auprès desquels nous avons trouvé toutes les facilités qui nous ont permis de mener à bien ce projet.*

# Dédicace Ishak

En tout début et avant toutes chose je remercie الله qui ma donné la force et l'ambition pour accomplir ce modeste travail.

Je dédie ce modeste travail à :

Mes parents qui grâce à leurs sacrifices j'ai pu concrétiser mon rêve.

Ma Mère qui m'a toujours aidé, encouragé.

Mon père qui est un symbole pour moi.

Mon défunt cousin MHAMED, baba SIMOUH, khalil MHAMED, ami AËK, khalil RACHID, HAMID, Ami AHMED, TOUFIK, KHEIRO. Et toute la famille E LMOKHFI et ANNANE.

Mon frère Yakoub, mes sœurs et surtout ma nièce Meriem que je l'adore.

A mon Binôme & copain : Raouf et a toutes sa famille.

À toute la ière Promo LMD : Melayka, Ali, Housseem, Harrachi, Mustafa, Aksil, Aniss, Abdellah, Abdel Baki

Et mon seul ami de cette promo, CHENOUI Salah Eddine

Mes amies Boumaza mohamed, Mehdi kadi, Larasse mohamed.

Mon voisin et frère TSIMOULT Samir.

# Dédicace Mohamed raouf

En tout début et avant toutes chose je remercie الله qui ma donné la force et l'ambition pour accomplir ce modeste travail.

Je dédie ce modeste travail

Aux Mes parents que Dieu les protègent, qui m'ont toujours soutenu, encouragé et aidé dans ma vie et mes études :

Mon père Mr. *Aksil Farid*

Ma mère Mme. *Ben Youcef Samya*

A mes chers frères : *Ismail et Lamine*

A mon unique sœur : *Leïla*

A tous mes amis qui m'ont soutenu de prêt ou de loin à la réalisation de ce modeste projet sans oublié bien sur tous les villageois du pavillon 1 et à leurs familles.

A mon Binôme & copain : *Ishak* et a toutes sa famille.

A toutes mes collègues de classe Master et à ceux qui ont déjà soutenue et qui travaillent en ce moments : *Housseem, Nabile, El Hadi, Walid, Manel et Yasmine.*

Et à tous ceux qui me connaissent, je citerais : les jumeaux *Walid* et *Billel*, *Amir*, *Mourad*, *Amdjad*, *Keltoum*, *Nadia*, *Djamila*, *Lahcene*, *Raouf Matmour* et une Spécial Dédicace pour *Sara Bouguena.*

# Table des matières

---

- Glossaire
- Liste des figures
- Liste des diagrammes
- Liste des tableaux
- Résumé

## Introduction générale

Contexte générale.....	2
Problématique .....	3
Objectifs .....	4
Organisation mémoire .....	5

## Chapitre 1 : Etat de l'art

<b>1. Introduction .....</b>	<b>7</b>
<b>2. Qu'est-ce qu'un système décisionnel .....</b>	<b>7</b>
<b>3. Bref historique.....</b>	<b>8</b>
<b>4. Chronologie .....</b>	<b>9</b>
<b>5. Positionnement du décisionnel au sein du système d'information .....</b>	<b>10</b>
<b>6. Qui a besoin au décisionnel.....</b>	<b>11</b>
<b>7. L'entrepôt de données (DataWarehouse) .....</b>	<b>11</b>
7.1 C'est quoi un DataWarehouse .....	11
7.2 Architecture globale d'un DataWarehouse.....	12
7.2.1 Définition concepts.....	13
7.3 Fonctions processus DW .....	15
7.3.1 Sélection des données sources.....	15
7.3.2 Extraction des données .....	16
7.3.3 Transformation.....	16
7.3.4 Chargement .....	17
<b>8. Les modèles des données .....</b>	<b>18</b>
8.1 Définition modèle .....	18

## Table des matières

---

8.2 Le modèle des données des DW .....	18
8.2.1 Caractéristiques d'utilisation OLAP vs OLTP .....	18
8.2.2 Les techniques de modélisation .....	21
8.2.2.1 Le modèle de données normalisé .....	22
8.2.2.2 Dé-normalisation pour la décision.....	23
8.2.2.3 La modélisation dimensionnelle .....	24
8.3 Composants d'un modèle dimensionnel.....	25
8.3.1 Table de fait.....	25
8.3.2 Les faits (mesures) .....	25
8.3.3 Table de dimension .....	25
8.4 La modélisation multidimensionnelle.....	27
8.4.1 Modèle en étoile .....	27
8.4.2 Modèle en flocon .....	29
8.4.3 Modèle en constellation.....	31
<b>9. Outils de restitution .....</b>	<b>32</b>
9.1 Technologie OLAP (Online Analytical Processing).....	32
9.1.1 Concept de Cube .....	32
9.1.2 Les technologies OLAP .....	33
9.1.3 Navigation dans le cube .....	34
9.1.4 Les Règles OLAP.....	35
9.2 Le Datamining .....	36
9.2.1 Définition.....	36
9.2.2 Principe du Datamining .....	36
9.2.3 Pour quelles applications utiliser le Datamining ? .....	36
9.2.4 Les techniques du datamining .....	36
9.2.5 Comment utiliser le Datamining ? .....	37
9.3 Le tableau de bord.....	37
<b>10. Comparaison entre système opérationnel et système décisionnel .....</b>	<b>38</b>
<b>11. Conclusion .....</b>	<b>39</b>

# Table des matières

---

## Chapitre 2 : Analyse des besoins

<b>1. Introduction</b> .....	<b>41</b>
<b>2. Présentation de La Structure d'accueil</b> .....	<b>41</b>
2.1 Historique du SPE (service public de l'emploi) .....	41
2.2 Organisation de L'ANEM .....	41
2.3 Mission et objectifs de l'ANEM .....	42
2.3.1 Mission .....	42
2.3.2 Objectifs .....	43
2.4 Limitation du champ d'étude .....	43
2.4.1 Missions et objectifs DREM .....	43
2.4.2 Organigramme DREM .....	44
<b>3. Définition des besoins</b> .....	<b>44</b>
3.1 Approche de définition des besoins .....	44
3.2 La définition des besoins .....	45
3.2.1 Recherche documentaire .....	45
3.2.2 Les entretiens .....	46
3.2.3 Les Objectifs .....	46
3.3 Extraction d'Indicateurs et Axes d'analyse .....	47
3.3.1 Indicateurs (faits) .....	47
3.3.2 Axes d'analyse (dimensions) .....	48
<b>4. Conclusion</b> .....	<b>49</b>

## Chapitre 3 : Conception

<b>1. Introduction</b> .....	<b>51</b>
<b>2. Solution Informatique</b> .....	<b>51</b>
2.1 Pourquoi une nouvelle Base de production .....	51
2.2 Pourquoi un DataWarehouse .....	52
<b>3. Système opérationnel</b> .....	<b>52</b>
3.1 Pourquoi modéliser .....	52
3.2 Choix d'outil Conception .....	53

# Table des matières

---

3.2.1 Introduction a UML .....	53
3.2.2 Pourquoi UML .....	54
3.2.3 Diagrammes UML .....	55
3.3 Choix d'outil de Modélisation .....	55
3.4 Diagrammes .....	56
3.4.1 Diagramme de cas d'utilisations .....	57
3.4.2 Diagrammes d'activités .....	63
3.4.3 Diagramme de séquences .....	67
3.4.4 Diagramme de classes.....	70
<b>4. Système Décisionnel .....</b>	<b>75</b>
4.1 Les diagrammes de cas d'utilisation .....	75
4.1.1 Analyse « Flux de la demande » .....	75
4.1.2 Analyse «la demande enregistrée » .....	76
4.1.3 Analyse « Offre » .....	76
4.1.4 Analyse « Placement » .....	77
4.1.5 Analyse « Placement Formation Professionnel » .....	77
4.1.6 Analyse « Placement Enseignement supérieur » .....	77
4.2 La modélisation multidimensionnelle .....	78
4.2.1 Flux de demande d'emploi .....	79
4.2.2 Demandes enregistrées .....	80
4.2.3 Nombre d'offre .....	81
4.2.4 Répartition des demandes d'emploi placées .....	82
4.2.5 Placements des diplômées issus de la FP .....	83
4.2.6 Placements des diplômées issus de ES .....	84
4.2.7 Schéma de l'entrepôt de données .....	85
4.3 Conception ETL .....	86
4.3.1 Système "ETL" .....	87
4.3.2 Diagrammes d'activité ETL .....	87
<b>5. Conclusion .....</b>	<b>89</b>

# Table des matières

---

## **Chapitre 4 : Mise en œuvre et déploiement**

<b>1. Introduction</b> .....	<b>91</b>
<b>2. Description de l'environnement de travail</b> .....	<b>91</b>
2.1. Système opérationnel .....	91
2.2. Système décisionnel.....	93
<b>3. Définition de l'architecture physique et technique</b> .....	<b>94</b>
3.1. Architecture physique .....	94
3.1.1. Diagramme de composants .....	95
3.1.2. Diagramme de déploiement .....	97
3.2. Architecture technique.....	97
<b>4. Mise en œuvre du système</b> .....	<b>99</b>
4.1. Authentification .....	99
4.2. Ajouter Demandeur .....	99
4.3. L'entrepôt de données .....	100
4.4. Un rapport final .....	101
4.4.1. Description .....	102
<b>4. Conclusion</b> .....	<b>102</b>
<b>Perspectives et suggestions</b> .....	<b>104</b>
<b>Conclusion générale</b> .....	<b>106</b>
<b>Bibliographie</b> .....	<b>109</b>

**Annexe**

**Glossaire de la BI (Business Intelligence) :**

■ **Aide à la décision :**

L'aide à la décision est l'activité de celui qui aide à obtenir des éléments après réponses aux questions que pose un intervenant dans un processus de décision, élément concourant à éclairer la décision et à recommander un comportement de nature à accroître la cohérence entre l'évolution du processus et les objectifs de cet intervenant.

■ **Analyse multidimensionnelle :**

Analyse des valeurs d'un indicateur selon plusieurs angles. L'utilisateur visualise, par exemple, le chiffre d'affaire mensuel par produit et par vendeur, mais par produit tout vendeur confondu (le CA et le cumul de celui des vendeurs), et enfin le chiffre d'affaires annuel par produit (qui représente le cumul des mois).

■ **Axe d'analyse :**

Thème fonctionnel au groupant Les informations permettant d'analyser les résultats de l'entreprise. Typiquement, les axes reflètent la perception que l'entreprise a de son environnement : client, produit, géographie, force de vente, temps (liste non exhaustive). Un axe d'analyse est organisé autour des différents concepts appelés dimension.

■ **Business Intelligence (BI) :**

La business intelligence est un des 1000 Concept forgé en 1993 par le Gartner Group pour désigner l'ensemble des moyens et des méthodes permettant de rassembler, consolider, analyser et rendre accessibles les données d'une entreprise au plus grand nombre d'acteurs possible, dans une perspective d'aide à la décision. Le business intelligence doit être un levier à l'augmentation des performances.

■ **CCR :**

Centre de Contrôle Régional.

■ **Data mart :**

Il s'agit d'un ensemble de données (magasin) à l'usage d'un groupe d'utilisateurs qui partagent le même métier ou la même activité (marketing, finance, ressources humaines).

■ **Datamining :**

Outil ou brique logicielle servant à effectuer de la recherche : corrélation, tendance. Une des finalités concerne la segmentation pour la gestion de la relation client. Les outils de datamining utilisent des techniques d'intelligence artificielle comme par exemple les réseaux de neurone.

■ **Datawarehouse :**

Le SID est structuré autour d'un espace de données unique appelé entrepôt de données ou encore Datawarehouse permettant de consolider l'ensemble des informations potentiellement utiles aux analyses décisionnelles.

■ **Décisionnel :**

Terme générique qui désigne le système d'information décisionnel (SID). Ce dernier regroupe l'ensemble des processus de collecte, de mesure, et d'analyse des données de l'entreprise à des fins de pilotage.

■ **Dimension :**

Concept ou information élémentaire permettant de structurer l'analyse des données. Les dimensions sont regroupées en axes d'analyse qui présentent en fait le thème fonctionnel d'analyse.

■ **Drill down / Drill up :**

Fonction offerte par les services de navigation des systèmes décisionnels. Le drill down est par exemple un zoom sur une donnée pour obtenir un niveau de détail plus fin. Le drill up est l'opération inverse.

■ **Entrepôt de données :**

Désignation française de Datawarehouse. Il s'agit de la base qui archive en un endroit unique les indicateurs de plusieurs domaines fonctionnels.

- **ETL (*Extract Transform Load*) :**

Logiciel d'alimentation de l'entrepôt de données ou des Data mart à partir des extractions effectuées dans le système d'information.

- **Etoile (modèle) :**

Cf. Modèle en étoile.

- **Fait :**

Terme technique désignant au sein des données de l'entreprise un concept évènementiel reflétant l'activité de celle-ci (exemple : les ventes, les mouvements de stock). Les tables de faits sont les tables de l'entrepôt relatives aux faits.

- **Flocon (modèle) :**

Cf. Modèle en flocon.

- **Flux de données:**

Un flux de données est un transfert d'un ensemble d'information d'une source (de données) vers une cible.

- **Hiérarchie :**

Les axes d'analyse sont organisés suivant une ou plusieurs hiérarchies. A chaque niveau de la hiérarchie, une agrégation est possible pour l'ensemble des indicateurs l'utilisant.

- **Historisation :**

C'est le fait de conserver l'historique des modifications subies par une même information au cours du temps.

- **Historiser :**

Action de mettre en place une historisation pour une information donnée.

- **H-OLAP (*Hybrid OLAP*) :**

C'est un mixte de SQL et d'OLAP, l'un passant le relais à l'autre en fonction des opérations effectuées.

- **Hypercube :**

Cf. OLAP.

**■ Indicateur(s) :**

Information quantitative permettant de mesurer la performance d'une activité de l'entreprise. En général, cette information est numérique (chiffre d'affaires, quantité en stock).

- ✓ Les indicateurs élémentaires (ou bruts) : données directement stockées dans l'entrepôt de données et caractéristiques d'une table de fait : quantité vendue, prix de vente (pour les ventes).
- ✓ Les indicateurs calculés: données qui sont évaluées à partir des indicateurs élémentaires comme le chiffre d'affaire. Celles-ci sont parfois stockées en base (pour des raisons d'optimisation) mais l'évaluation s'effectue, le plus souvent, au moment où elles sont demandées.

**■ Infocentre :**

Base de données dédiée à l'analyse décisionnelle par simple réplique des données des systèmes d'information opérationnels dans un environnement à part.

**■ Mesure :**

Une mesure est une quantité présente dans la table de fait qui permet de mesurer les faits.

**■ Métadonnée :**

Les métadonnées, c'est-à-dire des données concernant des données, constituent une base de connaissance permettant d'expliquer le contenu des autres bases (et du SID en particulier).

**■ Modèle de données :**

Appellation générique qui désigne l'organisation des données. Il peut être conceptuel, logique ou physique.

**■ Modèle en étoile :**

Technique de modélisation dimensionnelle qui représente l'ensemble des données d'un axe d'analyse dans une seule table. La table de faits est placée au centre du modèle, les tables d'axe sont mises en relation autour d'elle. Ce modèle représente une étoile.

- **Modèle en flocon :**

Technique de modélisation dimensionnelle, variante de la modélisation en étoile, dont la représentation visuelle s'apparente à un flocon car les axes d'analyse sont composés chacun de plusieurs tables organisées de façon arborescente.

- **Modèle dimensionnelle :**

Technique de modélisation qui diffère de celle des systèmes d'information, consistant à organiser les données (indicateurs) autour de dimensions. On dit que les données sont dimensionnées. Les dimensions sont communes à plusieurs indicateurs. Ce type de modélisation permet aux outils de restitution d'effectuer des agrégations sur les données brutes. Les niveaux d'agrégation possibles dépendent de la manière dont sont conçus les dimensions.

- **M-OLAP :**

Cf. OLAP.

- **Multidimensionnel :**

Cf. modélisation dimensionnelle.

- **ODS (*Operational Data Store*) :**

Espace technique de données en amont de l'entrepôt, dans la chaîne de gestion de contenus. L'ODS permet, entre autre, de synchroniser les multiples flux de données issus du système d'information opérationnel.

- **OLAP (*On Line Analytical Processing*):**

Le terme OLAP désigne souvent les outils d'analyse qui s'appuient sur une base de données organisée de façon dimensionnelle (cf. modélisation dimensionnelle), voire sur une base de données multidimensionnelle.

On parle dans les premiers cas de R-OLAP (Relational OLAP) et dans le second de M-OLAP (Multidimensional OLAP). En outre, il existe la notion d'H-OLAP qui s'appuie, suivant les besoins sur ces deux types de technologies.

- **OLTP (*On Line Transactional Processing*):**

Système transactionnel permettant aux utilisateurs de passer des transactions (mode interactif) comme la commande en ligne d'un billet d'avion.

■ **Reporting :**

Nom anglais, signifie la génération et la publication des rapports.

■ **Requête :**

Demande formulée à une base de données dans un langage informatique structuré.

■ **Requêteur :**

Outil graphique permettant d'interroger le contenu des bases de données d'un SID. Il génère automatiquement d'une manière transparente pour l'utilisateur les requêtes (commandes SQL) afin d'exécuter ces dernières.

■ **R-OLAP (*Relational On Line Analytical Processing*) :**

Architecture de système d'aide à la décision destinée à effectuer des analyses multidimensionnelles (cf. Multidimensionnel), uniquement sur des données relationnelles.

■ **Roll-up :**

Roll-up est le contraire de drill-down. C'est donc faire de l'agrégation (ou résumé) des données.

■ **SI (*Système d'Information*) :**

Le système d'information de l'entreprise définit l'ensemble des moyens par lesquels les acteurs créent, échangent, gèrent et stockent les informations dont ils ont besoin pour l'exercice de leurs activités. Il est également appelé SIO (Système d'Information Opérationnel).

■ **SIAD (*Système Informatique d'Aide à la Décision*), EIS (*Executive Information System*) :**

Terme générique qui symbolise le «système d'information du cadre ou de décideur ». Les indicateurs sont très synthétiques et fortement agrégés.

■ **SID (*Système d'Information Décisionnel*) :**

Sous-ensemble du SI spécialisé dans l'aide à la décision.

■ **SIO (*Système d'Information Opérationnel*) :**

Cf. SI.

## 7 Glossaire

- **Table de fait :**

Cf. Faits.

- **UML (*Unified Modeling Language*) :**

Langage de modélisation de troisième génération, normalisé par l'OMG début 1997, permettant de décrire une application en fonction des méthodes et objets avec lesquels elle a été construite. Utilisé pour la modélisation des SI.

### Glossaire de la Direction DREM :

- **ALEM :** Agence locale de l'emploi.
- **ANEM :** Agence Nationale de l'Emploi.
- **AWEM :** Agence de wilaya de l'emploi.
- **Demandeur DFP :** Demandeurs d'emplois enregistrés issus de la formation professionnelle
- **Demandeur DES :** demandeurs diplômés sortant de l'enseignement supérieur et équivalent.
- **DREM :** Direction Régionale de l'Emploi.
- **ONAMO :** Office national de la main d'œuvre.
- **Placement DFP :** Placements des diplômés issus de la formation professionnelle et équivalent.
- **Placement DES :** Répartition des placements des diplômés sortant de l'enseignement supérieur et équivalent.
- **SPE :** service public de l'emploi.
- **STR1 :** Ayant déjà travaillé.
- **STR2 :** N'ayant jamais travaillé.

## Liste des figures

Figure	Désignation	page
Figure 1.1	Schéma du système décisionnel	8
Figure 1.2	Prise de la décision	10
Figure 1.3	Le décisionnel au sein du système d'information	11
Figure 1.4	Architecture globale d'un DataWarehouse	12
Figure 1.5	Le modèle de données normalisé	22
Figure 1.6	Le modèle de données dé-normalisé	23
Figure 1.7	Table de fait	26
Figure 1.8	Table de dimension	26
Figure 1.9	Modèle en étoile	28
Figure 1.10	Modèle en flocon	30
Figure 1.11	Modèle en constellation	31
Figure 1.12	Cube de données	32
Figure 1.13	Tableau de bord	38
Figure 2.1	Organigramme DREM	44
Figure 3.1	Fenêtre principale de MS Visio 2010	56
Figure 3.19	Modèle en étoile de l'activité « Flux de la demande d'emploi	79
Figure 3.20	Modèle multidimensionnel en étoile de l'activité « demandes enregistrées »	80
Figure 3.21	Modèle multidimensionnel en étoile de l'activité « Offre »	81
Figure 3.22	Modèle multidimensionnel en étoile de l'activité « Placements »	82
Figure 3.23	Modèle multidimensionnel en étoile de l'activité « Placement_FP »	83
Figure 3.24	Modèle multidimensionnel en étoile de l'activité « Placement_ES »	84
Figure 3.25	schéma en constellation	86
Figure 4.1	Solution Business Intelligence SQL server 2008 R2	94
Figure 4.4	Architecture technique du système décisionnel	98
Figure 4.5	Interface d'authentification - système opérationnel	99
Figure 4.6	Interface ajouter un demandeur - système opérationnel	100
Figure 4.7	schéma de l'entrepôt de données - système décisionnel	101
Figure 4.8	Histogramme flux de demande généré par Excel	102

## Liste des Diagrammes

Figure	Désignation	page
Figure 3.2	Diagramme de cas d'utilisation Administrateur I	58
Figure 3.3	Diagramme de cas d'utilisation Administrateur II	59
Figure 3.4	Diagramme de cas d'utilisation Administrateur simplifiée	60
Figure 3.5	Diagramme de cas d'utilisation de l'agent de saisie	62
Figure 3.6	Diagramme de cas d'utilisation de l'agent de saisie simplifiée	63
Figure 3.7	Diagramme d'activité «Créer Demandeur»	64
Figure 3.8	Diagramme d'activité «Modifier Demandeur»	65
Figure 3.9	Diagramme d'activité «Ajouter Offre»	66
Figure 3.10	Diagramme de séquence «Modifier Demandeur».	67
Figure 3.11	Diagramme de séquence «Consulter Offre»	69
Figure 3.12	Diagramme de classe	71
Figure 3.13	Diagramme de cas d'utilisations d'analyse Flux de la demande	75
Figure 3.14	Diagramme de cas d'utilisations d'analyse demande enregistrée	75
Figure 3.15	Diagramme de cas d'utilisations d'analyse de l'offre.	76
Figure 3.16	Diagramme de cas d'utilisations d'analyse Placement	76
Figure 3.17	Diagramme de cas d'utilisations d'analyse PlacementFP	77
Figure 3.18	Diagramme de cas d'utilisations d'analyse PlacementES	77
Figure 3.26	Diagramme d'activité du système ETL	88
Figure 4.2	Diagramme de composant pour le système globale	96
Figure 4.3	Diagramme de déploiement du système global	97

## Liste des tableaux

<b>Tableau</b>	<b>Désignation</b>	<b>page</b>
<b>Tableau 1.1</b>	Technologies OLAP	33
<b>Tableau 1.2</b>	OLTP vs OLAP	39
<b>Tableau 2.1</b>	Les faits	47
<b>Tableau 2.2</b>	Les dimensions	48
<b>Tableau 3.1</b>	le scénario de modification demandeur	68
<b>Tableau 3.2</b>	le scénario de Consultation offre	69

# ملخص

يندرج هذا العمل في إطار تصميم و تنفيذ نظام معلومات مساعد على بناء توجهات و تصورات المؤسسة، و ذلك بإنتاج تقارير ذكية بشكل دوري. هذا العمل تمّ خلال فترة تربص لدى الوكالة الجهوية للتشغيل بالبيدة

لقد ركّزنا جهدنا في مشروع التخرج هذا على التقارير ، من أجل هدف واضح و محدد و هو تحليل كل بيانات الوكالة، و على هذا الأساس انقسم عملنا إلى جزئين ، الجزء الاول و هو تصميم نظام معلومات كلاسيكي تكون مهمته الأولى توفير كل ما يحتاجه مخزن المعلومات لإنتاج التقارير المطلوبة

من أجل قاعدة البيانات استعملنا SQL SERVER 2008 لتطويرها ، كما استعملنا الـ ASP.NET بلغة الفيجوال بيسك لتطوير موقع الويب الذي يحوي هذه القاعدة ، كما استعملنا أدوات تطوير مايكروسوفت البصرية لإنشاء المحمل ETL (SSIS,SSRS) و كذا إنتاج التقارير النهائية المطلوبة و التي هي هدف هذا المشروع.

## Résumé

Ce projet s'inscrit dans le cadre de conception et mise en œuvre d'un système de Reporting décisionnel avec la conception d'un entrepôt de données. Ce travail a été fait durant un stage pratique au niveau de l'Agence nationale de l'emploi.

Nous nous intéressons de le cadre de ce PFE essentiellement au Reporting dans un but précis qui est l'analyse et de ce fait, notre travail consiste à créer un système opérationnelle (base de données opérationnelle) pour les processus de routine, c.-à-d. l'ajout, la suppression, la mise à jour etc... Et cette base de données sera la source principale du data warehouse que nous allons réaliser par la même occasion, et ce dernier sera la clé pour réaliser le processus de Reporting attendu par l'entreprise. Notre projet se divise en deux parties, la première c'est la partie opérationnelle et la deuxième se sera la partie décisionnelle. Pour la partie opérationnelle, Notre base de données, sera implémentée sur le SGBD SQL serveur, quand au développement de l'application, il se fera sous ASP plate-forme .NET. Et en ce qui concerne le décisionnel, il sera composé des deux systèmes «ETL » et « Reporting ».

### **Mots Clés:**

Reporting, Data warehouse, source de données, SID, Flux de la demande d'emploi, OLAP, Cube de données, Système opérationnel.

## Abstract

This project is a part of the design and implementation of a decisional reporting system with the design of a data warehouse. This work was completed during an internship at the Regional Employment Agency.

In the scope of this PFE, we are mainly interested by the Reporting for a particular purpose which is the analysis, and therefore, our job consists to create an operational systems (operational database) for routine processes, i.e. add, delete, update etc...

This database will be the main source of data warehouse that we will achieve in the same time, and it will be the key to realizing the process of reporting expected by the company. Our project is divided into two parts, the first is the operational part and the second will be decisional.

For the operational part, our database will be implemented on the SQL Server DBMS, for the application development we will use ASP.NET platform. Regarding the decisional part, this will consist of two systems "ETL" and "Reporting".

### **Key Words:**

Reporting, Data Warehouse, data sources, SID, employment application flow, OLAP, Data cube, Operating System

# INTRODUCTION GENERALE

- Contexte générale
- Problématique
- Objectifs
- Organisation du mémoire

## Introduction générale

---

### 1. Contexte Générale :

Dans le monde turbulent et instable qui est le nôtre aujourd'hui, il est assez délicat de fonder le pilotage de l'entreprise sur des prévisions et des planifications à plus ou moins long terme. Avec la rapidité du changement et la constance des imprévus, les plans se révèlent rapidement inadéquats.

Déployer la stratégie exige un système d'information décisionnel cohérent, il est assez paradoxal de constater l'écart abyssal existant entre la stratégie poursuivie par une organisation et la structure du système d'information décisionnel. Que la qualité de prise de décision soit le véritable moteur de réussite de l'entreprise est désormais une évidence pour tout un chacun. Nul besoin de démontrer une fois de plus. Les temps d'incertitude que nous vivons au quotidien depuis déjà quelques décennies ont propulsé le pilotage de la performance au premier plan des préoccupations. L'erreur de direction ne pardonne plus.

Quel que soient l'organisation, son secteur d'activité ou son modèle économique, la réussite de la stratégie dépend étroitement de sa cohérence avec le système d'information décisionnel. Pour être réactif et prospectif, donc efficace, chaque décideur, expert en son domaine d'intervention, doit disposer impérativement d'un tableau de bord à part entière, composé d'indicateurs minutieusement sélectionnés par ses soins. Le tableau de bord, instrument incontournable de l'aide à la décision, est l'assistant du décideur dans sa démarche d'amélioration continue de la performance au sens de la stratégie poursuivie.

Pour bien concevoir le système d'information décisionnel, il faut se débarrasser de l'idée simpliste qu'il puisse exister quelque part une information déterminée qu'il suffirait de la débusquer pour décider à coup sûr. Ce mythe est encore aujourd'hui entretenu par les services marketing des éditeurs de produits « business intelligence », amalgamant allègrement quantité de données collectées et pertinence de l'information. La problématique est un peu plus délicate que cela. Là encore, les longues explications

## Introduction générale

---

ne sont guère nécessaires. Il suffit que chacun d'entre nous réfléchisse à une situation vécue de prise de décision, professionnelle ou non. Une situation un peu complexe, Lorsque aucune des alternatives candidates ne se détache franchement du lot. En temps d'incertitude, c'est le cas le plus courant. Comment alors évaluer chacun des choix possible ? Comment s'engager ? Comment apprécier le risque de chaque alternative ? Toute décision est une prise de risque. C'est là le point d'articulation du projet décisionnel.

Le premier chapitre consistera à présenter un état de l'art concernant le décisionnel en général et notamment le business intelligence tout en spécifiant leurs composants et en donnant les définitions des différents concepts et en citant l'historique de ces derniers.

Quand au deuxième chapitre, nous allons présenter notre structure d'accueil en citant ses objectifs et sa principale mission, ensuite nous allons citer les indicateurs les plus pertinents que nous avons pu tirer.

Dans le troisième chapitre, le plus important « conception », nous allons définir les différentes démarche de modélisation, ensuite nous allons spécifier nos différents diagrammes (diagramme de classe, diagramme de cas d'utilisation, etc..), et les diagrammes du système décisionnel (table de fait, schéma en constellation, etc..).

Et enfin arrive le chapitre qui dévoile le fruit de notre effort « Mise en œuvre et déploiement » dans lequel nous allons spécifier les différents composants de notre applications (diagramme de composant, diagramme de déploiement), les futurs interfaces de notre système.

## **2. Problématique et Objectifs :**

### **2.1. Problématique :**

L'agence nationale de l'emploi reçoit chaque jour des centaines de demandes et offre d'emplois, et donc à la fin du mois elle va accumuler un nombre assez important des données. Et quand les parties dirigeantes auront besoin de rapports mensuels et des fois

## Introduction générale

---

quotidiens, les agences locales seront obligées de préparer ces rapports via l'exploration de cette base de données massive, et tout ça va consommer beaucoup de temps et effort, et les agences wilaya et régionale à son tour va regrouper ces rapports pour constituer un rapport exploitable qui peuvent contenir facilement des erreurs.

### 2.2. Objectifs :

On estime à arriver à la fin du projet à mettre à la disposition de chaque ALEM un outil pour saisir les données entrant à la production des rapports demandées chaque mois et un outil pour DREM et AWEM pour la meilleur exploitation de ces données. Et voici quelques types des demandes attendue à la fin du ce travail :

1. Le flux de la demande d'emploi.
2. Le flux d'offre d'emploi.
3. Placements des diplômés de la formation professionnelle.
4. Placements des diplômés sortant de l'enseignement supérieur.
5. Répartition des placements effectués.

Notre travail consiste à:

- **Concevoir et mettre en œuvre un système opérationnel** permettant :  
D'archiver l'ensemble des informations concernant le flux de la demande d'emploi, La manipulation des données existantes et leurs mises à jour.
- **La mise en place d'un entrepôt de données** ayant pour rôle :
  - De fournir des présentations claires et adaptées des données servant pour l'analyse.
  - D'analyser dynamiquement les résultats des requêtes.
  - Et d'historier l'information à des fins de statistiques et d'études.
- **Mise en œuvre d'un référentiel d'extraction, de transformation et de chargement -ETL-**, qui a pour but d'assurer l'alimentation régulière du DataWarehouse, à partir des bases de production qu'on déjà été créé.

## Introduction générale

---

- **Déploiement d'une plateforme décisionnelle**, regroupant un ensemble d'outils de « *Business Intelligence* », permettant la génération des rapports, et l'analyse multidimensionnelle des données du DataWarehouse.

### 3. Organisation mémoire :

Notre mémoire sera organisé en deux parties, une partie Contient les notions et concepts théoriques du domaine des SID, et une étude préalable, et une partie pratique comportant la conception et la mise en œuvre du système pour résoudre la problématique.

Donc Ce mémoire est présenté en quatre chapitres, plus une introduction, une conclusion et des perspectives :

#### ■ Chapitre 1 : Etat de l'art

Dans cette partie, nous effectuons un état de l'art concernant les principes du décisionnel.

#### ■ Chapitre 2 : Etude de l'existant

Nous présentons dans ce chapitre l'organisme d'accueil, et définir les besoins.

#### ■ Chapitre 3 : Conception

Présente la conception des divers composants du système aux différents niveaux d'abstractions

#### ■ Chapitre 4 : Mise en œuvre et déploiement

C'est le fruit de ce projet, décrit les choix techniques et architecturales plus la présentation du système développé.

# 1 ETAT DE L'ART

- Qu'est-ce qu'un système décisionnel
- Brève historique
- Chronologie
- Positionnement du décisionnel au sein du système d'information
- Qui a besoin au décisionnel
- L'entrepôt de données (DataWarehouse)
- Les modèles des données
- Technologie OLAP (Online Analytical Processing)

## 1. Introduction :

Avec la généralisation de l'informatique, dans tous les secteurs d'activités, les entreprises produisent et manipulent de très importants volumes de données. Celles-ci sont stockées dans le système opérationnel de l'entreprise, au sein de bases de données, fichiers, ... etc.

L'exploitation de ces données dans un but d'analyse et de support à la prise de décision s'avère difficile; elle est réalisée le plus souvent de manière imparfaite par les décideurs, avec des moyens classiques (Requêtes SQL).

Face à cette inadéquation, il est fondamental de mettre en place une nouvelle informatique décisionnelle, pour obtenir une meilleure compréhension de la valeur des informations disponibles, en définissant les indicateurs pertinents pour faciliter la prise de décision.

## 2. Qu'est-ce qu'un système décisionnel :

Le système d'information décisionnel (SID) est un ensemble de données organisées de façon spécifique, facilement accessibles et appropriées à la prise de décision ou encore une représentation intelligente de ces données au travers d'outils spécialisés. La finalité d'un système décisionnel est le pilotage de l'entreprise.

Les systèmes de gestion sont dédiés aux métiers de l'entreprise pour les assister dans leurs tâches de gestion quotidiennes, et directement opérationnels car maintenus par les utilisateurs sur le terrain.

Les systèmes décisionnels sont dédiés au management de l'entreprise pour l'aider au pilotage de l'activité, et indirectement opérationnels car n'offrant que rarement le moyen d'appliquer les décisions. Ils constituent une synthèse d'informations opérationnelles, internes ou externes, choisies pour leur pertinence et leur transversalité fonctionnelles, et sont basés sur des structures particulières de stockage volumineux (*DataWarehouses*, *bases OLAP*).

Le Principal intérêt d'un système décisionnel est d'offrir au décideur une vision transversale de l'entreprise intégrant toutes ses dimensions. [1]

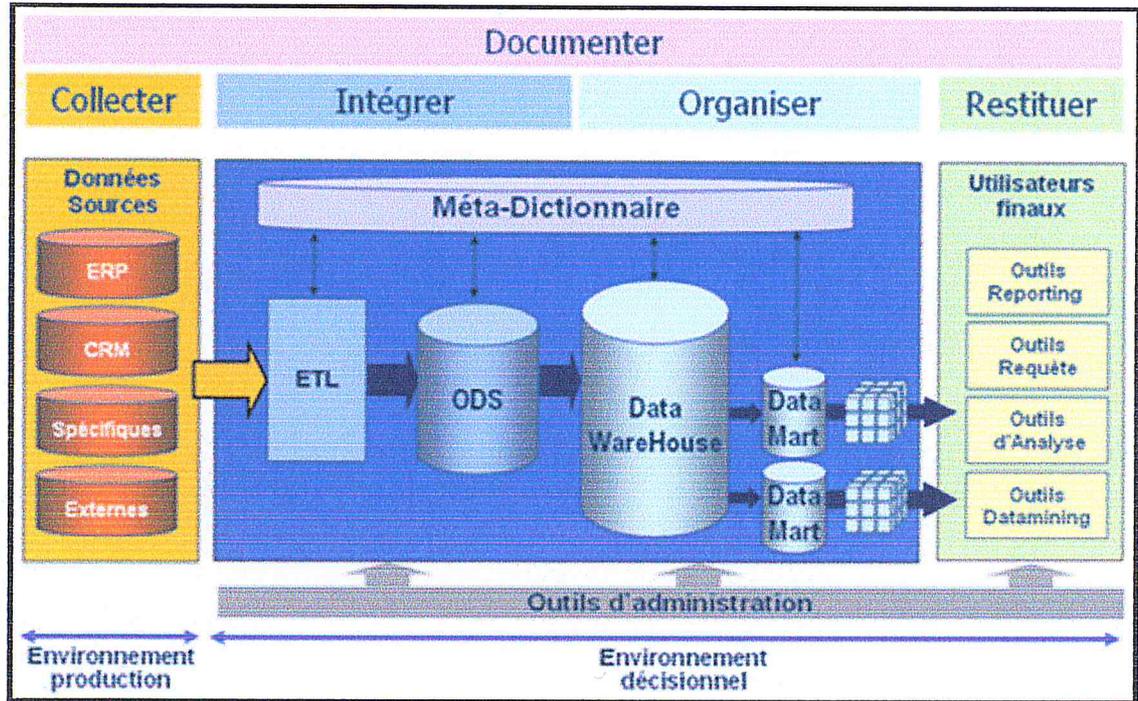


Figure 1.1 : Schéma du système décisionnel. [2]

### 3. Bref historique :

Dans ses débuts, l'informatique décisionnelle s'est contentée tout d'abord de dupliquer les bases de données des systèmes de gestion, afin d'isoler les requêtes d'analyse de données des requêtes opérationnelles. Les requêtes d'analyse étant souvent très lourdes, l'objectif était surtout de préserver les performances des systèmes opérationnels. Ensuite cette base de données dédiée aux requêtes et à l'analyse ce sont progressivement muté et se sont organisé.

Portant du constat qu'il est difficile de croiser des données contenues dans des bases de données distinct, le plus simple a été regrouper ces données éparses. Le Concept de la

base unique pour centraliser les données de l'entreprise est plus que jamais d'actualité. Il s'agit du concept d'entrepôt de données (DataWarehouse).

S'il est plus simple d'analyser ces données une fois qu'elles sont dans l'entrepôt de données, il n'en reste pas moins qu'il faut tout de même remplir l'entrepôt de données.

L'extraction et le croisement des données des différents systèmes opérationnels puis le chargement dans l'entrepôt de données, ont fait émerger des outils dédiés à cette tâche, avec des concepts métiers qui leur sont propres : les outils d'ETL (Extract-Transform-Load).

Si au début, les requêtes d'analyses portaient sur une base relationnelle (dites OLTP pour *Online Transaction processing*), le concept de base multidimensionnelle (dites OLAP pour *OnLine Analytical Processing*) s'est démocratisé fin des années 90. Ce concept de bases de données offrait des performances très largement supérieures aux bases OLTP pour répondre à des requêtes d'analyse. Ces bases OLAP se sont alors couplées avantageusement avec l'utilisation de l'entrepôt de données. En effet, elles offraient à la fois un environnement plus performant, mais permettaient également aux utilisateurs finaux de bénéficier d'une interface simplifiée d'accès aux données, beaucoup plus intuitive qu'une base de données OLTP. On parle alors de méta-modèle.

Si l'informatique décisionnelle s'est contentée dans ses débuts d'une approche technicienne, elle progresse et converge de plus en plus rapidement vers le poste utilisateur et s'adapte aux métiers des utilisateurs. Nous sommes encore aujourd'hui dans cette phase de convergence.

#### 4. Chronologie :

- **L'infocentre** : l'infocentre est une copie à l'identique des bases de données dans un nouvel environnement (années 70 et 80).
- **L'EIS** : *Executive Information System* proposant les premiers tableaux de bord dans les années 1990.

- **Les entrepôts de données (ED):** qui sont considérés comme étant le lieu de stockage des gros volumes de données devant être analysés.
- **Les bases de données multidimensionnelles (OLAP) :** une base où chaque indicateur est analysé en fonction de plusieurs critères ou dimensions. Et enfin
- **La Business Intelligence:** qui regroupe les fonctions d'analyse des données et de Reporting.

### 5. Positionnement du décisionnel au sein du système d'information :

Le développement de nouveaux marchés des entreprises passe par la maîtrise de la connaissance des divers environnements sur lesquels pourront croître ces nouveaux marchés, et par la connaissance d'informations légales, juridiques, fiscales, politiques, techniques et d'informations nécessaires aux marketings mix. Toutes ces sources d'informations issues du monde entier, alimentent par tous les canaux possibles et imaginables toutes les strates de l'entreprise qui doit alors les digérer pour en extraire les informations les plus importantes.

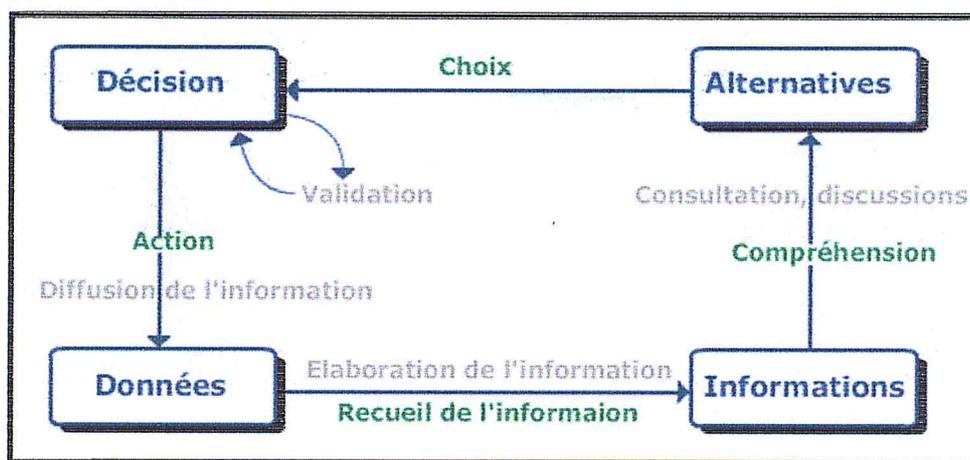


Figure 1.2: Prise de la décision. [3]

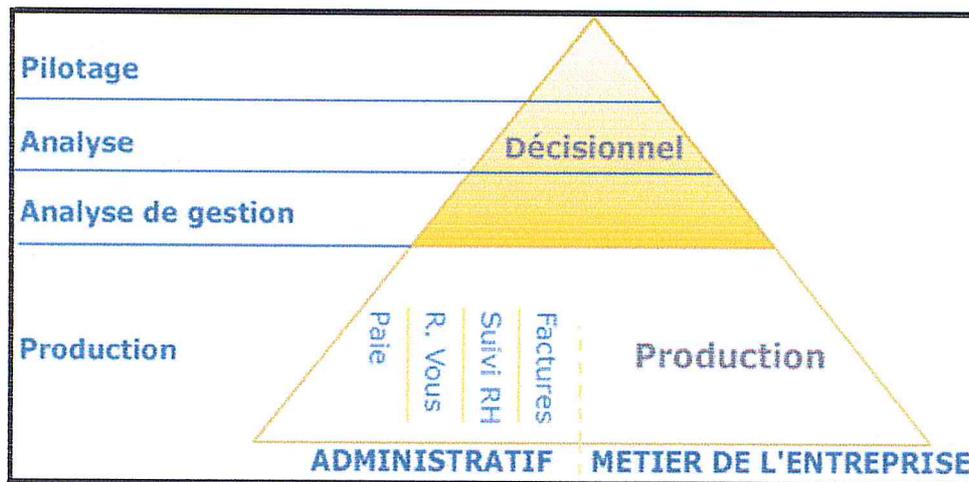


Figure 1.3 : Le décisionnel au sein du système d'information. [3]

## 6. Qui a besoin au décisionnel ?

Comme cela peut se deviner, les décideurs sont les principaux utilisateurs des systèmes décisionnels. Les décideurs sont généralement des " marqueteurs " ou analystes en général. Ces derniers établissent généralement des plans marketing qui leur permettent de mieux cibler leur clientèle, de les fidéliser etc... Et pour cela, ils ont besoin d'indicateurs et des données résumées de leur activités (ils n'ont souvent besoin de détail que pour des cas spécifiques).

Par exemple, contrairement aux systèmes relationnels (ou base gestion) où les utilisateurs chercheront à connaître leurs transactions pour faire un bilan, les systèmes décisionnels eux cherchent plutôt à donner un aperçu global pour connaître les tendances des clients (d'où l'opposition des deux modes (quantitatif contre qualitatif)). [4]

## 7. L'entrepôt de données (DataWarehouse):

### 7.1. C'est quoi un DataWarehouse :

Un entrepôt de données, ou DataWarehouse, est une vision centralisée et universelle de toutes les informations de l'entreprise. C'est une structure (comme une base de

données) qui a pour but, contrairement aux bases de données, de regrouper les données de l'entreprise pour des fins analytiques et pour aider à la décision stratégique.

L'entrepôt de données est l'élément central de l'informatique décisionnelle. En effet, l'entrepôt de données est le meilleur moyen que les professionnels ont trouvé pour modéliser de l'information pour des fins d'analyse, et il ne serait pas étonnant que d'ici quelques années un nouveau concept apparaisse pour révolutionner l'informatique décisionnelle. [5]

**7.2. Architecture globale d'un DataWarehouse:**

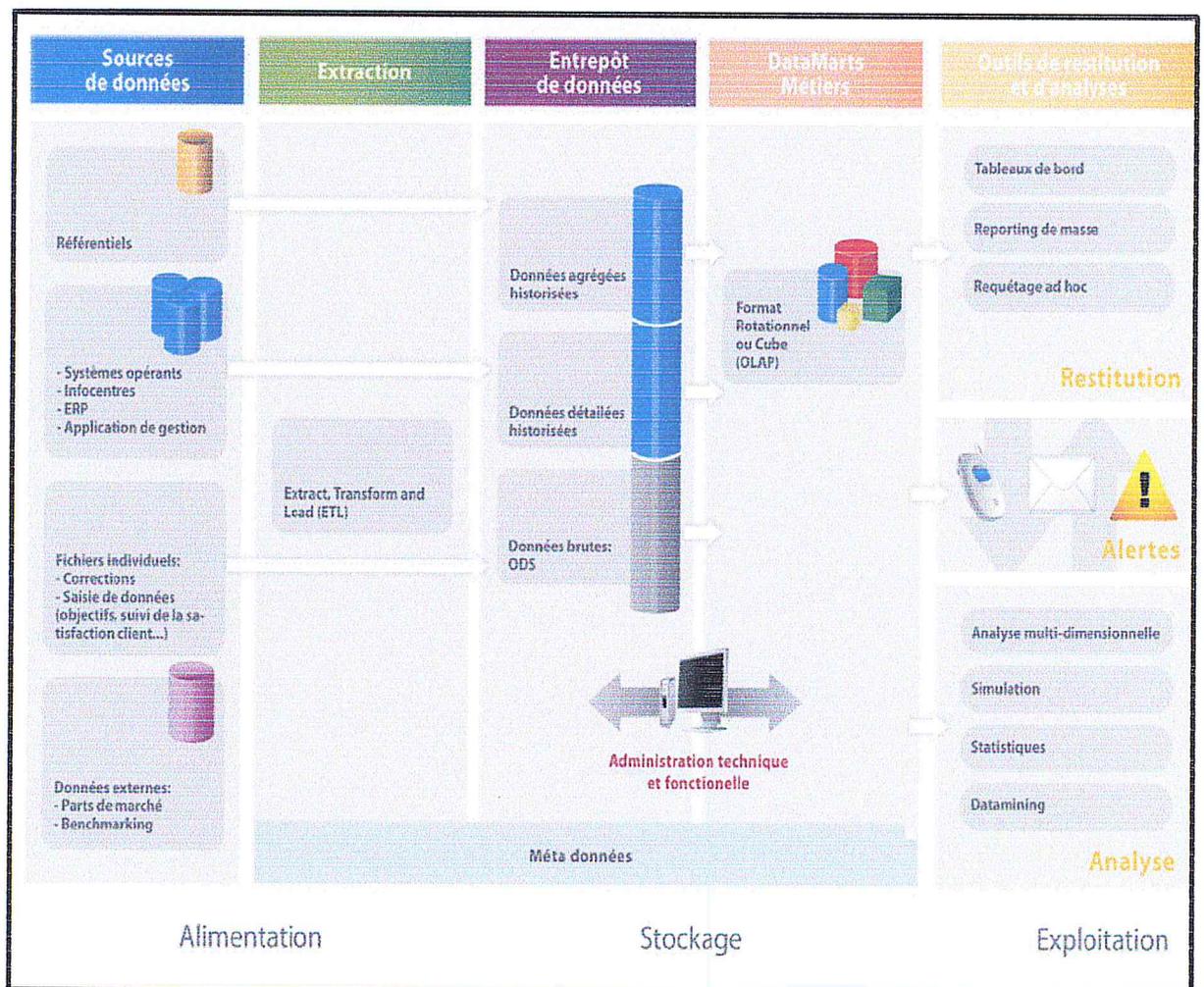


Figure 1.4: Architecture globale d'un DataWarehouse. [6]

### 7.2.1. Définition concepts :

- **Source de données (base de production) :** On appelle, d'une façon générale, bases de production toutes les sources (qu'il s'agisse de données de production, d'informations internes ou d'informations externes quel que soit leur mode de stockage « BD, Fichiers plat ou XML ...») dont il va falloir extraire des données pour alimenter le DW. On y retrouvera notamment, les données commerciales, les données du système comptable, les données sur le personnel, les informations sur la concurrence, des informations externes produites par divers organismes nationaux ou internationaux de statistiques et autres fournie par des services web.
- **Extraction :** Il s'agit de la première étape de récupération des informations dans l'environnement de l'entrepôt de données. L'extraction comprend la lecture et la compréhension de la source de données, ainsi que la copie des parties nécessaires à une exploitation ultérieure dans la zone de préparation.
- **Entrepôt de données :** Toutes les données qu'elles proviennent du système de production de l'entreprises, des sources externes, vont devoir être organisées, coordonnées, intégrées et enfin stockées pour donner à l'utilisateur une vue intégrée et orientée métier dans un « *entrepôt de données* » ou bien « *DataWarehouse* ». L'entrepôt de données est alimenté par la zone de préparation des données. En outre, l'entrepôt de données fréquemment mis à jour en tenant compte de la charge au fur et à mesure que les données sont corrigées, que les instantanés (vue matérialisé) s'accumulent et que les statuts et les libellées évoluent.
- **DataMart :** Sous ensemble logique d'un entrepôt de données. Il représente un projet réalisable. Au-delà de cette définition relativement simple, on considère le datamart comme la réduction de l'entrepôt de données à un seul processus où on a un groupe de processus, ciblant un groupe métier spécifique.
- **Outils ANALYSE :** permettent de manipuler les données suivant des axes d'analyses. L'information est visualisée au travers d'interfaces interactives et fonctionnelles dédiées à des décideurs souvent non informaticiens (directeurs, chefs de services,...).

Exemple : analyse multidimensionnel, statistique, simulation, datamining.

- **Outils Restitution** : C'est le bout de la chaîne. Il s'agit du point d'utilisation du système par les utilisateurs. La satisfaction de ceux-ci dépend de la capacité des outils de restitution à répondre à leur besoin en information et aide à la décision. Les types de restitution possible sont : Reporting de masse, requetage *ad hoc*, tableaux de bord.
- **Métadonnées**: Métadonnées signifie correctement « des données sur des données ». Une métadonnée est une donnée qui a pour but de décrire une autre donnée. Les métadonnées peuvent être des informations complémentaires, nécessaires à la compréhension d'une autre information ou dans le but de permettre une utilisation pertinente. L'un des grands principes de l'entrepôt de données est de conserver la trace des données produites, après agrégation, consolidation et application des règles de gestion. Cette fonction est assurée par les métadonnées. Dans ces métadonnées seront stockées des informations telles que le nom de la base de production dont la donnée est extraite, la date et l'heure de la dernière extraction, la fréquence de mise à jour de cette information, ... etc.
- **Données brutes ODS (Operational Data Store)** : Un stockage de données partagées entre applications système d'information décisionnel en préservant le système de production (ou système opérationnel). Il centralise la collecte de l'information issue des applications de gestion, il les transforme et les valorise puis il les injecte dans l'entrepôt de données. L'ODS gère aussi les rejets ainsi qu'un journal d'opération pour que l'administrateur puisse prendre en compte et corrige anomalies rencontrées. L'ODS s'appuie généralement sur une solution progiciel de type ETL (Extraction, Transformation and Loading) et sur une base de données relationnelle permettant de structurer les tables de références et de travail nécessaire à son fonctionnement.
- **OLAP (OnLineAnalyticalprocessing (Cube))** : La finalité d'un Datawarehouse est d'obtenir des vues multidimensionnelles. Ces vues sont représentées sous la forme d'un cube en trois dimensions sachant qu'une base multidimensionnelle peut comporter de nombreuses dimensions. Les systèmes OLAP mettent en

œuvre des technologies permettant de rassembler, gérer, traiter et présenter des données multidimensionnelles à des fins d'analyse et de décision. Un outil OLAP est capable de fournir une information multidimensionnelle partagée pour l'analyse rapide.

### 7.3. Fonctions processus DW : [7]

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

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

#### 7.3.1. Sélection des données sources :

Quelles sont les données de production qu'il faut sélectionner pour alimenter l'ED ?

Toutes les données sources ne sont forcément pas utiles ;

La synthèse des données sources a pour but de les enrichir ;

La sélection des données utiles à partir des BD de production n'est pas simple à faire

Les données sont :

- **hétérogènes** (différents SGBD et différentes méthodes d'accès);
- **diffuses** (différents environnements matériels et différents réseaux interconnectés ou non);
- **complexes** (différents modèles logiques et physiques principalement orientés vers les traitements transactionnels).

La définition de la granularité dépend du niveau de raffinement de l'information qu'on veut obtenir.

Il existe plusieurs niveaux de données :

- Les données sont parfois assemblées avant d'être injectées dans l'ED permettant une vision intégrée et transversale de l'entreprise.

- Elles peuvent être agrégées (Elles constituent ainsi un autre niveau de détail)
- Elles seront par la suite structurées dans des espaces d'analyse : l'Entrepôt de données, des cubes de données, des Datamarts.
- Elles seront finalement à un niveau de présentation, où elles peuvent avoir plusieurs formes (tableaux, graphiques, tableaux de bord, règles de connaissances...).

### **7.3.2. Extraction des données :**

La première phase de la construction d'un Datawarehouse consiste à extraire les données utiles des systèmes opérationnels, qui sont dans de nombreux cas hétérogènes (avec plusieurs gestionnaires de bases de données et plusieurs méthodes d'accès), diffuses (avec plusieurs environnements matériels et plusieurs réseaux interconnectés ou non), et complexes (avec plusieurs modèles logiques et physiques bien souvent orientés prioritairement vers les traitements transactionnels).

L'extraction comprend la lecture et la compréhension de la source de données, ainsi que la copie des parties nécessaires à une exploitation ultérieure dans la zone de préparation.

### **7.3.3. Transformation :**

La deuxième phase du processus ETL consiste à transformer les données extraites de manière à obtenir un ensemble homogène de données, qui deviennent ainsi comparables, additionnables, etc.

Les données peuvent être transformées de nombreuses manières, par exemple nettoyées (correction orthographique, résolution de conflits de domaine, traitement du problème des éléments manquants, conversion à des formats standard), combinées à partir de sources multiples, dédoublées s'il y a lieu et pourvues de clés propres à l'entrepôt de données. Ces transformations sont le prélude au chargement dans la zone de présentation de l'entrepôt de données.

On rencontre plusieurs étapes de transformation :

- **Nettoyage des données** : à savoir correction des fautes d'orthographe, résolution des conflits de domaine (noms de villes incompatibles avec le code postal, par exemple), résolution des cas d'informations manquantes, et conversion en format standard.
- **Le dé doublonnage** : Dès lors que l'on va extraire d'un ensemble de bases des données en vue de les agréger, on va être confronté au problème des redondances. On trouvera par exemple la référence client tant au sein de la base commerciale qu'au sein de la base marketing. On retrouvera le produit tant au niveau des bases de production, qu'au niveau commercial, qu'au niveau comptable qu'au niveau marketing, etc. Ces redondances figureront au sein d'une même source ou au sein des diverses sources. Le dé doublonnage interviendra sur chacun des niveaux requis pour garantir la cohérence et donc la qualité du Datawarehouse.
- **Le formatage et la restructuration** : Vise à normaliser les informations et à préparer leur injection dans le système cible en les convertissant au format cible. Cette normalisation ou dé-normalisation suivant les cas, s'applique fréquemment sur des codes (tel que le code INSEE dont l'ensemble des digits pris séparément a une signification particulière) ou sur des adresses. Le cas des adresses est typique car on découpera par exemple le champ source « adresse » en n champs cibles : « numéro », « rue », « nom rue », « particularités », « code postal », « ville », « pays ».
- **Création de clés** : de substitution pour chaque enregistrement dimensionnel afin d'éviter de dépendre des clés définies dans le système source. Dans ce cas, c'est le processus de génération des clés qui assure l'intégrité référent entre les tables dimensionnelles et les tables des faits.

#### 7.3.4. Chargement :

À la fin du processus de transformation, les données prennent la forme d'images d'enregistrement prêtes à être chargées. En règle générale, le processus de chargement

des données dans l'entrepôt s'effectue par réplique des tables des faits et des tables dimensionnelles, qui seront ensuite présentées aux services de chargement en masse de chaque datamart destinataire. Le chargement en masse est une fonctionnalité très importante à mettre en opposition avec le chargement ligne à ligne, qui est un processus nettement plus lent. Le datamart de destination doit alors indexer les données afin d'optimiser les performances de requête, si ce n'est déjà fait.

## **8. Les modèles des données :**

### **8.1. Définition modèle :**

Un modèle est la représentation d'un objet, d'un système ou d'une idée sous une forme quelconque autre que celle de l'entité représentée elle-même. Sa fonction est de nous aider à expliquer, à comprendre, ou à améliorer un système. Le modèle d'un objet peut être une réplique exacte de cet objet, ou une abstraction des propriétés saillantes de l'objet. [8]

### **8.2. Le modèle des données des DW :**

Un DataWarehouse est une base de données. Ainsi, le modèle de données du DataWarehouse est le cœur du système décisionnel. La modélisation d'un système décisionnel nécessite des approches spécifiques car l'utilisation dont ce dernier va faire l'objet différera radicalement de celle des systèmes d'information plus classiques.

#### **8.2.1. Caractéristiques d'utilisation OLAP vs OLTP :**

Les techniques couramment utilisées pour modéliser les données ont initialement été conçues pour qu'elles s'adaptent à des problématiques qui n'existent pas dans le cadre d'un système décisionnel.

Dans la mise en œuvre des systèmes d'informations, nous maîtrisons des approches centrées sur des méthodologies telles que *Merise*. Dans leurs composantes liées à la modélisation des données, ces méthodes sont précises, standardisées, puissantes et assez peu contestées.

Le modèle « *entité-association* » est le plus utilisé, permettant la création d'un modèle logique relationnel. Toutes les théories liées à ces modèles sont largement utilisées dans les entreprises.

Ces techniques sont apparues alors que l'informatique était destinée à l'automatisation des processus à caractère transactionnel. Ces applications sont communément nommées *OLTP*.

Cependant, l'informatique de décision, que certains désignent par *OLAP*, justifie une remise en cause des méthodes de conception d'un modèle de données.

○ **Caractéristiques d'un contexte OLTP :**

Dans la plupart des systèmes transactionnels, le rôle d'un modèle est de garantir la persistance des données. De fait, la base de données est conçue pour garder la trace d'événements survenus dans l'entreprise.

Dans un contexte transactionnel, le modèle de données est destiné à minimiser les redondances, pour préserver la fiabilité et la cohérence du système. Des concepts, tels que les formes normales, les clés uniques, les clés étrangères ou de contraintes d'intégrité référentielle, permettent de garantir constamment l'intégrité de la base de données. L'origine de ce souci de minimisation des redondances découle principalement de ce que les systèmes transactionnels effectuent leurs mises à jour en ligne, éventuellement au travers d'un ensemble d'applications partageant le même modèle de données.

Dans un système transactionnel, la conception est orientée processus et le modèle de donnée intervient en support de ceux-ci. De point de vue de l'utilisateur, le modèle de données est totalement transparent.

Les requêtes sont toujours prévisibles car elles sont effectuées au travers d'une application le plus souvent développée par la même équipe que celle qui a la charge du modèle de données.

Les données sont généralement accédées par des clés, notamment des clés uniques. Une bonne indexation permet de garantir des temps de réponse dépendant davantage du volume de données à traiter pour réaliser la transaction que du volume global de la base de données. Les volumes de données qui doivent être accédés pour traiter une transaction ou retournés en résultat de celle-ci sont limités. Il est très rare qu'une requête transactionnelle nécessite de rassembler ou d'agréger des informations issues d'un grand nombre de tables.

○ **Dans un monde décisionnel (OLAP) :**

Dans un contexte décisionnel, les requêtes sont complexes, les redondances plus difficiles à maîtriser ; l'optimisation consiste à anticiper sur les chemins d'accès aux données qui sont fréquemment employés plutôt qu'à faire des optimisations requête par requête.

Un Datawarehouse est une base dédiée au décisionnel, l'information est mise à la disposition des utilisateurs mais les mises à jour ne sont jamais faites en ligne.

Les seules mises à jour effectuées sur le Datawarehouse émaneront des systèmes de production, lors des phases de chargement (processus d'acquisition de données).

Il devient donc envisageable d'introduire des redondances, à condition de les maîtriser dans le processus d'alimentation. Dans un contexte décisionnel, les requêtes manipulent régulièrement des ensembles.

Elle effectue des sélections ou des restrictions de population, des regroupements, des calculs, des agrégations, etc.

Pour répondre aux besoins des utilisateurs, même si le résultat des requêtes peut n'être constitué que de quelques lignes, il faudra très souvent manipuler des volumes importants. Dès lors, obtenir des temps de réponse proportionnels au volume de données résultat d'une requête est beaucoup plus difficile qu'en transactionnel.

Il convient d'optimiser les requêtes effectuées fréquemment en prédefinisant physiquement des sous-ensembles de données, moins importants en taille que les données plus détaillées, mais suffisants pour résoudre les requêtes les plus courantes.

Une autre caractéristique du décisionnel veut que les utilisateurs cherchent à mettre en relation des éléments qui a priori ne sont pas corrélés au départ. Pour y parvenir, des requêtes complexes sont nécessaires, interrogeant un nombre important de tables. Face à cette complexité, le Datawarehouse doit pouvoir réagir dans des délais raisonnables.

Un Datawarehouse vise à répondre aux besoins des utilisateurs en termes d'informations et non en termes d'applications.

Dans un contexte décisionnel, du point de vue de l'administrateur de base de données, une des plus grosses difficultés qui se posent est de gérer l'imprévisible. En effet, les requêtes sont le plus souvent *ad hoc*, générées par l'utilisateur au travers d'un outil, et il est donc impossible d'optimiser chacune de celles-ci au cas par cas.

La dernière caractéristique du monde Datawarehouse est qu'il permet le plus souvent de mettre en place un modèle de données intégré, qui entend être transversal à l'entreprise. Ce modèle se constitue le plus souvent de manière incrémentale, au fur et à mesure des réalisations successives des projets décisionnels de l'entreprise.

### 8.2.2. Les techniques de modélisation : [9]

Cinq axes permettent de qualifier un modèle de données décisionnel :

- La lisibilité du point de vue de l'utilisateur final.
- Les performances au chargement.
- Les performances liées à l'exécution des requêtes.
- L'administration : ce n'est pas tant construire le Datawarehouse que le faire vivre qui pose des problèmes aux entreprises. Il faudra tracer les requêtes et identifier celles qui sont lancées fréquemment, maîtriser et industrialiser tous les processus d'extraction.

- L'évolutivité qui permet de faire en sorte que le développement d'un DataWarehouse soit incrémental, et pas seulement itératif : un développement itératif peut en effet amener à définir plusieurs modules applicatifs, indépendants les uns des autres. Par développement incrémental, l'intégration de chacun des modules doit être considérée dans la mise en œuvre itérative, afin de s'assurer de ce que l'homogénéité globale du système est prise en compte.

### 8.2.2.1. Le modèle de données normalisé :

Soit le modèle de données normalisé présenté dans la figure suivante :

En décisionnel, ce modèle permet par exemple de ventiler les chiffres d'affaires par produit, par client, etc. Ce type d'approche est courant dans les entreprises en matière de modélisation. Si le système à mettre en œuvre permet à la fois des sélections et des mises à jour en ligne, elle est adaptée. Dans un contexte décisionnel où les mises à jour en ligne ne sont pas d'actualité, sa pertinence doit être reconsidérée. D'un point de vue décisionnel, la sémantique de ce modèle est faible. Les informations intéressantes pour l'utilisateur n'existent pas *a priori*, elles doivent être extrapolées. Les indicateurs devront être recalculés dynamiquement à chaque requête.

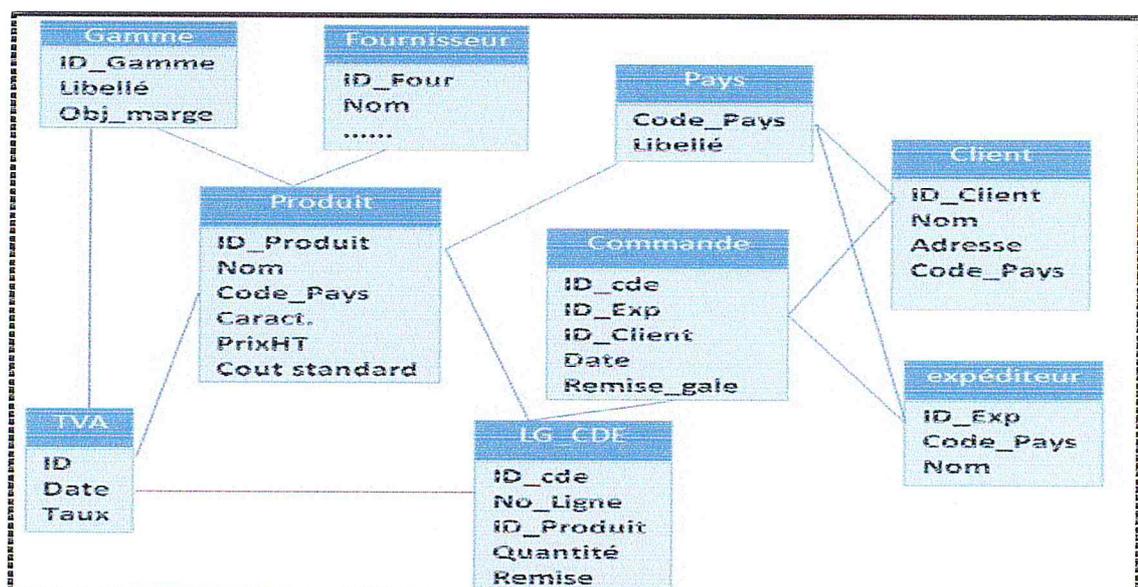


Figure 1.5 : Le modèle de données normalisé.

Le modèle est en revanche très complet. Il laisse une marge d'autonomie très forte à l'utilisateur.

En réalité, un modèle d'entreprise pourra contenir des centaines ou des milliers d'entités, et donc d'autant de tables au niveau physique. Une requête utilisera peut-être des dizaines de tables et qui sera très complexe à formuler pour l'utilisateur et à traiter pour l'optimiseur de la base de données. Les performances seront donc au mieux médiocres et pire inacceptables. Dans des systèmes décisionnels simples, où un nombre réduit d'utilisateurs lancent peu de requêtes sur un modèle de données de petite envergure, ce type d'approche peut fonctionner. Dans les autres cas, il faut absolument recourir à d'autres techniques.

### 3.2.2.2. Dé-normalisation pour la décision:

Cette approche vise à adapter le modèle précédent aux besoins liés au décisionnel. La transformation consiste à dé-normaliser et à *pré-calculer certains agrégats*, donc à introduire des redondances. Aucune technique formelle de dé-normalisation n'existe dans le domaine public. L'approche doit être pragmatique. Aboutir à un tel modèle découle d'une analyse précise des besoins des utilisateurs. Ainsi, le modèle précédent devient :

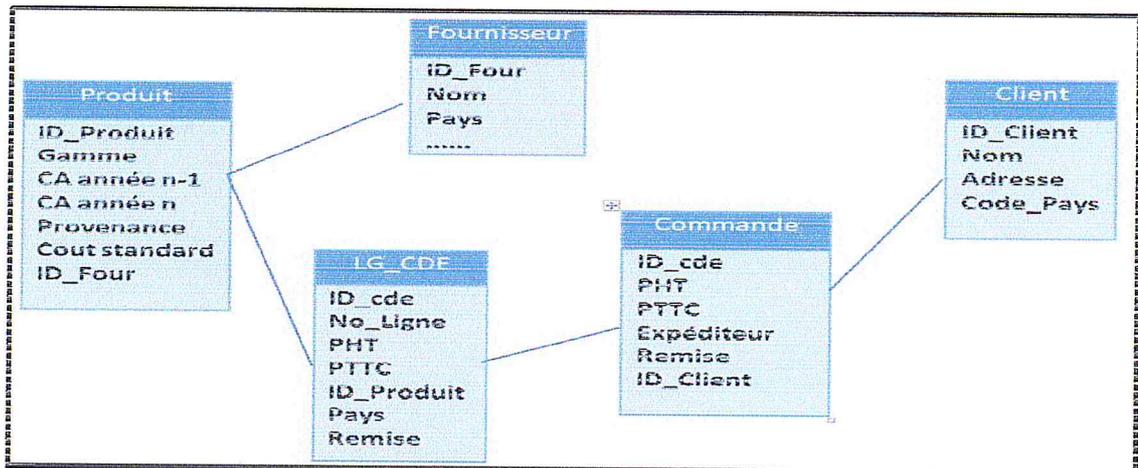


Figure 1.6 : Le modèle de données dé-normalisé.

Remarquez que le modèle contient un nombre de tables plus restreint, chaque table étant associée à un sujet d'intérêt. L'orientation « sujet » rapproche le modèle des besoins des utilisateurs. Le modèle présente un certain nombre d'informations agrégées très fréquemment demandées.

Ce modèle est moins complexe que le précédent. Le nombre de tables a été diminué d'un facteur de 2.

En appliquant le même facteur pour un modèle normalisé de 200 tables, on aboutit à une centaine de tables, ce qui reste complexe et peu lisible. Le gain en performance par rapport au modèle normalisé est également très relatif. Le nombre de tables a diminué, donc aussi le nombre de jointures nécessaires pour les requêtes décisionnelles. En contrepartie, les tables sont plus grosses. Ainsi, les requêtes seront plus simples mais porteront sur des tables plus volumineuses. [10]

#### **8.2.2.3. La modélisation dimensionnelle :**

La modélisation dimensionnelle est une discipline de modélisation des données qui se positionne comme solution de rechange à la modélisation entité/relation. Un modèle dimensionnel contient les mêmes informations qu'un modèle entité/relation, mais présente les données dans un format asymétrique dont les objectifs sont :

- ✓ La bonne compréhension de l'utilisateur,
- ✓ La performance des requêtes,
- ✓ L'adaptation au changement.

Il y a trop d'entrepôts de données qui se sont effondrés à cause de conception relationnelles trop complexes. Les techniques de modélisation dimensionnelles ont été appliquées avec succès dans certaines situations au cours des quinze dernières années.

### **8.3. Composants d'un modèle dimensionnel :**

#### **8.3.1. Table de fait :**

La table de fait est la table principale de tout modèle dimensionnel, destinée à héberger des données permettant de mesurer l'activité (les mesures).

Une table de fait est une table qui contient les données à analyser. Elle est distinguée grâce à sa grande taille et à son emplacement au centre du schéma.

La table des faits contient les clés des différentes dimensions qui sont reliées à elle ainsi que les mesures de l'activité. Ces dernières sont généralement numériques additives et valorisées de façon continue.

La clé primaire de la table de fait est la concaténation des clés primaires de toutes les tables dimensions reliées à elle. [9]

#### **8.3.2. Les faits (mesures) :**

Un fait représente un sujet d'analyse. Il est constitué de mesures relatives au sujet traité. Ces mesures sont numériques et généralement valorisées de façon continue.

La table de faits contient des champs qui ne sont pas des clés étrangères, ce sont les faits. Ils doivent être valorisés de façon continue et être additifs (chiffre d'affaire, quantité vendue, etc.).

Un fait, une mesure, ou encore un indicateur, est une information déterminée par la combinaison de deux ou plusieurs entités, susceptible de constituer le résultat ou un élément de résultat d'une requête. [8]

#### **8.3.3. Table de dimension :**

Ce sont des tables qui entourent la table de fait dans le schéma en étoile. Petites par rapport à la table de faits, une table de dimension contient une clé et des attributs généralement textuels ou numériques discrets plus ou moins statique décrivant une dimension de l'activité.

Chaque dimension est définie par sa clé primaire, qui assure l'intégrité référentielle avec la ou les tables des faits à laquelle elle est liée.

Une dimension est un ensemble des valeurs décomposables. Les valeurs d'une dimension sont généralement organisées en hiérarchie.

Une hiérarchie est une décomposition d'une dimension en niveaux, afin de permettre à l'utilisateur d'examiner ses indicateurs à différents niveaux de détail, allant du niveau global au niveau le plus fin. On aura alors une vision pyramidale des données. La base de la pyramide représentant le niveau le plus détaillé, le plus haut est le niveau global.

**Remarque:**

- Une dimension peut avoir plusieurs hiérarchies.
- L'attribut d'une dimension peut appartenir à plusieurs hiérarchie ou à aucune.

Table de faits des ventes journalières
Clé date (CE)
Clé Produit (CE)
Clé magasin (CE)
Quantité vendue
Montant des ventes

**Figure 1.7:** Table de fait.

Table de dimension produit
Clé Produit (CE)
Description du produit
Numéro US
Description de la marque
Description de la catégorie
Description du rayon
Description du type d'emballage
Taille de l'emballage
Description matière grasses
Description type de régime
Poids
Unités de mesures de poids
Type de stockage
Type de durée étagère
largeur sur étagère
Hauteur sur étagère
Profondeur sur étagère
..., et bien d'autre attributs.

**Figure 1.8:** Table de dimension.

#### 8.4. La modélisation multidimensionnel :

Le modèle entité-association est souvent utilisé pour la conception de la base de données. Ce modèle permet d'éliminer les redondances en introduisant un grand nombre de nouvelles entités. De ce fait l'accès aux données devient plus compliqué et le schéma difficile à lire. L'utilisation de ce modèle pour la conception des entrepôts de données n'est donc pas appropriée. [11]

La modélisation multidimensionnelle est une méthode de conception associée aux entrepôts de données. Elle consiste à considérer un sujet analyse comme un point dans un espace multidimensionnel. Le constructeur fondamental des modèles multidimensionnels est le cube de données.

Les données sont organisées de manière à mettre en évidence le sujet analysé et les différentes perspectives d'analyse. [12]

Cette modélisation a donné naissance aux concepts de fait et de dimension. [13]

Il existe 3 formes de modèles multidimensionnels :

1. Le modèle en étoile.
2. Le modèle en flocon de neige.
3. Le modèle en constellation.

##### 8.4.1. Modèle en étoile :

Un schéma en étoile ne comporte, en plus de la table de faits, qu'une table par dimension. Cette simplification est obtenue au prix d'une forte dé-normalisation.

Le modèle en étoile est générateur d'une forte redondance, et c'est là, son principal défaut. Mais ce défaut est sans grande conséquence ici car :

- La redondance des données ne compromet pas la cohérence d'une base de données destinée à la consultation et ne subissant pas de mises à jour transactionnelles.

- L'espace occupé par les tables dimensionnelles étant insignifiant par rapport au volume de la table de faits, la redondance dimensionnelle n'a qu'un effet négligeable sur l'encombrement total de la base de données.

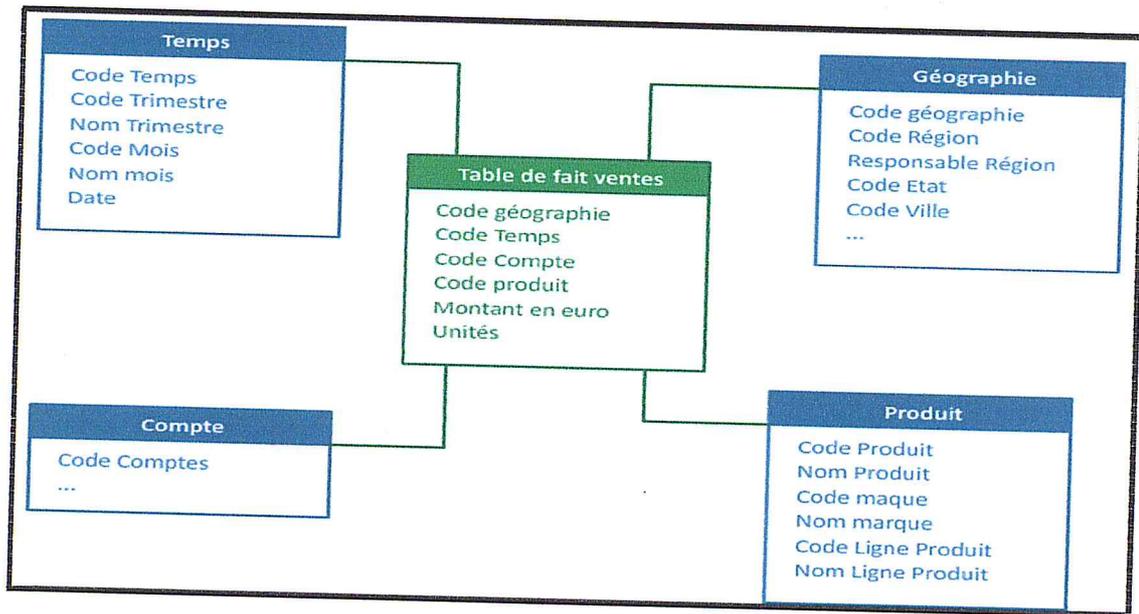


Figure 1.9 : Modèle en étoile [14]

L'avantage technique procuré en contrepartie est évident. Toutes les tables dimensionnelles ont une liaison directe avec la table de faits. Le nombre de tables pouvant être impliquées dans une requête, en plus de la table de faits, est inférieur ou égal au nombre de dimensions du contexte, quelle que soit la complexité des dimensions. [8]

Ce modèle présente des avantages et des inconvénients qui sont :

Avantage :

- Facilité de navigation.
- Performances : nombre de jointures limité.

Inconvénients :

- Toutes les dimensions ne concernent pas les mesures.
- Redondances dans les dimensions.

#### 8.4.2. Modèle en flocon :

Un réseau de tables jointes selon une figure particulière dite *schéma en flocon*. Dans ce mode de représentation, l'association conceptuelle qui contient les faits devient une table, dite *table de faits*, et chacune des entités dimensionnelles devient une table distincte.

La table de faits, outre les indicateurs significatifs qu'elle comporte par définition, possède dans sa structure un ensemble de *clés étrangères* dont chacune assure la liaison avec la table du niveau le plus fin de chaque dimension.

Lors de l'exécution d'une requête, les critères de sélection fournis par l'application portent sur les tables dimensionnelles et les résultats, par le jeu des jointures, sont extraits de la table de faits.

La structure en flocon présente l'avantage de ne laisser place à aucune redondance dans les données dimensionnelles.

Mais cet avantage de principe ne compense pas quelques inconvénients majeurs dans un environnement ouvert aux consultations complexes :

- Les requêtes invoquant des conditions sur des propriétés situées à un niveau élevé dans une hiérarchie (i.e. logiquement éloignées de la table de faits) sont sensiblement pénalisées, puisqu'elles impliquent une navigation plus longue. Or la complexité et le temps de traitement d'une requête, toutes choses égales par ailleurs, augmentent en raison directe du nombre de tables impliquées dans la jointure.
- Le nombre de clés techniques à générer, pour jalonner correctement les chemins hiérarchiques, est important, ce qui complique la tâche des programmes de chargement de la base de diffusion.
- Dans certains cas, la représentation d'un niveau hiérarchique par une table spécifique ne fait qu'alourdir le schéma sans aucun avantage technique. [8]

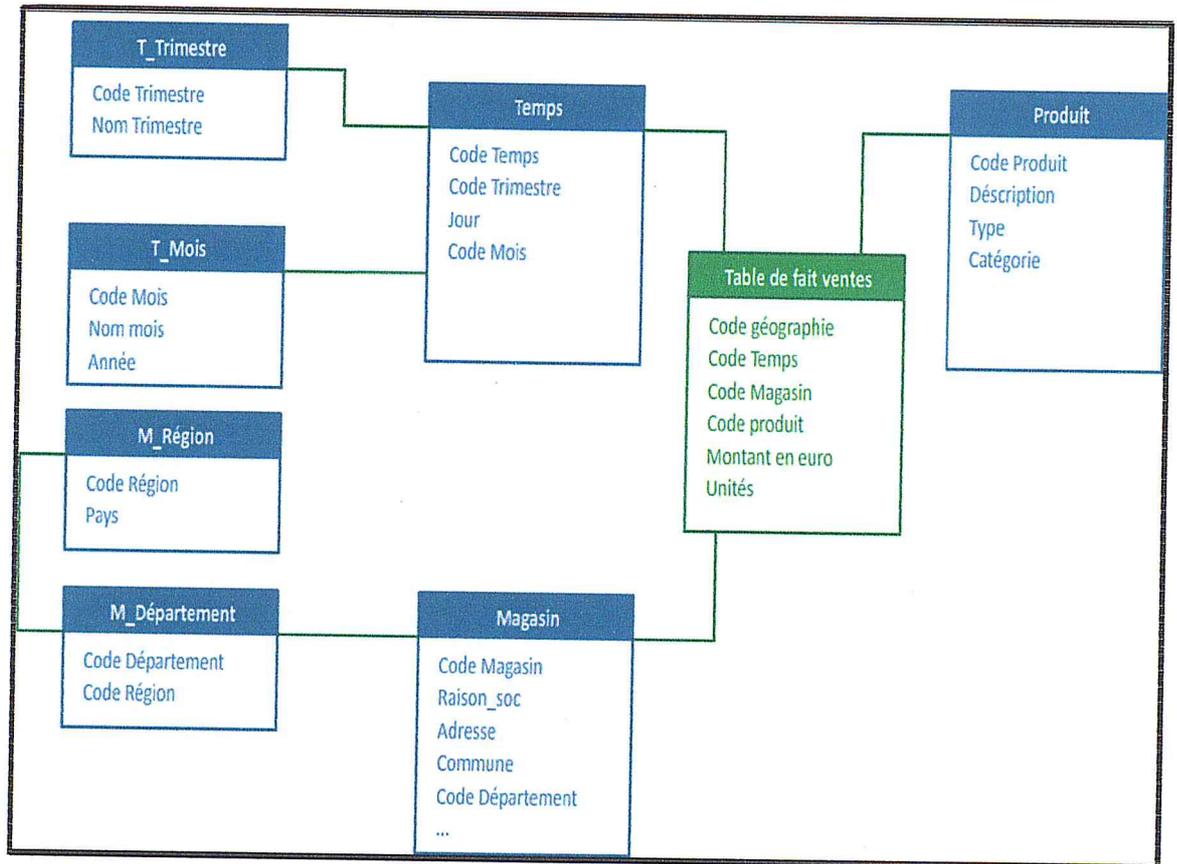


Figure 1.10 : Modèle en flocon. [14]

Comme le modèle en étoile, le modèle en flocon présente lui aussi les avantages et les inconvénients suivants :

Avantage :

- réduction du volume.
- permettre des analyses par pallier (drill down/Roll up) sur la dimension hiérarchisée.

Inconvénients :

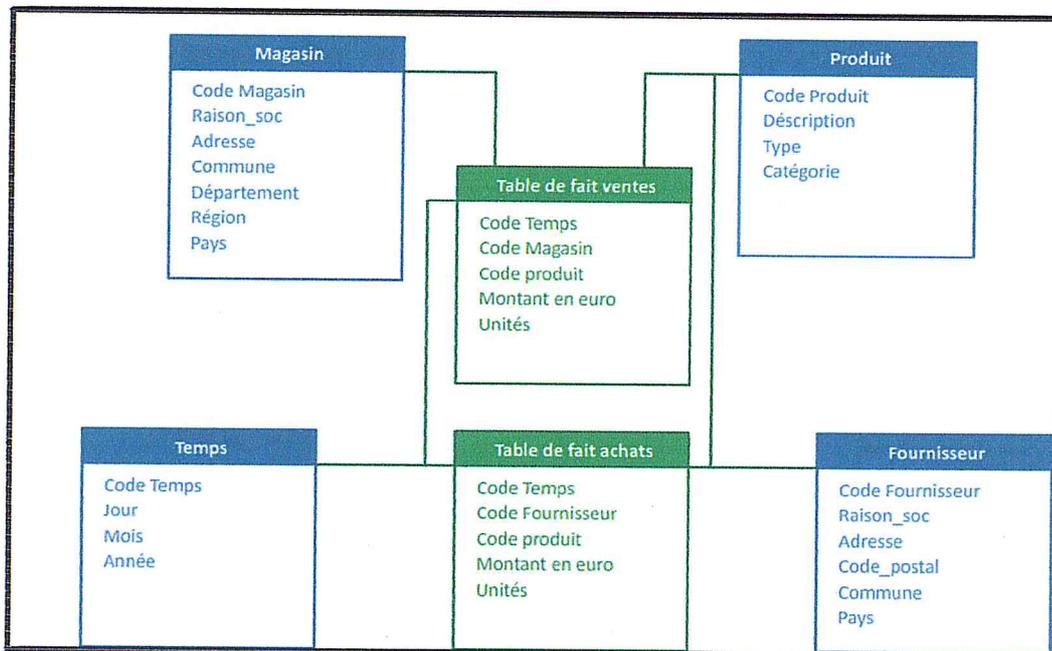
- navigation difficile.
- nombreuses jointures.

**8.4.3. Modèle en constellation :**

Une autre technique de modélisation, issue du modèle en étoile, est la modélisation En constellation.

La modélisation en constellation consiste à fusionner plusieurs modèles en étoile qui peuvent utiliser des dimensions communes.

Un modèle en constellation comprend donc plusieurs tables de faits et des tables de dimensions communes ou non à ces tables de faits.



*Figure 1.11 : Modèle en constellation. [14]*

## 9. Outils de restitution :

### 9.1. Technologie OLAP (Online Analytical Processing) :

#### 9.1.1. Concept de Cube :

OLAP propose une approche multidimensionnelle ce qui nous amène à la notion de cube. Un cube représente un ensemble de mesures organisées selon un ensemble de dimensions. Chaque case du cube représente une valeur. Les dimensions sont indiquées sur les arêtes du cube. Un plan de cube correspond à toutes les valeurs pour seule position d'une des trois dimensions.

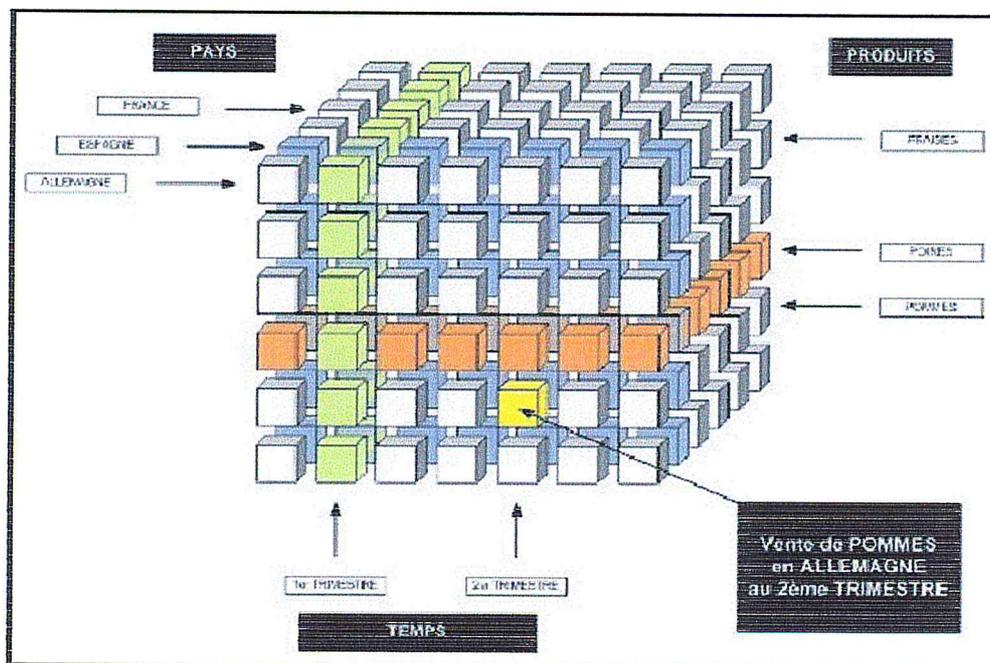


Figure 1.12 : Cube de données. [15]

La technologie matricielle présente trois avantages majeurs :

- ✓ Les données sont représentées sous une forme qui reflète directement le modèle conceptuel et rend la manipulation libre et plus intuitive.
- ✓ L'accès aux données dans l'hyper-cube est beaucoup plus directe, le temps de réponse est donc plus court.
- ✓ L'administration d'un schéma matriciel est plus simple. [16]

### 9.1.2. Les technologie OLAP :

Kimball pense que les datamarts de type OLAP peuvent être des composants à part entière du bus décisionnel s'ils sont conçus autour de faits conformes et de dimensions conformes.

1) **MOLAP** (*Multidimensionnel OLAP*): connu aussi sous le nom OLAP tout court, c'est un ensemble d'interfaces utilisateurs, d'applications et de technologies de bases de données propriétaires dont l'aspect dimensionnel est prépondérant.

2) **ROLAP** (*Relational OLAP*): ensemble d'interfaces utilisateurs et d'applications qui donnent une vision dimensionnelles des bases de données relationnelles.

3) **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. Ce sont les trois principales technologies, mais d'autres technologies OLAP existent également:

WOLAP (Web-based OLAP).

DOLAP (Desktop OLAP).

RTOLAP (Real-Time OLAP).

SOLAP (Spatial OLAP).

	ROLAP	MOLAP	HOLAP
<b>Stockage des données de base</b>	BD relationnelle	BD Multidimensionnelle	BD relationnelle
<b>Stockage des agrégations</b>	BD relationnelle	BD Multidimensionnelle	BD multidimensionnelle
<b>Structure de la Base de Données</b>	Modèle particulier (étoile, flocon, etc.)	Structure propriétaire au logiciel utilisé	Croisement des architectures ROLAP et MOLAP
<b>Fonctionnement</b>	Le serveur extrait les données par des requêtes SQL et les interprète selon une	Le serveur MOLAP extrait les données de l'hypercube et	Accède aux deux BD et les présente au module client selon leur méthode

	vue multidimensionnelle avant de les présenter au module Client.	les présentes directement au module client.	respective
<b>Performance des requêtes</b>	Le moins performant	Le plus Performant	Performance moyenne

*Tableau 1.1: Technologies OLAP. [17]*

### 9.1.3. Navigation dans le cube :

La navigation permet à l'analyste de visualiser les informations contenues dans le cube et passer d'un niveau d'agrégat à un autre afin de connaître le détail des données qui ont initialement servi à le construire. [3]

Les outils OLAP utilisent des opérateurs particuliers pour la navigation dans les hyper-cubes :

- **Roll-up** : Passage de mesures détaillées aux mesures résumées en remontant dans la hiérarchie de la dimension.
- **Drill-down** : Descendre dans la hiérarchie de la dimension.
- **Rotate** : Rotation des axes du cube pour fournir une vue alternative des données.
- **Slicing** : Extraction d'une tranche d'informations : Sélection d'une dimension pour passer à un sous-cube.
- **Dice** : Extraction d'un bloc de données : Sélection de deux ou plusieurs dimensions.
- **Drill-across** : Exécution de requêtes impliquant plus d'un cube ayant une dimension commune.
- **Drill-through** : Passage d'une mesure à l'autre ou d'un membre d'une dimension à un autre.

**9.1.4. Les Règles OLAP : [17]**

Douze règles de base2 permettant de qualifier une base décisionnelle :

- 1) **Multi-dimensionnalité** : le système doit permettre une vue multidimensionnelle des données.
- 2) **Transparence** : Le système doit être transparent à l'utilisateur qui doit accéder à la base via des outils standards (tableurs,...)
- 3) **Accessibilité** : Le système doit donner accès aux données nécessaires aux analyses demandées. Les outils OLAP doivent avoir leur propre schéma logique de stockage des données physiques hétérogènes et doivent accéder aux données et réaliser n'importe quelle conversion afin de présenter à l'utilisateur une vue simple et cohérente. Ils doivent aussi reconnaître de quel type de systèmes proviennent les données.
- 4) **Accès stable** : Le modèle, les dimensions et les niveaux d'agrégation changent sans remettre en cause son fonctionnement
- 5) **Client/serveur** : Respecter l'architecture Client/ Serveur
- 6) **Dimensionnalité générique** : Toutes les dimensions doivent être accessibles pour chacune des données.
- 7) **Gestion des matrices creuses** : quand on construit le cube multidimensionnel, le produit cartésien des divers axes réserve les espaces nécessaires au stockage des informations même si celles-ci n'existent pas, générant ainsi ce qu'on appelle « la matrice creuse ». Le système OLAP doit gérer ce problème.
- 8) **Multi-utilisation** : accès simultané de plusieurs utilisateurs.
- 9) **Croisement des données** : Toutes les tranches du cube doivent rester visualisable.
- 10) **Manipulation des données** : Une navigation simple et intuitive.
- 11) **Souplesse d'affichage**.
- 12) **Nombre illimité de dimensions et de niveaux d'agrégation.**

## **9.2. Le Datamining : [18]**

### **9.2.1. Définition :**

Le Datamining est en fait un terme générique englobant toute une famille d'outils facilitant l'analyse des données contenues au sein d'une base décisionnelle de type DataWarehouse ou DataMart. Les techniques de datamining sont particulièrement efficaces pour extraire des informations significatives depuis de grandes quantités de données.

### **9.2.2. Principe du Datamining :**

En peu de mots, le Datamining présente l'avantage de trouver des structures originales et des corrélations informelles entre les données.

Il permet de mieux comprendre les liens entre des phénomènes en apparence distincts et d'anticiper des tendances encore peu discernables.

### **9.2.3. Pour quelles applications utiliser le Datamining ?**

Au contraire des méthodes classiques d'analyses statistiques, le Datamining est particulièrement adapté au traitement de grands volumes de données. Avec l'augmentation de la capacité de stockage des supports informatiques, un maximum de renseignements sera capté, ordonnés et rangés au sein du DataWarehouse. Comportement des acheteurs, caractéristiques des produits, historisation de la production, désormais plus rien n'échappe à la collecte. Avec le Data Mining, ces "tera-nesque" bases de données sont exploitables.

### **9.2.4. Les techniques du datamining :**

Les outils de datamining proposent différentes techniques à choisir en fonction de la nature des données et du type d'étude que l'on souhaite entreprendre

- Les méthodes utilisant les techniques de classification et de segmentation
- Les méthodes utilisant des principes d'arbres de décision assez proches des techniques de classification.

- Les méthodes fondées sur des principes et des règles d'associations ou d'analogies.
- Les méthodes exploitant les capacités d'apprentissage des réseaux de neurones.
- Et pour les études d'évolution de populations, les algorithmes génétiques.

### 9.2.5. Comment utiliser le Datamining ?

Trois conseils :

1. Accordez le soin nécessaire à la phase de collecte des données. La collecte des données est une étape fondamentale. Elle est longue et coûteuse. Il est indispensable de disposer de la garantie de la fiabilité des données avant de lancer la moindre analyse.
2. L'étude des résultats mérite toute votre attention. Il ne faut pas hésiter à présenter différemment le problème et à tester d'autres techniques d'analyse avant de s'engager sur un résultat.
3. Le Datamining n'est pas un oracle absolu. Une absence de réponse du système ne doit pas être systématiquement considérée comme une négation. Il faut quelquefois prendre la précaution d'aborder le problème sous un autre angle avant de s'engager.

### 9.3. Le tableau de bord :

Un tableau de bord est qu'une forme particulière d'état de rapport, on appelle un tableau de bord un rapport qui est:

- **Synthétique** : tout doit tenir sur une feuille A4, ou sur un écran d'ordinateur
- **Composite** : présenté plusieurs indicateurs, ce qui offre une représentation complète de l'activité de l'entreprise.
- **Personnalisé** : chaque acteur consulte un rapport qui concerne son activité.



Figure 1.13 : Tableau de bord. [19]

## 10. Comparaison entre système opérationnel et système décisionnel: [20]

Le BI est le résultat d'une évolution des besoins de la part des décideurs et analystes des entreprises. Le but du BI est d'aider à la décision et de permettre des analyses précises, complexes et de grande envergure dans les entreprises.

Les systèmes opérationnels font tourner l'entreprise. Ils assistent la production et la vie quotidienne de celle-ci.

Les systèmes décisionnels voient tourner l'entreprise. Ils permettent de générer de la connaissance à partir des données, et donc, d'aider à faire des décisions stratégiques.

Les différences entre le monde opérationnel et décisionnel peuvent être résumées ainsi :

Système Opérationnel (OLTP)	Système Décisionnel (OLAP)
Petits volumes de données à gérer.	Gros volumes de données à gérer.
Utilisé par toute l'entreprise.	Nombre d'utilisateur restreint (décideurs, analystes).
Processus fermés, transactionnels, le but est de donner le moins de marge de manœuvre possible.	Processus ouverts pour permettre la génération de connaissance.
Données en lecture - Écriture.	Données en lecture seule.
Réponses très rapides.	Rapidité moyenne comparée aux systèmes opérationnels.
Niveau de granularité fin.	Niveau de granularité très grand (on peut avoir des résumés sur ce qui c'est passé durant les 10 dernières années par exemple).
Décentralisés.	Centralisés (on veut avoir toutes les données de l'entreprise dans une seule structure).

*Tableau 1.2 : OLTP vs OLAP. [20]*

## 11. Conclusion :

Tous l'enjeu des systèmes d'information de demain consistera à concilier les finalités divergentes, pour ne pas dire contradictoire que sont la délégation sur le terrain de la prise de décision avec la nécessité de cohérence qui se traduit par une intégration croissante supposée mieux gérer la complexité et garantir les objectifs financiers. Ces deux pressions ne s'excluent pas mais se complètent.

# 2 ETUDE DE L'EXISTANT

- **Présentation de la structure d'accueil Brève historique**
- **définition des besoins**
  - *Approche de définition des besoins*
  - *La définition des besoins*
  - *Extraction d'Indicateurs et Axes d'analyse*
  - *Pourquoi produire une nouvelle BDD*

**1. Introduction:**

Dans cette partie nous présenterons l'organisme d'accueil, et nous allons récupérer toutes les informations nécessaires pour définir les principaux besoins de cet organisme en passant par plusieurs méthodes de recueils.

**2. Présentation de La Structure d'accueil :****2.1. Historique du SPE (service public de l'emploi):**

Dans tous les pays du monde la gestion et la régulation du marché du travail relèvent des prérogatives de l'état par un service public.

Le SPE est apparu dans les pays industrialisés vers la fin du 19ème siècle en raison des préoccupations relatives à l'impact économique et social du chômage.

A la fin de l'année 2004 une loi relative au placement des travailleurs et au contrôle de l'emploi est promulguée en Algérie. Il s'agit de la loi 04 / 19 du 25 décembre 2004 qui définit les contours de cette organisation en précisant ce qui suit, En 1919, l'OIT recommande aux états membres la mise en place du SPE à travers la convention N° 02 sur le chômage.

Cette mission est mise en œuvre par l'Agence Nationale de l'Emploi (ANEM) en succession à l'Office national de la main d'œuvre (ONAMO). Elle est Instituée par décret exécutif n° 06-77 du 18 février 2006 fixant les missions, l'organisation et le fonctionnement de l'agence nationale de l'emploi.

**2.2. Organisation de L'ANEM:**

L'agence nationale de l'emploi ou ANEM est une entreprise publique, son rôle se limite à un travail de normalisation, de conseil et de contrôle technique. L'information sur le marché de l'emploi est réunie par les structures extérieures, rassemblée et traitée par l'agence elle même. L'agence se divise en plusieurs agences et directions :

- **Direction Régionale de l'Emploi (DREM):** Celle-ci constitue le relais entre le siège et les structures qui lui sont inférieures (Agences Wilayas, Agences Locales).
- **Agence de wilaya de l'emploi (AWEM):** Véritable cheville ouvrière de l'ANEM cette agence est placée en première ligne face aux usagers et à leurs multiples besoins.
- **Agence locale de l'emploi (ALEM):** Considéré comme des annexes de wilayas caractérisées par une grande concentration de la population et des activités les ALEM constituent le dernier palier de l'organisation. Implémentée au niveau communal, peuvent se spécialisées dans la prospection des offres cependant elles sont tenues d'orienter leurs activités à destination des populations résidentes.  
[Voir ANNEXE A]

### 2.3. Mission et objectifs de l'ANEM :

#### 2.3.1. Mission:

L'Agence Nationale de l'Emploi (ANEM) a pour mission d'organiser et d'assurer la connaissance de la situation et de l'évolution du marché nationale de l'emploi et de la main d'œuvre et de garantir à tout demandeur et à toute entreprise un service de recrutement efficace et personnalisé. A ce titre, l'ANEM est chargée de :

- Mettre en place un système d'information permettant de renseigner de manière précise régulière et fiable sur les fluctuations du marché de l'emploi et de la main d'œuvre.
- Procéder à toute analyse et expertise en matière d'emploi et de main d'œuvre.
- Entreprendre toutes études et enquêtes liées à l'accomplissement de sa mission.
- Développer et normalisé les instruments et outils permettant le développement de la fonction observation du marché de l'emploi.
- Recueillir et mettre en relation l'offre et la demande de travail et à ce titre elle est chargée de :

1. Assurer l'accueil l'information l'orientation et le placement des demandeurs d'emploi.
2. Procéder à la prospection et la collecte d'offre d'emploi auprès des organismes employeurs.
3. Organiser la compensation des offres et des demandes d'emploi au niveau national régional et local.
4. Favoriser la mobilité géographique et professionnelle des demandeurs d'emploi.

### **2.3.2. Objectifs:**

L'objectif de l'Agence Nationale de l'Emploi(ANEM) est d'assurer une mise en marché maximale des candidatures des demandeurs, et de maximiser la réussite des démarche de recrutement des entreprises et de les épauler efficacement dans le recrutement de candidates et de candidats. Car nous croyons que le capital humain d'une entreprise sera un élément majeur pour l'accroissement de sa compétitivité et son développement.

### **2.4. Limitation du champ d'étude :**

C'est au niveau de la DREM que notre projet a été proposé et même suivi. Et plus précisément au niveau du «Service informations et gestion informatique».

#### **2.4.1. Missions et objectifs DREM:**

- Assurer une gestion décentralisée des fonctions et des moyens basés sur une politique de partage et de responsabilité entre le siège et les structures extérieures.
- Fournir un appui technique ou de gestion aux agences wilayas et locales afin de faire face aux problèmes qu'elles rencontrent quotidiennement afin d'assurer un bon fonctionnement des activités.

### 2.4.2. Organigramme DREM:

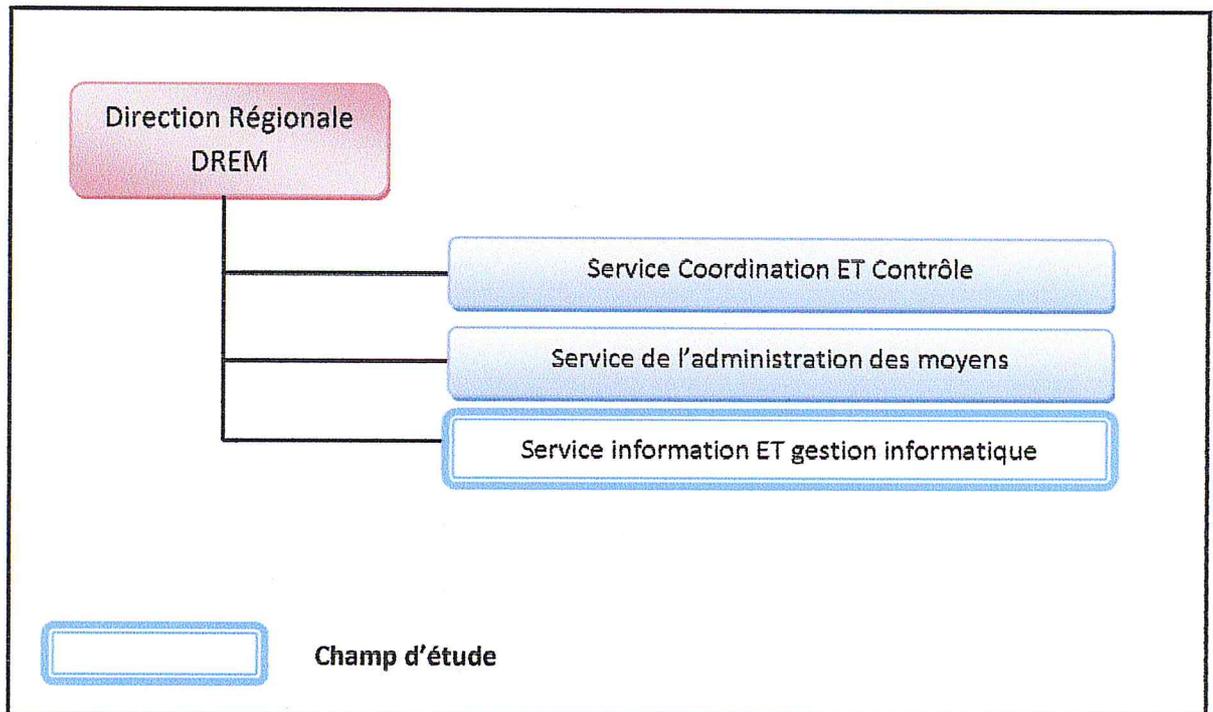


Figure 2.1 : Organigramme DREM.

## 3. Définition des besoins :

Les utilisateurs, ainsi que leurs besoins, affectent presque toutes les décisions prises au cours de l'implémentation du Data Warehouse. Pour cela, la bonne définition des besoins est une chose primordiale, qui va nous permettre de répondre aux différentes questions concernant le but du système, ses utilisateurs, les services qu'il va rendre, etc.

### 3.1. Approche de définition des besoins :

Afin de mieux assimiler les besoins, on a suivi l'approche de "Ralph Kimball". Celle-ci consiste à commencer par parler avec les utilisateurs de l'entreprise qui occupent le centre d'univers du Datawarehouse, en ce qui concerne leurs travail, leurs

enjeux et leurs objectifs en essayant de comprendre leur manière de prendre les décisions actuelles et celles à venir. Aussi est-il nécessaire de rencontrer les acteurs clés du service informatique pour intégrer quelques réalités relatives aux données.

### 3.2 La définition des besoins :

En s'appuyant sur deux méthodes, on a pu définir les besoins de l'entreprise. La première c'est lire la documentation mise à notre disposition concernant les activités de la DREM. La seconde est de planifier des entretiens avec les dirigeants suivant :

- **Département ressources humaines** : en particulier le chef de département ... etc.
- **Service système informatique** (La plupart des applications sont développées au sein de ce service) : En particulier le chef de service Etude et Développement des systèmes informatique au niveau du DREM.
- **Agence locale de l'emploi (ALEM)** : En particulier le chef d'agence plus les agents qui prépare les rapports au niveau d'agence et l'agent de saisie aussi.

#### 3.2.1. Recherche documentaire :

L'ANEM a mis à la disposition de chaque ALEM des rapports vides à remplir pour ces agences locales, à partir de ces rapports on peut récupérer facilement nos besoins en données pour citer les faits (indicateurs) et les dimensions.

Ces rapports remplis seront transférer par la suite à L'ANEM via DREM pour construire les rapports finaux qui ont aider les parties dirigeante à prendre des décisions stratégiques.

Mais dans ces rapports on peut récupérer les données qu'on va saisir via notre système opérationnel, donc on sera obligé de consulter d'autres documents :

- Fiche d'identification de l'entreprise.
- Fiche offre d'emploi.
- Fiche carte demandeur. [Voir ANNEXE B]

Et nous allons devoir aussi consulter l'application pour récupérer tous les données concernant les demandeurs, offres, placement....etc.

### 3.2.2. Les entretiens :

Nous avons effectué des entretiens avec un directeur d'agence ALEM et le responsable service informatique Au niveau de la DREM et les responsables des offres et emploi au niveau d'ALEM et un agent de saisie.

L'entretien avec les responsable emploi et offre (qui préparent généralement les rapports) était pour comprendre toutes les informations et abréviations entrant dans les rapports.

Mais l'entretien avec le directeur ALEM et responsable informatique plus le directeur ressources humaines était pour citer les besoins et les résultats attendus de notre projet.

Parmi les questions que notre questionnaire à soulever :

- ✓ Quels types d'analyse sont effectués ? A partir de quelles sources ? et sur quels outils ?
- ✓ Quels problèmes affrontez-vous principalement aujourd'hui ?
- ✓ Quelles analyses souhaitez-vous faire ?
- ✓ Quelle méthode de diffusion envisagez-vous ?

### 3.2.3. Les Objectifs :

On peut les résumés dans les points suivants :

- Extraire les indicateurs stratégiques ainsi que les axes d'analyse correspondants à chaque indicateur.
- Spécification des applications utilisateur.
- Choix des modes de visualisation des rapports.

### 3.3. Extraction d'Indicateurs et Axes d'analyse :

Les deux tableaux suivants représentent les indicateurs et les dimensions que nous avons extraits, après l'étude des documents et les entretiens avec les utilisateurs de DREM :

#### 3.3.1. Indicateurs (faits) :

Grains	Indicateurs (Faits)	Description
<b>Demande</b>	demande	Flux de la demande
	Demandes disponible	Demandes disponible à la fin du mois
<b>DemandeursD</b>	Demandeur DFP	Demandeurs d'emplois enregistrés issus de la formation professionnelle
	Demandeur DES	demandeurs diplômés sortant de l'enseignement supérieur et équivalent
<b>Offre</b>	Offre	Répartition Offre d'emploi
	Offre disponible	Offre disponible à la fin du mois
<b>Placements</b>	Placement	Répartition des demandes d'emplois placées
<b>PlacementsD</b>	Placement DFP	Placements des diplômés issus de la formation professionnelle et équivalent
	Placement DES	Répartition des placements des diplômés sortant de l'enseignement supérieur et équivalent

*Tableau 2.1: Les faits.*

## 3.3.2. Axes d'analyse (dimensions) :

Dimension	Description
Tranche d'âge	16 a 19 ans, 20 a 24 ans, ...,50 ans et plus
Sexe (genre)	Homme, Femme
Groupe de métiers	Agriculture, Textiles et confection. etc.
Raison d'inscription	STR2, STR1, Autre demandeur
Niveau d'instruction	Sans instruction, primaire, moyen, .... etc.
Ancienneté d'inscription	Moins de 3 mois, 3 à 6 mois, ... etc.
diplôme	CFPS, CAP, CMP, BT, BTS, Autres
Filière de formation	Sciences exactes, technologie, Architecture, ... etc.
Secteur Juridique	Public, Prive (National, Etranger)
Nature poste proposée	Permanent, Temporaire
Branche d'activité économique	Eau et énergie, commerce, chimie et plastiques
catégorie	1, 2, 3, 4, 5,6
Niveau de qualification	Tentative de placement, placement effectués

Tableau 2.2 : Les dimensions.

**STR2** : N'ayant jamais travaillé.

**STR1** : Ayant déjà travaillé.

**Autre demandeur** : Ayant un travail lors de son inscription, retraites a la recherche d'un emploi, ... etc.

Et après les entretiens avec le directeur agence et responsable informatique nous avons proposé d'ajouter quelque dimension comme :

- ✓ Situation militaire.
- ✓ Situation familiale.
- ✓ Etat de santé (handicapé).

Et d'autres qui sont en cours d'élaboration.

**5. Conclusion :**

Nous avons présenté l'organisme d'accueil, son organisation interne, ses missions et ses activités. Nous avons également limité le champ d'étude et cité les différents systèmes opérationnels de l'établissement.

La phase « Définition des besoins » consiste à raffiner les besoins des utilisateurs finaux de façon à pouvoir entamer la conception avec des données qui existent réellement et qui sont bien définies.

Dans cette phase, nous avons pu analyser les sources de données que nous allons utiliser pour alimenter notre Data Warehouse. Nous avons réalisé cette étape par un « reverse engineering » des sources de données et cela pour définir leurs schémas conceptuels afin de pouvoir utiliser ceux qu'on en a besoins dans le schéma du Data Warehouse.



# 3 CONCEPTION

- **Solution informatique**

- **Système opérationnel**

- *Pourquoi modéliser ?*

- *Choix d'outil Conception*

- *Choix d'outil de Modélisation*

- *Les diagrammes*

- **Système décisionnel**

- *Diagrammes de cas d'utilisation*

- *La modélisation multidimensionnelle*

- *schéma de l'entrepôt de données*

## **1. Introduction :**

Ce chapitre est le cœur de notre travail. Dans ce dernier on présentera la conception et la modélisation du système. On commencera par la méthode que nous avons suivi pour modéliser le système, puis la schématisation des différents composants du système.

## **2. Solution Informatique :**

Notre projet consiste à créer une application web qui se déploie sur un réseau permettant de gérer la partie opérationnel du système et de créer un Data Warehouse pour la partie décisionnel et surtout pour faire le Reporting.

Pour avoir une meilleure gestion des flux de la demande d'emploi, et une bonne maîtrise de leurs paramètres, on a apporté les solutions suivantes:

- Conception et la mise en œuvre d'un système opérationnel pour tous ce qui est archivage mise à jour et suivi.
- Conception, modélisation et déploiement d'un système décisionnel à des fins d'analyse et d'automatisation de la procédure de calcul des paramètres d'exploitation liés aux flux de la demande et de l'offre d'emploi.

### **2.1. Pourquoi une nouvelle Base de production :**

Nous avons jugé nécessaire de créer une nouvelle base de production pour les raisons que nous allons citer ci-dessous :

En premier lieu, la direction ANEM a jugé nécessaire de ne pas divulguer les informations concernant les candidats donc ne nous a pas autorisé à utiliser leur base de données pour créer notre Data Warehouse. Donc, pas de source de données et c'est pour cela qu'il faut impérativement que l'on crée une nouvelle base de production pour pouvoir poursuivre notre travail adéquatement.

Une autre raison qui nous a poussés à créer une nouvelle base de données qui est le fait d'avoir un problème de connexion à internet, et vue que la base de données est répartie, un problème d'hétérogénéité se créer. La base de données ne sera plus la même dans les différentes directions.

Et enfin, malgré la présence de l'application, les employés de l'ANEM ne saisissent pas les données sur la base de façon régulière, ils préfèrent parfois de saisir sur des fichiers EXCEL et de négliger la base de données existante malgré qu'elle soit performante.

## **2.2. Pourquoi un Data Warehouse :**

Nous avons décidé de créer un Data Warehouse parce qu' il nous permet de consolider l'ensemble des informations potentiellement utiles à l'analyse et notamment au Reporting qui est l'aspect principal de notre projet d'une part, d'autre part, puisque les données d'un Data Warehouse sont archivées, historiées et orientées sujet (Thématique), cette caractéristique nous permet de faire des comparaisons entre les rapports (mensuel/annuel) établis concernant le flux de la demande d'emploi, les offres d'emploi etc..., ensuite nous allons obtenir par la suite une vision globale du déroulement de l'activité de l'entreprise et cela permettra au dirigeant de prendre les bonnes décisions.

## **3. Conception Système opérationnel :**

### **3.1. Pourquoi modéliser ? [21]**

Un modèle est une représentation simplifiée d'une réalité. Il permet de capturer les aspects pertinents pour répondre à un objectif défini *a priori* le modèle s'exprime sous une forme simple et pratique pour le travail.

Quand le modèle devient compliqué, il est souhaitable de le décomposer en plusieurs modèles simples et manipulables.

L'expression d'un modèle se fait dans un langage compatible avec le système modélisé et les objectifs attendus. Dans le cas du logiciel, l'un des langages utilisés pour

la modélisation est le langage UML. Il possède une sémantique propre et une syntaxe composée de graphique et de texte et peut prendre plusieurs formes (diagrammes).

Les modèles ont différents usages :

- Ils servent à circonscrire des systèmes complexes pour les dominer.
- Ils optimisent l'organisation des systèmes. La modélisation de la structure d'une entreprise en divisions, départements, services, etc. permet d'avoir une vision simplifiée du système et par là même d'en assurer une meilleure gestion.
- Ils permettent de se focaliser sur des aspects spécifiques d'un système sans s'embarrasser des données non pertinentes.
- Ils permettent de décrire avec précision et complétude les besoins sans forcément connaître les détails du système.
- Ils facilitent la conception d'un système.
- Ils permettent de tester une multitude de solutions à moindre coût et dans des délais réduits et de sélectionner celle qui résout les problèmes posés.

La modélisation objet produit des modèles discrets permettant de regrouper un ensemble de configurations possibles du système et pouvant être implémentés dans un langage de programmation objet. La modélisation objet présente de nombreux avantages à travers un ensemble de propriétés (classe, encapsulation, héritage et abstraction, paquetage, modularité, extensibilité, adaptabilité, réutilisation) qui lui confère toute sa puissance et son intérêt.

### **3.2. Choix d'outil Conception :**

#### **3.2.1. Introduction a UML:**

UML (*Unified Modeling Language*) est un langage unifié de modélisation et non pas une méthode.

Ce langage est né de la fusion de trois méthodes d'analyse orientées objet :

- la méthode OOD (Object Oriented Design).
- la méthode OMT (Object Modeling Technique).
- La méthode OOSE (Object Oriented Software Engineering).

A l'initiative de la société Rational Software, les auteurs principaux de ces trois méthodes se sont mis d'accord sur un langage de modélisation visant à devenir une référence.

### 3.2.2. Pourquoi UML ?[22]

UML a été conçu pour permettre la modélisation de tous les phénomènes liés à l'activité de l'entreprise, indépendamment des techniques d'implémentation (système automatisé ou non, langage de programmation...).

UML offre une manière claire de représenter le système selon différentes vues complémentaires grâce aux diagrammes qu'il fournit:

- Diagramme de Classes.
- Diagramme d'Objets.
- Diagramme de Cas d'Utilisation.
- Diagramme de séquences.
- Diagramme de collaborations.
- Diagramme d'états.
- Diagramme d'activités.
- Diagramme d'implémentation.
- Diagramme de déploiement.

Ces diagrammes UML fournissent les informations sur un problème et sa solution. Les Combinaisons de diagrammes représentent les vues du système.

### 3.2.3. Diagrammes UML :[23]

Nous n'évoquons ici que les diagrammes utilisables pour la description des processus:

#### ❖ Diagrammes de cas d'utilisation :

Il permet de décrire l'interaction entre les trois concepts suivants :

- **Acteur** : est un utilisateur type qui a toujours le même comportement vis-à-vis d'un cas d'utilisation.
- **Cas d'utilisation** : correspond à un certain nombre d'actions que le système devra exécuter en réponse à un besoin d'un acteur.
- **Interaction** : correspond à l'échange entre un acteur et un cas d'utilisation.

#### ❖ Diagramme d'activités :

Le diagramme d'activités est attaché à une catégorie de classes et décrit le déroulement des activités de cette catégorie. Le déroulement s'appelle "flot de contrôle".

Il donne une vision des enchaînements des activités propre à une opération ou à un cas d'utilisation. Il indique la part prise par chaque objet dans l'exécution d'un travail. Il sera enrichi par les conditions de séquence.

#### ❖ Diagrammes de séquences :

Les diagrammes de séquences présentent les messages échangés entre des objets. Il Permet de mieux visualiser la séquence des messages par une lecture de haut en bas. L'axe vertical représente le temps et l'axe horizontal représente les objets qui collaborent.

Il permet de décrire les scénarios de chaque cas d'utilisation en mettant l'accent sur la Chronologie des opérations en interaction avec les objets.

### 3.3. Choix d'outil de Modélisation : (Ms Visio 2010)

Visio 2010 de Microsoft offre un modèle de processus métier qui permet de visualiser, d'analyser et de communiquer des informations complexes grâce à une bibliothèque de diagrammes professionnels prédéfinis : diagrammes UML, diagrammes de flux, diagrammes d'architecture, et organigrammes.

La figure ci-dessous représente l'interface principale MS Visio 2010 pour les diagrammes UML. Celle-ci se compose:

- D'une barre de menus en haut.
- D'une barre d'outils offrant des raccourcis pour les fonctions des menus juste en dessous,
- D'une fenêtre de navigation à gauche qui propose une structure arborescence du projet, avec les rubriques de tous les diagrammes UML: (Activité, Collaboration, Composant,..., Séquence, et Cas d'utilisation).

Enfin, d'une grande fenêtre destinée à visualiser et éditer les diagrammes.

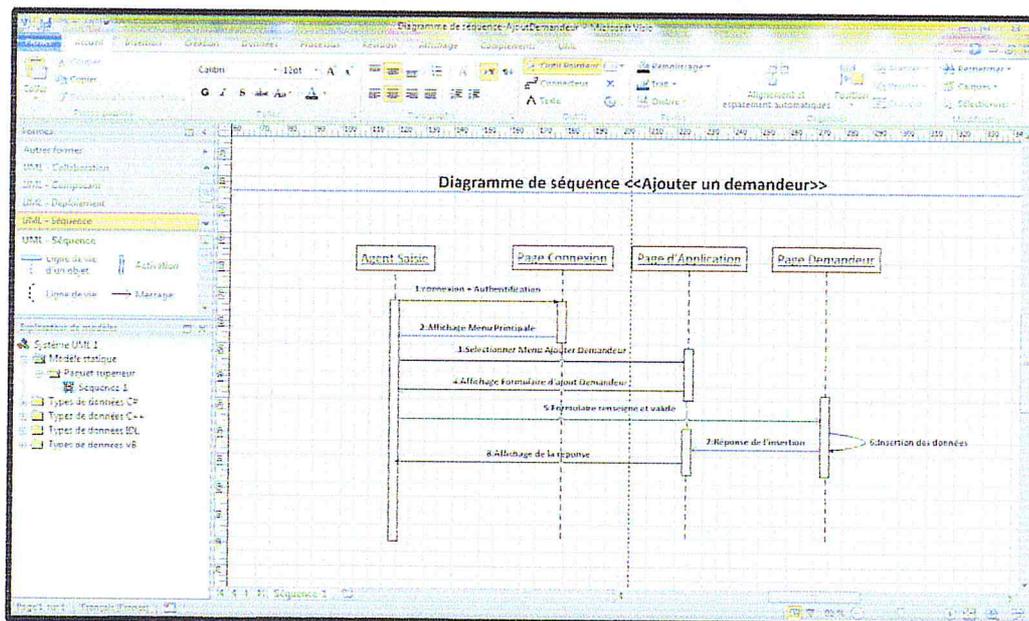


Figure 3.1 : Fenêtre principale de MS Visio 2010.

### 3.4. Diagrammes :

#### 3.4.1. Diagrammes de cas d'utilisation :

##### ❖ Administrateur I :

L'administrateur agit directement sur l'outil en créant, modifiant, ou supprimant des données, notamment sur les clients et sur les utilisateurs.

**S'authentifier** : l'administrateur s'authentifiera par un nom et un mot de passe pour pouvoir accéder à l'outil.

**Ajouter utilisateur** : création d'un nouvel utilisateur à travers un formulaire pour les autoriser à se servir de l'outil.

**Modifier utilisateur** : modifier les données propres à un utilisateur après l'avoir sélectionner.

**Supprimer utilisateur** : désactivation d'un utilisateur après l'avoir sélectionner.

**Consulter tableau de bord** : consultation des tableaux de bord après avoir sélectionner les données les concernant.

**Imprimer tableau de bord** : impression des tableaux de bord après les avoir sélectionner.

Ces cas d'utilisation incluent obligatoirement le cas d'utilisation «S'authentifier». En UML, cette relation d'inclusion obligatoire est formalisée en flèche de dépendance entre le cas d'utilisation de base et le cas inclus, nommée avec le mot-clef «inclus», comme indique sur le schéma suivant.

##### **La notion « gérer » :**

La Notion gérer dans le diagramme de cas d'utilisation est venu pour simplifier les cas d'utilisation complexe, elle comprend dans son package trois opération (Ajouter, Supprimer, Modifier). En effet, lorsqu'on trouve ces trois opération en même temps qui concerne un seul acteur à la fois, on peut les remplacer avec le cas d'utilisation gérer.

Par exemple les trois cas d'utilisation : Ajouter utilisateur, Supprimer utilisateur, Modifier utilisateur concernant le même acteur (Admin par exemple), pourrons être

remplacé par le cas d'utilisation Gérer utilisateur et de cette façon, on a pu rendre notre diagramme plus simple et plus compréhensif.

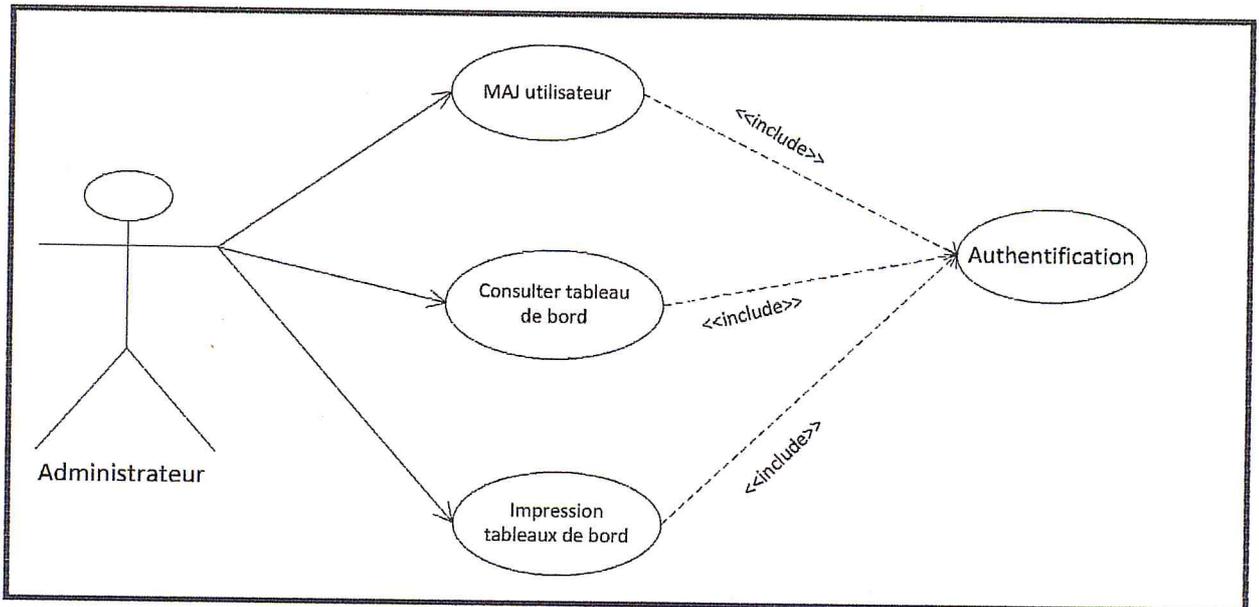


Figure 3.2 : Diagramme de cas d'utilisation Administrateur I.

#### ❖ Administrateur II :

**S'authentifier** : l'administrateur s'authentifiera par un nom et un mot de passe pour pouvoir accéder à l'outil.

**MAJ niveau de qualification** : La mise à jour(MAJ) et l'action qui consiste à mettre à niveau, c'est-à-dire, ajouter, supprimer ou modifier niveau de qualification et bien sûr, la suppression ou la modification se fait après avoir sélectionner les éléments concernés.

**Consulter niveau de qualification** : consulter niveau de qualification après avoir sélectionner les données propres à ce dernier.

**MAJ groupe de métier** : Ajout, suppression ou modification d'un groupe métier.

**Consulter groupe de métier** : Consulter groupe métier après avoir sélectionner les données propres à ce dernier.

**MAJ raisons d'inscription :** Ajout, suppression ou modification raison d'inscription.

**Consulter raisons d'inscription :** Consulter raisons d'inscription après avoir sélectionner les données la concernant.

**MAJ niveau d'instruction :** Ajout, suppression ou modification niveau d'instruction.

**Consulter niveau d'instruction :** Consulter niveau d'instruction après avoir sélectionner les données le concernant.

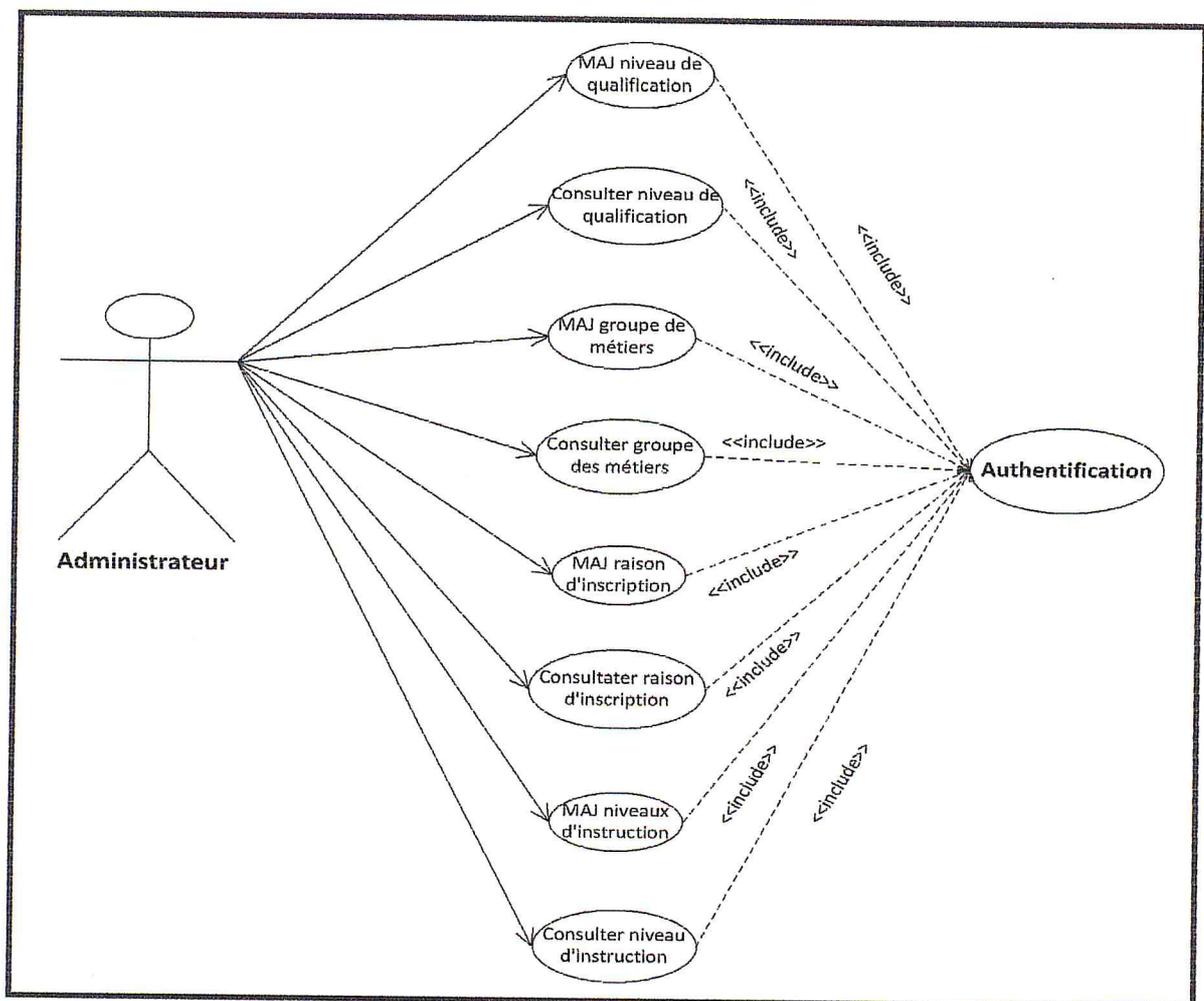


Figure 3.3 : Diagramme de cas d'utilisation Administrateur II.

Ce diagramme de cas d'utilisation pourra être simplifié à son tour, le tableau suivant comprendra les différents cas d'utilisation mis en package.

Package	Cas d'utilisation
Gérer niveau de qualification	MAJ, Consulter niveau de qualification
Gérer groupe de métier	MAJ, Consulter groupe de métier
Gérer raisons d'inscription	MAJ, Consulter raisons d'inscription
Gérer niveau d'instruction	MAJ, Consulter niveau d'instruction

On obtiendra le diagramme suivant :

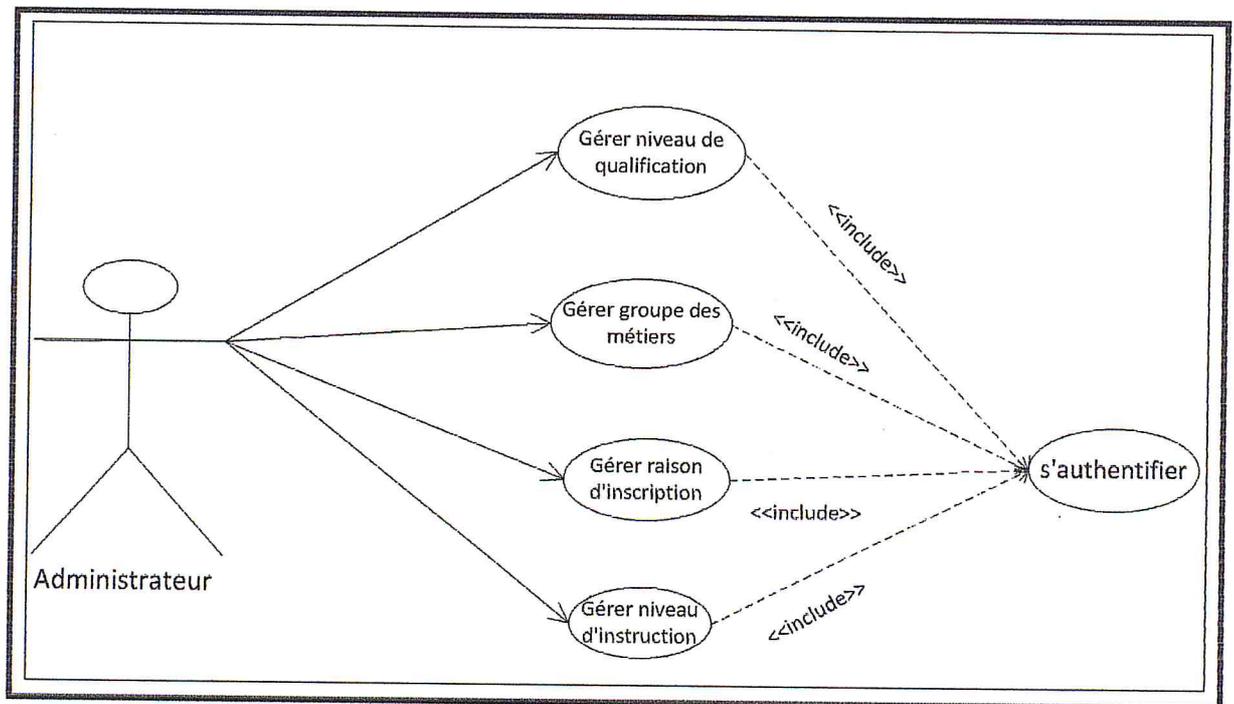


Figure 3.4 : Diagramme de cas d'utilisation Administrateur simplifiée.

**❖ Agent de saisie :**

L'agent de saisie est un acteur agissant directement sur l'outil en créant, modifiant ou supprimant des données, mais par contre, ce dernier a un accès limité, il peut seulement agir sur les données concernant les demandeurs et les placements.

**S'authentifier** : l'agent s'authentifiera par un nom et un mot de passe pour pouvoir accéder à l'outil.

**Ajouter demandeur** : création d'un nouveau demandeur à travers un formulaire.

**Modifier demandeur** : modifier les données propres à un demandeur après l'avoir sélectionner.

**Supprimer demandeur** : supprimer les données propres à un demandeur après l'avoir sélectionner.

**Consulter demandeur** : consultation des données d'un demandeur après l'avoir sélectionner.

**Ajouter Placement** : création d'un nouveau placement à travers un formulaire.

**Modifier Placement** : modifier les données propres à un placement après l'avoir sélectionner.

**Supprimer Placement** : supprimer les données propres à un placement après l'avoir sélectionner.

Ces cas d'utilisation incluent obligatoirement le cas d'utilisation «S'authentifier». En UML, cette relation d'inclusion obligatoire est formalisée en flèche de dépendance entre le cas d'utilisation de base et le cas inclus, nommée avec le mot-clef «inclus», comme indique sur le schéma suivant.

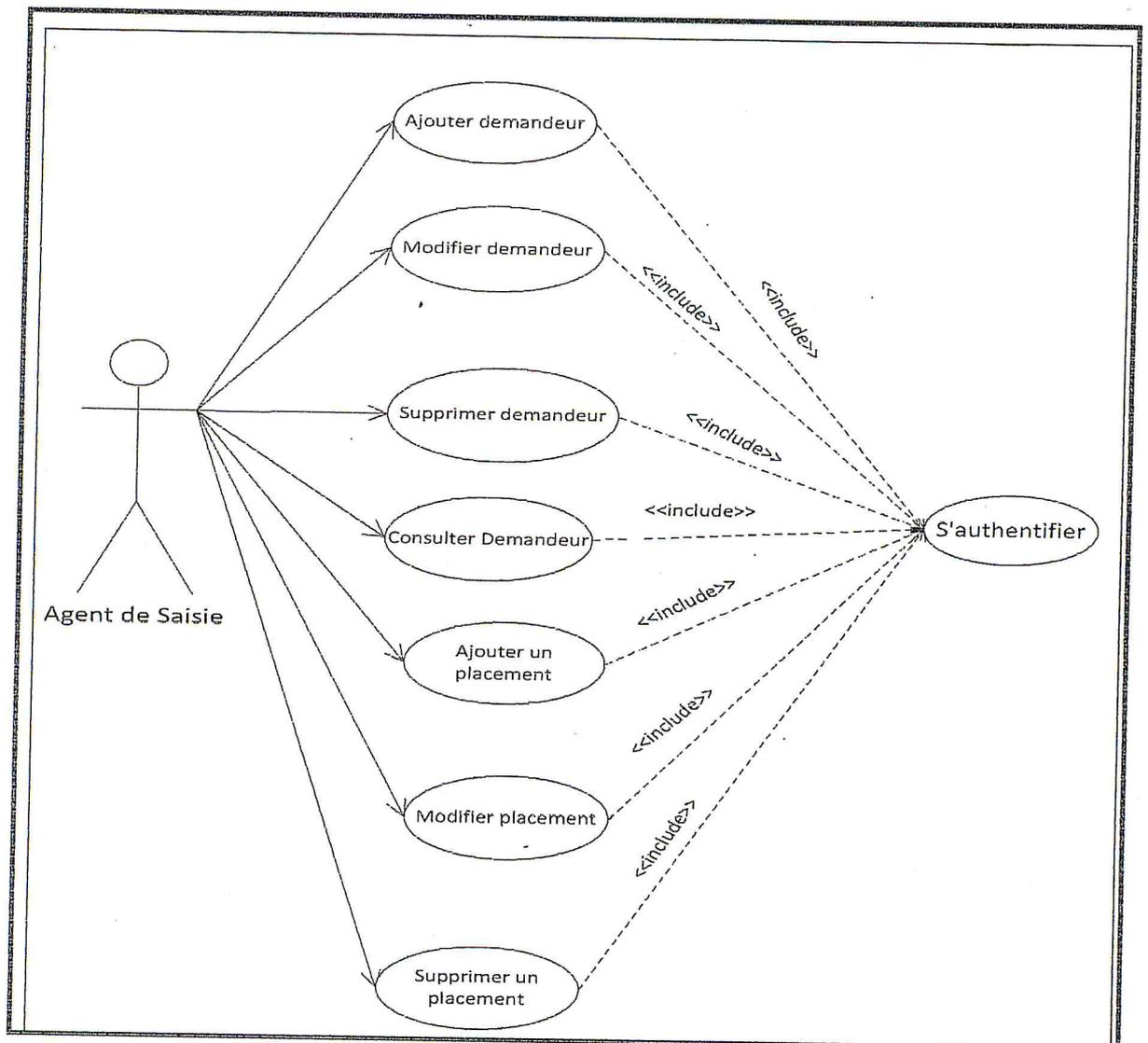


Figure 3.5 : Diagramme de cas d'utilisation de l'agent de saisie.

Afin de simplifier le diagramme de cas d'utilisation de « l'agent de saisie », le tableau suivant comprendra les cas d'utilisation mis en package :

package	Casd'utilisation
Géredemandeur	Ajouter, Modifier, Supprimer et consulter demandeur
Gérer placement	Ajouter, Modifier et Supprimer placement

Concernant le cas d'utilisation « s'authentifier », ce dernier suit le même raisonnement que le diagramme de cas d'utilisation de l'administrateur.

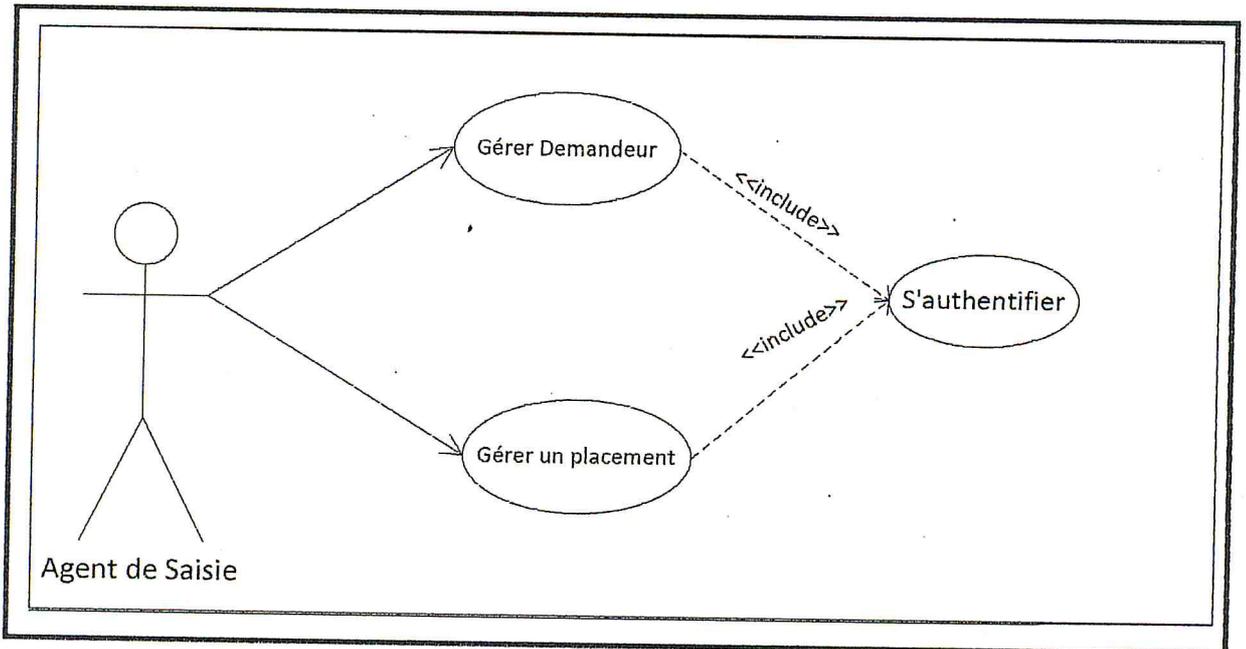


Figure 3.6 : Diagramme de cas d'utilisation de l'agent de saisie simplifiée.

### 3.4.2. Diagramme d'activité :

Certains diagrammes d'activités sont similaires entre eux, donc nous ne présenterons que les cas les plus pertinents, et nous précisons les diagrammes d'activités qui lui sont similaires.

Remarques sur la représentation des schémas :

Etat activité

Etat action

#### ❖ Créer Demandeur :

Ce diagramme d'activité pourrait aussi, de par son principe de fonctionnement, correspondre au cas d'utilisation «Créer un utilisateur».

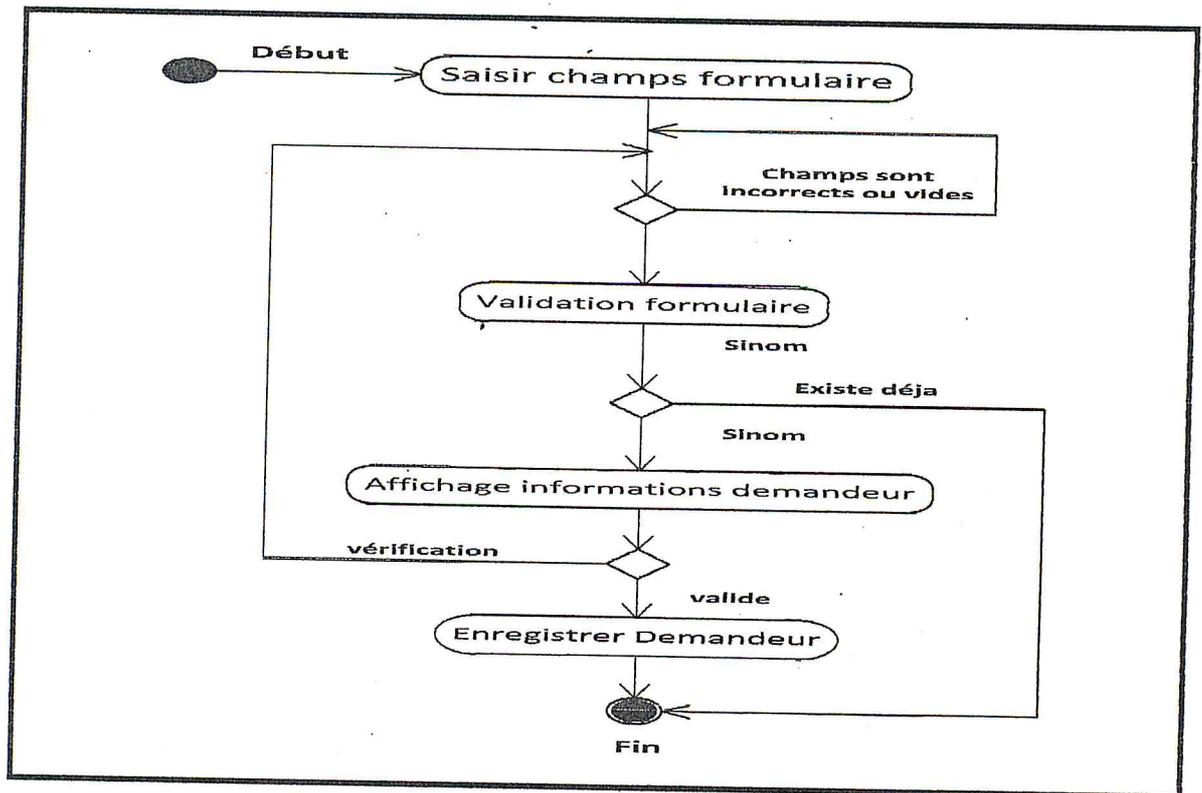


Figure3.7 : Diagramme d'activité «Créer Demandeur».

#### ❖ Modifier Demandeur :

Le cas d'utilisation «modifier demandeur» comporte plusieurs étapes. Dans un premier temps, une sélection est effectuée sur le choix du demandeur, par la suite toutes les informations sur ce dernier sont récupérées et affichées dans un formulaire pour pouvoir être modifiées ou non. Après la validation de la modification, un rappel des modifications est affiché. Ce diagramme d'activité pourrait aussi, de par son principe de fonctionnement, correspondre au cas d'utilisation «Modifier un utilisateur»

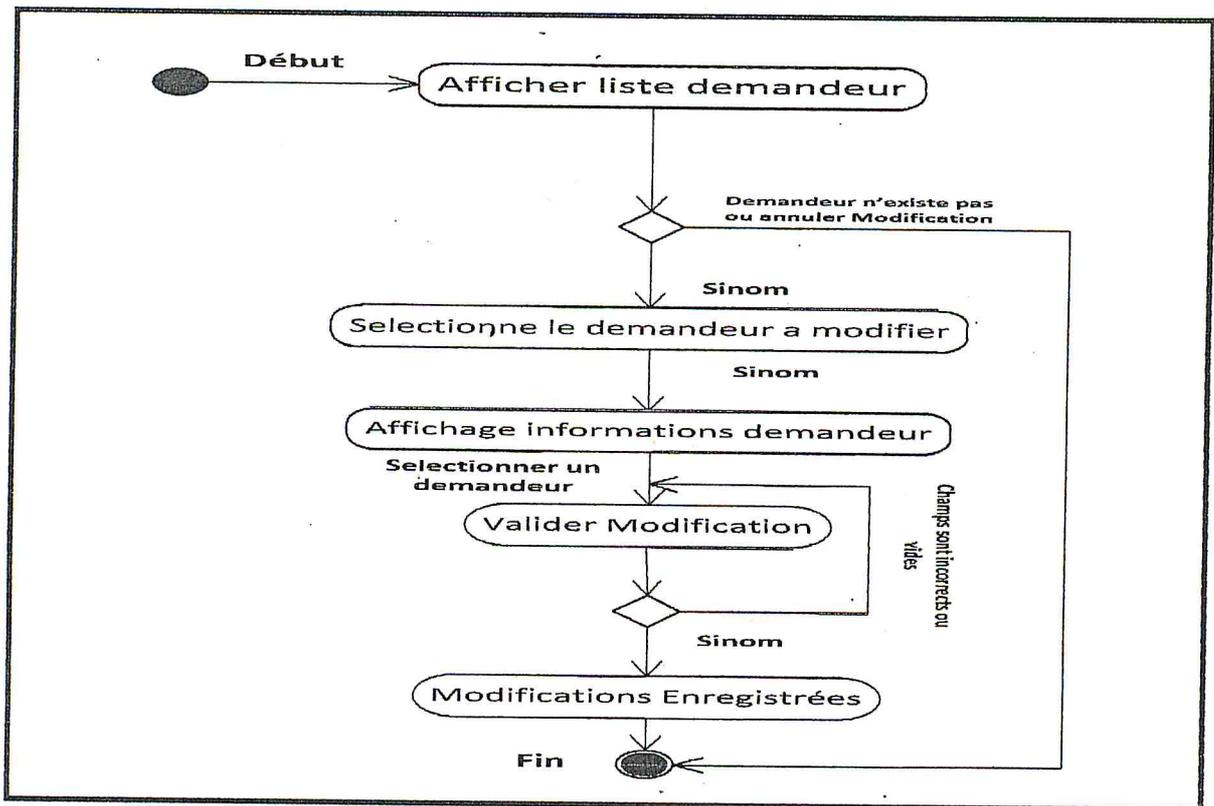


Figure 3.8 : Diagramme d'activité «Modifier Demandeur».

#### ❖ Ajouter Offre d'emploi :

Dans un premier temps, le formulaire de saisi des données est affiché, l'agent saisira les informations concernant l'offre d'emploi. Ces derniers subiront une verification, si les champs du formulaire sont vides ou bien incorrects le système nous l'indiquera et va nous demandé de refaire l'action correctement, sinon, si l'entreprise qui a proposé l'emploi n'existe pas le système l'indiquera et va nous demandé de l'ajouter. Dans le cas ou les champs du formulaire saisi sont corrects et que l'entreprise offrante existe, l'étape de validation va avoir lieu. Après la validation le système va verifié si l'offre existe déjà, si c'est le cas le système l'indiquera et l'activité prendra fin. Dans le cas contraire, le système recupera les informations de l'offre puis elle sera enregistrés puis l'activité prendra fin.

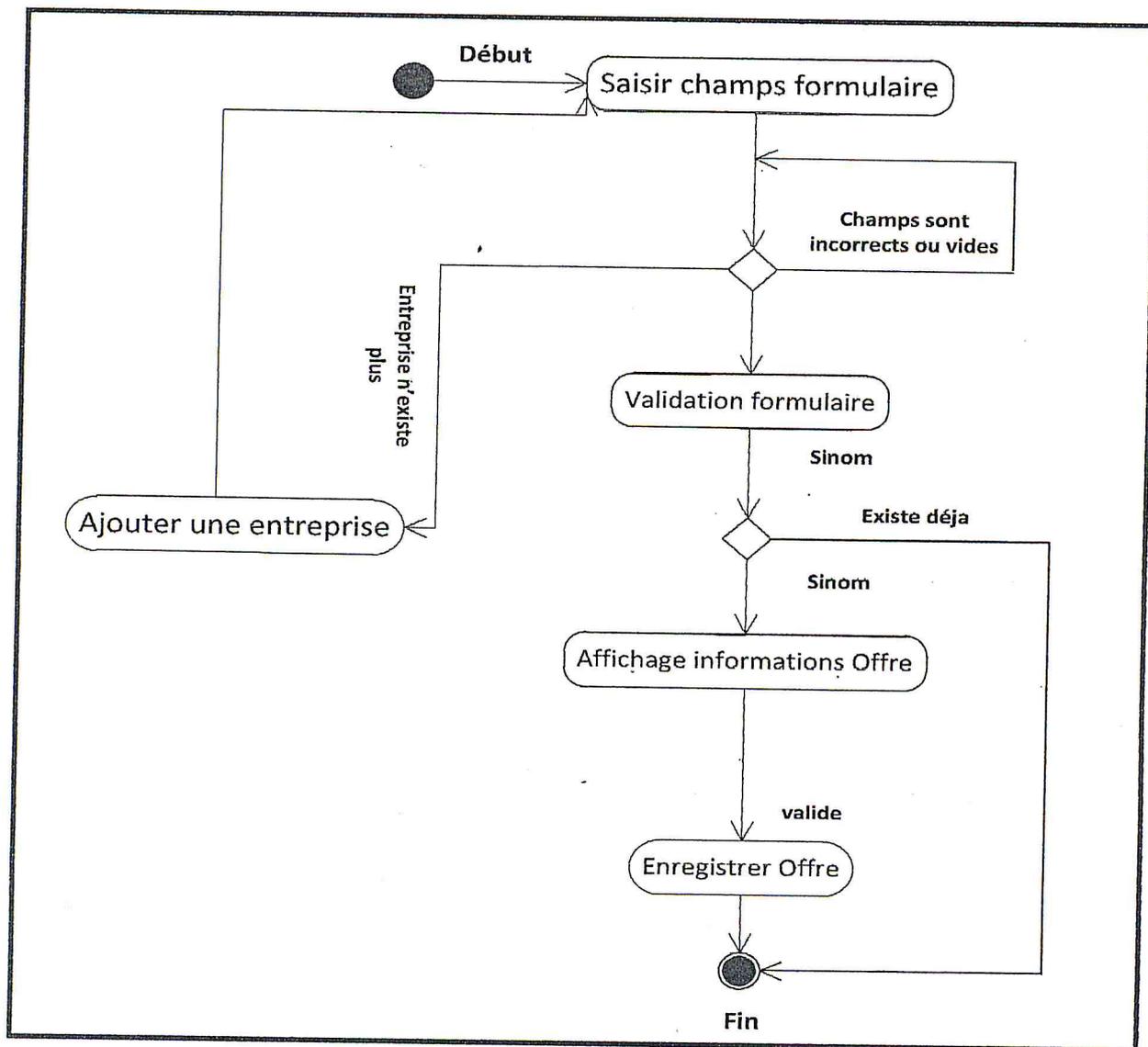


Figure 3.9 : Diagramme d'activité «Ajouter Offre».

3.4.3. Diagramme de séquence :

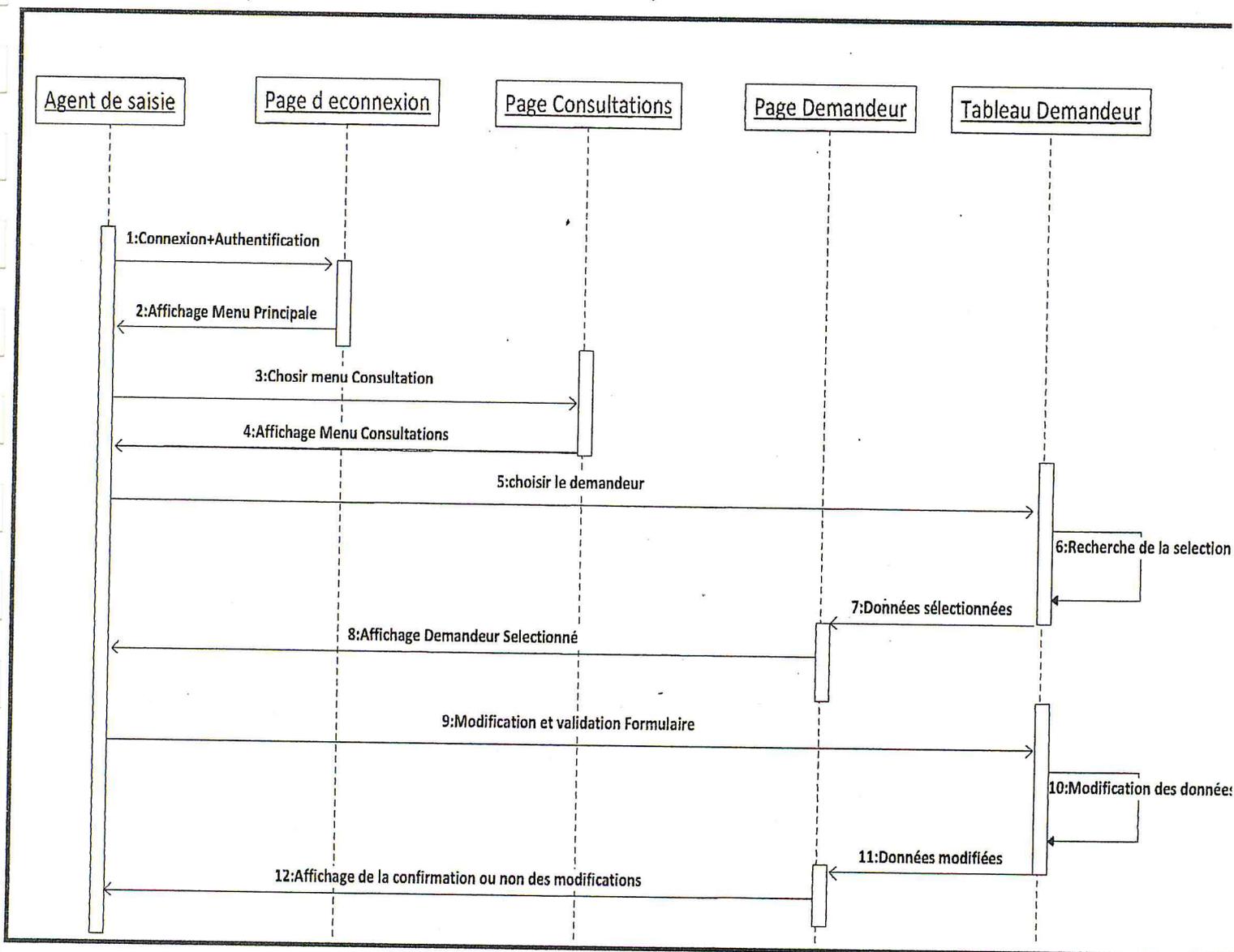


Figure 3.10 : Diagramme de séquence « Modifier Demandeur ».

❖ **Modifier Demandeur :**

Le tableau suivant résume le scénario de modification demandeur :

Etape	Description
1. Connexion + authentification.	L'agent de saisie doit d'abord s'authentifier pour pouvoir accéder au menu principal de l'outil. Une fois l'authentification réussie, le page d'accueil est automatiquement affiché.
4. Affichage menu consultation.	Le menu de consultation s'affiche après l'avoir sélectionner.
5. Choisir le demandeur.	L'agent de saisie choisi dans la liste (menu déroulant) qui affiche les demandeurs existant.
8. Affichage demandeur sélectionné.	Affichage de tous les renseignements concernant le demandeur choisi.
9. Modification et validation du formulaire.	L'agent de saisie doit modifier comme il peut ne rien faire et valide. Ensuite, les champs modifier seront contrôlés afin d'éviter les erreurs.
12. Affichage de la confirmation.	Après la validation du formulaire d'insertion des données, l'application envoie le résultat de la requête en affichant un message de confirmation ou d'erreur.

*Tableau 3.1 : le scénario de modification demandeur.*

❖ Consulter Offre :

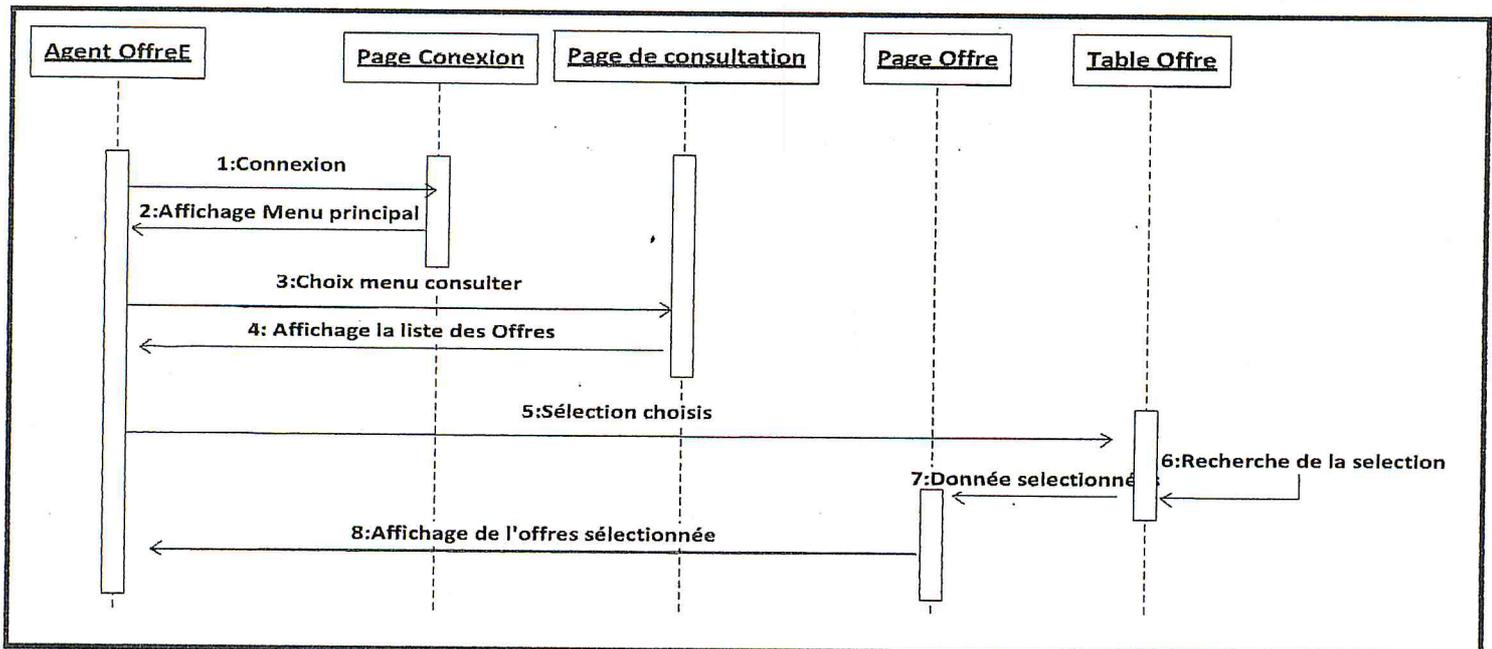


Figure 3.11 : Diagramme de séquence «Consulter Offre».

Le tableau suivant résume le scénario de consultation d’offres :

Etape	Description
<b>1. Connexion</b>	L’agent de saisie doit d’abord s’authentifier pour pouvoir accéder au menu principal de l’outil. Une fois l’authentification réussie, la page d’accueil est automatiquement affichée.
<b>3. Choix menu consultation.</b>	Le menu de consultation s’affiche après l’avoir sélectionner.
<b>4. Affichage de la liste des offres</b>	L’agent de saisie choisi dans la liste (menu déroulant) qui affiche les offres existant.
<b>8. Affichage de l’offresélectionnée</b>	Affichage de tous les renseignements concernant l’offre sélectionnée.

Tableau 3.2 : le scénario de Consultation offre.

#### 3.4.4. Diagramme de classe :

❖ **Les règles de gestion :**

- Chaque demandeur a le droit de choisir trois métiers préférés parmi le groupe de métiers.
- Chaque offre propose un ou plusieurs offres d'emploi.
- Une offre d'emploi est proposée par un seul offreur.
- Chaque demandeur a un seul niveau de qualification.
- Un niveau de qualification concerne un ou plusieurs demandeurs.
- Un niveau de qualification concerne une et une seule offre.
- Un demandeur peut avoir une ou plusieurs formations de l'enseignement supérieur.
- Un demandeur a un seul niveau d'instruction.
- Un niveau d'instruction concerne un ou plusieurs demandeurs.
- Un demandeur appartient à une seule raison d'inscription.
- Un demandeur peut avoir une ou plusieurs formations.
- Une formation concerne un ou plusieurs demandeurs.
- Une formation peut être professionnel ou enseignement supérieure.
- Un demandeur peut être concerné par un seul placement.
- Un placement concerne par un ou plusieurs demandeurs.
- Un placement peut être concerné par une seule offre d'emploi.
- Une offre d'emploi concerne un ou plusieurs placements.
- Les conditions de travail sont contenues par plusieurs offres.
- Une offre d'emploi a plusieurs conditions de travail.
- Une branche d'activité appartient à une seule offre d'emploi.

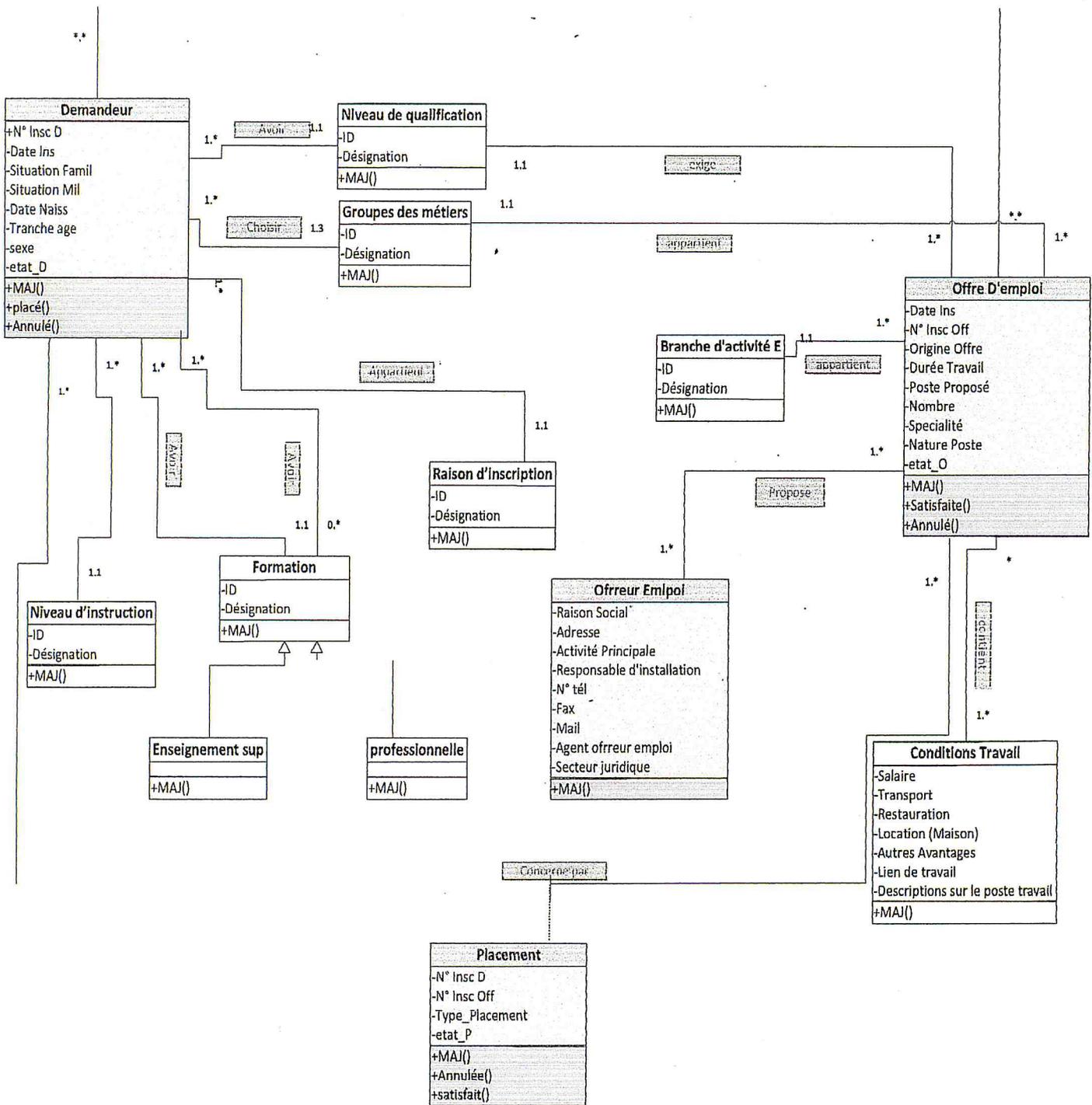


Figure 3.12 : Diagramme de classe.

❖ Description des classes :

• Demandeur :

Classe	Attribut	Désignation
Demandeur	N° insc D	Numéro d'inscription du demandeur
	Date Insc	Date de l'inscription
	Situation famil	la situation familiale du demandeur
	Situation mil	la situation militaire du demandeur
	Date Naiss	la date de naissance du demandeur
	Tranche age	tranche d'âge du demandeur
	sexe	sexe du demandeur
	etat_D	état de la demande

• Offre :

Classe	Attribut	Désignation
Offre	Date Ins	Date de réception offre
	N°Insc Off	Numéro d'inscription de l'offre
	Origine Offre	l'origine de l'offre
	Durée Travail	La durée du contrat de travail
	Poste proposé	Désignation du poste de travail
	Nombre Poste	Nombre de poste disponible
	Spécialité	Spécialité requise pour le poste
	Nature Poste	Nature du poste
	Etat_O	l'état de l'offre

• Placement :

Classe	Attribut	Désignation
Placement	N° insc D	Numéro d'inscription du demandeur
	N°Insc Off	Numéro d'inscription de l'offre
	Type Placement	Type Placement (Classique,DAIP)
	Etat_P	Etat du placement

- **Offreur :**

Classe	Attribut	Désignation
Offreur Emploi	Raison Sociale	nom de société civile
	Adresse	Adresse du siège de l'entreprise
	Activité Principale	Activité principale de l'entreprise
	Responsable d'installation	Responsable ou directeur RH de l'entreprise
	N°tél	Numéro de téléphone de l'entreprise
	Fax	Numéro Fax de l'entreprise
	Mail	l'Adresse électronique de l'entreprise
	Agent offreur emploi secteur juridique	Agent responsable des offres d'emploi Publique ou privée

- **Formation:**

Classe	Attribut	Désignation
Formation	ID	Identifiant
	Désignation	Désignation de la formation

- **Niveau d'instruction:**

Classe	Attribut	Désignation
Niveau d'instruction	ID	Identifiant
	Désignation	Désignation du niveau d'instruction (sans instruction, primaire, Secondaire, BAC, Formation professionnelle, Université etc...)

- **Raisons d'inscription:**

Classe	Attribut	Désignation
Raisons d'inscription	ID	Identifiant
	Désignation	STR1, STR2

NB : STR1 (Licenciement, démission, fin de contrat, Autre demandeur) -  
STR2 (il n'a jamais travaillé).

- Niveau de qualification:

Classe	Attribut	Désignation
Niveau Qualification	ID	Identifiant
	Désignation	Désignation du niveau de qualification (Personnel sans qualification, personnel d'aide, personnel qualifié, personnel hautement qualifié, technicien supérieur, cadres et cadres supérieur etc...)

- Branche d'activité E (E pour économique) :

Classe	Attribut	Désignation
Branche d'activité E	ID	Identifiant
	Désignation	Pêche, Agriculture, sylviculture et exploitation forestière, Eau et énergie, hydrocarbures, industries diverses,... etc.

- Groupe des Métiers

Classe	Attribut	Désignation
Groupe des Métiers	ID	Identifiant
	Désignation	Agriculture, Mine et géologie, Textiles et confection, Industrie, Alimentaire et tabac...etc.

**4. Système Décisionnel :**

Les besoins sont collectés et les données sont auditées. Nous voilà prêt à lancer la conception logique de notre entrepôt de données.

**4.1. Les diagrammes de cas d'utilisation :**

Afin de simplifier la compréhension des décideurs, Nous avons fait les diagrammes de cas d'utilisations suivants :

❖ **4.1.1. Analyse « Flux de la demande » :**

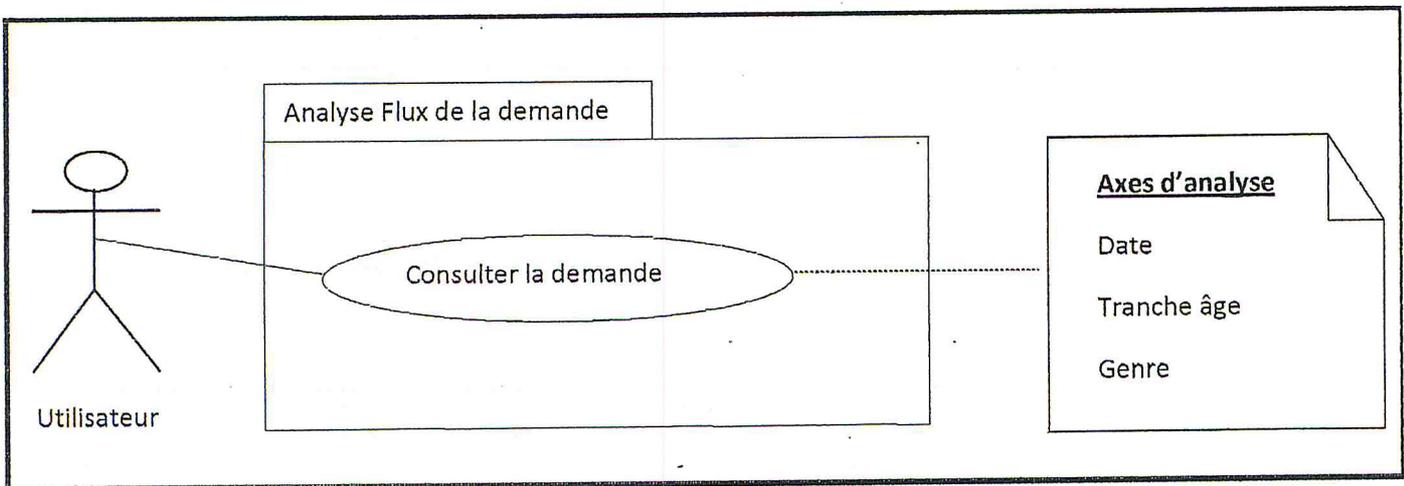


Figure 3.13 : Diagramme de cas d'utilisations d'analyse Flux de la demande.

❖ **4.1.2. Analyse « la demande enregistrée » :**

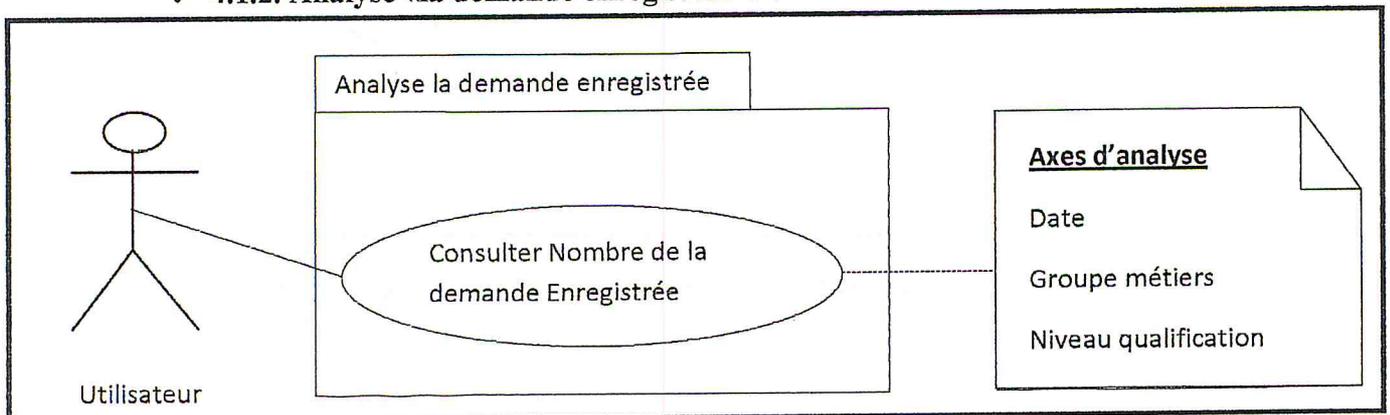


Figure 3.14 : Diagramme de cas d'utilisations d'analyse demande enregistrée.

#### 4.2. La modélisation multidimensionnelle :

La modélisation multidimensionnelle de notre entrepôt de données se compose de :

- ❖ six tables de faits :
  - ✓ Flux demande
  - ✓ Demande enregistré
  - ✓ Offre
  - ✓ Placement
  - ✓ Placement\_FP
  - ✓ Placement\_ES
  
- ❖ Sept tables de dimensions :
  - ✓ Date
  - ✓ Tranche Age
  - ✓ Genre
  - ✓ Groupe métier
  - ✓ Diplôme
  - ✓ Niveau Qualification
  - ✓ Filière

Sachant qu'on va utiliser le modèle en étoile dans cette modélisation multidimensionnelle, et que l'alimentation de notre entrepôt de données se fait chaque début de mois.

## 4.2.1. Flux de demande d'emploi par tranche d'âge et sexe:

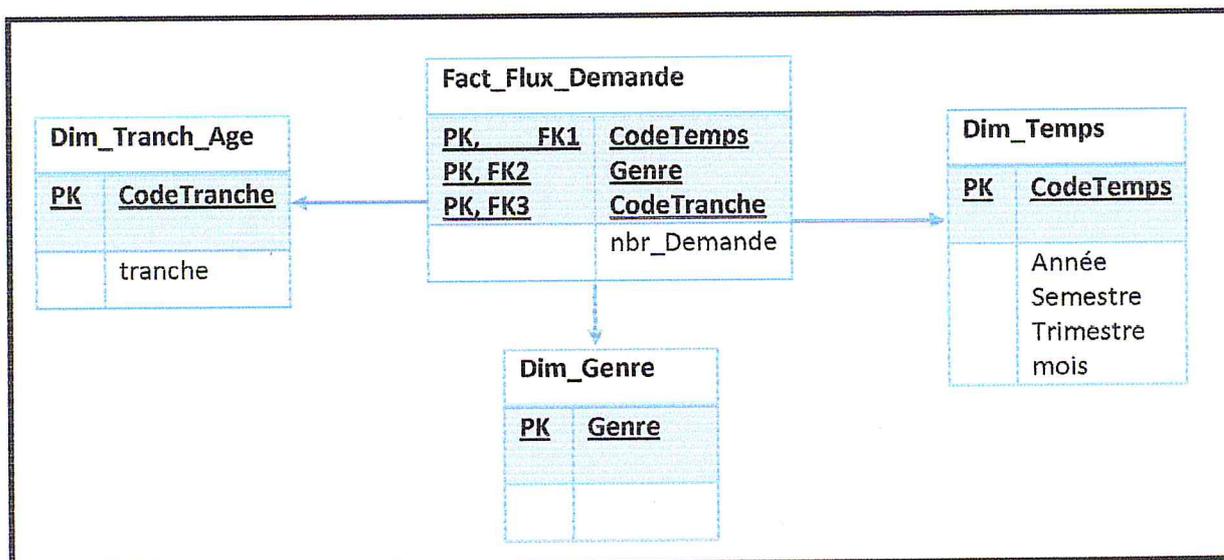


Figure 3.19 : Modèle en étoile de l'activité « Flux de la demande d'emploi ».

- **Le processus d'activité :**

Ce processus nous montre la répartition des flux de demande d'emploi par genre (sexe) et tranche d'âge.

- **Le grain du processus d'activité :**

Il est important de connaître mensuellement le nombre de demande effectués et de faire une comparaison entre chaque mois et le même mois de l'année qui précède.

- **Les dimensions:**

Nous avons déterminé les dimensions suivantes : Temps, Tranche d'âge, Sexe

- **Les faits mesurés:**

Le fait est : Le nombre de la demande d'emploi.

#### 4.2.2. Demandes enregistrées par niveau de qualification, par groupe de métier :

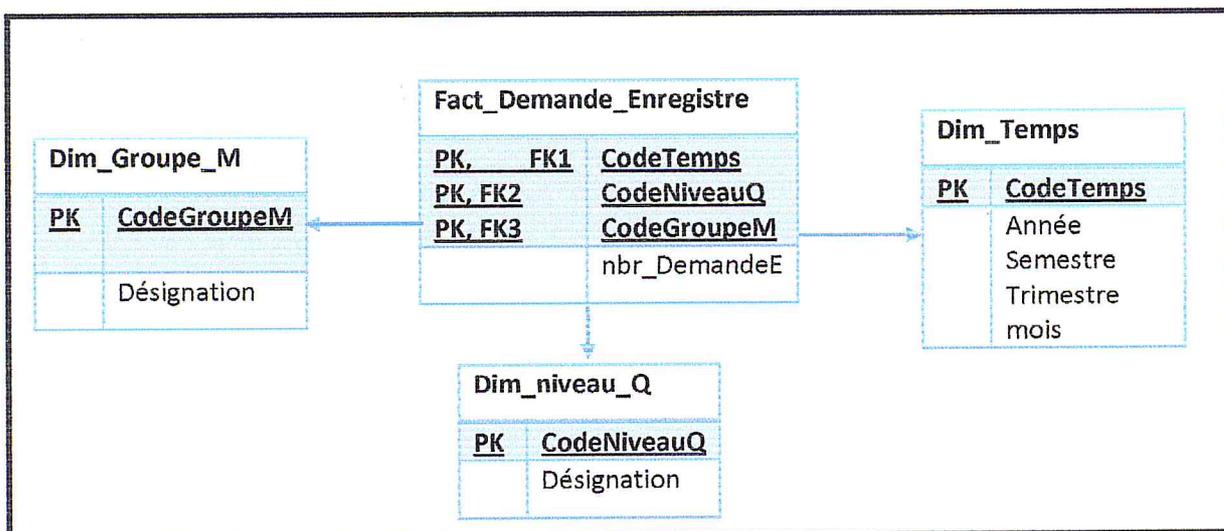


Figure 3.20 : Modèle multidimensionnel en étoile de l'activité « demandes enregistrées ».

##### ■ Le processus d'activité :

Ce processus nous décrit le rapport qui nous donne le nombre de demandes d'emploi enregistrées à la fin du mois par niveau de qualification et groupe de métier.

##### ■ Le grain du processus d'activité :

Dans ce processus, il est important de connaître mensuellement le nombre de demandes enregistrées chaque fin du mois, par niveau de qualification et groupe de métier et de faire une comparaison entre chaque fin du mois et le même mois de l'année qui précède.

##### ■ Les dimensions:

Nous avons déterminé les dimensions suivantes : Temps, Groupe de métier, Niveau de qualification.

##### ■ Les faits mesurés:

Le fait est Le nombre de demande d'emploi enregistrée.

### 4.2.3. Nombre d'offre par niveau de qualification et par groupe de métier :

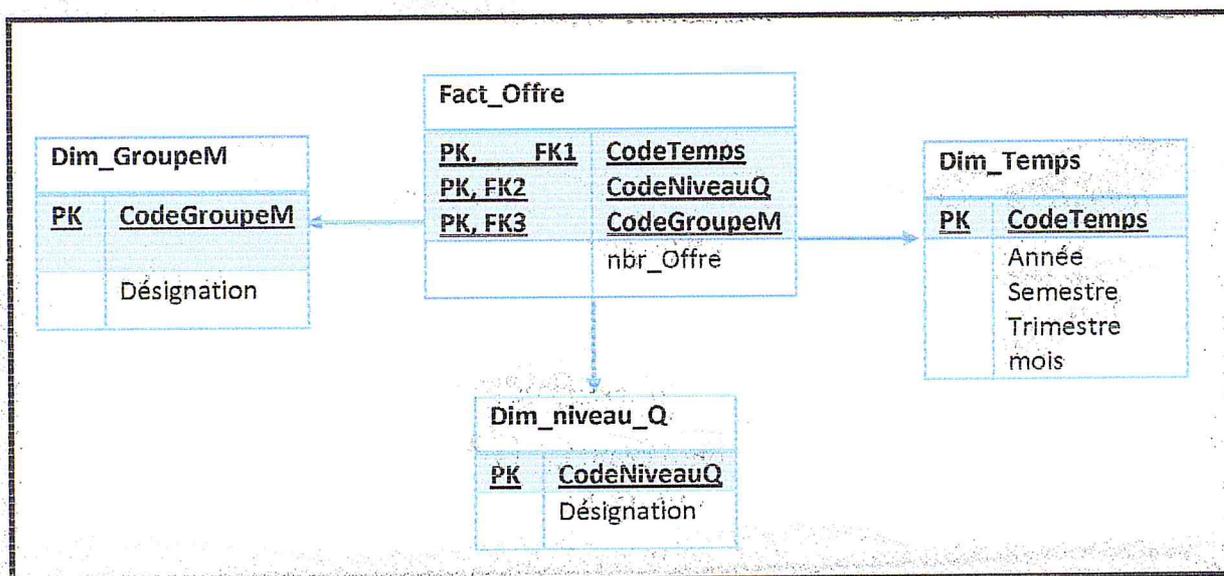


Figure 3.21 : Modèle multidimensionnel en étoile de l'activité « Offre ».

#### ■ Le processus d'activité :

Ce processus nous décrit l'établissement des rapports concernant les offres d'emploi réparties par groupe de métier et niveau de qualification.

#### ■ Le grain du processus d'activité :

Dans ce processus, il est important de connaître mensuellement le nombre de d'offre par niveau de qualification et groupe de métier et de faire une comparaison entre chaque mois et le même mois de l'année qui précède.

#### ■ Les dimensions:

Nous avons déterminé les dimensions suivantes : Temps, Groupe de métier, Niveau de qualification.

#### ■ Les faits mesurés:

Le fait est Le nombre d'offres.

#### 4.2.4. Répartition des demandes d'emploi placées par sexe, par groupe de métier, niveau de qualification :

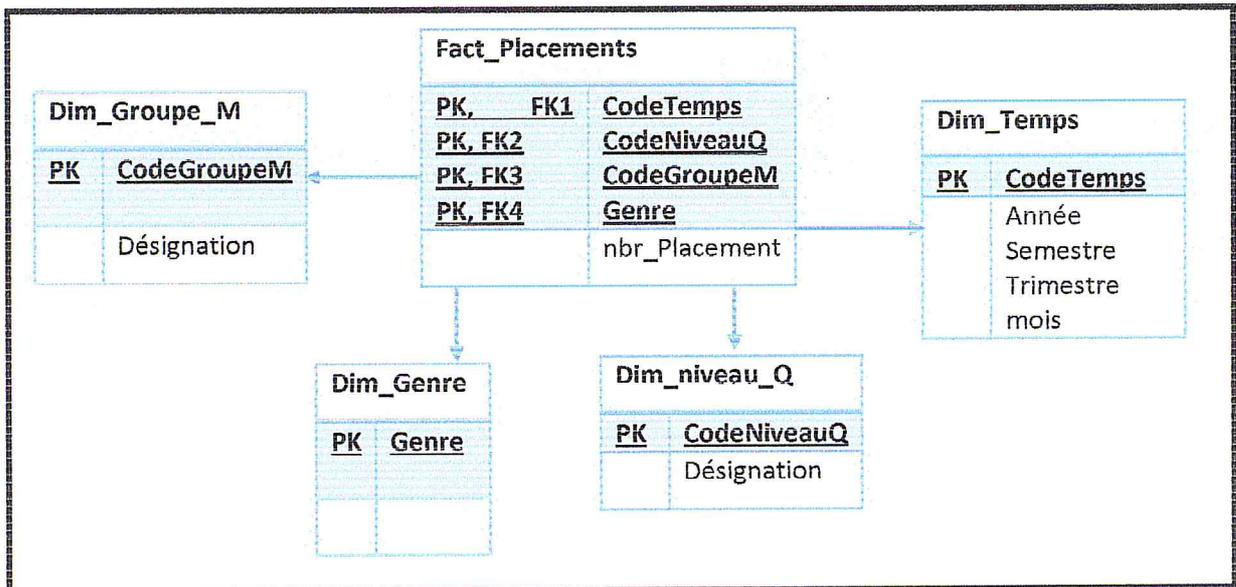


Figure 3.22 : Modèle multidimensionnel en étoile de l'activité « Placements ».

##### ■ Le processus d'activité :

Ce processus explique les rapports établis mensuellement concernant les placements par genre (sexe), groupe de métier et niveau de qualification.

##### ■ Le grain du processus d'activité :

Dans ce processus aussi, il est important de connaître mensuellement le nombre de placement par niveau de qualification, groupe de métier et sexe et de faire une comparaison entre chaque mois et le même mois de l'année qui précède.

##### ■ Les dimensions:

Nous avons déterminé les dimensions suivantes : Temps, Groupe de métier, Niveau de qualification, Sexe.

##### ■ Les faits mesurés:

Le fait est Le nombre de placement par Groupe de métier, sexe et Niveau de qualification effectués.

#### 4.2.5. Placements des diplômées issus de la formation professionnelle par diplômes, par groupe de métier, genre (sexe):

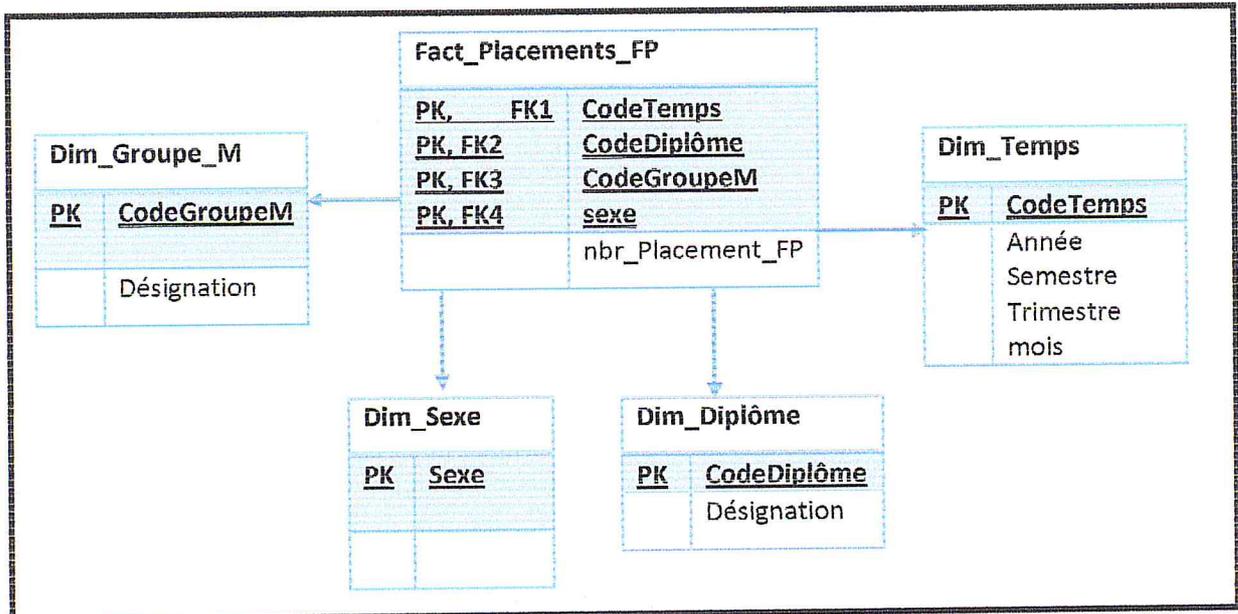


Figure 3.23 : Modèle multidimensionnel en étoile de l'activité « Placement\_FP ».

##### ■ Le processus d'activité :

Ce processus nous décrit les rapports établis chaque fin de mois concernant les placements des demandeurs d'emploi sortant de la formation professionnelle répartis par groupe de métier, genre (sexe) et diplôme.

##### ■ Le grain du processus d'activité :

Il est important de connaître mensuellement le nombre des placements issus de la formation professionnelle par Diplôme, groupe de métier et sexe et de faire une comparaison entre chaque mois et le même mois de l'année qui précède.

##### ■ Les dimensions:

Nous avons déterminé les dimensions suivantes : Temps, Groupe de métier, Diplôme, Sexe.

##### ■ Les faits mesurés:

Le fait est : Le nombre de placements issus de la formation professionnelle.

#### 4.2.6. Placements des diplômées issus de l'enseignement supérieur par filière, par tranche d'âge, genre (sexe):

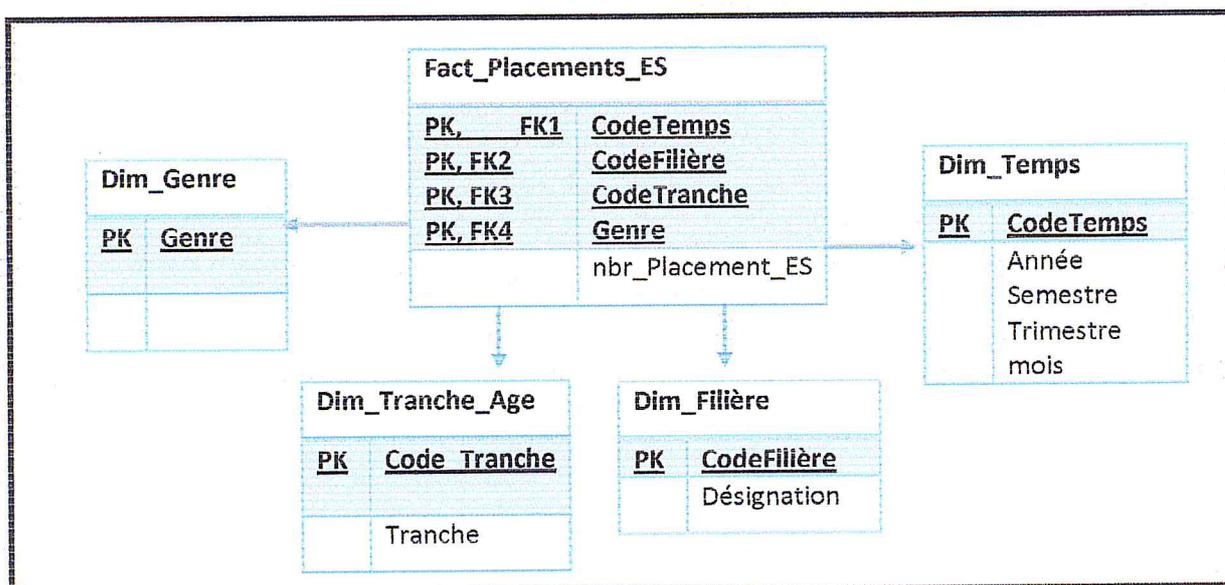


Figure 3.24 : Modèle multidimensionnel en étoile de l'activité « Placement\_ES ».

##### ■ Le processus d'activité :

Ce processus nous décrit les rapports établis chaque fin de mois concernant les placements des demandeurs d'emploi sortant de l'enseignement supérieur répartis par tranche d'âge, genre (sexe) et filière.

##### ■ Le grain du processus d'activité :

Il est important de connaître mensuellement le nombre des placements issus de l'enseignement supérieur par Filière, tranche d'âge et sexe et de faire une comparaison entre chaque mois et le même mois de l'année qui précède.

##### ■ Les dimensions:

Nous avons déterminé les dimensions suivantes : Temps, Tranche d'âge, sexe, filière.

##### ■ Les faits mesurés:

Le fait est : Le nombre de placements issus de l'enseignement supérieur.

**4.2.7. Schéma de l'entrepôt de données (DREM) :**

Le schéma relationnel de notre base de l'entrepôt de données (Schéma en constellation) regroupe tous les schémas en étoile définis précédemment. Chaque entité de notre modélisation sera une table de notre base de données relationnelle.

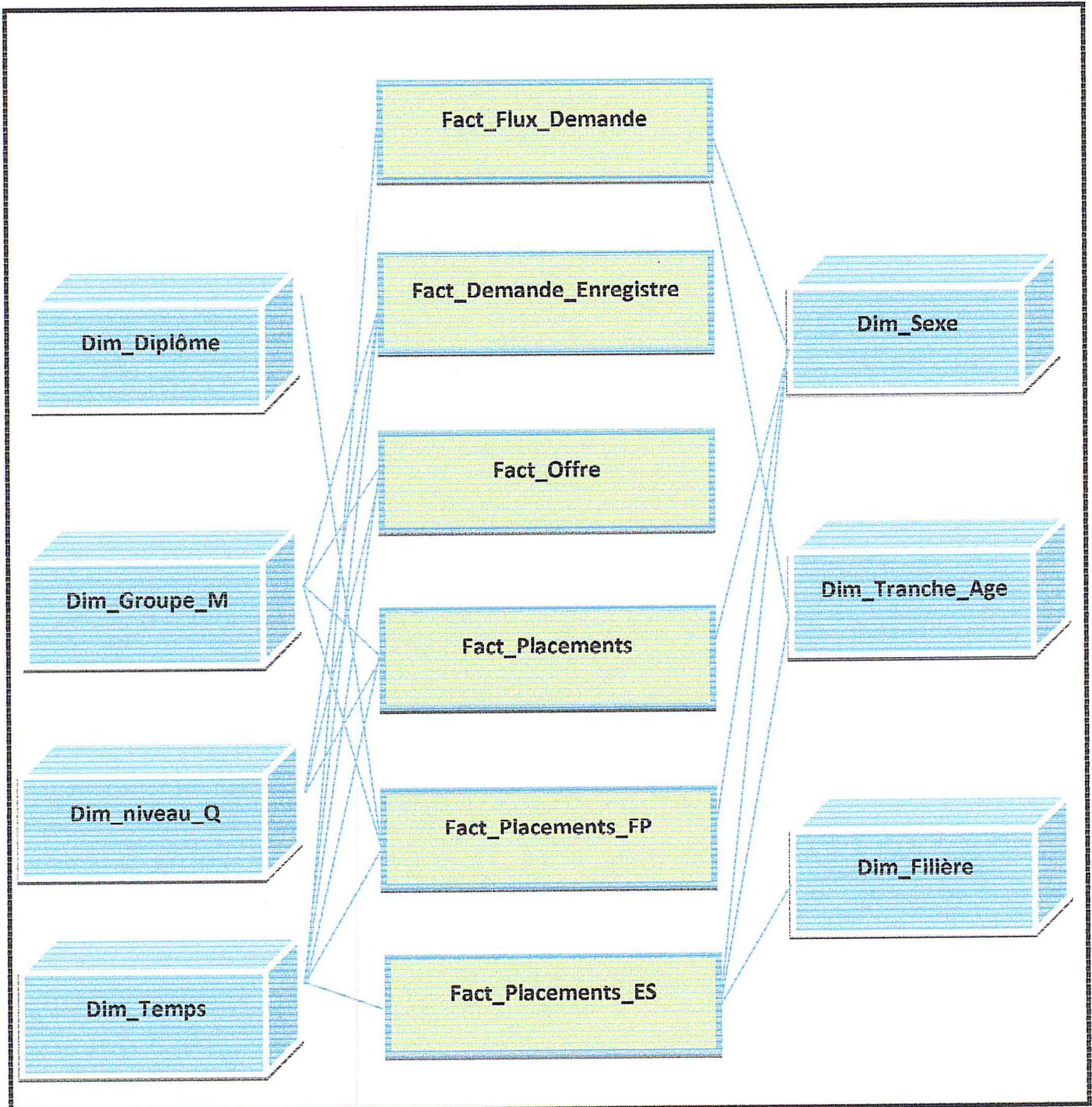


Figure 3.25 : schéma en constellation.

### 4.3. Conception ETL :

#### 4.3.1. Système "ETL" :[24]

On entend par là : *Extraction, Transformation, Loading*. C'est un système par lequel vont passer toutes les données des systèmes opérationnels avant d'arriver dans la forme souhaitée dans l'entrepôt.

Ce système est une sorte de moulinette qui fait passer toutes les données de l'entreprise. Les données en sortie (passées à la moulinette) seront nettoyées, purifiées, contextualisées (les données des différents systèmes s'homogénéiseront) et prêts à être reçus dans l'entrepôt.

Le système d'ETL est la partie la plus importante d'un projet décisionnel. Car c'est avec l'ETL que les systèmes seront mis en relation, les erreurs détectées, les calculs complexes effectués, etc. On peut dire que la solidité d'un ETL détermine la viabilité du projet.

Pour alimenter notre entrepôt de données, nous allons passer par trois étapes nécessaires :

- **Une étape d'extraction** : Il s'agit d'extraire les données pertinentes depuis les bases de production.
- **Une étape de Transformation** : Il s'agit de faire le nettoyage, le formatage et l'agrégation des données extraites.
- **Une étape de chargement** : Charger les données propres dans la base décisionnelle.

#### 4.3.2. Diagrammes d'activité ETL:

Le diagramme d'activité représente le comportement d'une méthode ou d'un cas d'utilisation sous forme de flux.

Le diagramme d'activité doit représenter l'ensemble des actions à réaliser par le système avec tous les branchements conditionnels et toutes les boucles possibles.

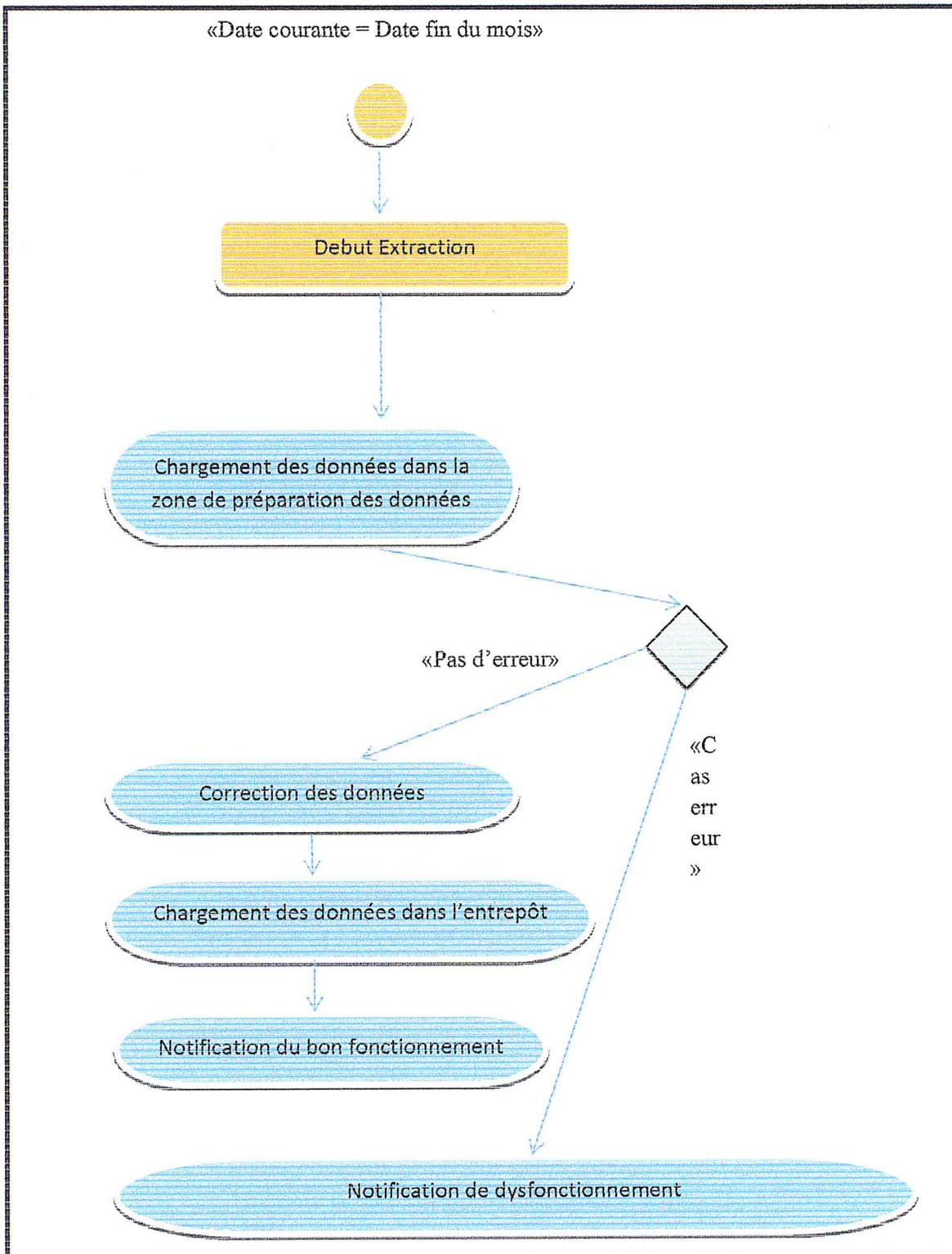


Figure 3.26 : Diagramme d'activité du système ETL [25]

**5. Conclusion :**

Dans ce chapitre, nous avons modélisé notre entrepôt de données. Le modèle de données est basé sur un schéma en étoile comportant plusieurs faits ainsi qu'une hiérarchie détaillée des dimensions.

Notre modélisation multidimensionnelle est décrite tout en respectant la démarche de « Ralph Kimball » de la conception multidimensionnelle.

# **4 MISE EN OEUVRE ET DEPLOIEMENT**

- Description de l'environnement de travail
- Définition de l'architecture physique et technique
- Mise en œuvre du système

## 1. Introduction:

La dernière phase de ce travail est la mise en œuvre du logiciel. Pour cela nous avons opté pour une architecture spécifique et utilisé différents outils de développement de notre système logiciels

La première partie consiste à concevoir une application transactionnelle (Système opérationnel) pour la saisie des informations nécessaires aux rapports.

Notre base de données, sera implémentée sur le SGBD SQL serveur, quant au développement de l'application, il se fera sous ASP plate-forme .NET.

Et pour la partie système décisionnel, elle sera composé des deux systèmes «ETL » et « Reporting ».

## 2. Description de l'environnement de travail:

### 2.1. Système Opérationnel :

#### ■ La plateforme .NET et Framework :

.NET (prononcer «Dotent») est un standard proposé par la société Microsoft, pour le développement d'applications d'entreprises multi-niveaux, basées sur des composants. Microsoft .NET constitue ainsi la réponse de Microsoft à la plate-forme *J2EE* de Sun. La plate-forme .NET a été élaborée en s'appuyant sur une communauté d'utilisateurs et a abouti à l'élaboration de spécifications. Ces spécifications ont été ratifiées par un organisme international de standardisation, l'ECMA (*European Computer Manufacturer Association*), ce qui en fait un standard. Ainsi l'effort de standardisation a permis l'émergence de plates-formes portées par des entreprises tierces et disponibles sous un grand nombre de systèmes d'exploitation.

On parle généralement de «Framework» (traduit par «socle») pour désigner l'ensemble qui constitue les services (API) offerts et de l'infrastructure d'exécution. Le framework.NET comprend notamment :

- ✓ L'environnement d'exécution :
  - un moteur d'exécution, appelé CLR (*Common Language Runtime*), permettant de compiler le code source de l'application en un langage intermédiaire, baptisé MSIL (*Microsoft IntermediateLanguage*) et agissant telle la machine virtuelle Java. Lors de la première exécution de l'application, le code MSIL est à son tour compilé à la volée en code spécifique au système grâce à un compilateur JIT (*Just In Time*).
  - un environnement d'exécution d'applications et de services web, appelé *ASP .NET*.
  - un environnement d'exécution d'applications lourdes, appelé *WinForms*.
- ✓ Des services, sous forme d'un ensemble hiérarchisé de classes appelé Framework Class Library (*FCL*). La FCL est ainsi une librairie orientée objet, fournissant des fonctionnalités pour les principaux besoins actuels des développeurs. Le SDK (*Software Development Kit*) fournit une implémentation de ces classes.

#### ■ Le Visual-Basic :

Le VB est directement dérivé du BASIC et permet le développement rapide d'applications, la création d'interfaces utilisateur graphiques, l'accès aux bases de données, ainsi que la création de contrôles ou d'objets ActiveX.

#### ■ Le Visual Studio :

Visual Studio .NET ne fait pas partie du framework .NET. Il s'agit tout simplement d'un environnement de développement intégré proposé par Microsoft pour développer des applications conformes aux spécifications de .NET.

---

### ■ Microsoft SQL Server :

Microsoft est l'un des leaders des éditeurs de SGBD avec sa solution SQL Server. Selon des études faites par *Garthner*, IDC, et *ForresterWave*, SQL Server se positionne troisième derrière Oracle et IBM en terme du part de marché, premier en termes de croissance avec 28 % et aussi premier en terme de sécurité selon une étude de ESG en 2006.

### 2.2. Système Décisionnel :

#### ■ Microsoft SQL server 2008 R2 :

SQL Server est une plateforme décisionnelle (Business Intelligence) complète fournissant quatre composantes majeures :

- ✓ Des solutions d'entrepôts de données (Data Warehouse)
- ✓ Un ETL
- ✓ Un moteur OLAP et Datamining
- ✓ Une solution complète de Reporting

#### ■ Microsoft Excel :

Microsoft Excel, dont le nom officiel actuel est Microsoft Office Excel, est un tableur de la suite bureautique Microsoft Office qui est écrite et distribuée par l'éditeur Microsoft, il est destiné à être utilisé sur des plateformes Microsoft Windows ou Macintosh Mac OS X.

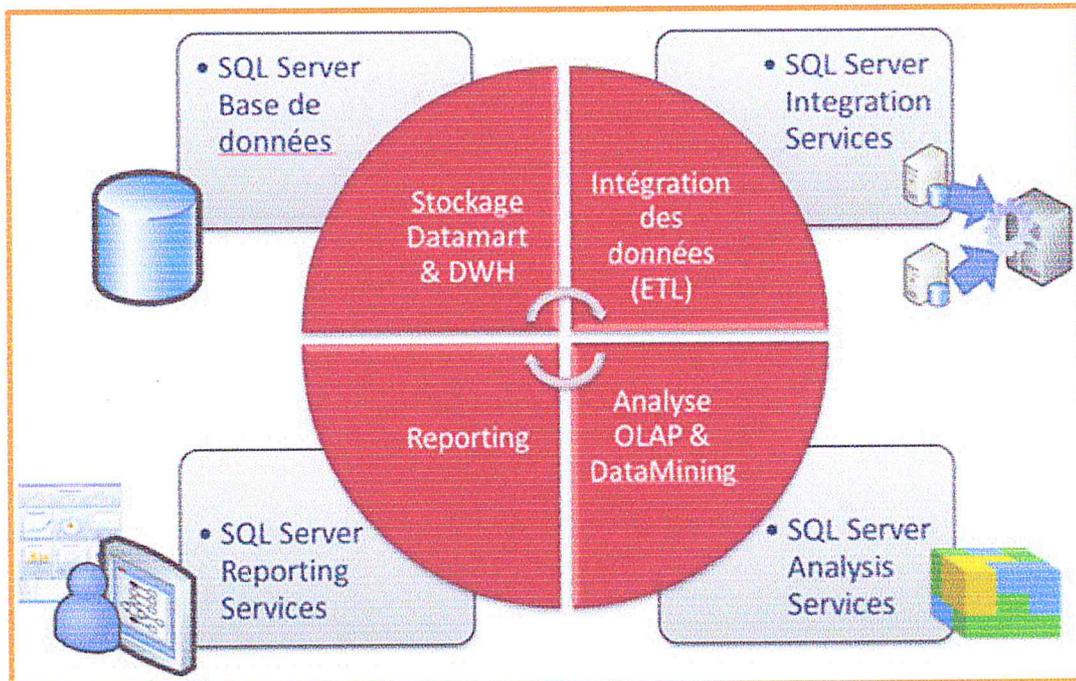


Figure 4.1 : Solution Business Intelligence SQL server 2008 R2. [26]

### 3. Définition de l'architecture physique et technique :

Avant la mise en œuvre de notre système, la présentation d'une architecture physique et technique est nécessaire.

#### 3.1 Architecture physique :

L'architecture physique (matérielle) retenue est définie en quatre tiers (quatre niveaux) :

- 1) Station utilisateur
- 2) Serveur base de données
- 3) Serveur d'application service web
- 4) Serveur DataWarehouse

On va représenter cette architecture (physique) par un diagramme de composant et une diagramme de déploiement.

**3.1.1. Diagramme de composant :**

Le diagramme de composant permet de décrire l'architecture physique et statique d'une application. Autrement dit, il permet l'organisation et la dépendance des composants de l'application de différentes natures.

Parce que notre système est composé de deux parties, opérationnel et décisionnel donc le diagramme de composant sera devise vers deux sous-systèmes.

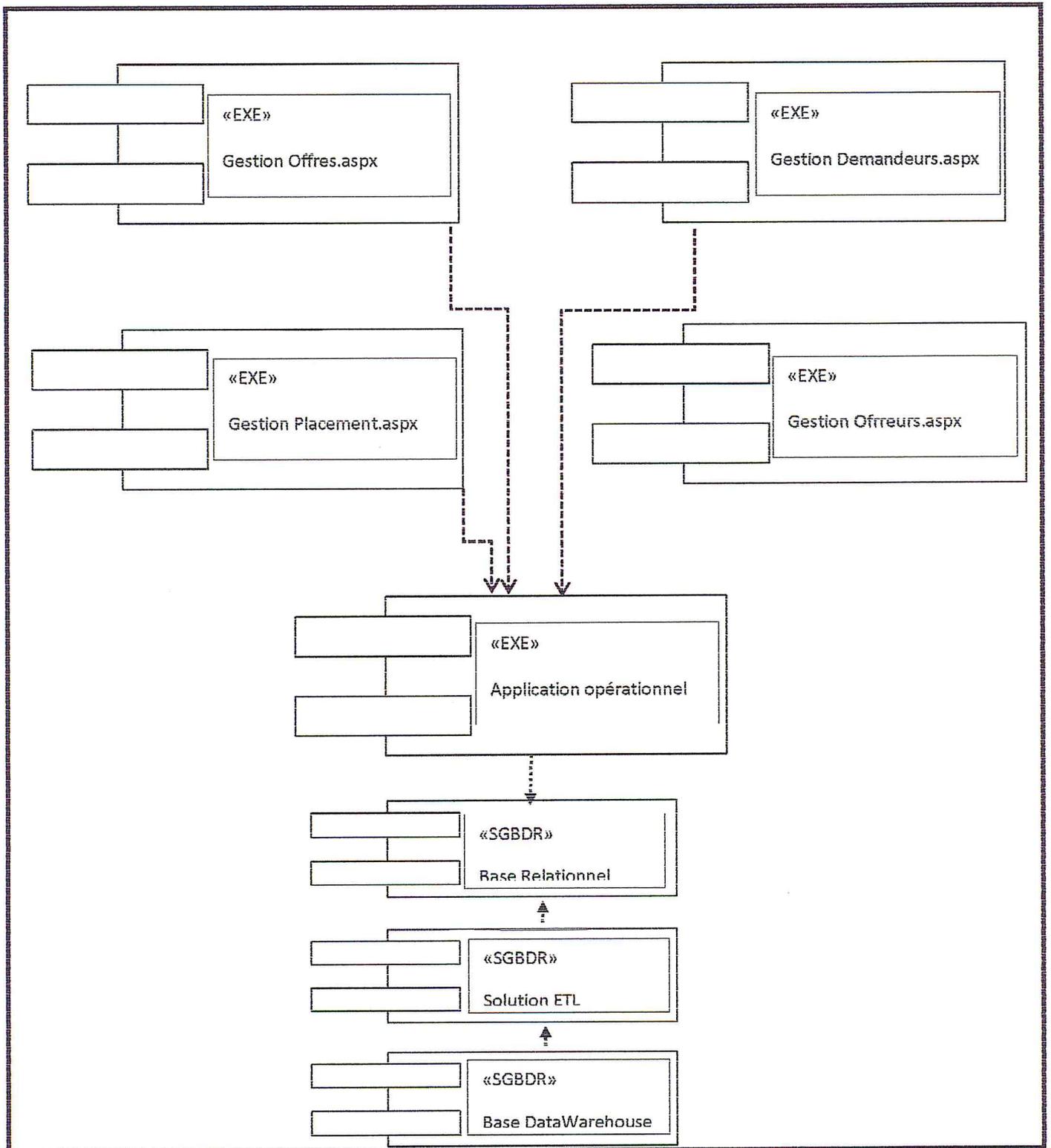


Figure 4.2: Diagramme de composant pour le système globale.

### 3.1.2. Diagramme de déploiement :

Le diagramme de déploiement décrit la disposition physique des ressources matérielles qui compose le système.

Le diagramme de déploiement de notre système est représenté comme suit :

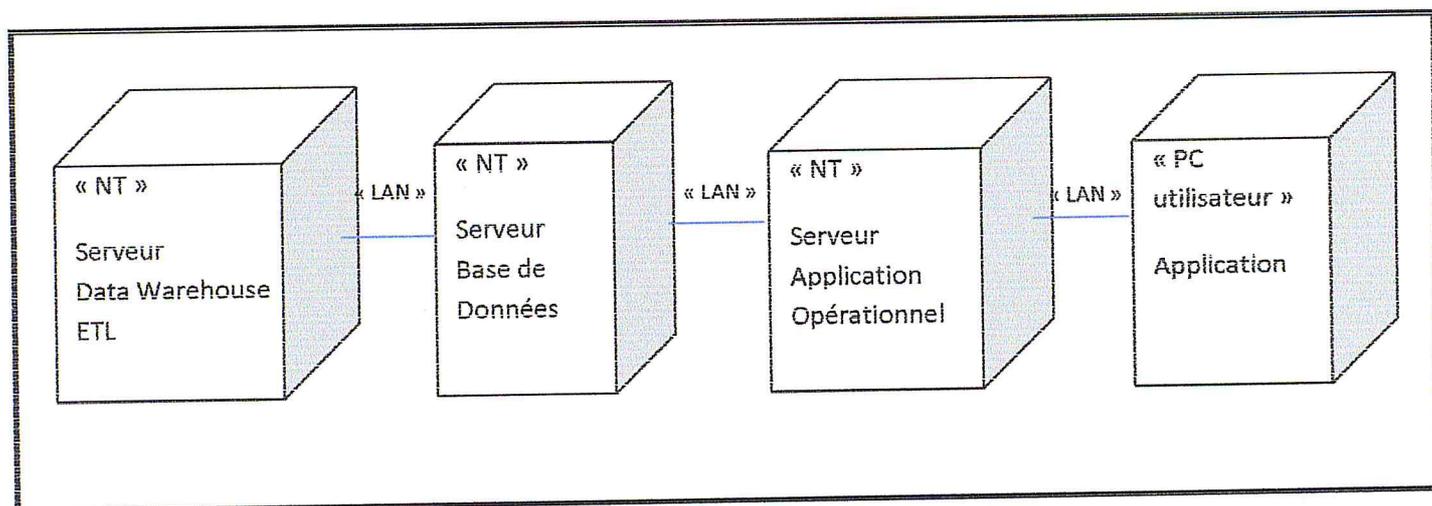


Figure 4.3 : Diagramme de déploiement du système global.

### 3.1 Architecture technique :

Les besoins des entreprises diffèrent beaucoup concernant l'analyse de données et la manière de les représenter. Nous avons opté pour la solution Microsoft qui offre une plate-forme Business Intelligence (BI) couvrant l'intégralité de la gamme des besoins.

Elle comprend outre, la base de données relationnelle, des composants visant à : l'intégration et à l'analyse des données ainsi qu'une infrastructure complète de Reporting.

Dès que les données proviennent de plusieurs sources de données, elles doivent être regroupées au cours du processus d'intégration.

Microsoft dispose de *SQL SERVER INTEGRATION SERVICES* qui nous permet d'importer les données à partir de nombreuses bases de données (SQL SERVER, ORACLE, DB2, ACCESS, etc.) également à partir de fichiers textes et documents XML pour les charger en mémoire, les transformer et enfin les propulser dans notre Data Warehouse créé avec le SGBD *SQL SERVER 2005* et installé dans le serveur DW.

Les données une fois présentes dans le Data Warehouse, elles vont être mises sous forme de CUBES par l'intermédiaire d'*ANALYSIS SERVICES* pour assurer l'analyse multidimensionnelle des données.

Pour présenter les données sous une forme compréhensible pour l'utilisateur, nous avons utilisé *REPORTING SERVICES* installé dans le serveur d'application IIS, qui est en mesure de lire les données à partir de nombreuses sources de données et de les publier dans les formats les plus divers (HTML, PDF, Excel, etc.).

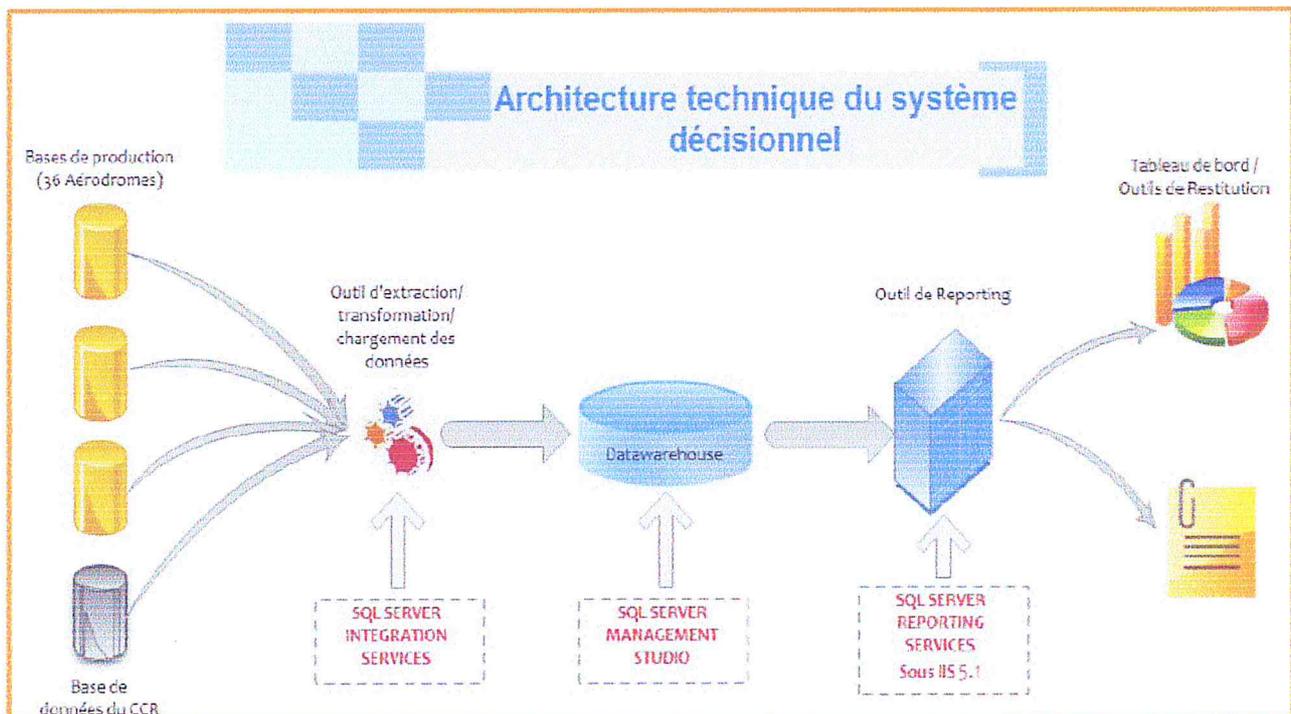


Figure 4.4 : Architecture technique du système décisionnel.[27]

#### 4. Mise en œuvre du système :

Dans cette partie nous allons décrire notre système opérationnel et l'outil de Reporting décisionnel en présentant quelques interfaces.

##### 4.1. Authentification : (système opérationnel)

La figure 4.5 montre la première page qu'un utilisateur voit s'afficher lorsqu'il accède à l'application.

Après la saisie du nom utilisateur et mot de passe, dans le cas où les champs sont valides, cette page va faire appel à la page « Accueil.aspx » qui va donner la main à l'utilisateur de faire tout ce qui est autorisé pour lui bien sûr.

The image shows a web application interface for authentication. At the top, there is a blue header with a house icon on the left, the text "DREM" in large letters, and "Direction régionale de l'emploi" below it. To the right of the header is a circular refresh icon. Below the header is a navigation bar with four items: "Accueil", "Demandeur", "Placement", and "Offres". On the left side, there is a vertical sidebar with four items: "Gestion demanqueur", "Gestion Offres", "Gestion placement", and "Gestion Entreprise". The main content area features a cartoon illustration of a person sitting at a computer. To the right of the illustration are two input fields: the first is labeled "Utilisateur" and the second is labeled "Mot de passe". Below these fields is a blue button labeled "Connecter". At the bottom of the page, there is a footer with the text "2010-2011".

Figure 4.5 : Interface d'authentification - système opérationnel.

##### 4.2. Ajouter un demandeur : (système opérationnel)

Pour ajouter un demandeur, l'agent saisi les informations de celui-ci, pour enfin valider la création du demandeur (AjouterDemandeur.aspx).



Figure 4.6 : Interface ajouter un demandeur - système opérationnel.

### 4.3. L'entrepont de données : (système décisionnel)

Le schéma de l'entrepôt regroupe les schémas en étoile définie dans la modélisation dimensionnelle. La base de données qui correspond à ce schéma global, s'appelle la base décisionnelle (par opposition aux bases de production) elle sera illustrée dans la figure suivante :

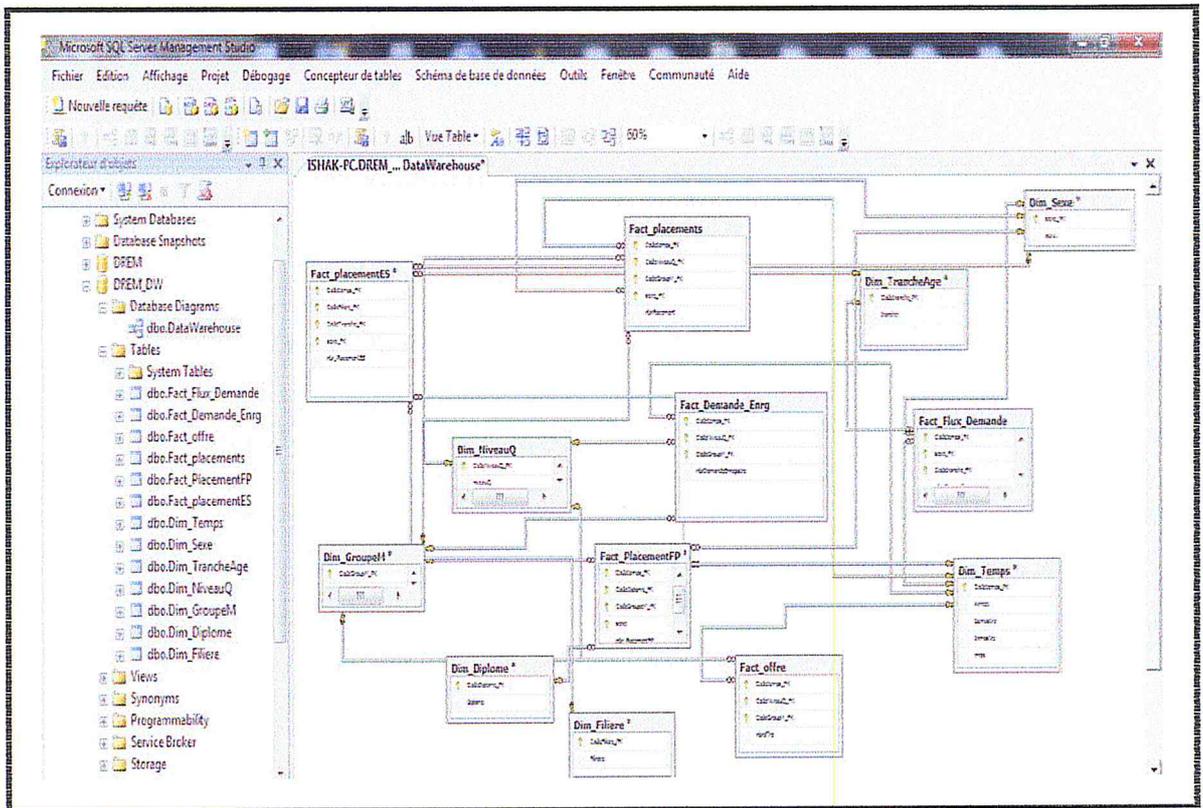


Figure 4.7 : schéma de l'entrepôt de données - système décisionnel.

#### 4.4. Un rapport final : (système décisionnel)

Un outil d'analyse très puissant les graphes, Excel propose une multitude de modèles entre autres les histogrammes, courbes, secteurs, barres, aires, bulles... etc.

Ce graphe représente le tableau récapitulatif du flux de la demande par tranche d'Age et par sexe.

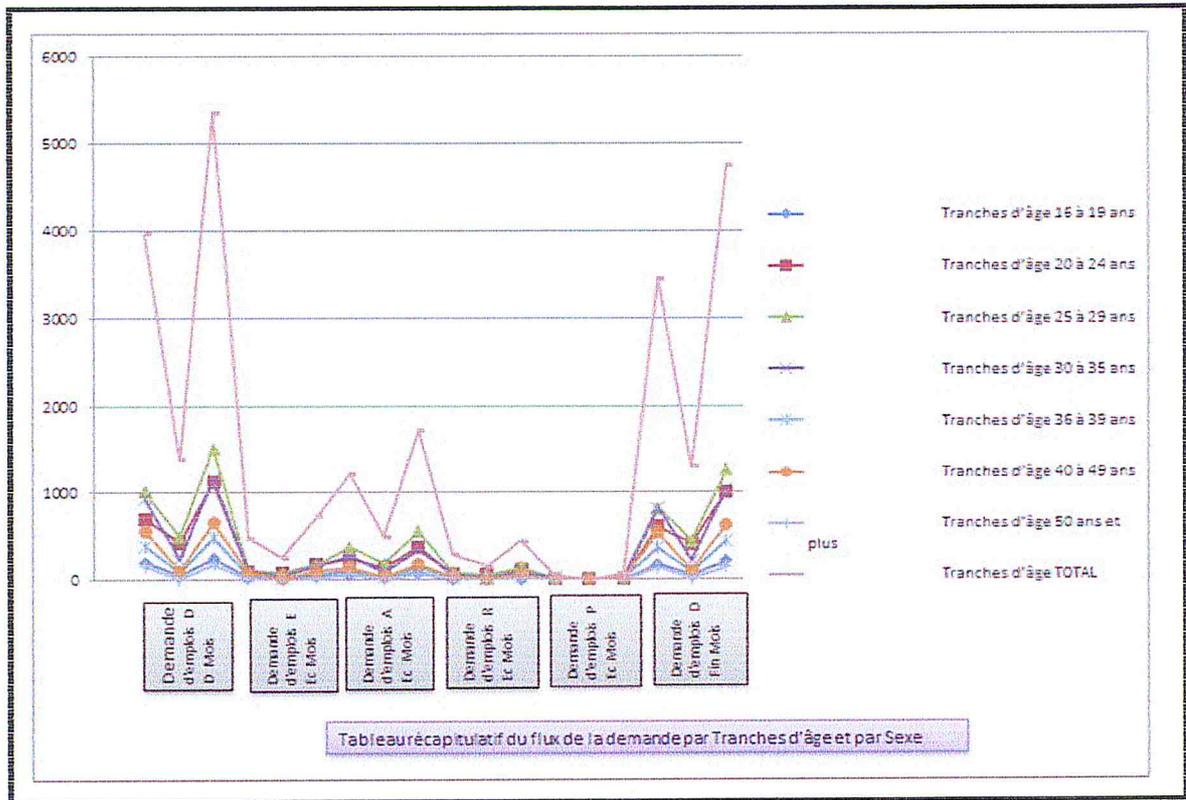


Figure 4.8 : Histogramme flux de demande généré par Excel.

#### 4.4.1. Description :

Ce graphe est généré par Microsoft Excel depuis la source de données « Cube flux de la demande ».

Cet histogramme représente le flux de la demande d'emploi enregistré disponible, revalider et annuler au début de mois, au cours du mois et la fin du mois par tranche d'âge et par sexe de chaque trimestre

#### 5. Conclusion :

Dans ce chapitre, nous avons vu les interfaces principales de nos outils opérationnels et décisionnels décrivant l'environnement de travail de chacune.

**PERSPECTIVES**

**ET**

**SUGGESTIONS**

## Perspectives et suggestions

### **Perspectives et suggestions :**

Etant donné son importance et dans le but d'offrir à DREM un tableau de bord facilitant le pilotage de l'ensemble de ses activités, nous souhaitons améliorer et enrichir notre système par :

- Intégrer les deux systèmes d'informations (système classique, et système DAIP)
- Sécurisé la plateforme.
- Implémentation d'une application WEB pour la gestion à distance de l'ETL.
- Application des méthodes de Datamining sur les données de l'entrepôt.
- Centralisé l'application au niveau de l'ANEM.
- installation d'un système de gestion des privilèges pour les analystes, les administrateurs, les agents de saisie et les décideurs.
- Extension du système pour les autres activités de l'entreprise.

# **CONCLUSION GENERALE**

## Conclusion générale

---

### **Conclusion générale :**

Dans le cadre de notre projet de fin d'études s'intitulant « Conception et réalisation d'un système de Reporting décisionnel », pour la direction régionale de l'emploi DREM nous avons dû passer par plusieurs étapes afin de réaliser nos objectifs.

Le but de notre travail intervient en aval, i.e. dans la phase exploitation des entrepôts de données par des applications d'analyse et de Reporting. Cependant, ces applications pourront intervenir et accéder à différents niveaux pour restituer les informations/connaissances : entrepôt de données, magasin de données (data mart), cube de données, etc...

Afin d'atteindre nos objectifs nous sommes passés par les étapes suivantes :

Dans un premier temps nous avons fait un descriptif de l'état de l'art sur tout ce qu'est décisionnel, en effet, c'est l'axe sur lequel est basé notre travail.

Dans la deuxième partie nous avons présenté l'organisme d'accueil, récupéré toutes les informations nécessaires pour définir les besoins principaux de cet organisme passant par plusieurs méthodes de recueils. Sachant que nous avons consacré pour cela un temps considérable reflétant l'importance de cette partie principale qui est une introduction à la partie conception.

Dans cette dernière, qui est le cœur de notre projet, nous avons modélisé le système opérationnel et décisionnel.

La dernière phase de ce travail est la mise en œuvre du logiciel sous la plateforme windows basé sur Microsoft SQL Server comme serveur de base de données, le service SSIS pour l'implémentation de l'outil d'alimentation du système et Microsoft Visual basic et ASP.net pour le système opérationnel.

Notre travail nous a permis d'enrichir nos connaissances dans l'informatique décisionnelle ainsi que dans le monde de l'emploi en Algérie, à savoir :

## Conclusion générale

---

- Recherche profonde dans le domaine du Data Warehousing et l'étude des solutions d'ETL et de Reporting.
- Connaissance du monde de l'emploi (politiques préconiser, l'importance qu'occupe l'ANEM quotidiennement dans le marché de l'emploi, les obstacles ...etc.).

Comme tout travail, nous avons rencontré quelques problèmes au niveau de :

- La définition des besoins, puisque les utilisateurs ont eu des difficultés pour cerner leurs besoins et les exprimer de manière claire ;
- La collecte des indicateurs et axes d'analyse, faute de culture dans le domaine des DataWarehouses au sein de nos entreprises, les gens ont toujours les réflexes des systèmes opérationnels ;
- L'apprentissage d'une nouvelle méthode et l'utilisation de nouveaux outils, vu le changement radical du raisonnement vers le décisionnel, nous avons pris un temps que nous aurions pu consacrer à la mise en œuvre.

# **BIBLIOGRAPHIE**

## Bibliographie

---

### **Bibliographie :**

- [1] : Haciane ahmed, Conception d'un datawarehouse CRM orienté objet, mémoire magister en informatique ,2007.
- [3] : Goglin et Jean-François, La construction du DataWarehouse, HERMES, 2001.
- [4] : Mohamed taslimanka sylla, Initiation au décisionnel (Business Intelligence, DataWarehouse, OLAP), www.developpez.com, 2007
- [6] : Inès Gam El Golli, Concepts, Modèles et Processus La méthode CADWE, mémoire Ingénieure en informatique, Octobre 2008.
- [7] : Josiane Mothe, Claude Chrisment et Joel Alaux, Visualisation globale de collections de documents sous forme d'hypercube, Hermès, 2002.
- [8] : Jean-Marie Gouarne, Le projet Décisionnel : Enjeux, modèles et Architecture du DataWarehouse, Eyrolles, 1999.
- [9] : R. Kimball, L. Reeves, M. Ross, W. Thornwaite, Concevoir et déployer un datawarehouse : guide de conduite de projet, Eyrolles; 2000.
- [10] : J.M. Franco, S. De Lignerolles, Piloter l'entreprise grâce au datawarehouse, Eyrolles; 2000.
- [11] : Ralph Kimball. Entrepôt de données. International Thomson Publishing, 1997.
- [12] : Olivier Teste, Modélisation et manipulation d'entrepôt de données complexes et historisés. Thèse de doctorat, 2000.
- [13] : Ralph Kimball. The data Warehouse toolkit, Jhone Wiley, 1996.
- [14] : Ralph Kimball et Margy Ross. Entrepôt de données Guide pratique de modélisation dimensionnelle. Vuibert Informatique, 2003
- [15]: Erik Thomsen , OLAP Solutions Building Multidimensional Information Systems, Jhone Wiley , 2002

## Bibliographie

---

- [16] : F. Ravat, O. Teste, G. Zurfluh. Algèbre OLAP et langage graphique Toulouse IRIT, 2006
- [17] : Mahfoud Bala, Modèles et Techniques utilisés dans le processus de développement des Datawarehouses, Cours Décisionnel, USDB, 2009
- [21] : livre : UML 2 pratique de la modélisation (<http://www.eyrolles.com/Informatique/Livre/uml-2-pratique-de-la-modelisation-9782744072871>)
- [22] : Islem El Felli Plate-Forme de Génération d'applications Web, mémoire Ingénieure en informatique, 2005
- [25] : ADEL mehdi, SAFER imène, Réalisation d'un tableau de bord pour la navigation aérienne, mémoire ingénieure en informatique, 2009.
- [27] : Sébastien FANTINI , Business intelligence avec SQL server 2008 R2 , ENI , 2010

## Webographie :

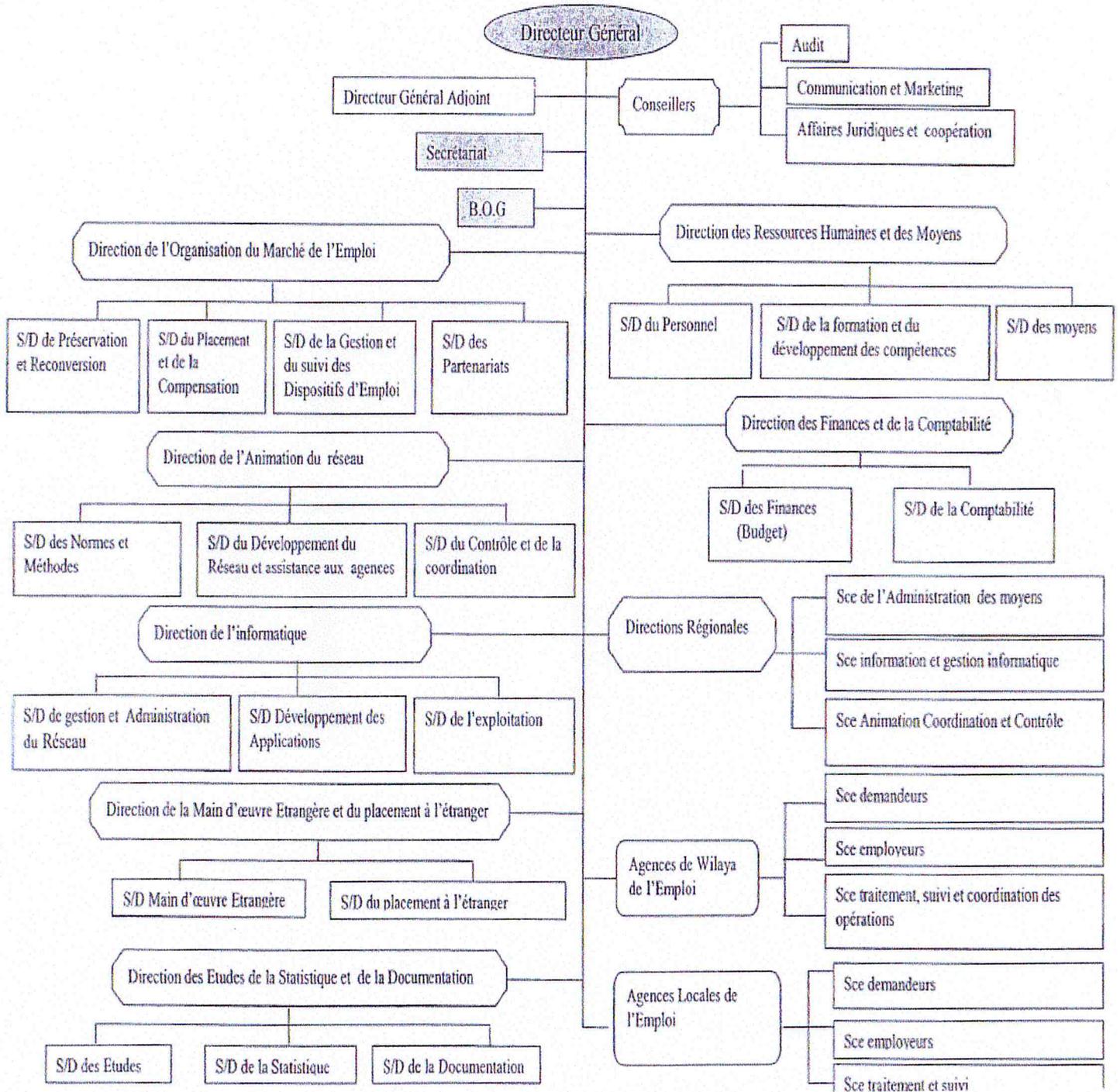
- [2] : [http://www.crawdesign.com/business\\_intelligence/presentation.html](http://www.crawdesign.com/business_intelligence/presentation.html)
- [5]:<http://grim.developpez.com/cours/businessintelligence/concepts/conceptiondatawarehouse/#LIIA>
- [18] : <http://www.piloter.org/business-intelligence/datamining.htm>
- [19] : <http://www.hackersafe.nl/info/?groep=fr&pageid=1000000226>
- [20] :<http://grim.developpez.com/articles/concepts/bi-intro/>
- [23] : [www.uml.free.fr/cours/i-p21.html](http://www.uml.free.fr/cours/i-p21.html)
- [24] : <http://grim.developpez.com/articles/concepts/bi-intro/>
- [26] : <http://www.microsoft.com/france/serveur/sql/decisionnel.aspx>

# ANNEXE

# Annexe

## ANNEXE A : L'organigramme général de l'ANEM.

### ORGANIGRAMME DE L'ANEM



# Annexe

**ANNEXE B** : un tableau vide à remplir par les agents.

**Tableau n.5:**

Demandeurs d'emplois enregistrés issus de la formation professionnelle et équivalent, par diplômes, par groupes de métiers et

Groupes de Métiers	Diplômes		CFPS		CAP		CMP		BT		BTS		Autres		Sous-Total		TOTAL	Primo dema		
	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F		M	F	
1-Agriculture															0	0	0			
2-Mine et Géologie															0	0	0			
3-Mécanique et Construction Métallique															0	0	0			
4- Production, Transformation Métaux															0	0	0			
5-Electricité, Electronique, Télécom															0	0	0			
6-Industrie de la															0	0	0			
7-Industrie, Alimentaires et Tabac															0	0	0			
8-Textiles et Confection															0	0	0			
9-Métiers de Cuir et de Peau															0	0	0			
10-Métiers du Bois et d'Ameublement															0	0	0			
11-Métiers: Industrie, Graphique-Photos															0	0	0			
12-Métiers: De type Commercial, Touristique.															0	0	0			
13-Transport et Communication															0	0	0			
14-B.T.P Génie Civil et Hydraulique															0	0	0			
15-Métiers: de type Economie, Administration, juristes.															0	0	0			
16-Culture, Informatique, Enseignement, Recherche scientifique, Information															0	0	0			
17-Métiers du Verre et la Poterie															0	0	0			
18-Métiers de Type Médical															0	0	0			
19-Services de Protection et Sécurité															0	0	0			
20-Autres Métier. (Manœuvre, Femme de ménage, Agent de réception...)															0	0	0			
<b>Sous-total</b>			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0
<b>Total</b>			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0

**CFPS:** Certificat de Formation Professionnel Spécialisé.

**CAP:** Certificat d'Aptitude Professionnelle.

**CMP:** Certificat de Maîtrise Professionnelle.

**BT:** Brevet Technicien.

**BTS:** Brevet Technicien Supérieur.