

11A.2017 08 1

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

Université Saâd DAHLAB, Blida 1



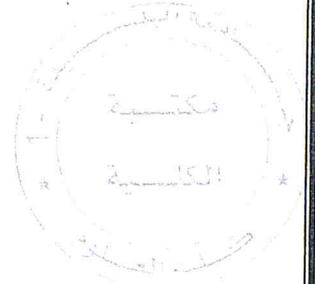
Faculté des sciences

Département d'Informatique

Mémoire de fin d'études

Pour l'obtention du diplôme de Master en informatique

Option : Ingénierie de logiciel



Thème :

**Vers un modèle dynamique d'aide à la décision
basé sur un entrepôt de données
pour la gestion de compétences dans la Gendarmerie Nationale**

Réalisé par :

BOUILOUT Fouad

KOUIDER EL OUAHED Abdellah

Président : *Dr Boumahdi*

Examinateur : *Mme Djekid.L.*

Promotion : 2016/2017

Encadreur :

Dr. M. BALA

Co-Encadreur :

Dr. A. BOUKORÇA

Mr. A. MEGHAOUI

MA-004-508-1

Dédicace

Alhamdou lillah d'avoir finalisé ce modeste travail que je dédie à tous ceux qui me sont chers en reconnaissance du soutien qu'ils m'ont apporté, particulièrement :

À mes très chers parents, sans leurs prières et conseils je n'aurais jamais pu progresser, ni devenu ce que je suis. Puissent-ils trouver en cet humble travail toute la gratitude et la reconnaissance d'un fils dévoué.

À ma femme et à mes enfants adorables :

Nouha qu'Allah tout-puissant la guérisse et Mouhammed El Amine qu'Allah le protège.

À mes chères sœurs et à ma belle-famille pour leur encouragement et sympathie.

À tous mes collègues et amis du corps et à tous les militaires.

À la promotion IL 2016/2017 de l'université Saâd DAHLAB.

Fouad

Remerciements

Nous présentons notre sincère reconnaissance au haut commandement de la Gendarmerie Nationale pour nous avoir donné cette opportunité de poursuivre nos études en cycle de Master II en Informatique.

Nous tenons à remercier Mr le Directeur de la Télématicque et Mr le Directeur de la Resource Humaine ainsi que le Chef du Service Informatique de la Gendarmerie Nationale pour leur assistance et soutien permanent.

Nos remerciements s'adressent aussi à :

- Notre enseignant et promoteur Mr BALA Mahfoud pour avoir assuré le suivi de ce projet et aussi pour ses conseils et sa disponibilité.
- Tous les enseignants du département informatique de la faculté des sciences de l'université de BLIDA.
- Nos collègues de travail qui nous ont aidé pour l'accomplissement de notre projet de fin d'étude : Meghaoui Ali, Ghezal Abdesslem, Zabich Khalil, Boukorça Ahcène, Boudraa Abdelmajid et Chaïb Ainou Ammar. Nos collègues de la DRH : Laiouar Réda, Djeghader Hichem, Bounouala Fethi, Korti Chakib, Bayadha Karim et Guasmi Amer. Ainsi que nos collègues de la DSPE : Bouchaghchough Rabeh et Khemissi H.
- Nos parents, nos familles et nos amis pour leur soutien moral tout au long du cursus.
- Toutes ces personnes sur internet qui ont partagé leurs connaissances et expériences et répondu à nos questions, ils nous ont été d'une grande aide.

Enfin, nous remercions tous ceux qui ont aidé de près ou de loin à l'élaboration de ce mémoire et à la réussite de ce projet.

ملخص

من أجل تحسين السيطرة على سلطة اتخاذ القرار، تولي المؤسسات والشركات اهتماما خاصا لجمع البيانات وتحليلها من مصادر متعددة ومتنوعة وغير متجانسة. يتم تخزين هذه البيانات في مستودعات البيانات مع النمذجة الخاصة باحتياجات صنع القرار. يهدف نظام صنع القرار، الذي يطبق على مجال الموارد البشرية، الذي يدير مفتاح نجاح أي منظمة، إلى حل عدد من المشاكل المعقدة التي تسمى إدارة الكفاءات. والمشاكل الرئيسية المتوخاة هي إدارة الحياة الوظيفية، وتعيين الشخص المناسب في الوظيفة المناسبة، ورصد تدريب الموظفين وتقييمهم، وما إلى ذلك.

والدرك الوطني مؤسسة تعنى بالأمن العمومي، حيث يرتبط نجاح مهامها ارتباطا قويا بنوعية موظفيها. وبالنظر إلى عدد الموظفين المعنيين، وتنوع المهام وامتداد الإقليم الوطني، تصبح إدارة الكفاءات في الدرك الوطني عنصرا أساسيا. والجانب الآخر الذي يعقد التحكم في الكفاءات هو التغيير المتكرر في سياق الأمن العام (الجيوستراتيجية، الأوضاع الأمنية في البلدان المجاورة، الوسائل والأساليب المستخدمة في الجريمة، تنوع مناطق البلد، وما إلى ذلك).

في هذه المذكرة، نقترح نموذجا ديناميكيا لدعم اتخاذ القرار في إدارة الكفاءات المستند إلى مستودع البيانات لأقصى قدر من المطابقة منصب / مرشحين من خلال إدخال مفهوم السياق لنموذج التغيير في المؤشرات.

تم وضع نموذجنا المقترح على نظام الدرك الوطني باستخدام أدوات أوراكل.

كلمات مفتاحية: الأنظمة المساعدة في اتخاذ القرار، إدارة الكفاءات، التصميم بالأبعاد، مستودع البيانات، السياق، الترجيح.

Résumé

Dans le but de maîtriser davantage leur pouvoir décisionnel, les institutions et les entreprises donnent un intérêt particulier à la collecte et l'analyse des données provenant de multiples sources variées et hétérogènes. Ces données sont entreposées dans des entrepôts de données avec des modélisations spécifiques aux besoins décisionnels. Appliqué au domaine de la ressource humaine, qui gère l'élément clé de la réussite de n'importe quelle organisation, le système décisionnel vise à résoudre plusieurs problèmes complexes, appelés gestion des compétences. Les principaux problèmes envisagés sont la gestion de la carrière, l'affectation de la bonne personne au bon poste, le suivi de la formation et de l'évaluation des personnels, etc.

La Gendarmerie Nationale est une institution de sécurité publique, où la réussite de ses missions est fortement liée à la qualité de son personnel. Vu le nombre de personnels impliqués, la diversité des missions et l'étendue du territoire national, la gestion des compétences au sein de la GN devient un élément primordiale. Un autre aspect qui complique la maîtrise des compétences est le changement fréquent du contexte lié à la sécurité publique (géopolitique, moyens et modes opératoires empruntés dans la criminalité, état sécuritaire des pays voisins, diversité des régions du pays, etc.).

Dans ce mémoire, nous proposons un modèle dynamique d'aide à la décision pour la Gestion des Compétences basé sur un entrepôt de données en vue d'une adéquation maximale poste / candidat, en introduisant le concept de contexte pour modéliser le changement dans les indicateurs.

Notre modèle proposé a été concrétisé sur le système de la GN, en utilisant des outils d'oracle.

Mots clés : Système d'aide à la décision, Gestion des compétences, Modélisation dimensionnelle, Entrepôt de Données, contexte, Pondération.

Abstract

In order to better control their decision-making power, institutions and companies pay particular attention to the collection and analysis of data from multiple, diverse and heterogeneous sources. These data are stored in data warehouses with modeling specific to the decision-making needs. Applied to the human resource domain, which manages the key to the success of any organization, the decision-making system aims to solve a number of complex problems, called competencies management. The main problems envisioned are career management, assignment of the right person to the right job, monitoring of training and evaluation of personnel, etc.

The National Gendarmerie is a public security institution, where the success of its missions is strongly linked to the quality of its personnel. Given the number of staff involved, the diversity of missions and the extent of the national territory, the management of competencies within the National Gendarmerie becomes an essential element. Another aspect that complicates competence management is the frequent change in the context of public security (geopolitics, means and methods used in crime, security situations in neighboring countries, diversity of regions of the country, etc.).

In this paper, we propose a dynamic decision support model for Data Warehouse-Based Competency Management for maximum position / candidate matching by introducing the concept of context to model change in indicators.

Our proposed model was concretized on the National Gendarmerie system, using oracle tools.

Key words: Decision-making system, Competency management, Dimensional modelization, Data warehouse, Context, Weighting.

Contenu

INTRODUCTION GENERALE.....	13
Contexte.....	13
Problématique.....	14
Objectifs	16
Contenu du mémoire	16
Partie 1 : ETAT DE L'ART	18
Chapitre I L'INFORMATIQUE DECISIONNELLE	19
I.1. Introduction	19
I.2. Définition.....	19
I.2.1. Décisionnel vs Transactionnel.....	20
I.1. Objectifs du Business Intelligence (BI).....	21
I.2. Processus du BI	21
I.3. Architecture d'un système décisionnel.....	22
I.4. Étapes du processus décisionnel.....	22
I.4.1. La phase de collecte.....	22
I.4.2. La phase d'intégration	23
I.4.3. La phase de structuration	24
I.4.4. La phase de restitution	24
I.4.5. Fonctionnalités transverses.....	26
I.5. La BI et le domaine de la Ressource humaine (RH)	26
I.5.1. Indicateurs de performance utiles aux RH.....	26
I.5.2. Objectifs du BI-RH.....	27
I.5.3. Le reporting dans la RH.....	27
I.6. Conclusion	28
Chapitre II LES ENTREPÔTS DE DONNÉES	29
II.1. Introduction	29
II.2. Définitions et caractéristiques.....	29
II.2.1. Entrepôt de données (ED) ou Data Warehouse (DW):.....	30
II.2.2. Magasins de données ou Data Marts (DM)	30
II.2.3. Choix d'un type de table.....	31
II.2.1. Granularité	31
II.2.2. Agrégats.....	32

II.3.	Structure des données d'un Data Warehouse	32
II.4.	Les éléments d'un Data Warehouse	33
II.5.	Cycle de conception d'un Data Warehouse.....	33
II.5.1.	Planification.....	33
II.5.2.	Conception et implémentation.....	33
II.5.3.	Maintenance et évolution.....	44
II.6.	Conclusion.....	44
Chapitre III	Introduction au Clustering	45
III.1.	Introduction	45
III.2.	Domaines d'application.....	45
III.3.	Mesures de similarité.....	46
III.3.1.	Mesures de similarité entre objets	46
III.3.2.	Mesures de similarité entre cluster	46
III.4.	Algorithmes de clustering.....	47
III.4.1.	Algorithmes de clustering par partitionnement	47
III.4.2.	Algorithmes hiérarchiques.....	48
III.4.1.	Conclusion.....	48
Partie 2	CONTRIBUTION	49
Chapitre IV	ÉTUDE D'EXISTANT ET BESOINS.....	50
IV.1.	Organisme d'accueil.....	50
IV.2.	Etat des lieux de la ressource humaine dans la GN.....	52
IV.2.1.	Système d'Information Ressource Humaine SI-RH.....	53
IV.2.2.	Rubriques de l'application de saisie / consultation des personnels	54
IV.3.	Etat des lieux de la Formation	55
IV.3.1.	Système d'Information de la Formation SI-F	55
IV.4.	Etat des lieux du métier : Police Judiciaire (PJ)	56
IV.4.1.	Système d'Information de la Police Judiciaire SI-PJ.....	56
IV.5.	Expression des besoins	61
IV.5.1.	Description de la démarche d'étude des besoins.....	61
IV.5.2.	Etude de cas.....	62
IV.6.	Conclusion.....	63
Chapitre V	MODELISATION DE LA SOLUTION	64
V.1.	Introduction :	64
V.2.	Compétence et évaluation.....	65

V.3.	L'approche du modèle proposé :	66
V.3.1	Motivation d'évaluation des compétences.....	67
V.3.2	Architecture du modèle proposé.....	68
V.3.3	Processus	68
V.4.	Conception du Data Warehouse	69
V.4.1	Processus de la modélisation dimensionnelle.....	69
V.4.2	Volet : Evaluation explicite du candidat :.....	70
V.4.3	Volet : Quantification de l'activité du candidat.....	78
V.4.4	Le modèle en constellation de l'évaluation des candidats.....	83
V.5.	Conception de la zone d'alimentation	83
V.5.1.	Processus de chargement	84
V.6.	Contraintes du modèle.....	85
V.7.	Modélisation du système de contexte.....	85
V.8.	Conclusion.....	93
Chapitre VI	Réalisation & Déploiement.....	94
VI.1.	Environnement de développement	94
VI.2.	Architecture technique de la solution	95
VI.3.	Les outils technologiques choisies.....	95
VI.3.1.	Oracle Data Base 10g r2.....	95
VI.3.2.	Oracle Warehouse Builder.....	95
VI.3.3.	Oracle Business Intelligence.....	97
VI.4.	Implémentation des DataMarts.....	99
VI.5.	Déroulement de l'étude de cas.....	109
VI.6.	Implémentation du BI.....	115
IV.5.1.	Création de référentiel	115
IV.5.2.	Le chargement de référentiel dans le BI	118
VI.7.	Conclusion.....	119
	CONCLUSION GENERALE	120
	REFERENCES BIBLIOGRAPHIQUE	122
	Annexe A Installation & Configuration du OWB (Oracle Warehouse Builder 11gR2)	127
	Annexe B Installation & Configuration du OBIEE	130
	Annexe C Bilan PJ. 1^{er} Semestre 2015 : « Lutte contre le crime et le crime organisé »	133

Liste des figures

Figure I-1 Le décisionnel au sein du Système d'information et les différentes composantes du décisionnel [6] ..20	20
Figure I-2 Processus de transformation de la donnée en action [7]	21
Figure I-3 Processus global de traitement des données d'un système décisionnel	22
Figure I-4 Processus détaillé de traitement des données d'un système décisionnel.....	22
Figure I-5 Processus détaillé de traitement de données d'un système décisionnel [11]	24
Figure II-1 Architecture globale d'un DW [14].....	31
Figure II-2 Structure des données d'un DW [12].....	32
Figure II-3 L'approche Top-Down selon le cycle de vie dimensionnel de Kimball [27]	35
Figure II-4 L'approche Bottom-up selon le cycle de développement du DW de Inmon, [18].....	36
Figure II-5 Illustration de l'approche hybride.....	36
Figure II-6 Modèle en étoile [28].....	37
Figure II-7 Modèle en flocon [28]	37
Figure II-8 Modèle en constellation [28]	37
Figure II-9 Principe de l'architecture MOLAP [31]	39
Figure II-10 Considération d'un sujet d'analyse comme un cube à plusieurs dimensions [32].....	39
Figure II-11 Principe de l'architecture ROLAP [31]	40
Figure II-12 Principe des trois architectures OLAP [33].....	40
Figure II-13 Exemple de Slice et de Dice	42
Figure II-14 Exemple de Roll up [35].....	42
Figure II-15 Exemple de Drill-Down [35].....	43
Figure IV-1 portée de l'étude par rapport à l'Organigramme de la GN.....	51
Figure IV-2 Quelques aspects du système de production RH.....	54
Figure IV-3 Quelques aspects du système de production Formation.....	55
Figure IV-4 Classification de l'infraction.....	57
Figure IV-5 Modélisation de l'Incident dans le SI-PJ	57
Figure IV-6 Portail du SI-PJ	59
Figure VI-7 page web de la recherche multicritères des affaires de Police Judiciaire.....	60
Figure IV-8 Page web de consultation de l'affaire avec aperçu du personnel en charge.....	60
Figure IV-9 Page web de consultation des détails relatifs aux crimes liés à l'affaire.....	60
Figure IV-10 Page web des statistiques sur les affaires	61
Figure V-1 Schéma global du système de gestion des compétences	64
Figure V-2 Schéma global de combinaison des pondérations aux évaluations des candidats	65
Figure V-3 Architecture du modèle proposé.....	68
Figure V-4 Processus de l'approche proposée.....	68
Figure V-5 Analyse des priorités du cas « Gestion des compétences »	70
Figure V-6 La Dimension Temps du volet « Evaluation directe»	71
Figure V-7 Aperçu de la liste des 22 compétences recensées.....	73
Figure V-8 La Dimension Epreuve du volet «Evaluation directe»	74

Figure V-9 Le modèle en étoile d'évaluation explicite des compétences	76
Figure V-10 La dimension Infraction du volet « Evaluation par l'activité»	79
Figure V-11 La dimension Lieu de service « Evaluation par l'activité»	79
Figure V-12 Schéma décisionnel de l'activité par candidat, lieu de service, infraction et date.....	81
Figure V-13 Schéma décisionnel de l'évaluation globale des candidats	83
Figure V-14 Assemblage de critères par fiche de poste.....	86
Figure V-15 Modèle décisionnel étendue avec le modèle du contexte.....	91
Figure V-16 Modèle décisionnel étendue avec le modèle conceptuel du contexte.....	92
Figure V-17 Schéma du modèle de pondération et référentiel de fiches de postes (contextes).....	92
Figure VI-1 Architecture technique de la solution.....	95
Figure VI-2 L'interface d'OWB	96
Figure VI-3 L'interaction des principaux composants du logiciel OWB [63]	97
Figure VI-4 Interface de BI Publisher	98
Figure VI-5 Interface de BI Answers.....	98
Figure VI-6 BI Dashboard	98
Figure VI-7 L'ajout d'une nouvelle dimension	99
Figure VI-8 Spécification des attributs de la dimension.....	100
Figure VI-9 Spécification des niveaux	100
Figure VI-10 Spécification des attributs des niveaux	101
Figure VI-11 L'ajout d'un nouveau mapping.....	101
Figure VI-12 La sélection des tables sources.....	102
Figure VI-13 Regroupement des données avec l'opérateur JOINER	102
Figure VI-14 Spécification de la condition de jointure.....	103
Figure VI-15 Liaison entre JOINER et « DIM_INFRACTION »	103
Figure VI-16 Schéma final du mapping INFRACTION.....	104
Figure VI-17 Schéma final du mapping de la dimension Date.....	104
Figure VI-18 Déploiement et démarrage du « MAP_INFRACTION ».....	104
Figure VI-19 Les données de la dimension « INFRACTION »	105
Figure VI-20 Mapping de cube.....	106
Figure VI-21 Condition de la jointure du cube	106
Figure VI-22 L'agrégation du cube	106
Figure VI-23 Schéma final du cube.....	107
Figure VI-24 Déploiement des cubes	107
Figure VI-25 Affichage des données du datamart	108
Figure VI-26 Affichage des données du datamart global	108
Figure VI-27 La couche physique.....	115
Figure VI-28 Indication du type d'agrégation	116
Figure VI-29 Création d'une dimension logique	116
Figure VI-30 Création d'un niveau.....	117
Figure VI-31 L'introduction des informations sur le niveau	117

Figure VI-32 Dimension logique infraction finale.....	117
Figure VI-33 Administration final	118
Figure VI-34 Le montage du référentiel	118
Figure VI-35 Le redémarrage des services du BI	119

Liste des tableaux

Tableau I-1 Comparatif entre les systèmes transactionnels et les systèmes décisionnels	20
Tableau II-1 Tableau comparatif entre les tables de faits et les tables de faits et les tables de dimensions [19] ...	31
Tableau II-2 Avantages et inconvénients de l'approche « Besoins d'analyse » [21]	35
Tableau II-3 Avantages et inconvénients de l'approche « sources de données » [18].....	36
Tableau V-1 Tableau descriptif de la dimension « Temps ».....	71
Tableau V-2 Tableau descriptif de la dimension « Candidat »	73
Tableau V-3 Aperçu des domaines (non ordonnés) inhérents à une compétence donnée	74
Tableau V-4 Illustration d'indicateurs pour l'évaluation sur épreuve par rapport aux niveaux de difficulté	75
Tableau V-5 Liste des agrégats potentiels pour le volet «Evaluation explicite»	77
Tableau V-6 Descriptif des agrégats utiles du volet «Evaluation explicite».....	77
Tableau V-7 Descriptif de la dimension « TD_Infraction»	79
Tableau V-8 Descriptif de la dimension « TD_LieuService ».....	80
Tableau V-9 Liste des agrégats potentiels pour le volet «évaluation par activité»	82
Tableau V-10 Tableau descriptif des agrégats utiles du volet «évaluation par activité»	82
Tableau V-11 Evaluation des candidats pour les m domaines durant une période de service	87
Tableau V-12 L'évaluation des candidats pour les épreuves couvrant une période	87
Tableau V-13 Quantification de l'activité des candidats par type d'infraction.....	88

I NTRODUCTION G ENERALE

Contexte

Face au développement alarmant de la criminalité dans toutes ses formes, les services de sécurité sont appelés, plus que jamais, à reconsidérer leurs pratiques et leurs procédures. En outre, la prise de décision en rapport avec le potentiel humain pour les organes de sécurité est de plus en plus critique notamment au vu de la diversité du territoire national et de la fréquence élevée des mutations du personnel.

Désormais, il ne s'agit plus d'examiner des postes et d'y rattacher du personnel, mais plutôt d'identifier et de traiter les compétences propres des employés, quitte à les rapprocher d'un métier ou d'une spécialité.

En effet, on ne raisonne plus en termes de qualification de poste, classique et cloisonnée, mais en termes de qualification de l'individu. Les deux concepts se différencient :

« La qualification du poste est déduite des opérations requises par les postes de travail et est mesurée par des outils d'analyse spécialisés qui permettent, à l'aide de critères, d'évaluer et de quantifier le poste [1].

La qualification des personnes, quant à elle, fait référence aux qualités acquises par l'individu soit au travers de la formation et du diplôme, soit par l'expérience professionnelle»

A cet effet, les directions des ressources humaines doivent procurer au commandement un aperçu à la fois synthétique et raisonnablement détaillée de la situation de la ressource humaine de l'institution.

Pour cela, il y a lieu d'utiliser les systèmes décisionnels basés sur des entrepôts de données dotés d'une panoplie adaptée d'outils d'extraction et d'analyse de données. Ces entrepôts de données permettent particulièrement l'intégration des informations émanant de sources multiples et hétérogènes afin de les exploiter via des couches de représentation.

En s'inscrivant dans cette optique, la Gendarmerie Nationale(GN) nous a confié, dans le cadre de notre projet de fin d'étude de Master en Ingénierie de Logiciels, de proposer un modèle d'aide à la décision pour la Gestion des Compétences basé sur un entrepôt de données en vue d'une adéquation maximale poste/candidat.

La GN est un service de sécurité publique au service des citoyens, chargé de missions diverses à caractère opérationnel, technico-logistique et administratif. Chacun des trois volets est régi par un ensemble d'applications disparates formant un sous-système plus ou moins intégré persévérant des informations concernant, entres autres, le facteur humain de l'institution. Dans le cadre des missions administratives, la GN veille sur la gestion de sa ressource humaine en visant à exploiter toutes les données et les informations inhérentes à la bonne gouvernance de ses ressources et de son capital humain.

Par ailleurs, il faut préciser qu'avoir les systèmes d'information les plus comblés en termes de données sur son personnel (évaluations périodiques et activité quotidienne) n'est pas, en lui-même, le but à convoiter, il faut surtout pouvoir les analyser afin d'en restituer les éléments essentiels pour une bonne prise de décision.

Problématique

Dans le domaine de la ressource humaine, on observe toujours un recours très restreint aux solutions informatiques pour intégrer des données disparates, en effectuer les analyses nécessaires et tirer les éléments de réponse essentiels afin de se prononcer sur la disposition, l'aptitude et la primauté d'une compétence pour un poste donné.

D'autre part, comme stipulé dans le contexte, une disjonction et une hétérogénéité caractérisent les sources de données qui, dans l'absence d'une couche décisionnelle, rendent leur exploitation plus difficile.

En outre, la spécificité classique a voulu que les compétences soient liées à un diplôme (**savoir**), qui va suivre l'employé tout au long de sa carrière, or le diplôme ne renseigne que sur le potentiel cognitif de l'individu et ne met en jeu que les connaissances et savoirs fonctionnels nécessaires pour accéder à un rôle et conserver son aptitude à l'assurer. Où sont alors passées les qualités comportementales (**savoir-être**), la personnalité et la capacité d'apprentissage. Une compétence doit concerner la mise en œuvre intégrée d'aptitudes, de traits de personnalité et aussi de connaissances acquises pour mener à bien une mission

complexe dans le cadre de l'institution et dans l'esprit de sa stratégie. Il ne s'agit donc pas d'une logique de diplôme, mais une logique de **savoir-faire** opérationnel dans des situations concrètes. La description d'un poste au sein de l'organisation ne se penche pas sur une analyse des tâches à réaliser, mais sur les compétences requises pour y accéder.

En réalité, le besoin d'apporter une solution décisionnelle pour la Gestion des compétences au sein de la GN est provoqué par maintes difficultés, entre autres :

- Le besoin de gérer un effectif, en nombre croissant, catégorisé en tranches de grades et avec des profils différents à distinguer selon le besoin ;
- Le besoin de prendre en charge les aspects : humain, intellectuel et métier pour une éventuelle spécialisation ce qui nécessite de réunir les référentiels de critères objets d'appréciation ;
- La diversité géographique du territoire national, fait paraître une spécificité de zone, de risque, etc. pour des évaluations et des affectations plus équitables ;
- L'étendue du territoire national multipliée par un déploiement à grande échelle des unités de la GN augmente considérablement l'effort consenti à la gestion des compétences souhaitée ;
- l'astreinte et le caractère militaire de l'institution augmente l'effet des contraintes socio-médicales de l'individu ou des personnes à sa charge ;
- Lecture unique et linéaire des évaluations portées dans les bases de données. Il n'y a pas de possibilité d'en faire des lectures différentes sans avoir à passer par des requêtes alambiquées ;
- Une subjectivité souvent reprochée aux appréciations personnelles du fait qu'elles reposent principalement sur le facteur humain. En effet, tous les chefs ne sont pas de la même intransigeance ni ont la même moralité.
- Rechercher de la flexibilité au niveau de la gestion de production (de service), qui impose aux organisations la connaissance du savoir-faire des individus, plutôt que celle de leur niveau de poste. Les organisations peuvent ainsi augmenter leur réactivité face à un environnement de plus en plus instable et dynamique.
- La logique « poste » ne répond plus aux besoins des organisations. En effet, les métiers ne cessent d'évoluer et il est important de se construire un référentiel de compétences, qui permettra d'être en adéquation avec les attentes des décideurs et le contexte sécuritaire.

Objectifs

Dans le cadre du présent projet et afin de répondre aux problématiques susmentionnées, notre objectif, est de modéliser un système décisionnel pour le personnel basé sur un entrepôt de données, destiné à l'échelon commandement, qui doit permettre d'appliquer des contextes, d'effectuer des analyses, et de représenter les données sous forme de tableaux et de graphiques illustratifs.

Aussi, nous projetons éventuellement d'appliquer des techniques de fouille de données, afin de donner la possibilité de détecter et de faire ressortir des liens insoupçonnés dans les bases de données, et ce pour des fins de prospection et d'analyse.

Pour cela, nous envisageons de :

- Réduire le temps consacré à l'analyse.
- Permettre un accès facile et rapide à l'ensemble des données en les centralisant.
- Permettre d'établir un diagnostic spécifique aux besoins des décideurs en donnant l'accès à des indicateurs pertinents, fiables, interprétables et mesurés dans le temps.
- Offrir à l'utilisateur un système de contexte qui lui permet de créer et de personnaliser une fiche de poste cible.
- Offrir à l'utilisateur la possibilité de naviguer dans les données de manière interactive à l'aide des cubes OLAP.
- Mettre à la disposition des utilisateurs, dans des délais très courts, des états de sorties multiples et personnalisés (rapport, graphes, tableau de bord ...).
- Fournir une interface dynamique et fluide avec des chartes et des graphes pour simplifier l'analyse.

Contenu du mémoire

Après avoir introduit le contexte de notre étude et fixé la problématique ainsi que les objectifs du projet, la suite de ce mémoire est organisée en deux parties structurées de la manière suivante.

La partie I consiste en l'état de l'art. Il s'agit d'une partie théorique où nous présentons les notions relatives à notre projet dans les domaines des systèmes décisionnels et de la modélisation dimensionnelle ainsi que quelques notions sur la fouille de données (Data mining). En effet, en chapitre I nous avons décrit les étapes du processus décisionnel, qui se

compose d'un entrepôt de données alimenté depuis les sources opérationnelles en utilisant des outils d'extraction, transformation et chargement (*ETL : Extracting-Transforming-Loading*). Cet entrepôt va contenir des données résumées qui vont être exploitées à l'aide des outils de restitution. Dans le chapitre II, après l'introduction des principales notions liées à l'entreposage de données, nous avons détaillé la notion d'entrepôt de données, son architecture et son cycle de conception. Le chapitre III étant consacré aux techniques et algorithmes de classification (*clustering*) dans le domaine de la fouille de donnée.

Dans la partie II réservée à la contribution apportée, nous avons étalé en chapitre IV une présentation de l'organisme d'accueil, une étude de l'existant en termes de systèmes opérationnels puis l'expression du besoin. Dans le chapitre V nous avons abordé la partie pratique de notre projet en présentant notre conception de la solution qui consiste en un système décisionnel qui s'articule autour d'un entrepôt de données avec une couche de pondération modélisant le principe de contexte d'analyse. Enfin nous avons entamé le dernier chapitre par la présentation de l'environnement de déploiement et la description des outils utilisés puis nous avons procédé à la réalisation, la mise en œuvre ainsi que le déroulement du processus proposé.

Nous clôturons notre document avec une conclusion générale où nous synthétisons notre travail et dressons une liste de perspectives en guise d'inspirations pour d'éventuels travaux complémentaires.

Partie **1** : ETAT DE L'ART

Cette partie sera consacrée à un recueil bibliographique sur les systèmes décisionnels, les Data Warehouse, ainsi que des notions sur la fouille de données.

Chapitre I L'INFORMATIQUE DECISIONNELLE

I.1. Introduction

À l'ère de l'information, les organisations sont disposées à générer quotidiennement des données en masse. Cependant, cette génération est souvent réalisée de manière peu coordonnée. De ce fait, les décideurs ont du mal à avoir une vue synthétique de leur potentiel et de leur activité à un moment donné pour prendre les bonnes décisions. C'est là que prend part l'informatique décisionnelle également l'intelligence d'affaires (*BI : Business Intelligence*).

En effet, les décideurs de l'organisation ont besoin d'avoir une meilleure vision de leur environnement et de son évolution, ainsi, que des informations auxquelles ils peuvent se fier, en mettant en place des indicateurs « métiers » clairs et pertinents permettant la sauvegarde, l'utilisation de la mémoire de l'organisation et d'offrir aux décideurs la possibilité de se reporter à ces indicateurs pour une bonne prise de décision.

I.2. Définition

Parmi les définitions du décisionnel on trouve : « *Le Décisionnel est le processus visant à transformer les données en informations et, par l'intermédiaire d'interrogations successives, transformer ces informations en connaissances.* » [3]. La BI est « *outil d'aide à la décision, basé sur une base de données fédérant et homogénéisant les informations des différents services d'une organisation.* » [4]. Elle est aussi un ensemble d'outils, de concepts et de méthodes permettant la maîtrise de l'information et des connaissances nouvelles. Ces méthodes analysent les informations recueillies, les traitent puis les mémorisent et diffusent des informations stratégiques aux analystes et aux décideurs, pour les aider à détecter les opportunités et les menaces, et à améliorer par conséquence la prise de décision.

Pour mieux cerner la finalité des systèmes décisionnels, nous devons les mettre dans leurs contextes et rappeler ce qu'est un système d'information.

« **Le système d'information** est l'ensemble des méthodes et moyens de recueil de contrôle et de distribution des informations nécessaires à l'exercice de l'activité en tout point de l'organisation. Il a pour fonction de produire et de mémoriser les informations, de l'activité du système opérant (système opérationnel), puis de les mettre à disposition du système de décision (système de pilotage) » [5].

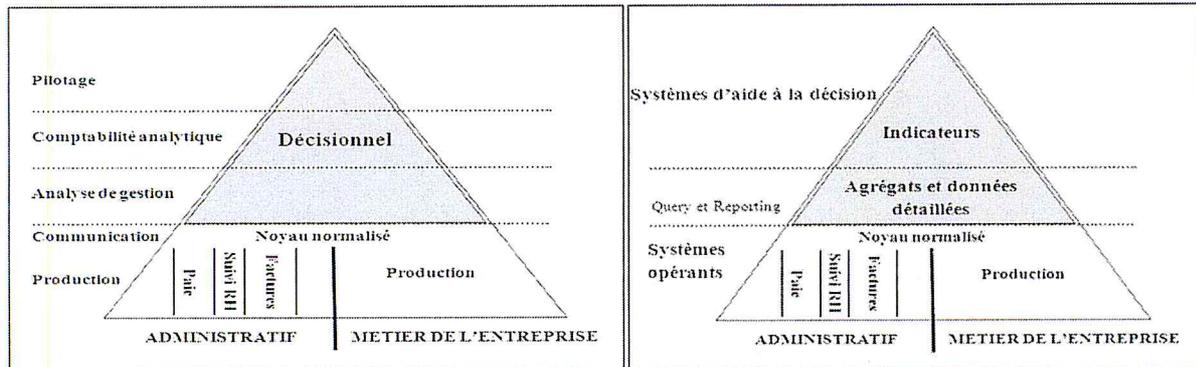


Figure 1-1 Le décisionnel au sein du Système d'information et les différentes composantes du décisionnel [6]

I.2.1. Décisionnel vs Transactionnel

Le tableau suivant dresse une comparaison non exhaustive des deux systèmes :

CRITERE	SYSTEMES TRANSACTIONNELS	SYSTEMES DECISIONNELS
Objectif	Exécution de processus métiers	Evaluation de la performance des processus métiers
Vocation	Métiers (Assure l'activité au quotidien)	Analyse et prise de décision (vue transversale)
Usagers	Les opérationnels	Les décideurs
Priorités	Performance, disponibilité	Souplesse, autonomie
Données	Orienté applications	Orienté thèmes et sujets
Situation de Données	Instantanée	Historique et transformée
Nature des données	Détaillées, codées et non redondantes	Agrégées (dérivées) cohérentes souvent redondante
Données affectées	Faibles volumes par transaction (Enregistrements unitaires)	Large volume manipulé (données en masse)
Fréquence de mise à jour	Synchronisée au processus métier (temps réel)	Synchronisée périodiquement (stables); mode batch
Volatilité des données	Elevée	Faible
Temps de réponse	Immédiat	Moins critique
Mode d'interaction	Insertion, mise à jour, suppression et sélection de données	Sélection et insertion de données
Mode d'utilisation	Stable, prédéfini et prévisible	Dynamique, non prédéfini et imprévisible
Requêtes	Complexité faible	Complexité élevée (jointure)
Normalisation de données	Haute (3 ^{ème} Forme Normale)	Forme dénormalisée

Tableau 1-1 Comparatif entre les systèmes transactionnels et les systèmes décisionnels

I.1. Objectifs du Business Intelligence (BI)

La BI rassemble des technologies de l'information pour apporter une aide à la décision aux professionnels avec des rapports et tableaux de bord de suivi des activités de l'organisation à la fois analytiques et prospectifs.

Comme tout système d'aide à la décision, les outils du BI ont pour objectifs de :

- Anticiper les changements ;
- Accroître la force de réactivité ;
- Avoir une information analysée sous différents angles ;
- Offrir une vision transversale en intégrant toutes les dimensions ;
- Intégrer les données, analyser les indicateurs et faciliter les prises de décisions.

I.2. Processus du BI

Les concepts nécessaires à la mise en place d'une solution d'aide à la décision, depuis l'extraction des données jusqu'à leur présentation sous formes agrégées et synthétisées sont:[7]

Données : nombres, mots, événements existants en dehors d'un cadre conceptuel de référence; en conséquence, et en l'absence de contexte, les données prises n'ont pas une grande signification. L'accumulation de données n'est donc pas « information ».

Information : ensemble de données, validées et confrontées, qui commencent à avoir un sens. L'accumulation d'informations n'est pas « connaissance ».

Connaissance : ensemble d'informations interprétées par l'entreprise qui lui permet de prendre des décisions. L'accumulation de connaissance n'est pas sagesse « intelligence ».

Intelligence : elle apparaît lorsque les principes fondamentaux qui sont à l'origine de la connaissance sont compris. L'accumulation de sagesse « intelligence » n'est pas vérité.

Ces notions entrent en action dans le cycle de l'intelligence économique.

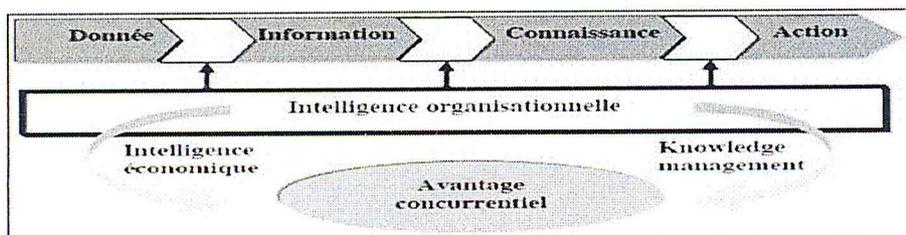


Figure I-2 Processus de transformation de la donnée en action [7]

I.3. Architecture d'un système décisionnel

L'architecture d'un système décisionnel se compose de trois composantes principales :

- 1) **Les sources de données** : Souvent hétérogènes, elles constituent les informations à intégrer (BD, fichiers plats ou XML, sources externes à l'organisation, etc.) ;
- 2) **La structuration** : C'est l'extraction des données, leur nettoyage et leur chargement dans une structure centrale ou dans des magasins de données;
- 3) **L'exploitation** : ou interprétation des requêtes et passage des résultats aux outils d'aide à la décision et aux outils de formatage (tableaux, graphiques, statistiques,...).

I.4. Étapes du processus décisionnel

Pour profiter pleinement et efficacement d'une plateforme BI, il est nécessaire de mettre en place les quatre (04) étapes suivantes dans une chaîne décisionnelle: [8]

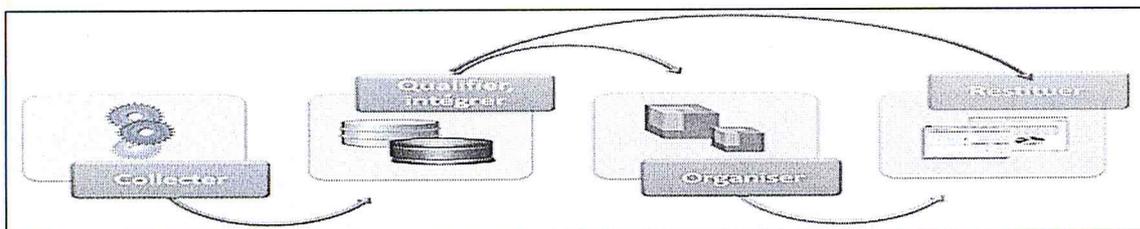


Figure I-3 Processus global de traitement des données d'un système décisionnel

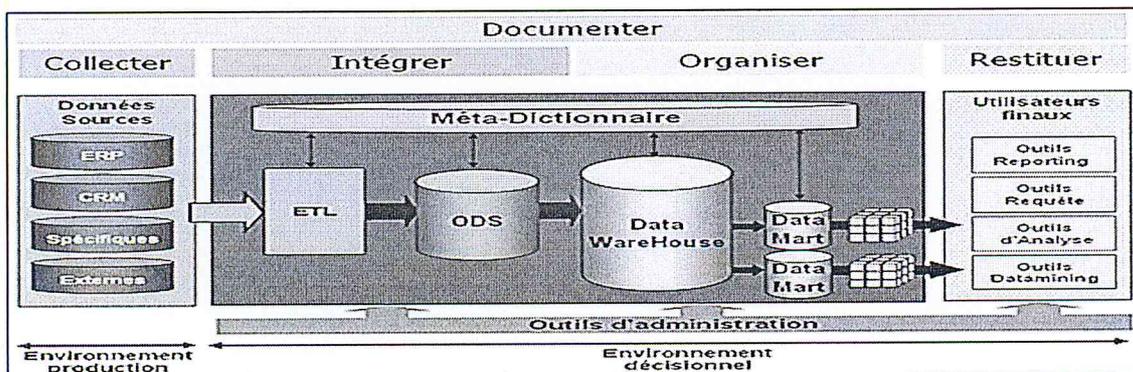


Figure I-4 Processus détaillé de traitement des données d'un système décisionnel

I.4.1. La phase de collecte

Elle s'effectue à partir de données **sources** qui, généralement **hétérogènes**, peuvent se présenter sous différents formats de fichiers (plats, CSV, XML, ASCII, etc), mais aussi de systèmes de bases de données (MySQL, PostgreSQL, DB2, ORACLE...).

I.4.2. La phase d'intégration

C'est à ce niveau qu'apparaît la première couche logicielle de l'environnement décisionnel à savoir l'ETL. Très coûteuse en termes de temps [9] et de ressources [10], cette couche offre des fonctions d'extraction de données issues de différents systèmes, de transformation de ces données (homogénéisation, filtrage, calcul) et de leur chargement dans l'ED pour une exploitation immédiate sans recalculs. Il faut signaler que le passage des données par l'ODS¹ (figure I-5) n'est qu'éphémère, garantissant la délocalisation de la charge de calcul et une meilleure disponibilité des sources. En effet, les données y vont être manipulées, transformées, traitées et modifiées plusieurs fois avant d'être copiées dans le DW².

Plusieurs outils ETL ont été proposés, que ce soit dans le monde industriel comme Microsoft Data Transformation Services, Oracle Warehouse Builder, IBM Data Warehouse Center, Informatica Power Center, Talend open studio et Pentaho Data Integration ou bien dans le monde académique comme AJAX, ARKTOS, HUMER et Potter's Wheel.

I.4.2.1. Extraction des données

Les données seront propagées temporairement vers la DSA³ pour des manipulations.

- L'identification des données pertinentes se fera souvent à un moment ultérieur.
- Certaines transformations peuvent avoir lieu au cours de l'extraction.
- Les données extraites varient de centaines de kilo-octets aux giga-octets.
- Le temps entre extractions varie de jours, heures et minutes jusqu'à temps quasi-réel.
- Le rafraîchissement est effectué périodiquement[11].

I.4.2.2. Transformation des données

C'est l'étape d'homogénéisation, de filtrage et de calcul, où on :

- Nettoie les erreurs, corrige et complète les données manquantes.
- Fournit des mesures documentées de confiance dans les données.
- Consolide les données émanant de sources multiples pour être utilisées ensemble.
- Structure les données pour être utilisables par les outils de l'utilisateur final.

I.4.2.3. Chargement des données

Le chargement dans les structures dimensionnelles accessibles par l'utilisateur final.

¹ ODS: Operational Data Store.

² DW: Data Warehouse.

³ DSA: Data Staging Area.

I.4.3. La phase de structuration

Cette phase permet de stocker les données dans un entrepôt. Les données disponibles sont finalisées et rassemblées de sorte à pouvoir être extraites et traitées facilement.

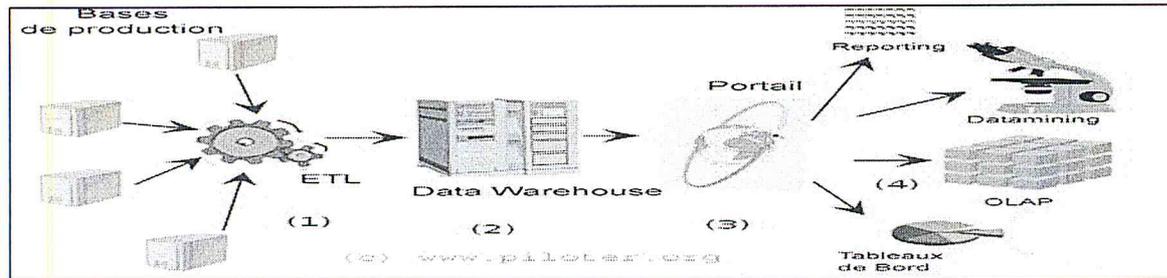


Figure I-5 Processus détaillé de traitement de données d'un système décisionnel [11]

Dans cette phase les données sont stockées sous une forme adaptée pour être analysées. En effet, l'ED est chargé de centraliser ces données dans des cubes, où la navigation des données peut être faite soit en utilisant le langage MDX (Multidimensional Expressions), soit à travers des interfaces ne nécessitant pas de connaissance spécifique.

I.4.4. La phase de restitution

Dans cette étape les utilisateurs finaux interviennent et analysent les informations qui leurs sont fournies sous forme de reporting, des portails d'accès à des tableaux de bord, des outils de navigation dans des cubes, ou des outils de statistique et de fouille.

I.4.4.1. Les outils de reporting et de requêtes

Ces outils permettent la mise à disposition de rapports périodiques, pré-formatés et paramétrables. Ils offrent une couche d'abstraction orientée métier pour faciliter la création de rapports par les utilisateurs eux-mêmes en interrogeant l'ED grâce à des analyses croisées. Ils permettent également la production de tableaux de bord avec des indicateurs de haut niveau ainsi que des graphes et des rapports d'activité en fonctions des dimensions de l'ED (temps, service, employé, citoyen, etc.) afin d'évaluer la performance [12].

Deux grands domaines se distinguent : le reporting **ad hoc** et le reporting **de masse**. Le reporting ad hoc offre la possibilité de créer soi-même le rapport souhaité à travers des vues métiers spécialement conçues en fonction de ses besoins. Le reporting de masse, quant à lui, permet de créer à l'avance des modèles de rapport réguliers et souvent demandés par les utilisateurs.

a) Tableaux de bord

« Le tableau de bord est un ensemble d'indicateurs peu nombreux conçus pour permettre aux gestionnaires de prendre connaissance de l'état et de l'évolution des systèmes qu'ils pilotent et d'identifier les tendances qui les influenceront sur un horizon cohérent avec la nature de leurs fonctions » [13].

b) Les indicateurs clés de performances (KPI : Key Performance Indicator)

Après l'analyse de sa mission, l'identification des intervenants et la définition de ses objectifs, il est nécessaire pour une organisation de mettre en place un ensemble de mesures afin de suivre le progrès vers l'atteinte de ces objectifs[14].

Un indicateur doit être simple, précis, mesurable, spécifique, accessible et valorisant.

1.4.4.2. Les outils d'analyse

Ils permettent de traiter, d'afficher et de naviguer les données. Plusieurs représentations d'un même résultat, sont possibles en une seule requête.

Afin de simplifier l'analyse, OLAP⁴ propose des opérateurs et mécanismes servant à naviguer dans les hiérarchies et les dimensions. Ils sont développés dans le 2^{ème} chapitre.

Introduit en 1993 par E.F. Codd [15], OLAP est qualifié par ces (12) règles de base :

1. **Vue conceptuelle multidimensionnelle**: permet la vision multidimensionnelle de données.
2. **Transparence** : l'utilisateur accède aux données, sans se préoccuper de leurs sources.
3. **Accessibilité** : les données doivent toutes être accessibles, sans ambiguïté.
4. **Performance consistante des rapports** : indépendamment des dimensions et de la BDD.
5. **Architecture Client/Serveur** : essentielle avec intégration facile des clients.
6. **Dimensionnalité générique** : par rapport à la structure et la capacité opérationnelle.
7. **Ajustement automatique du niveau physique** : au type du modèle et volume de données.
8. **Support multiutilisateur** : fournissant l'accès concurrents, l'intégrité et la sécurité.
9. **Opérations cross-dimensionnel libre** : Calculs possibles à travers toutes les dimensions.
10. **Manipulation intuitive des données** : à travers les cellules du model analytique.
11. **Rapports flexibles** : les dimensions peuvent être présentées de n'importe quelle manière.
12. **Dimensions et niveaux d'agrégation illimités.**

⁴ OLAP: On-Line Analytical Processing ou Traitement Analytique En Ligne.

I.4.4.3. Les outils de Datamining

Ils offrent une analyse plus poussée des données historisées permettant de découvrir des connaissances cachées dans les données comme la détection de corrélations et de tendances, l'établissement de typologies et de segmentations ou encore des prévisions [16]. Le Datamining est basé sur des algorithmes statistiques et mathématiques, et sur des hypothèses métier.

I.4.5. Fonctionnalités transverses

Deux grandes familles d'outils existent, à savoir le méta-dictionnaire et les outils d'administration. Le **méta-dictionnaire (méta-data)** décrit l'ensemble des données gérées depuis les sources jusqu'aux restitutions. Il permet aussi de gérer les données communes à toutes les couches. **Les outils d'administration** réalisent des sauvegardes ou des restaurations, assurent un suivi de l'utilisation et des performances (logs) ou encore gèrent les droits d'accès aux applications et aux données.

I.5. La BI et le domaine de la Ressource humaine (RH)

Autrefois centrée sur les questions financières et le marketing, la BI s'est petit à petit étendue à l'ensemble des grands domaines de l'organisation, de la gestion de la relation client à la gestion de la chaîne logistique en passant par les ressources humaines.

En effet, les Directions de Ressources Humaines ont de plus en plus recours à l'analyse prédictive de données (data mining) inhérente au BI, notamment pour optimiser les plans de GPEC (Gestion Prévisionnelle des Emplois et des Compétences)[2].

Appliquée à la gestion des ressources humaines, la Business Intelligence est un excellent outil de pilotage (recrutement, formation des collaborateurs, gestion des talents, repérage des profils à haut potentiel, optimisation de plan de carrière). Encore faut-il préparer les indicateurs de performance RH les plus pertinents.

I.5.1. Indicateurs de performance utiles aux RH

Grâce aux nouvelles technologies (Big Data, Cloud et le mode Software as a service SaaS), des masses colossales de données brutes sont disponibles et accessibles en temps réels.

Le défi aujourd'hui pour les RH consiste à les hiérarchiser pour en déduire les choix stratégiques pour l'organisation. Parmi les indicateurs RH de performance utiles on trouve: [2]

- la répartition des effectifs (par tranche d'âge, grade, sexe, zone géographique) ;
- le taux de turn-over (renouvellement d'effectifs pour la découverte de compétences) ;
- l'organisation des heures de travail ;
- la politique de rémunération (parts de salaires fixes et variables, commission, etc.) ;
- la structure de masse salariale (la totalité des dépenses engagées) ;
- le nombre d'heures de formation ;
- l'augmentation du taux de rendement en tenant en compte la vie sociale, le suivi de l'état de santé et les programmes de perfectionnement et de formations continues.

Les **solutions BI RH** s'organisent souvent autour de cinq principaux thèmes, à savoir : l'emploi, la rémunération, l'hygiène et la sécurité, les conditions de travail et la formation.

I.5.2. Objectifs du BI-RH

Une solution BI RH permet de disposer d'une **vue globale de l'effectif** et d'effectuer des **simulations** et des **prédictions** concernant le renouvellement d'effectifs, la gestion des compétences, la maîtrise de la masse salariale ou encore la fidélisation des talents.

Pratiquement, cela permet de **mieux identifier les talents**, d'anticiper et de faire évoluer **les compétences**, de **motiver** et de **fidéliser** les collaborateurs, et d'effectuer des **simulations** afin d'observer l'impact des décisions sur les effectifs ou sur la masse salariale.

I.5.3. Le reporting dans la RH

La **gestion des RH** nécessite la manipulation d'un **volume important de données de différentes sources**. Les responsables RH ont donc besoin de disposer de **tableaux de bord** répertoriant les principaux **indicateurs** nécessaires à l'**arbitrage** et à la **prévision** (suivi des effectifs, absentéisme ou encore de la rémunération). Ces tableaux de bords sont alors mis en œuvre dans le cadre d'une **démarche de BI**, qui consiste à synthétiser et à agréger des données obtenues par **reporting**, dans le but de leur **donner du sens** et de **répondre à des questions** ou d'**identifier des tendances** économiques, sécuritaire, financières et/ou concurrentielles.

I.6. Conclusion

L'informatique décisionnelle se révèle en élément stratégique pour les organisations, dont la réussite repose sur l'implication des directions métiers. Elle est devenue progressivement indispensable à l'ensemble des collaborateurs. Ceci permet ainsi d'établir que l'innovation métier est supportée par les technologies d'avenir, comme la BI, en étant intégrées directement dans les processus opérationnels.

Dans le domaine de la RH, la BI permet une connaissance plus fine des besoins des collaborateurs. Judicieusement organisées et analysées, les statistiques font émerger des informations plus qualitatives comme les leviers de motivation de tel ou tel employé collectés notamment lors des évaluations. Favorisant un suivi plus individualisé, la BI RH reste avant tout un outil au service de l'humain qui ne dispense en rien du face à face[2].

Chapitre **II** LES ENTREPÔTS DE DONNÉES

II.1. Introduction

Le but principal des SGBD⁵ est de stocker, organiser et permettre d'interroger au moyen de requêtes de grands volumes d'informations opérationnelles [17]. Ces données cachent souvent des connaissances et des informations stratégiques, utiles au processus décisionnel. L'informatique décisionnelle apporte des solutions nouvelles pour la modélisation, l'interrogation et la visualisation de données dans un objectif d'aide à la décision. Les modèles multidimensionnels ou entrepôts de données sont des modèles qui permettent de structurer les données pour l'analyse décisionnelle en explicitant la notion de dimension comme axe d'analyse et la notion de fait comme mesure ou indicateur.

II.2. Définitions et caractéristiques

L'origine du concept d'entrepôt de données (ED) remonte aux années 80, où un intérêt croissant au système décisionnel a vu le jour dû essentiellement à l'émergence des SGBD relationnels, la simplicité du modèle relationnel et la puissance offerte par le langage SQL [12].

Au début, l'ED était une base de données « infocentre » contenant une copie périodique des données du système opérationnel. Cette image a été dédiée au support à la prise de décision. Le motif principal étant de répondre aux requêtes des décideurs sans pour autant altérer les performances des systèmes opérationnels.

Actuellement, l'ED est devenu une autre source d'information, alimenté avec des données collectées et consolidées des différentes sources internes et externes.

⁵ SGBD : Systèmes de Gestion de Bases de Données.

II.2.1. Entrepôt de données (ED) ou Data Warehouse (DW):

« *Bill Inmon* » définit l'ED comme suit [18]:

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

- (i) **Orientées sujet** signifie organisées autour des sujets majeurs de l'organisation (employés, services, ventes, etc.) contrairement à l'approche transactionnelle utilisée dans les systèmes opérationnels. Ces derniers sont conçus autour d'applications et de fonctions (cartes bancaires, solvabilité client, etc.) ;
- (ii) **Intégrées** signifie uniformisées de toute hétérogénéité syntaxique et/ou sémantique due à la disparité des sources de provenance ;
- (iii) **Non volatiles** : du fait de l'historisation décrite plus bas, sauf le cas du rafraichissement, les opérations de mise à jour ou de suppression n'ont pas de sens pour un ED. les données ne sont accessibles qu'en mode consultation;
- (iv) **Evolutives dans le temps** : signifie la conservation les différentes valeurs d'une donnée pour d'éventuelles comparaisons et suivi d'évolution dans le temps. A l'inverse du système opérationnel où la valeur d'une donnée est simplement mise à jour, l'ED associe chaque valeur à un référentiel temporel;

« *Chaque structure clé dans l'ED contient - implicitement ou explicitement – un référentiel temporel*⁶ » [18].
- (v) **Organisées pour le support d'un processus d'aide à la décision** : Les données du DW doivent être agrégées et disponibles afin de faciliter leur interrogation et leur analyse favorisant le processus de prise de décision (Reporting, OLAP et Data Mining...).

II.2.2. Magasins de données ou Data Marts (DM)

Il s'agit de miniaturisations de DW dont l'ensemble forme l'entrepôt de données de l'organisation. Un DM répond à un besoin métier plus spécifique en étant construit autour d'un sujet précis d'analyse ou dédié à une hiérarchie départementale.

⁶ « *Every key structure in the data warehouse contains - implicitly or explicitly - an element of time.* ».

II.2.3. Choix d'un type de table

II.2.3.1. Table de fait

Les faits représentent les sujets de l'analyse. Une table de faits est la table centrale d'un modèle dimensionnel, où les mesures de performances sont stockées. Ces mesures sont généralement numériques, rarement textuelles. En effet, le concepteur doit faire son possible pour faire des mesures textuelles des dimensions, car elles peuvent y être corrélées efficacement avec les autres attributs textuels.

II.2.3.2. Tables de dimension

Elles accompagnent une table de faits et contiennent les descriptions textuelles de l'activité. Elles représentent, sous forme de listes hiérarchisées, les perspectives de l'analyse par exemple, analyse des séjours des patients selon : temps, services, pathologies. C'est grâce à la dimension que l'ED est compréhensible et utilisable [12].

II.2.3.3. Faits Vs Dimensions

Critère	Tables des faits	Tables des dimensions
Structure	Peu de colonnes et beaucoup de lignes	Peu de lignes et beaucoup de colonnes
Données	Mesurable, généralement numérique	Descriptives, généralement textuelles
Référentiel	Plusieurs clés étrangères	Une clé primaire
Valeur	Prend de nombreuses valeurs	Plus ou moins constantes
Manipulation	Participe à des calculs	Participe à des contraintes
Signification	Valeurs de mesure	descriptive
Rôle	Analyse un sujet particulier et assure les relations entre les dimensions	Assure l'interface homme/entrepôt de données

Tableau II-1 Tableau comparatif entre les tables de faits et les tables de faits et les tables de dimensions [19]

II.2.1. Granularité

Lorsqu'on interroge un DW pour des fins d'analyse on commence d'abord à un niveau de granularité plus élevé puis on se déplace vers le bas à des niveaux plus détaillés. Ainsi, il est judicieux de garder les données dans le DW sous formes agrégée à différents niveaux [19].

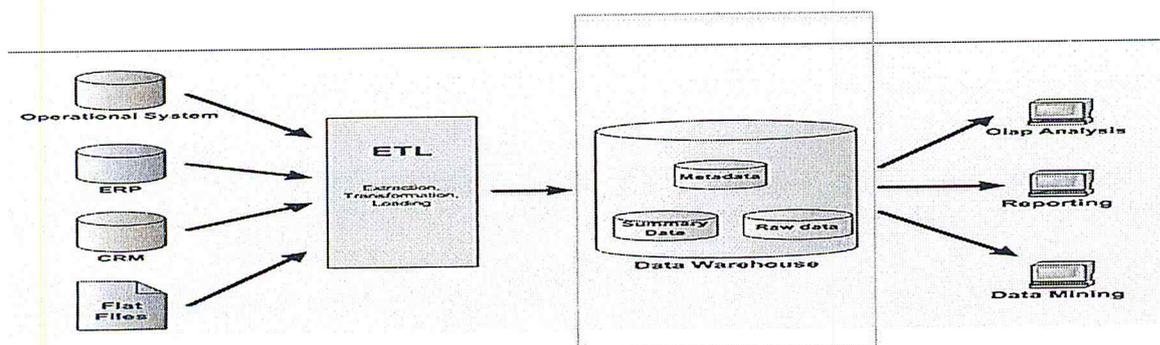


Figure II-1 Architecture globale d'un DW [14]

II.2.2. Agrégats

Les tables d'agrégats améliorent les performances en réduisant le nombre de lignes manipulés afin de répondre à une requête. Cela se fait grâce à l'agrégation des données des tables de faits détaillées et qui sont stockées dans de nouvelles tables de faits.

La construction des agrégats se base sur le modèle en étoile, et elle peut nécessiter la création de nouvelles dimensions dérivées ou la suppression de quelques dimensions. On peut aussi créer de nouveaux faits ou créer des tables pré-jointes d'autres tables.

II.3. Structure des données d'un Data Warehouse

Le DW a une structure bien définie, selon différents niveaux d'agrégation et de détail des données. Cette structure est définie par Inmon [18] comme suit :

- ❑ **Données détaillées** : reflètent des événements les plus récents, fréquemment consultées et généralement volumineuses car elles sont d'un niveau détaillé plus bas.
- ❑ **Données détaillées archivées** : anciennes, rarement sollicitées, généralement stockées dans un disque de stockage de masse ordinaire, elles sont du même niveau de détail.
- ❑ **Données agrégées** : sont des résumées à partir des données détaillées.
- ❑ **Données fortement agrégées**: données résumées avec un niveau d'agrégation plus élevé.
- ❑ **Métadonnées** : décrivant d'autres données, [20] pour la gestion physique et sémantique. Elles sont relatives à la structure des données, aux méthodes d'agrégation et au lien entre les données opérationnelles et celles du l'ED. On distingue trois grandes catégories[21]:
 - (i) **Métadonnées opérationnelles** renseignant sur les sources de données ;
 - (ii) **Métadonnées d'extraction et de transformation** (fréquences, méthodes et règles) ;
 - (iii) **Métadonnées de l'utilisateur final** facilitant la navigation dans l'ED.

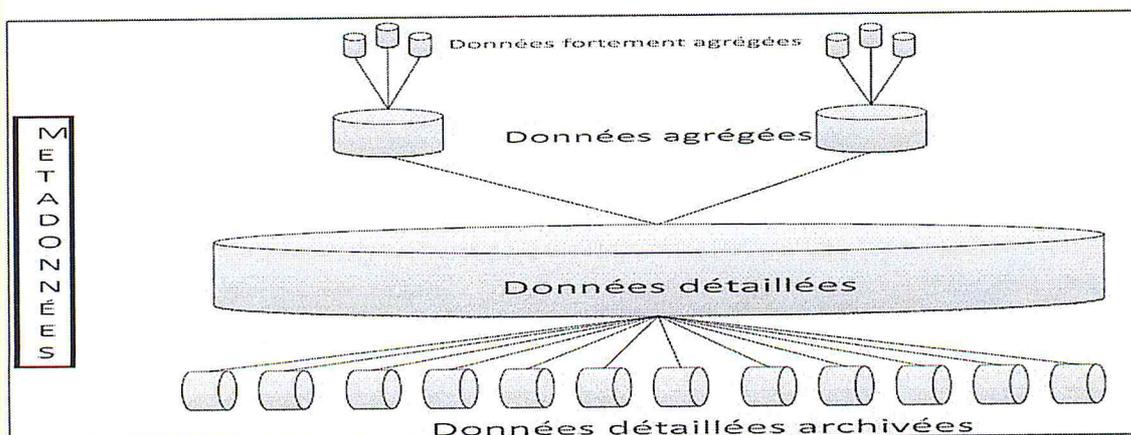


Figure II-2 Structure des données d'un DW [12]

II.4. Les éléments d'un Data Warehouse

L'environnement du DW est constitué essentiellement de quatre composantes : les **applications opérationnelles**, la **zone de préparation** des données, la **présentation** des données et les **outils d'accès** aux données [22].

- ❑ Les **applications opérationnelles** du système de production dont la priorité est d'assurer son fonctionnement et sa performance sont extérieures au DW.
- ❑ La **zone de préparation des données** englobe tout ce qui est entre les applications opérationnelles et la présentation des données. Il s'agit des processus ETL, où les données sont extraites pour subir des transformations avant leur chargement.
- ❑ La **zone de présentation des données** est l'entrepôt où les données sont organisées et stockées. C'est tout ce que l'utilisateur peut voir et toucher.
- ❑ La **zone d'outils d'accès** est l'ensemble des moyens fournis aux utilisateurs pour exploiter la zone de présentation des données en vue de la prise de décision.

II.5. Cycle de conception d'un Data Warehouse

Le cycle de vie d'un ED inclut trois (03) phases principales [23] : **Planification, Conception & implémentation et Maintenance & évolution.**

II.5.1. Planification

Cette phase consiste en trois (03) tâches de préparation pour la construction :

- Fixer les objectifs et les finalités de l'ED à construire;
- Evaluer techniquement et économiquement la faisabilité et la rentabilité;
- Identifier les utilisateurs de l'ED ainsi que leurs rôles.

II.5.2. Conception et implémentation

Cette phase consiste à concevoir le schéma de l'ED et à préparer les ressources nécessaires à son implémentation et à son déploiement;

Vu son importance, cette phase de conception est divisée en cinq (05) autres principales phases [24] : la définition des besoins, la modélisation conceptuelle, la modélisation logique, le processus ETL et la modélisation physique. Ceci constitue le cycle de conception de l'ED.

II.5.2.1. Définition des besoins

Comme dans tout projet informatique, la définition des besoins est une étape clé dans les projets de construction des ED. Elle permet d'identifier les différentes fonctions de l'ED ainsi que l'ensemble des informations qui doivent être disponibles [25]. Il existe deux types de besoins : [26]

- (i) **les besoins fonctionnels** : les services et les informations que l'ED doit fournir ;
- (ii) **les besoins non fonctionnels**: Contraintes et exigences implicites (sécurité, etc).

La définition des besoins comporte quatre (04) activités principales :

- La collecte des besoins** selon l'une des trois (03) approches suivantes:
 - (i) Axée sur les données disponibles au niveau des sources ;
 - (ii) Axée sur les utilisateurs exprimant leurs exigences ;
 - (iii) Axée sur les objectifs de l'organisation en impliquant ses cadres dans la conception.
- L'analyse des besoins** : détecter les conflits: contradiction, complémentarité, etc.
- La validation avec l'utilisateur** : garantir la conformité aux attentes de l'organisation.
- Le suivi de l'évolution des besoins** des utilisateurs et étudier son impact sur l'ED.

II.5.2.2. Modélisation conceptuelle

Elle consiste à définir le schéma conceptuel de l'ED. Ce dernier est soit élaboré à partir des besoins des utilisateurs, soit défini comme une vue sur les données des sources.

D'où, les deux approches les plus répandues dans la conception des DW sont :

- L'approche basée sur les besoins d'analyse. "Top-Down" de Kimball ; [21]
- L'approche basée sur les sources de données. "Bottom-up" de Inmon.

(1) Choix d'une approche

Aucune des deux approches citées n'est ni parfaite, ni applicable à tous les cas. Toutes deux doivent être étudiées pour choisir celle qui s'adapte le mieux à notre cas.

a) L'approche descendante (Top-Down)

Elle incite à construire l'ED en premier lieu, en considérant le besoin de l'organisation dans sa globalité. Dès que l'ED est mis en place, il est alors possible de créer des magasins de données contenant individuellement des parties de l'ED pour répondre aux besoins de services particuliers. Elle est illustrée par R. Kimball grâce à son cycle de vie dimensionnel :

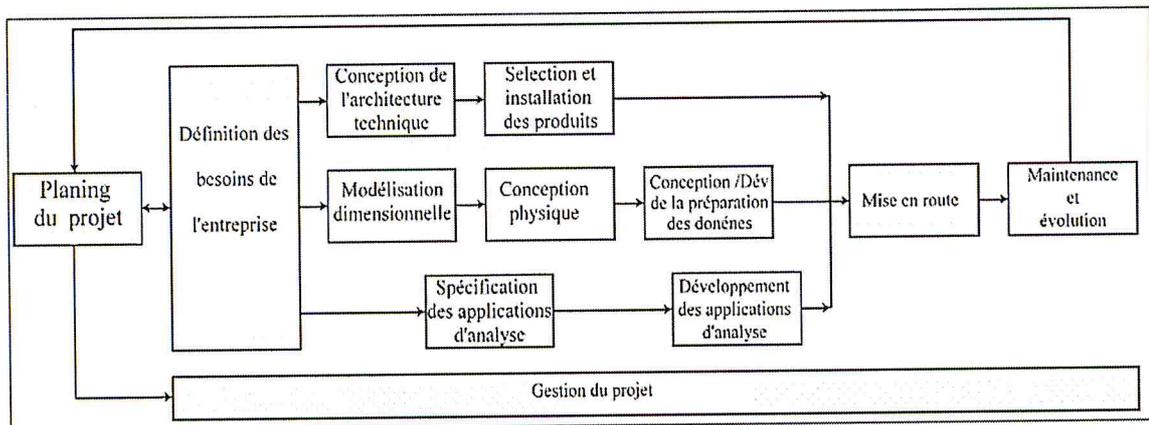


Figure II-3 L'approche Top-Down selon le cycle de vie dimensionnel de Kimball [27]

Avantages	Inconvénients
<ul style="list-style-type: none"> - Aucun risque de concevoir une solution obsolète avant d'être opérationnelle - Vision claire et globale des données ; - Architecture intégrée de l'ED ; - Stockage simple et centralisé des données; - Règles et un contrôle centralisé ; - Maintenance facile. 	<ul style="list-style-type: none"> - Pas de prise en compte de l'évolution des besoins de l'utilisateur. Nécessité de restructurer ED en cas de nouveau besoin ; - Difficulté de déterminer les besoins des utilisateurs ; - Longue et coûteuse à développer ; - Risque d'échec important ; - Résultats non immédiats ; - Négligence du système opérationnel.

Tableau II-2 Avantages et inconvénients de l'approche « Besoins d'analyse » [21]

b) L'approche ascendante (Bottom-Up)

Commence par la mise en place des magasins de données, ceux-ci sont rassemblés pour former l'ED pour l'ensemble de l'organisation. **Inmon** considère que l'utilisateur ne peut jamais déterminer ses besoins dès le départ, «*Donnez-moi ce que je vous dis que j'en ai besoin, et je vous dirai ce dont j'ai vraiment besoin* »⁷ [18], il considère que les besoins sont en constante évolution.

⁷ « Give me what I tell you I want, then I can tell you what I really want. » [18]

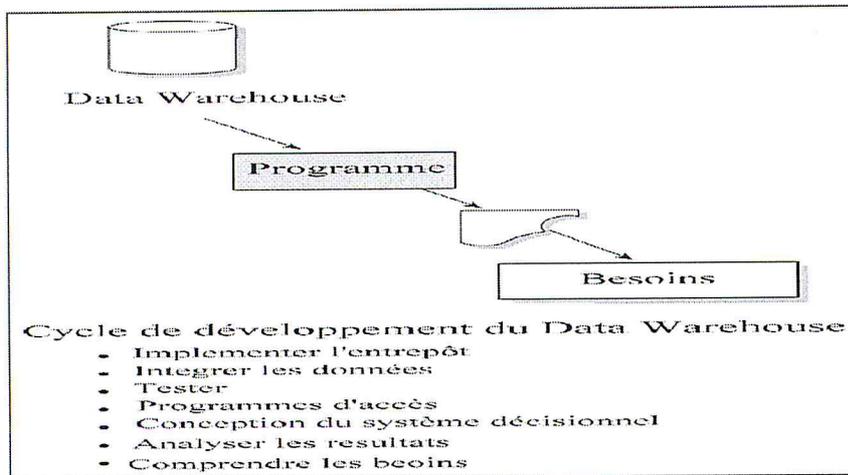


Figure II-4 L'approche Bottom-up selon le cycle de développement du DW de Inmon, [18]

Avantages	Inconvénients
<ul style="list-style-type: none"> - Meilleure prise en charge de l'évolution des besoins ; - Implémentation immédiate et peu coûteuse ; - Résultats rapides ; - Risque d'échec minime ; - Approche incrémentale (DMs importants d'abord). 	<ul style="list-style-type: none"> - Complexité des sources de données ; - Le DM a sa propre vision limitée des données ; - Incohérences et redondance entre les DMs ; - Fragmentation des données ; - Analyses inter DMs non optimisées; - Intégration importante des DMs en ED; - Risque d'obsolescence avant déploiement ; - Evolution du schéma des données source.

Tableau II-3 Avantages et inconvénients de l'approche « sources de données » [18]

c) Approche hybride

Issue de la combinaison des deux premières, cette approche peut se révéler efficace. Elle prend en considération à la fois les sources de données et les besoins des utilisateurs. Elle consiste en l'élaboration des schémas dimensionnels à base des structures des données du système opérationnel et la validation par rapport aux besoins analytiques. De ce fait, la présente approche cumule les avantages et quelques inconvénients des deux approches précédentes, telles que la complexité des sources de données et la difficulté quant à la détermination des besoins analytiques.

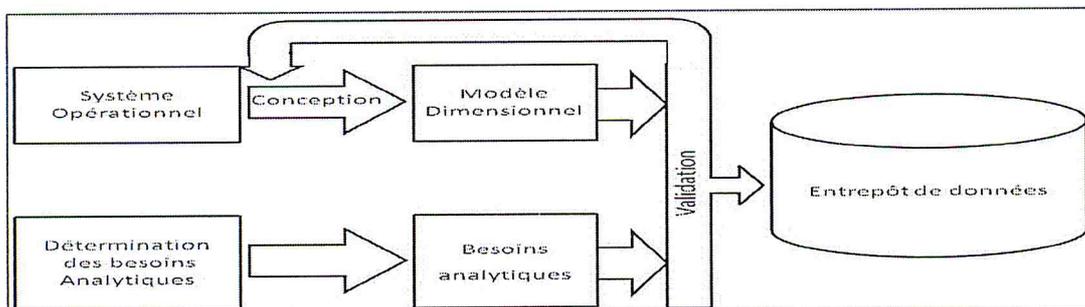


Figure II-5 Illustration de l'approche hybride

(2) Choix d'un schéma

La modélisation multidimensionnelle permet de recourir à un modèle de données simplifié et aisément compréhensible à travers les trois types de schéma suivants:

a) Schéma en étoile (Star schema)



Figure II-6 Modèle en étoile [28]

Le diagramme qui représente un modèle dimensionnel comporte une grande table centrale (**table de faits**) et un jeu de petites tables auxiliaires (**tables de dimensions**) disposées en étoile autour de la table centrale.

b) Schéma en flocon de neige (Snowflake schema)

Identique au modèle en étoile, sauf que les dimensions sont éclatées en sous-hiérarchies. Elle est généralement justifiée par l'économie d'espace de stockage, cependant elle peut s'avérer moins compréhensible et très coûteuse en termes de performances [16].

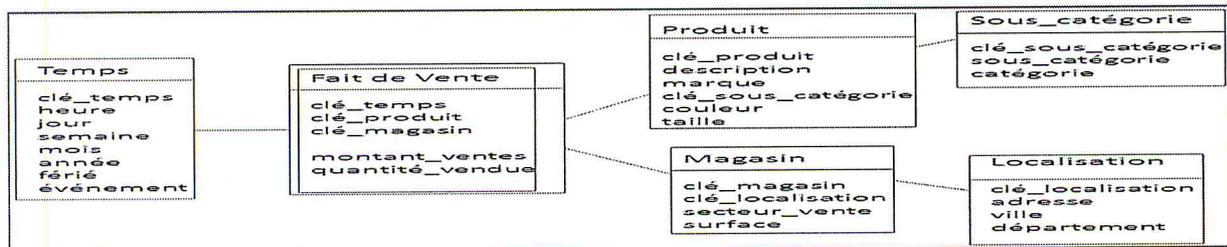


Figure II-7 Modèle en flocon [28]

a) Schéma en constellation de faits (Multi-star schema)

Consiste en plusieurs modèles en étoile liés entre eux par des dimensions communes.

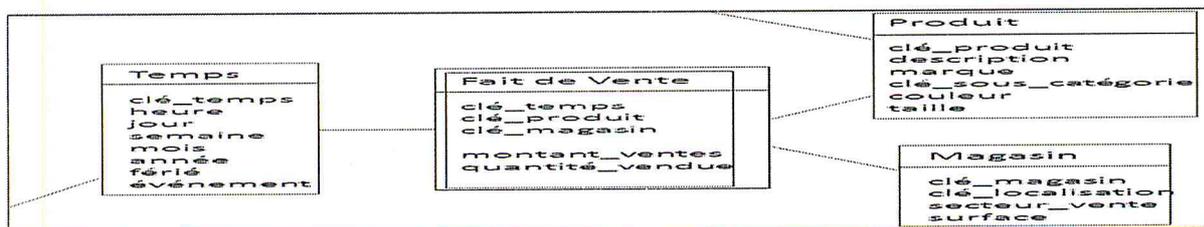


Figure II-8 Modèle en constellation [28]

Cette étape aboutit à l'établissement du modèle dimensionnel validé du DW.

II.5.2.3. Modélisation logique

La modélisation logique consiste à traduire le modèle conceptuel de l'ED en un modèle logique qui tient compte des spécificités du modèle d'implémentation choisi.

(1) Le concept OLAP

Les systèmes OLAP formalisent les opérateurs nécessaires à l'interrogation et à la navigation au sein des données multidimensionnelles. Il s'agit de parcours des hiérarchies de dimensions permettant de visualiser les mesures à différentes granularités de détail (par exemple les jours, les mois, les années, etc.).

«... le nom donné à l'analyse dynamique requise pour créer, manipuler, animer et synthétiser l'information par des modèles d'analyse de données selon des formules» [15].

En d'autres termes, il s'agit d'applications de modélisation descriptive et d'analyse exploratoire des données, conçues à des fins de prise de décision.

Nigel PENDSE, auteur d'OLAP Report (The BI Verdict), a jugé que les douze (12) règles, citées précédemment, étaient trop controversés et subjectives et l'a redéfini en cinq (05) mots : '**Analyse Rapide d'Information Multidimensionnelle Partagée**' (FASMI)⁸.

La modélisation multidimensionnelle est une approche proposée par Kimball [23] pour remplacer la modélisation traditionnelle relationnelle qui s'avérait inappropriée pour le support de la prise de décision et les analyses en ligne de type OLAP [29].

(2) Choix d'une architectures

Il existe trois (03) architectures de représentations logiques :

a) L'architecture MOLAP⁹

Elle est conçue exceptionnellement pour l'analyse multidimensionnelle. *Kimball* la définit comme étant un «*Ensemble d'interfaces utilisateur, d'applications et de technologies de bases de données propriétaires dont l'aspect dimensionnel est prépondérant*» [29].

Elle met en évidence les sujets analysés et les différentes perspectives de l'analyse.

⁸ FASMI: Fast Analysis of Shared Multidimensional Information.

⁹ MOLAP: Multidimensional On-line Analytical Processing.

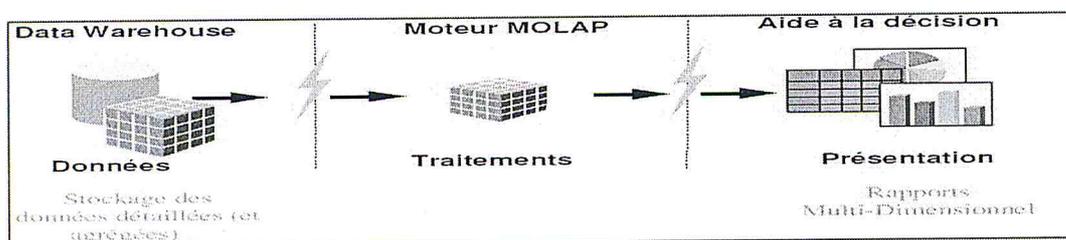


Figure II-9 Principe de l'architecture MOLAP [31]

La figure suivante illustre le modèle dimensionnel de l'activité vente, chaque cellule du cube contient les mesures (montants et quantités des ventes) permettant l'analyse selon les dimensions : temps, produit et ville.

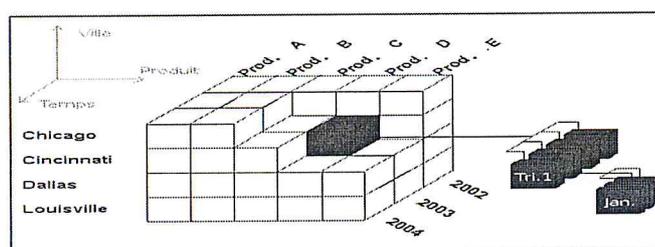


Figure II-10 Considération d'un sujet d'analyse comme un cube à plusieurs dimensions [32]

MOLAP consiste à représenter un sujet d'analyse comme un cube à dimensions (hyper-cubes), offrant des vues en tranches ou des analyses selon différents axes. Ces cubes permettent le calcul de mesures agrégées. Chaque dimension a la possibilité d'être hiérarchisée en fonction du besoin. Une dimension temps pourra par exemple utiliser la hiérarchie : Année, Trimestre, Mois, et jour. Il est ainsi possible d'accéder facilement et rapidement à l'information souhaitée en fonction de plusieurs dimensions.

MOLAP est principalement performante en termes de temps d'accès. Cependant, elle présente des limitations comme la complexité des mises à jour et l'indisponibilité de langage d'interrogation ce qui nécessite la redéfinition des opérations pour manipuler les structures multidimensionnelles ;

b) L'architecture ROLAP

ROLAP¹⁰ consiste en une représentation relationnelle du modèle de l'ED. Le fait est représenté en table contenant des attributs qui sont des mesures ou des clés étrangères des tables de dimension. La dimension est représentée également en table dont les attributs sont les paramètres. Les avantages de cette approche sont outre la grande capacité de stockage, l'utilisation d'un standard maîtrisé (le modèle relationnel).

¹⁰ **ROLAP**: Relationnel On-Line Analytical Processing.

ROLAP est un « Ensemble d'interfaces utilisateurs et d'applications qui donnent une vision dimensionnelle à des bases de données relationnelles » [29]. Il est une simulation du multidimensionnel en exploitant un SGBD relationnel. Il n'agrège rien ; Les règles d'agrégations sont créées au préalable dans une table relationnelle ce qui cause une lourdeur d'administration mais confère une certaine performance et un gage de cohérence.

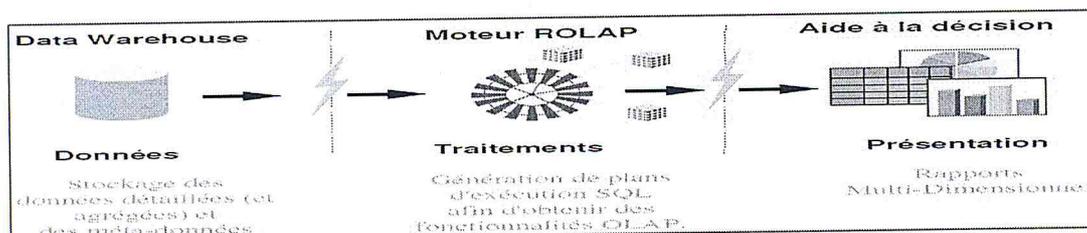


Figure II-11 Principe de l'architecture ROLAP [31]

c) L'architecture HOLAP

Cherchant à bénéficier des avantages des deux approches précédentes, **HOLAP¹¹** représente, de manière transparente, les données fréquemment utilisées dans un tableau multidimensionnel et représenter les autres données dans un schéma relationnel.

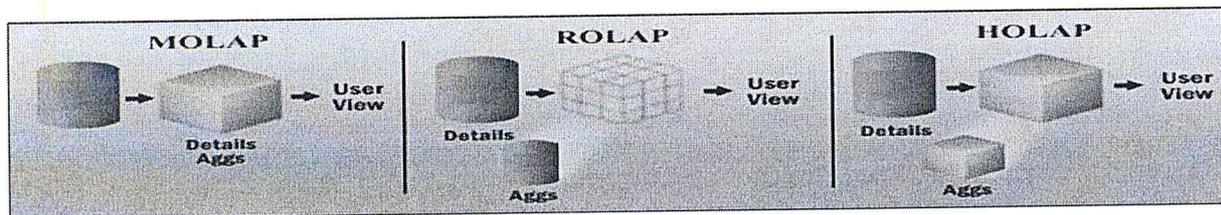


Figure II-12 Principe des trois architectures OLAP [33]

d) Autres architectures OLAP

Bien que les architectures déjà évoquées soient les plus répandues et les plus adoptées par les fournisseurs de solutions OLAP, d'autres systèmes se basent sur des architectures différentes telles que OOLAP¹², ou alors DOLAP¹³ qui décrit une catégorie de produits qui ne sont pas nécessairement connectés à un serveur et qui s'appuient sur une source de données (cube) construites, stockées et exploitées en local sur la machine de l'utilisateur.

(3) La navigation dans les données

Une fois que le serveur OLAP a construit/simulé le cube multidimensionnel, plusieurs opérations y sont possibles offrant la possibilité de naviguer dans les données qui le

¹¹ **HOLAP**: Hybrid On-Line Analytical Processing.

¹² **OOLAP**: Object On-line Analytical Processing.

¹³ **DOLAP**: Desktop On-line Analytical Processing.

constituent. Ces opérations de navigation « Data Surfing » doivent être, paradoxalement, assez complexes pour adresser l'ensemble des données et assez simples afin de permettre à l'utilisateur de circuler de manière libre et intuitive dans le modèle dimensionnel.

Afin de répondre à ces attentes, un ensemble de mécanismes est exploité, permettant une navigation par rapport à la dimension et par rapport à la granularité d'une dimension [34].

Il existe de nombreuses propositions concernant la définition d'opérations de manipulation OLAP dans laquelle les structures manipulées sont les cubes de données et les tranches du cube de données.

a) Opérations classiques de sélection

- **Slice** ou Tranche de cube: exprime une sélection de tranches du cube selon des prédicats et selon une dimension (projeter selon un axes d'analyse) « *filtrer une dimension selon une valeur* » [34].
- **Dice** ou Sous-cube: exprime une restriction sur les données d'un indicateur d'analyse.

b) Opérations agissant sur la granularité

- **Roll-up** ou forage vers le haut c'est d'analyser les données en fonction d'un niveau de granularité supérieur, moins de précision, selon la hiérarchie définie sur la dimension.
- **Drill-down** ou forage vers le bas permet d'analyser les données avec un niveau de granularité inférieur, donc sous une forme plus détaillée.

c) Opérations agissant sur la structure :

- **Rotation** : Réorientation d'une analyse en changeant l'un des axes en cours. La rotation de dimension c'est changer le sujet de l'analyse. Rotation de fait ou drill-across c'est de changer de perspective d'analyse (de hiérarchie).
- **Switch** (Permutation) Change la position des membres d'une dimension.
- **Split** (Division) réduit le nombre de dimensions en présentant chaque tranche du cube.
- **Nest** (Emboîtement) Imbrique des membres issus de dimensions différentes.
- **Push** (Enfoncement) Combine les membres d'une dimension aux mesures.
- **Pull** (retrait) transforme une mesure en paramètre en changeant le statut de mesures.
- **Fold** (Factualisation) consiste à transformer une dimension en mesure(s).
- **Unfold** (Paramétrisation) d'une mesure en paramètre dans une nouvelle dimension.
- **Cube** permet de calculer des sous-totaux et un total final dans le cube.

d) MDX (MultiDimensional eXpressions)

C'est un langage de requête pour les bases de données OLAP, analogue au rôle de SQL pour les bases de données relationnelles. C'est aussi un langage de calcul avec une syntaxe similaire à celle des tableurs. Il possède une syntaxe appropriée à l'interrogation et manipulation des données multidimensionnelles mémorisées dans un cube OLAP.

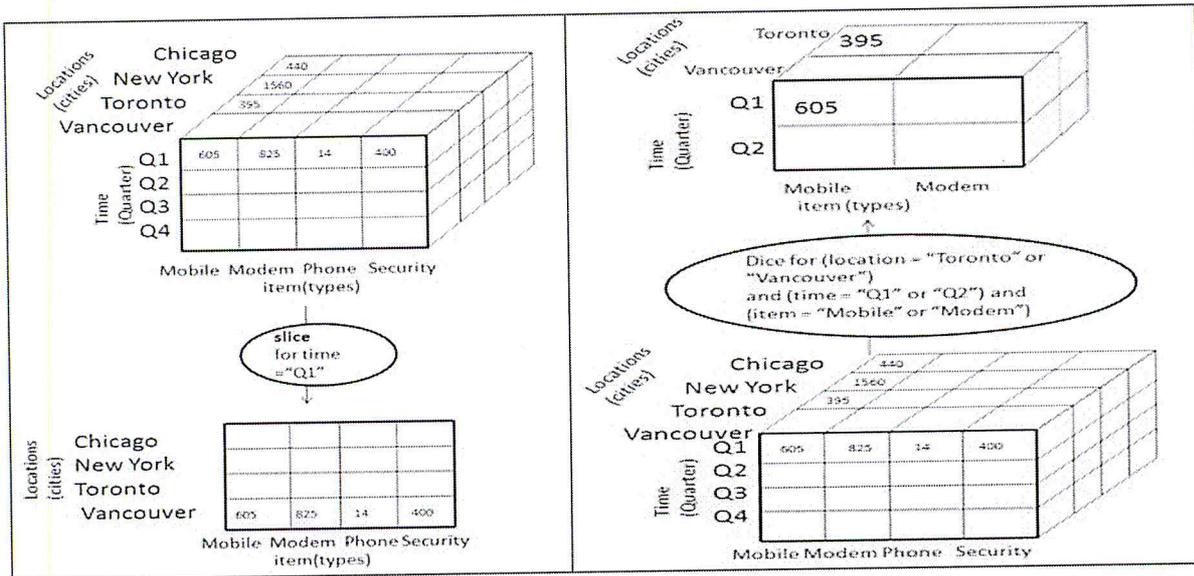


Figure II-13 Exemple de Slice et de Dice

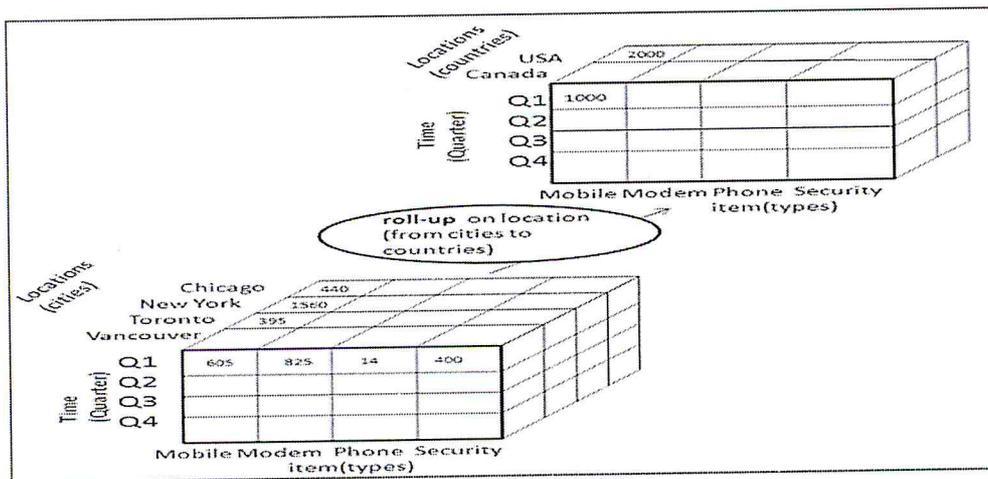


Figure II-14 Exemple de Roll up [35]

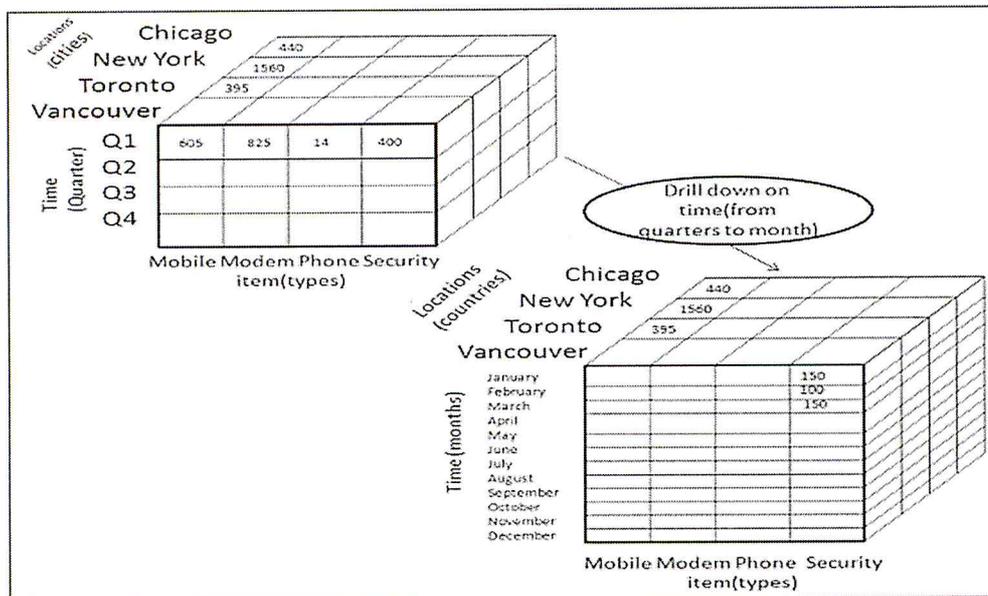


Figure II-15 Exemple de Drill-Down [35]

II.5.2.4. Alimentation du Data Warehouse (ETL)

Une fois le DW conçu, il faut l'alimenter et le charger en données. Cette alimentation n'est rien d'autre que le processus ETL cité dans le chapitre précédent. Ce processus décrit une mécanique cyclique qui a pour but de garantir l'alimentation du DW en données homogènes, propres et fiables.

II.5.2.5. Modélisation physique

Cette dernière phase correspond à l'implémentation physique du modèle logique de l'ED. Durant cette phase, l'administrateur choisit les techniques et structures d'optimisation adéquates qui doivent assurer les performances optimales en termes d'accès à l'ED. Quatre (04) tâches principales doivent être réalisées :

- Choisir les structures d'optimisation parmi plusieurs existantes ;
- Choisir le mode de sélection [35];
- Développer des algorithmes de sélection [36], [37] ;
- Valider les solutions d'optimisation.

II.5.3. Maintenance et évolution

Cette phase concerne l'optimisation et la mise à jour périodique de l'ED en fonction des changements pouvant se produire au niveau des sources de données ou des besoins des utilisateurs.

La mise en service du DW ne signifie pas la fin du projet, car un projet DW nécessite un suivi constant compte tenu des besoins d'optimisation de performance et ou d'expansion. Il est donc nécessaire d'investir dans les domaines suivants [22] :

- Support ;
- Formation ;
- Support technique ;
- Management de l'évolution.

Ces travaux d'expansion sont à prévoir de façon à faciliter l'évolution du schéma DW.

II.6. Conclusion

Ainsi, la modélisation multidimensionnelle repose sur les concepts de fait et de dimension. Leur association compose un schéma en étoile, en flocon ou en constellation. Dont l'implémentation se fait en R-OLAP ou en MOLAP. Plusieurs opérations permettent la navigation et la manipulation des données.

Afin de mettre en place un tel système, il est nécessaire de choisir et d'adopter une démarche précise qui tient compte des réalités de l'organisation et des contraintes du projet.

La modélisation de l'ED est multidimensionnelle et l'alimentation en données constitue l'étape à laquelle il faut accorder le plus d'attention et de temps.

En effet, elle est le garant d'une contenance d'ED fiables et correctes. Une fois l'alimentation terminée, l'exploitation des données peut alors se faire par différentes méthodes. L'utilisation d'outil OLAP reste, cependant, l'aspect le plus intéressant dans cette exploitation permettant la navigation dans les données de l'ED à la demande.

Chapitre III Introduction au Clustering

III.1. Introduction

Le grand succès de la classification non supervisées (clustering) est qu'elle permet d'extraire les liens cachés entre les données et de les regrouper en clusters ou classes homogènes, sans aucune connaissance préalable sur la formation de ces clusters. Elle n'exige ni une phase d'apprentissage, ni un travail d'un expert humain (procédure automatique), ce qui favorise son utilisation et rend l'implémentation de ses algorithmes plus facile.

Le **clustering** consiste à diviser ou séparer un ensemble d'objets en différents groupes, dits clusters, en fonction d'une certaine notion de similarité. Les objets qui sont considérés comme similaires sont assignés au même cluster tandis que ceux qui sont considérés comme différents sont placés dans des clusters distincts [38]. Le clustering est une problématique de recherche étudiée depuis plusieurs années dans différents domaines, tels que : machine learning, fouille de données, pattern recognition, statistiques, etc.

III.2. Domaines d'application

Le clustering est utilisé pour la compréhension de phénomènes décrits par des bases de données. Quelques domaines d'applications sont la reconnaissance de la parole, la segmentation d'images, la fouille de textes, la catégorisation des clients, etc[39].

Ces domaines peuvent être regroupées en trois groupes principaux : [40]

- **Segmentation** de base de données pour simplifier le traitement. Elle est appliquée dans l'imagerie, les bases de données spatiales, etc. ;
- **Classification** en sous-ensembles ayant des comportements ou caractéristiques similaires dans les bases de données clients. Elle est appliquée dans la gestion de la relation Client ;

- **Extraction des connaissances** en permettant la compréhension de la structure des données. Elle est utilisée dans la fouille du Web, la fouille du texte, la bioinformatique, l'analyse génétique, etc.

III.3. Mesures de similarité

D'une manière générale, l'objectif du clustering est de regrouper les objets (ou les données) similaires dans le même cluster et les objets dissimilaires dans des clusters différents en se basant sur un critère ou une mesure de similarité. Globalement, ces mesures de similarité peuvent être divisées en deux catégories : mesures de similarité entre objets et mesures de similarité entre clusters [41].

III.3.1. Mesures de similarité entre objets

Le concept de similarité ou de dissimilarité entre objets permet de montrer à quel point deux objets sont proches et donc assignés au même cluster. En pratique, la fonction de distance entre paires d'objets est la mesure de similarité la plus utilisée.

Soit X un ensemble d'objets, la fonction de distance (d) doit respecter les conditions suivantes [42]:

1. $d(x_i, x_j) \geq 0, \forall x_i, x_j \in X$.
2. $d(x_i, x_i) = 0, \forall x_i \in X$.
3. $d(x_i, x_j) = d(x_j, x_i), \forall x_i, x_j \in X$.
4. $d(x_i, x_k) + d(x_k, x_j) \geq d(x_i, x_j), \forall x_i, x_j, x_k \in X$.

Dans la littérature, les fonctions de distance les plus utilisées sont :

Distance Euclidienne :
$$d(x_i, x_j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$$

Distance de Hamming :
$$d(x_i, x_j) = \sum_{k=1}^n |x_{ik} - x_{jk}|$$

Distance de Minkowsky :
$$d(x_i, x_j) = \sqrt[p]{\sum_{k=1}^n (x_{ik} - x_{jk})^p}, \quad p > 0$$

Distance de Chebychev :
$$d(x_i, x_j) = \max_{1 \leq k \leq n} |x_{ik} - x_{jk}|$$

III.3.2. Mesures de similarité entre cluster

En plus de la similarité entre objets, certains algorithmes, tels que les algorithmes hiérarchiques, nécessitent d'évaluer la similarité entre les clusters. Souvent, le calcul de similarité entre deux clusters revient au calcul de similarité entre leurs objets.

Soient deux clusters C_p et C_q , dont le nombre d'objet dans ces deux clusters est n_p et n_q , respectivement. Les fonctions de mesure de distance (similarité), $dist(C_p, C_q)$, entre les clusters C_p et C_q les plus utilisées sont [41]:

Plus proche voisin (Nearest neighbor) : distance minimale entre toutes les paires d'objets

$$dist(C_p, C_q) = \min_{\substack{x \in C_p \\ y \in C_q}} \{d(x, y)\}$$

Voisin le plus éloigné (farthest neighbor): distance maximale entre toutes les paires d'objets

$$dist(C_p, C_q) = \max_{\substack{x \in C_p \\ y \in C_q}} \{d(x, y)\}$$

Moyen de groupe (Group mean) : distance moyenne entre toutes les paires d'objets :

$$dist(C_p, C_q) = \frac{\sum_{x \in C_p} \sum_{y \in C_q} d(x, y)}{n_p \times n_q}$$

Centroïde : valeur moyenne des objets d'un cluster. Si v_p est le centroïde du cluster C_p et v_q est le centroïde du cluster C_q , la distance entre C_p et C_q est la distance entre v_p et v_q

$$dist(C_p, C_q) = d(v_p, v_q) \quad \text{avec : } v_p = \frac{\sum_{x \in C_p} x}{n_p} \quad \text{et } v_q = \frac{\sum_{y \in C_q} y}{n_q}$$

Medoïde : l'objet le plus central dans le cluster. Si m_p est le medoïde du cluster C_p , et m_q est le medoïde du cluster C_q , la distance entre C_p et C_q est la distance entre m_p et m_q :

$$dist(C_p, C_q) = d(m_p, m_q)$$

$$\text{Avec : } m_p = \operatorname{argmin}_{x_i \in C_p} \sum_{x_j \in C_q} d(x_i, x_j) \quad \text{et} \quad m_q = \operatorname{argmin}_{x_j \in C_q} \sum_{x_i \in C_p} d(x_i, x_j)$$

III.4. Algorithmes de clustering

Globalement, les algorithmes de clustering peuvent être classés dans deux types : les algorithmes de partitionnement et les algorithmes hiérarchiques.

III.4.1. Algorithmes de clustering par partitionnement

Le principe de base des algorithmes de clustering par partitionnement est de regrouper les données en un ensemble fixe de clusters, tous dans le même niveau. Les plus connus sont :

❑ **Famille des algorithmes C-Moyenne :**

- Algorithme des K-Means [44] [45]
- Algorithme Fuzzy C-Means (FCM) [46]
- Algorithme Possibilistic C-Means (PCM)

❑ **Algorithme Quality Threshold (QT)**

III.4.2. Algorithmes hiérarchiques

Le clustering hiérarchique construit une hiérarchie de clusters représentée par un arbre de clusters appelé «dendrogramme». Deux types sont distingués: ascendant et descendant.

III.4.1. Conclusion

Du parcours de la littérature, nous avons vu qu'il existe plusieurs approches et algorithmes pour le clustering des données qui souvent produisent des résultats différents. Cette multitude d'algorithmes, bien que bénéfique d'une part, peut également poser de sérieux problème sur le choix de l'algorithme et les paramètres optimaux. Une solution possible à ce problème consiste en la combinaison de deux ou plusieurs algorithmes.

Partie 2 CONTRIBUTION

Dans cette partie, après l'exposition de l'existant et du besoin, nous allons concevoir graduellement notre modèle dynamique d'aide à la décision pour la gestion des compétences. Il s'agit d'un système décisionnel basé sur un entrepôt de données. En effet, nous allons alors commencer par concevoir ledit entrepôt, la zone d'alimentation, et la partie présentation en passant par la conception des cubes. Ensuite nous allons passer à la conception du système de pondération.

Chapitre IV ÉTUDE D'EXISTANT ET BESOINS



Gendarmerie Nationale

IV.1. Organisme d'accueil

Créée par l'ordonnance n°62/019, du 23 août 1962, la Gendarmerie Nationale est une partie intégrante de l'Armée Nationale Populaire. Ses missions et son organisation sont portées par le décret n° 99/77/AG/A1 en date du 23 octobre 1977. A travers son histoire et l'exercice de ses missions, la GN s'est forgée pour assumer convenablement sa noble charge d'institution militaire républicaine, garante de la sécurité des personnes et des biens, en luttant avec détermination contre le crime sous toutes ses formes.



Pour se faire, le seul recours aux nouvelles technologies est insuffisant ce qui exige de valoriser l'action des hommes par une formation continue et par l'instauration d'une veille active permettant d'atteindre l'art de la perfection dans tous les domaines impliquant les missions de la GN.

Néanmoins, la formation n'est pas envisageable sans considérer la gestion de la ressource humaine, notamment les profils de carrière et la spécialisation dans les postes étroitement liés aux compétences. Ceci impose une adéquation avec les normes de gestion afin de préserver la ressource formée et de prévenir tout dysfonctionnement.

Ces actions conduites à terme, contribueront à l'émergence d'un nouvel environnement basé sur la reconnaissance du mérite et permettra à la GN d'anticiper et de réagir aux menaces affectant le contexte sécuritaire national [CGN, 2015].

Le Commandement de la Gendarmerie Nationale a mis en place une stratégie prometteuse et ambitieuse dans le cadre de principes certains pour la modernisation des capacités de l'arme et l'affinement du professionnalisme, en harmonie avec le développement de l'Algérie dans tous les domaines et en fonction des concepts de lutte contre la criminalité.

Cette stratégie a généré la structuration illustrée dans l'organigramme ci-après.

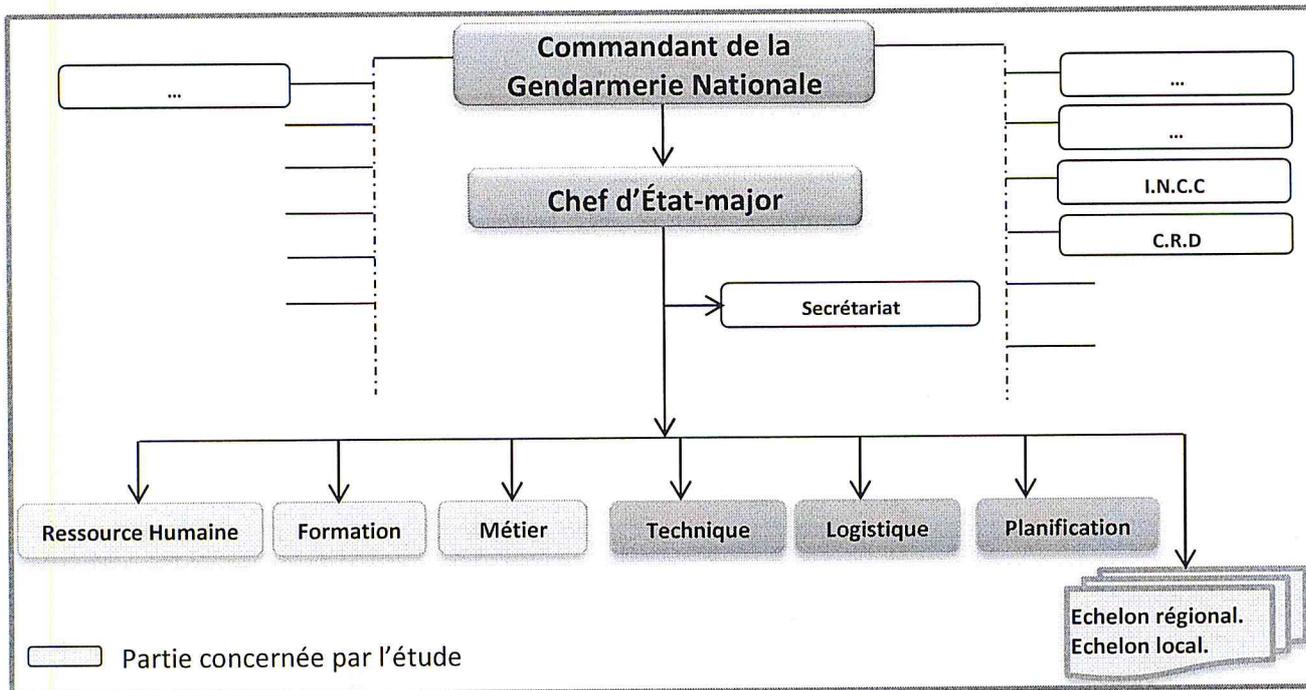


Figure IV-1 portée de l'étude par rapport à l'Organigramme de la GN

A l'échelon régional :

Six Commandements Régionaux correspondant aux six Régions Militaires.

A l'échelon local :

- Un groupement territorial au niveau de chaque wilaya, soit 48 groupements ;
- Des unités d'intervention articulées en groupements et escadrons autonomes ;
- Des compagnies territoriales au niveau Daïra,
- Des brigades territoriales au niveau de la commune, à travers le territoire national,
- Des unités spécialisées comprenant les sections de recherches, les unités de sécurité routière, et de brigades spécialisées.
- Des unités de garde-frontières,
- Des formations aériennes (d'hélicoptères).
- Le détachement spécial d'intervention spécialisé dans la neutralisation de malfaiteurs dangereux et intervient aussi contre le crime organisé et participe aux missions de secours et de sauvetage.

Dans ce chapitre, nous allons présenter seuls les aspects conceptuels des différents systèmes de production en rapport avec notre champ d'étude. Nous ne prétendons nullement que cette étude d'existant soit exhaustive.

Missions de la Gendarmerie Nationale¹⁴

La GN participe à la défense nationale conformément aux plans arrêtés par le Ministère de la Défense Nationale et à la lutte contre le terrorisme. Elle a pour charge l'exercice des missions de police judiciaire, de police administrative et de police militaire.

En matière de police judiciaire :

La GN lutte contre la criminalité et le crime organisé. Elle met en œuvre, à cet effet, des moyens d'investigations de police scientifique et technique et d'expertise criminalistique. Elle exerce cette mission conformément aux dispositions du code de procédures pénales. En effet, elle constate les infractions à la loi pénale, en rassemble les preuves et en recherche les auteurs.

En matière de police administrative :

La GN veille au maintien de l'ordre et de la paix publics par une action préventive caractérisée par une surveillance générale et continue, et assure la sécurité publique par la protection des personnes et des biens et la liberté de circulation sur les voies de communication. A ce titre, elle veille à l'application des lois et règlements régissant les polices générale et spéciale.

En matière de police militaire :

La GN assure la police judiciaire militaire conformément aux dispositions du code de justice militaire et la police générale militaire conformément aux règlements en vigueur au sein de l'Armée Nationale Populaire.

La Gendarmerie Nationale est chargée des missions de surveillance générale et continue du territoire national, de renseignement et d'information des autorités publiques et d'exercice de l'action préventive et répressive.

IV.2. Etat des lieux de la ressource humaine dans la GN

Le processus de développement de la GN a abouti à une transformation quasi-totale de sa composante humaine. Son effectif a connu une augmentation du personnel par rapport aux années précédentes injectant essentiellement une ressource plus jeune, plus qualifiée et plus spécialisée.

¹⁴ Récupérée du site : « http://www.mdn.dz/site_principal/index.php?L=fr » consulté le:17/09/2017

Les objectifs en matière de Ressources Humaines se recentrent sur:

- La normalisation continue de la gestion des RH, afin d'identifier judicieusement les compétences et de reconnaître les potentialités permettant une meilleure adéquation Gendarme/Emploi.
- L'accompagnement permanent sur le plan socioprofessionnel, psychologique des personnels et de leurs familles tout au long de leur carrière.
- L'administration, la gestion des carrières du personnel et l'automatisation de la gestion ;
- La participation au développement de la ressource humaine et la veille sur l'exécution de la politique prévisionnelle en matière de recrutement, de formation, d'avancement et de mise à la retraite conformément à la réglementation en vigueur ;
- La veille sur l'exécution et le suivi des visites médicales systématiques annuelles;
- Le suivi de la situation sanitaire des personnels et la veille à leur bonne prise en charge;
- La participation à l'opération d'assainissement des dossiers d'expertise médicale.

IV.2.1. Système d'Information Ressource Humaine SI-RH

La mise à jour des dossiers des personnels est une tâche de grande importance pour les gestionnaires. En effet, elle consiste à classer et à archiver tous les documents et informations y afférents dans des conteneurs normalisés ainsi qu'à les restaurer au moment voulu.

Le carnet de notes, où sont consignées toutes les références et les dates relatives au déroulement de la carrière du militaire, a été remplacé progressivement par la version numérisée. Considérée comme le noyau du SI-RH, cette dernière n'est rien d'autre que l'application de saisie et de consultation des personnels GN.

La mise au point du SI-RH actuel a nécessité un effort particulier de la part des gestionnaires RH et des informaticiens. Le SI-RH a été déployé au profit des échelons gestionnaires connectés au réseau pour la consultation et la mise à jour des données d'état civil, sanitaires et sociales des personnels et de leurs familles.

IV.2.2. Rubriques de l'application de saisie / consultation des personnels

- | | |
|----------------------------------------|-----------------------------|
| ▪ Identifications (recherche) | Recrutement et réengagement |
| ▪ Nominations dans le grade | Mutations successifs |
| ▪ Fonctions/positions administratives | Diplômes et formations |
| ▪ Stages et résultats | Concours et examens |
| ▪ Situation familiale | Décorations |
| ▪ Appréciations et notations annuelles | Récompenses et sanctions |
| ▪ Jugements et infractions | Cas sanitaires et sociaux |

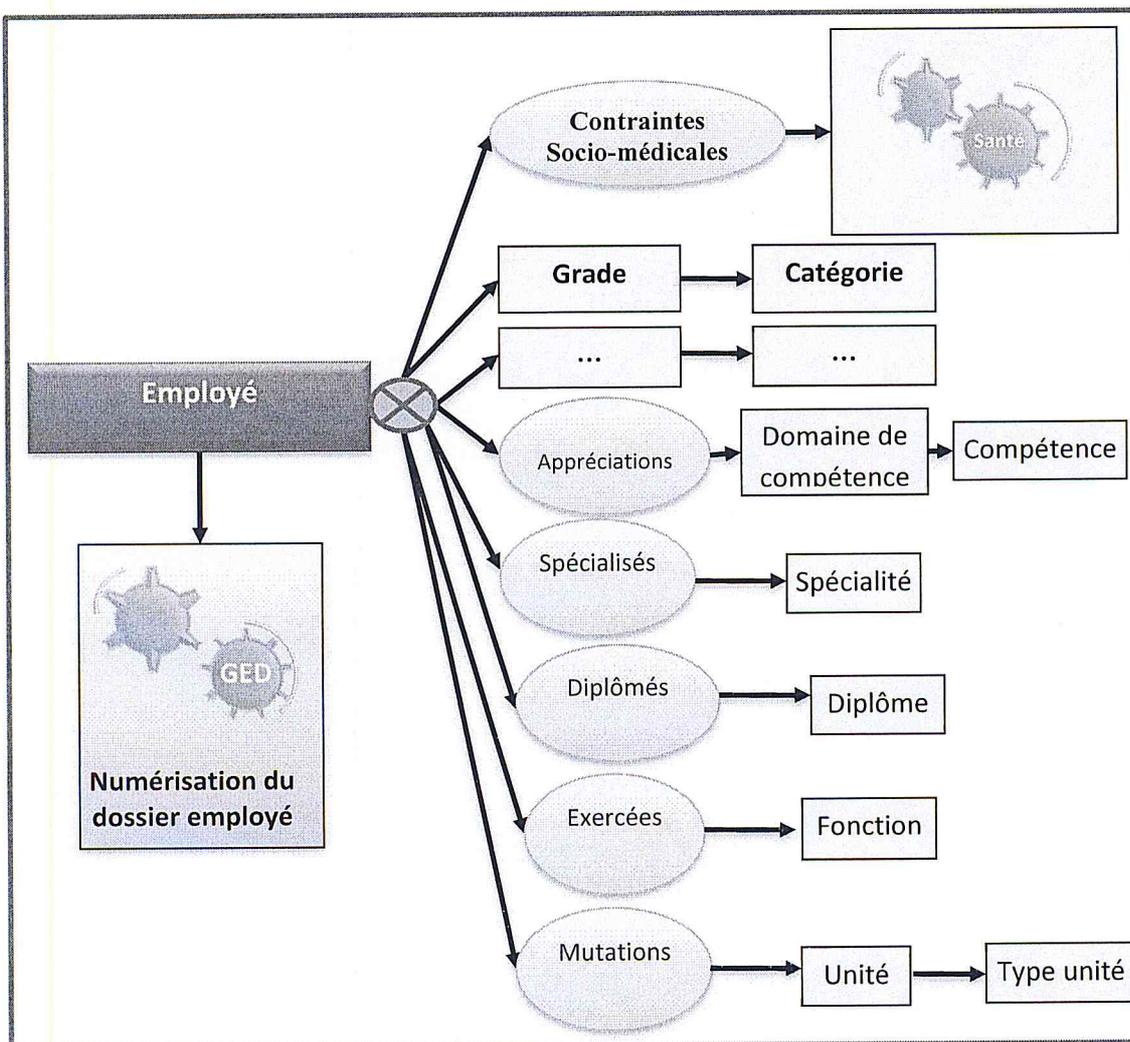


Figure IV-2 Quelques aspects du système de production RH

IV.3. Etat des lieux de la Formation

La formation dans la GN s'inscrit dans une logique de coordination et de conjugaison des actions communes, ayant pour objectif une gestion optimale des ressources et moyens afin de répondre à l'attente de la sécurité publique.

La synergie créée entre les structures centrales chargées de recrutements, d'examens et de formation avec les organes de formation participe à la cohésion et la complémentarité dans l'exécution des missions. Parmi ces missions :

- Assurer la cohérence de la continuité des enseignements entre les profils et spécialités.
- Organiser le recrutement, la sélection des candidats et l'orientation des stagiaires;
- Organiser les concours d'admission aux différentes formations.
- Gérer les différents cycles de formation et le soutien pédagogique.
- Superviser la préparation et le déroulement des examens des cycles de formation.
- Evaluer les performances et les résultats obtenus en fonction des programmes Arrêtés,
- Etablir des relations avec les établissements civils de formation.

IV.3.1. Système d'Information de la Formation SI-F

L'application phare du SI-F est le module « **Gestion des Stagiaires** ». Interconnectée avec le SI-RH, cette application vise à gérer les dossiers des militaires en stages dans les différents types d'établissements GN, ANP, civils et étrangers.

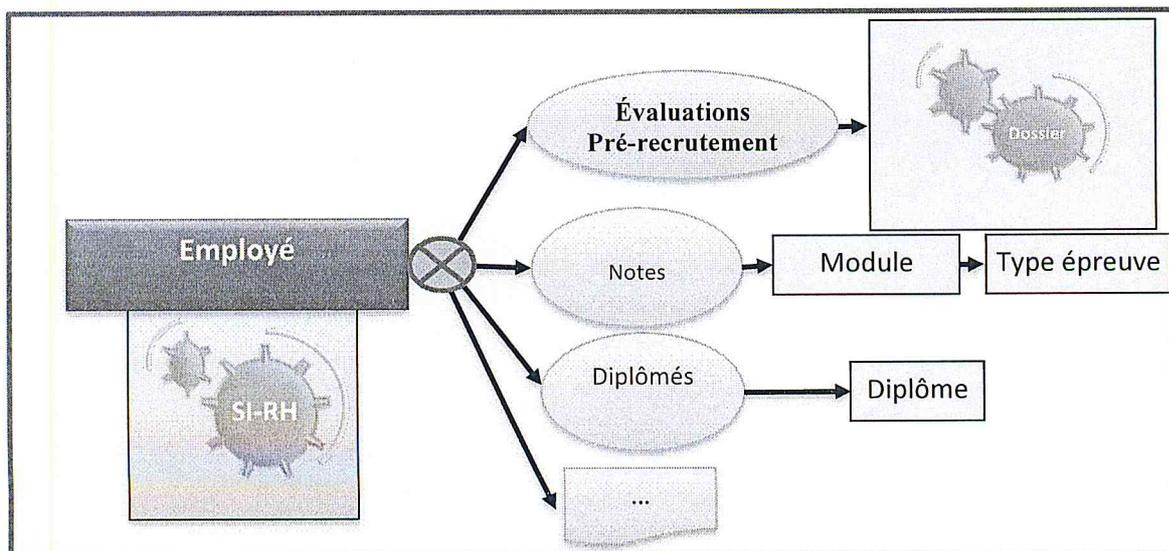


Figure IV-3 Quelques aspects du système de production Formation

IV.4. Etat des lieux du métier : Police Judiciaire (PJ)

La GN, Force militaire chargée du maintien et du rétablissement de l'ordre public, de la surveillance des voies de communications et de l'exécution des lois, œuvre continuellement à adapter au mieux sa stratégie de lutte contre les diverses formes de la criminalité et du crime organisé.

La réussite de toute stratégie de lutte passe obligatoirement par l'intégration des technologies modernes, à une solide formation des hommes alliés à leur engagement à protéger le citoyen dans sa personne et ses biens.

Parmi les principales missions relatives aux domaines de la PJ et de la Sécurité Publique (SP), nous citons :

- Suivre la ressource humaine spécialisée en police judiciaire.
- Diriger et contrôler l'emploi des personnels et l'activité des unités en matière de PJ.
- Concevoir, proposer et coordonner les perspectives de développement de la chaîne de PJ.
- Diriger et contrôler l'activité des unités, en matière de sécurité publique et de PJ.
- Evaluer l'activité des unités sur la qualité des services exécutés et des enquêtes diligentées et participer aux études et analyses relatives à la prévention et à la réduction de toute forme de criminalité.
- Evaluer les performances et les résultats obtenus en fonction des objectifs fixés.
- Analyser les phénomènes criminels et élaborer la carte de la criminalité, dans le but de dresser une analyse criminelle.
- Développer et promouvoir l'action de recherche, d'exploitation et de diffusion des renseignements.

IV.4.1. Système d'Information de la Police Judiciaire SI-PJ

Les services de sécurité en charge de la PJ ont recours à un SI-PJ dans un souci d'organisation, et de persistance des informations de plus en plus volumineuses, relatives à l'activité.

Un tel système constitue une banque importante d'informations, qui peut être exploitée à des fins d'analyse et de rapprochement criminel, mais aussi, à des fins d'analyse

décisionnelle et de reporting pour accompagner les opérations de sélection et de profiling quant aux agents et officiers de police judiciaire selon leur activité et leurs résultats réalisés.

V.4.1.1. Classes d'infraction

Les infractions sont classifiées et catégorisées, afin de faciliter l'analyse et d'orienter l'étude selon le type ou la classe d'infraction commise. Dans le code pénal, les infractions sont détaillées davantage, et sont regroupées en chapitres, et titres comme ci-contre.

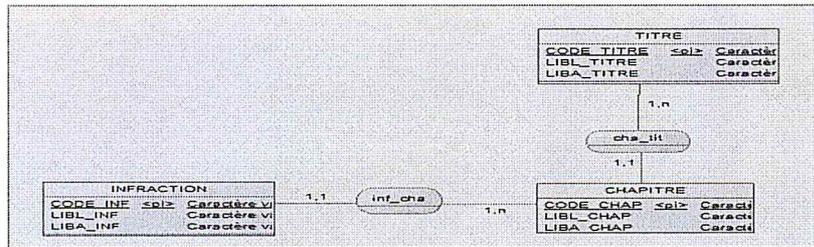


Figure IV-4 Classification de l'infraction

V.4.1.2. Livrables périodiques du SI-PJ

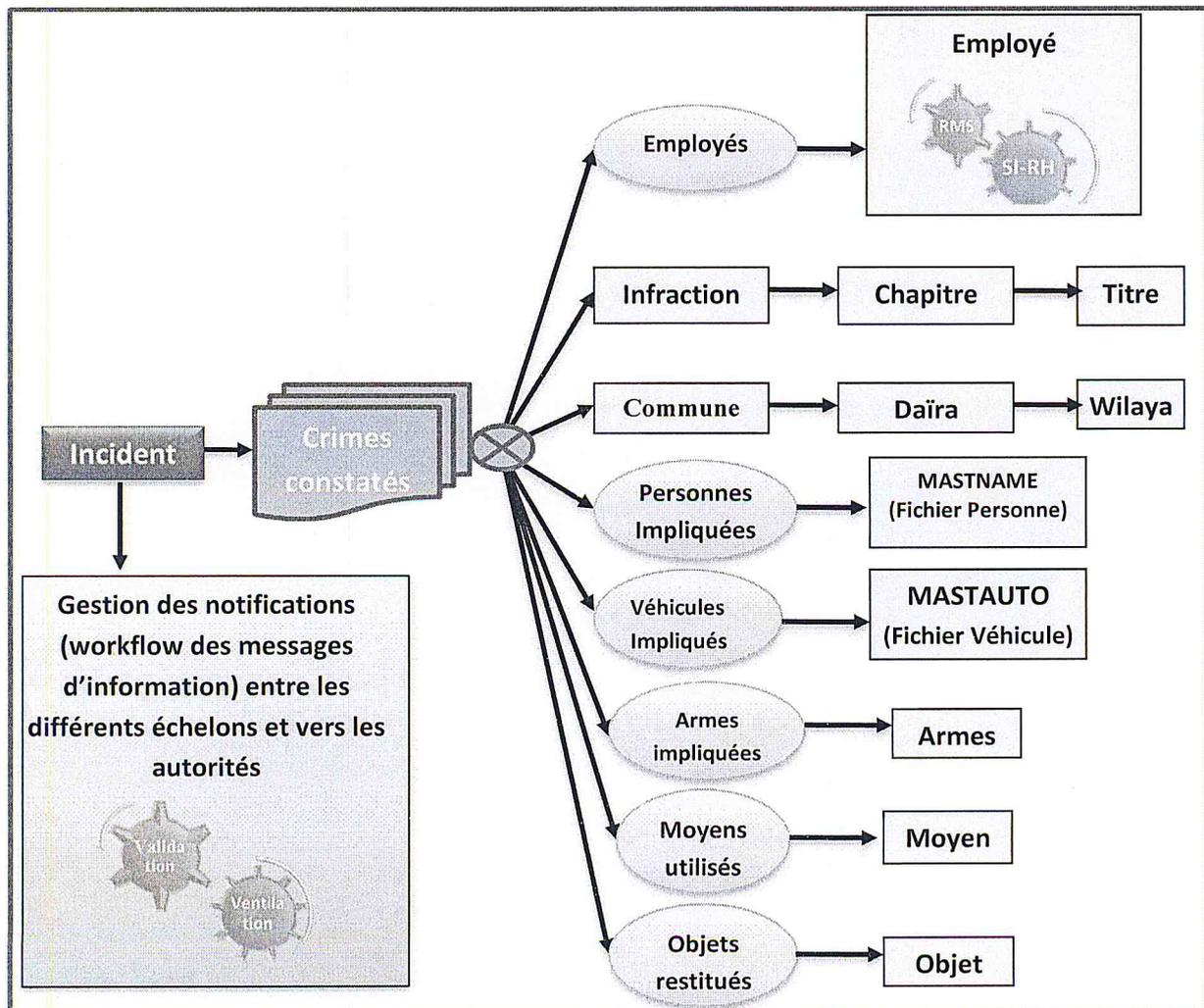


Figure IV-5 Modélisation de l'Incident dans le SI-PJ

(1) Au niveau central :

- Des rapports sur l'état de la criminalité et de la délinquance ;
- Des rapports sur les affaires résolues et/ou à suivre ;
- Des rapports par catégorie d'infraction (homicides, vols, stupéfiant, etc.) ;

(2) Au niveau local :

L'activité périodique des unités dans le domaine de la PJ :

Indicateurs d'agrégation :

- Nombre d'affaires traitées/en cours ;
- Nombre d'arrêtés écroués/libérés ;
- Nombre d'objets volés/restitués (camion, véhicule, moto, etc.) ;
- Quantité saisie ;
- La valeur vénale ;
- Taux d'élucidation des affaires ;

Ces indicateurs sont classifiés selon les types d'infractions, tel que :

- Vol de cheptel ;
- Vol de véhicule ;
- Agression des personnes suivie de vol ;
- Stupéfiant ;
- Contrebande ;
- Immigration clandestine ;
- Commerce illicite des boissons alcoolisées ;
- Commerce illicite du tabac.

Affaires importantes traitées :

- Date
- Unité
- Liste d'infractions
- Mode opératoire
- Auditions
- Arrestations
- Saisies
- Présentations
- Ecrous.

V.4.1.3. L'application de police judiciaire

L'application de la police judiciaire comporte trois (03) niveaux :

1- Le niveau opérationnel : Pour la saisie de l'activité de PJ, qui consiste en la saisie de toutes les affaires criminelles en cours de traitement, avec tout le détail u afférent tel que les personnes arrêtées, les véhicules, les produits ou les moyens saisis.

A noter que les personnes habilitées à mener ou agir dans des investigations judiciaires au sein de la Gendarmerie Nationale, sont les personnels ayant la qualité de police judiciaire, composée des officiers de police judiciaires (OPJ) et pour les assister, des agents de police judiciaire (APJ). Plusieurs types unités de la Gendarmerie Nationale ont pour missions, la police judiciaire. Elles peuvent ainsi exécuter des investigations judiciaires dans leurs compétence territoriale, tels que les brigades territoriales, les compagnies territoriales, les sections de recherche, les brigades de recherche, les brigades et les postes de police judiciaire.

2- Le niveau décisionnel : Ce niveau est réservé aux responsables en matière de police judiciaire tels que les Commandants de Groupement Territoriaux, etc.

3- Le niveau technique : d'administration et de maintenance du SI, par rapport aux ajouts, modifications et la création des états de sortie nécessaires aux autres niveaux.

4- Les interfaces de l'application : L'application est développée en mode web, avec l'outil Oracle Jdeveloper, et déployée sur un serveur d'application Oracle Weblogic. L'interface principale du module utilisateur, consiste en une interface de création et de consultation des affaires criminelles et de leurs états d'évolution dans l'investigation.



Figure IV-6 Portail du SI-PJ

IV.5.2. Etude de cas

Dans cette section, nous allons présenter notre étude de cas liée à la gestion de compétences du poste de Commandant de Groupement Territorial (Commandement au niveau wilaya). Nous allons procéder par une illustration démonstrative à travers quelques exemples d'expression du besoin. En fait, il s'agit de fixer des prérequis pour trois (03) cas de figure pour ledit poste. L'objectif de l'étude de besoins est d'avoir les indicateurs pertinents pour dégager des listes de candidats, selon un contexte donné, pour chaque poste.

IV.5.2.1. Description du Poste (commune)

- 1- Poste : **Commandant de Groupement Territorial ;**
- 2- Domaine d'emploi : **Opérationnel ;**
- 3- Famille fonctionnelle : **Sécurité publique ;**
- 4- Dominante : **Police Judiciaire ;**

IV.5.2.2. Conditions d'accès (commune)

- 1- Grade : **Lieutenant-colonel/Colonel ;**
- 2- Diplôme : **Etat-major ;**
- 3- Expérience professionnelle : **Avoir servi en unité territoriale ou de recherche ;**
- 4- Age minimum : **40 ans ;**
- 5- Aptitude physique : **Apte ;**

IV.5.2.3. Compétences principales :

Savoir

- 1- Management et Commandement - **Pilotage stratégique**
- 2- Gestion administrative du personnel - **Notation (aspects juridiques et procéduraux)**
- 3- Gestion administrative du personnel - **Procédure disciplinaire**

Savoir faire

- 1- **Communication interne**
- 2- **Communication externe en situation de crise**
- 3- Compétences transversales - **Travail en réseau/partenariat**
- 4- Compétences transversales - **Inspection/contrôle**
- 5- Compétences transversales - **Gestion de crise**
- 6- Sécurité publique - **Renseignement**
- 7- Sécurité publique - **Planification et conduite des opérations**
- 8- Management et Commandement - **Management d'équipe**

Savoir-être (Les qualités personnelles)

- 1- **Capacité à conceptualiser**
- 2- **Discrétion et confidentialité**
- 3- **Ouverture d'esprit et curiosité**
- 4- **Sens de l'analyse**
- 5- **Capacité à décider**
- 6- **Sens de la synthèse**

IV.5.2.4. Contexte du poste

Nous sommes intéressés par le contexte régional de wilaya, dont on s'intéresse, à priori, aux trois types de cas suivant :

- (1) **Contexte du poste n°1** : Zone frontalière Nord-ouest ;
- (2) **Contexte du poste n°2** : Zone haut plateaux (traitement par postes semblables) ;
- (3) **Contexte du poste n°3** : Zone frontalière Sud Est.

IV.6. Conclusion

Après examen des structures des bases de données, du flux de données entre les échelons, des procédures de saisie d'informations et des dossiers, il y a lieu de recommander, en matière de consolidation du SI-RH, la mise en place d'une couche d'aide à la prise de décision GRH adéquate et normalisée, permettant de quantifier et de valoriser toute entité qualitative intrinsèque à la personne ou à son emploi.

Dans ce chapitre, nous avons identifié le domaine de police judiciaire comme source principale de descriptifs métiers inhérents à notre étude. En effet, nous avons mené un travail d'étude et d'analyse du métier de la police judiciaire afin de comprendre le besoin qui est d'extraire à partir de l'application de la police judiciaire des indicateurs relatifs à l'activité de chaque employé. Afin de ne pas entraver le travail quotidien des unités opérationnelles, nous proposons d'utiliser le SI-PJ comme source de données pour la solution décisionnelle à mettre en œuvre.

Il convient aussi de souligner que l'action informatique visant à apporter une amélioration en matière de qualité de prise de décision au sein de la fonction RH, ne peut être entreprise qu'après la procuration d'outils de gestion nécessaire tels qu'un système d'appréciation chiffré, les fiches de postes et les socles ou les référentiels des compétences.

Chapitre V MODELISATION DE LA SOLUTION

V.1. Introduction :

Dans l'optique de faire concorder les compétences nécessaires et les compétences disponibles dans son capital humain vis-à-vis du savoir, savoir-être et savoir-faire, nous proposons un modèle dynamique basé sur du contexte pour concevoir un entrepôt de données tout en définissant l'ensemble de critères d'aide à la décision pour la gestion des compétences.

Cet entrepôt de données sera alimenté par les différentes bases de données de production citées précédemment.

Le modèle proposé va permettre de répondre aux besoins suivants :

- ❑ La possibilité de proposer des fiches de postes dynamiques, où les missions peuvent être classées par priorités selon le contexte du travail (régions, temps, l'évolution de la criminalité en Algérie, la situation géo-sécuritaire de l'Algérie et les orientations du commandement).
- ❑ Proposer les meilleurs candidats pour un poste donné, en tenant compte de leur expérience, de leur compétence et de tout autre critère jugé utile.

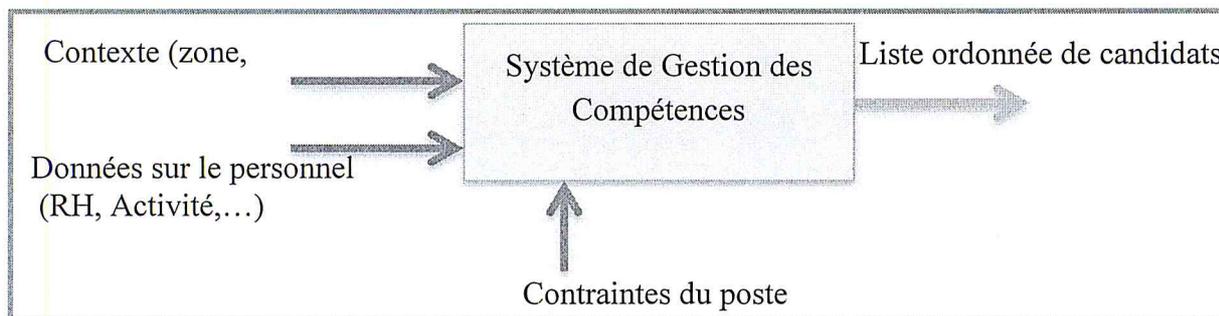


Figure V-1 Schéma global du système de gestion des compétences

Dans un système de gestion de compétences, les candidats pour un poste donné ont besoin d'être estimés et distingués en vue d'un emploi optimal. Pour se faire, ces candidats seront évalués continuellement selon tous les volets régissant l'activité de l'institution cible, entre autres le savoir, savoir-être et savoir-faire. A cet effet, un entrepôt de donnée va nous servir de conteneur pour ces évaluations.

Cependant, les critères d'évaluation n'ont pas tous le même degré d'importance pour toutes les instances d'un poste donné notamment par rapport à la vision du décideur, à la diversité du territoire national (aspect géographique) et au changement continu (aspect temporel) du contexte sécuritaire national, régional et international. Ceci exige l'introduction d'une couche supplémentaire qui donne une valeur contextuelle. Vu le temps réservé pour ce travail, nous allons nous limiter à modéliser le contexte par des pondérations. Néanmoins, le contexte peut être l'objet d'autres modélisations ou de pondérations différentes.

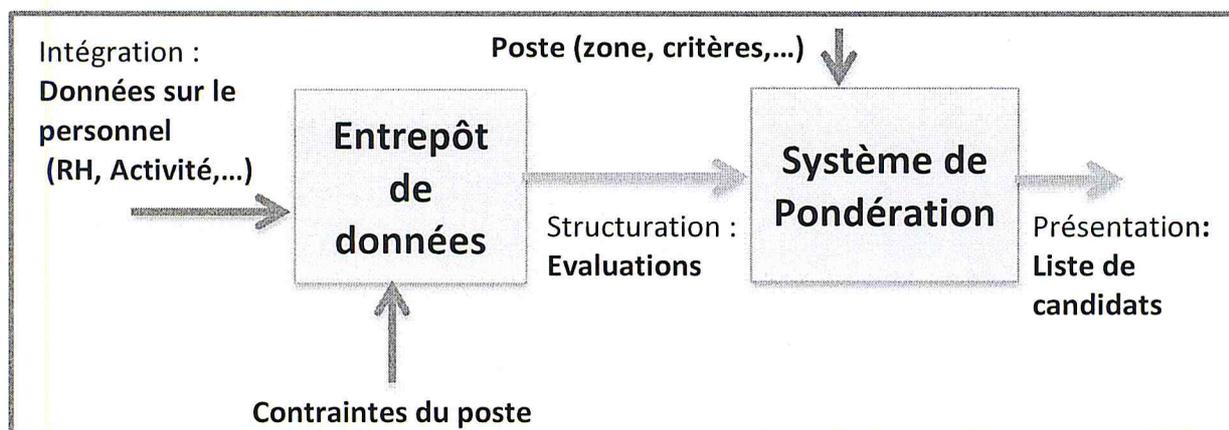


Figure V-2 Schéma global de combinaison des pondérations aux évaluations des candidats

Tous les contextes cités précédemment seront pondérés et combinés adéquatement pour pouvoir se prononcer sur l'adéquation des candidatures.

V.2. Compétence et évaluation

Les **compétences** sont des répertoires de comportements que certaines personnes maîtrisent mieux que d'autres, qui les rendent efficaces dans une situation donnée.

Ces comportements sont observables dans la réalité quotidienne du travail et également dans des situations-tests. Ils mettent en œuvre, de manière intégrée, des aptitudes, des traits de personnalité, des connaissances acquises.

Les compétences représentent donc un trait d'union entre les caractéristiques individuelles et les qualités requises pour mener à bien des missions professionnelles précises.

Plusieurs questions sont à poser comme : de quelles compétences s'agit-il ? et peut-on constituer des listes de compétences qui représentent une référence universelle dans les organisations ou doit-on, au contraire, adapter les compétences à chaque organisation, voire à chaque poste ou même à chaque situation ?

De l'analyse des réponses, nous avons constaté le manque de liaisons entre les deux aspects **caractéristiques individuelles** et **missions à accomplir**. Les méthodes d'analyse de poste traditionnelles visent soit les aptitudes et les traits de personnalité, soit les activités de travail, mais rarement les compétences. Les questionnaires en ressources humaines doivent donc construire eux-mêmes les référentiels de compétences de leur organisation.

V.3. L'approche du modèle proposé :

Notre approche prend en compte trois types de compétences : les compétences de base, les compétences d'activité, et celles induites par le diplôme. Dans ce troisième type sont incluses également toutes les contraintes requises.

Les compétences de base induisent la capacité de l'individu à s'adapter à son environnement et à acquérir les valeurs de l'institution (capacité d'analyse, les relations humaines, l'initiative et la communication écrite et orale).

Les compétences d'activité, quant à elles, varient d'un rôle à un autre. Elles relèvent de l'aptitude de chacun à utiliser des qualités personnelles ou interpersonnelles pour valoriser son rôle à un poste donné, et donc, aboutir à des résultats plus efficaces pour les mêmes objectifs donnés au préalable.

Le potentiel d'un individu sera donc défini comme un ensemble de comportements, signes de la possession de capacités et de compétences considérées elles-mêmes comme des indicateurs de ce potentiel recherché. Autrement dit, des résultats meilleurs de ceux d'autres candidats seront analysés comme le signe d'une plus grande capacité de vision, d'influence, d'esprit d'initiative...

La potentialité d'un individu est donc mesurée (évaluée) par la réunion de compétences de base, opérationnelles, mais surtout sur la capacité de valoriser ses connaissances.

V.3.1 Motivation d'évaluation des compétences

Le potentiel est une supposition d'évolution forte à propos d'un individu pour une durée déterminée dans une filière professionnelle. La compétence devient donc une capacité à évoluer dont la valeur change dans le temps. Pour assurer sa progression, l'organisation se transforme en fonction des décisions internes et des pressions de l'environnement.

La capacité d'adaptation des compétences de son personnel dépend en grande partie de l'aptitude de l'organisation à choisir les hommes, à les préparer à de nouvelles fonctions, en tenant compte de leurs aspirations et de leurs capacités potentielles. La gestion des carrières ou le développement du potentiel humain d'une organisation vise donc à atteindre le meilleur équilibre possible entre les besoins des structures et les potentiels. L'évaluation des compétences comporte plusieurs aspects qui vont conduire à des modifications de l'organisation. En effet, L'évaluation complète et objective de la compétence aide les organisations à :

- **Détecter le potentiel nécessaire pour accéder à un poste donné**, notamment dans une organisation comme la GN où le taux de mutation est important ;
- **Détecter les manques de compétences de certains individus**, qui auront un impact direct sur le programme de formation et de perfectionnement.
- **mettre en avant les compétences** des individus **non utilisées dans leur activité**, ce qui aide à mieux exploiter les ressources humaines dans une organisation ou entreprise.

V.3.2 Architecture du modèle proposé

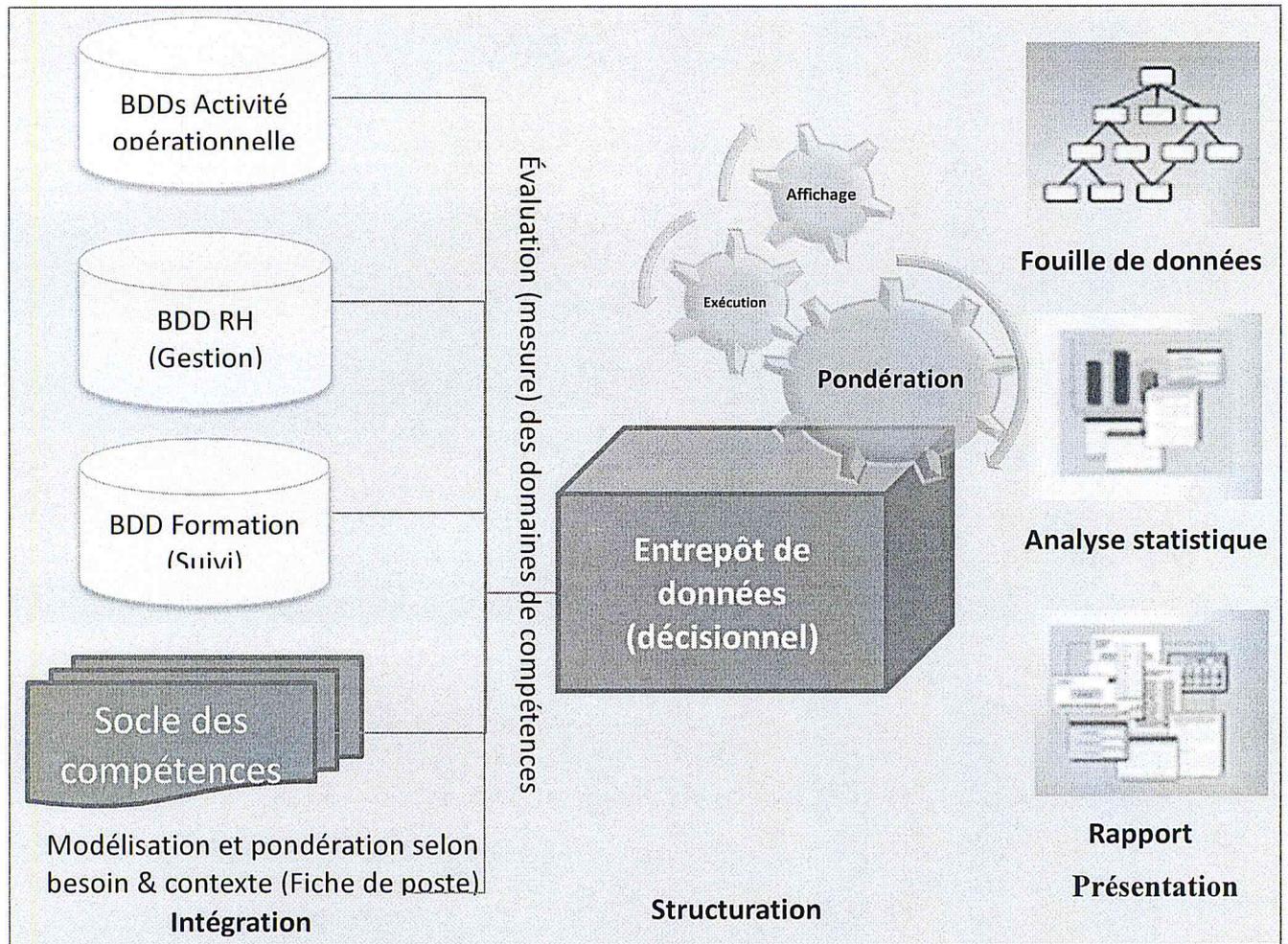


Figure V-3 Architecture du modèle proposé

V.3.3 Processus

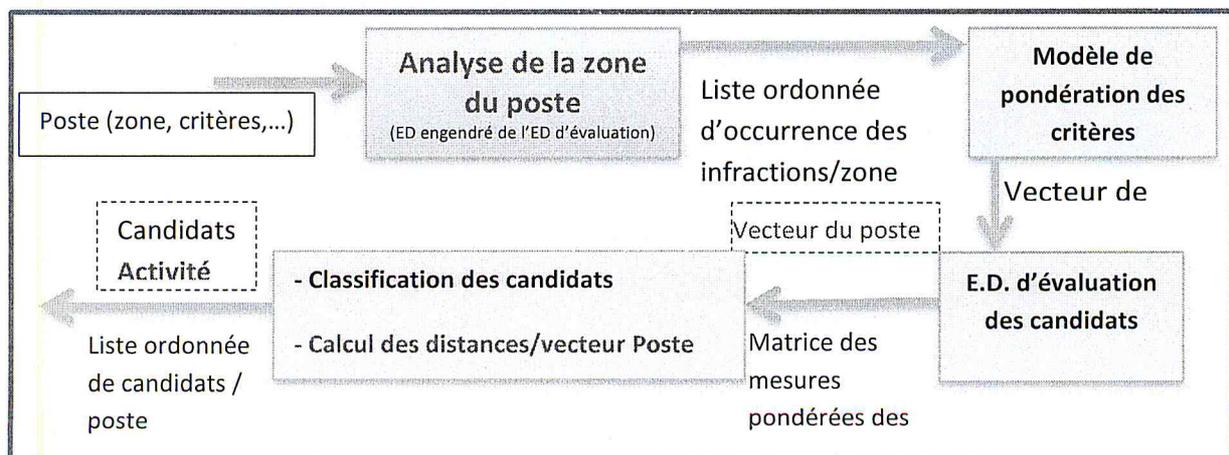


Figure V-4 Processus de l'approche proposée

Le défi à surmonter dans un tel système est d'explicitier, autant que possible, tous les critères et indicateurs permettant de formaliser le potentiel d'un candidat par des valeurs numériques. Cette formalisation peut être réalisée par plusieurs méthodes, à savoir : (i) l'évaluation explicite ou directe basée sur l'observation du chef hiérarchique qui peut, le cas échéant, prévoir un entretien notamment pour les nouveaux recrues, ou bien sur une épreuve sous forme de concours ou examens. (2) l'évaluation implicite basée sur la quantification des participations et contributions des individus aux différents services effectués particulièrement les affaires élucidées.

De ce qui précède, nous envisageons de concevoir notre système décisionnel composé des deux parties : évaluation et activité, tout en ayant recours à la modélisation dimensionnelle qui est souvent associée aux entrepôts de données compte tenu des avantages comme cités précédemment.

V.4. Conception du Data Warehouse

V.4.1 Processus de la modélisation dimensionnelle

La conception d'un modèle dimensionnel passe par cinq (05) étapes essentielles :

- 1) Choix de l'activité à modéliser ;
- 2) Définition de l'activité et de son grain ;
- 3) Définition des dimensions qui décrivent une ligne de la table de fait ;
- 4) Définir les mesurables du fait ;
- 5) Définir les agrégats.

V.4.1.1. Choix de l'activité à modéliser :

Il s'agit de classer les activités recensées par ordre moyennant une matrice dite *d'analyse des priorités* [27]. En réalité, le classement s'effectue en impliquant et les décideurs et les techniciens de l'organisation.

Des entretiens avec les gestionnaires, il ressort que les décideurs favorisent les systèmes décisionnels orientés le plus possible vers l'évaluation explicite des compétences (les grilles d'évaluation du candidat) que vers la quantification de l'activité. Ceci est motivé par la qualité supérieure et la disponibilité des données de production du fait de la responsabilisation accrue de l'évaluateur (le chef hiérarchique ou le chargé d'examen).

Certes les données dont repose la quantification de l'activité sont tributaires de la qualité de saisie dans le système opérationnel. En effet, l'expérience a montré qu'il est difficile de cerner la saisie quand les applications sont déployées à grande échelle (unités de base). Cependant, l'utilisation de telles mesures implicites, basées sur l'activité, diminue la subjectivité tant reprochée à l'évaluation explicite sans pour autant la remplacer.

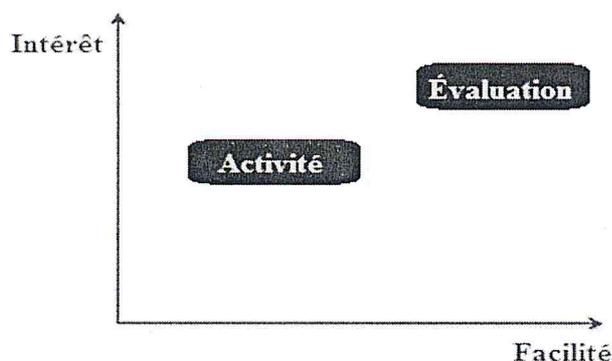


Figure V-5 Analyse des priorités du cas « Gestion des compétences »

Evaluation : l'évaluation explicite basée sur l'observation du chef ou bien sur une épreuve.

Activité : l'évaluation implicite basée sur la quantification des services effectués.

V.4.2 Volet : Evaluation explicite du candidat :

En réalité cette partie fusionne deux types d'évaluations à savoir l'évaluation directe par rapport à des critères et l'évaluation sur épreuve. Pour l'évaluation directe, Il est question d'énumérer d'abord tous les critères relatifs aux différents volets régissant l'organisme cible pour pouvoir façonner cette évaluation. Pour cela, on a exposé un socle de compétences qui est un recueil des compétences sollicitées par les différentes missions et responsabilités au sein du même organisme. Chacune de ces compétences englobent une liste de domaines plus fins qu'on assimile à des critères à évaluer chez un individu afin de le classer parmi l'ensemble des candidats.

Pour l'évaluation sur épreuve, les indicateurs sont issus des notes obtenues par le candidat dans les différents concours et examens passés le long de sa carrière. Ces notes ont été introduites dans le SI-Formation et transférées pour persistance dans le système SI-RH.

V.4.2.1. Grain de l'activité

Opter pour le grain le plus fin, i.e. le niveau de détail le plus bas, donne plus de souplesse. Pour les évaluations des compétences de base, le grain le plus fin correspond à une

opération d'évaluation aussi bien pour un critère du socle des compétences que pour un examen ou un concours, d'où une ligne de table de fait correspondant à :

Besoin à modéliser pour l'évaluation des compétences de base : Suivi de l'évaluation d'un candidat par domaine de compétence **ou** épreuve passée à une date donnée (jour, mois, trimestre, semestre et années).

V.4.2.2. Les dimensions du modèle :

Pour décrire le fait, nous essayons de dresser toutes les informations décrivant une évaluation et pouvant intéresser les décideurs.

La Dimension Temps

TD_Temps
<<pk>> ID_Temps
Date
Jour
Mois
Trimestre
Semestre
Année

Figure V-6 La Dimension Temps du volet « Evaluation directe »

Le niveau de détail le plus bas de cette dimension est la journée. En effet, les gestionnaires ont fait ressortir le besoin de suivre les évaluations, notamment pour les épreuves, au jour le jour et d'en garder l'historique de ces derniers.

Cette dimension utilise une clé artificielle comme clé primaire à fin de faciliter la manipulation de la dimension. Plus amples informations sur cette dimension sont données dans le tableau suivant :

Désignation	Détails
ID_Temps	Clé artificielle de la dimension temps.
Date	La date au format complet.
Jour	Position du Jour dans le mois.
Mois	Numéro du mois dans l'année.
Annee_mois	Année et mois (concaténation).
Trimestre	Trimestre de la date.
Trimestre_annee	Trimestre et année (concaténation).
Semestre	Semestre de la date.
Semestre_annee	Semestre et année (concaténation).
Annee	Année de la date.

Tableau V-1 Tableau descriptif de la dimension « Temps »

(1) La Dimension Candidat : du volet « Evaluation »

Le candidat s'impose comme un élément important dans l'analyse et intéresse les analystes et les décideurs de l'organisation. L'analyse des compétences de base du candidat peut aider à un emploi optimal.

Commune aux deux parties du système décisionnel préconisé, la dimension candidat contient, entre autres, les informations sur le grade, la fonction, le diplôme, la spécialité et l'unité. Vue la spécificité de cette dernière par rapport aux besoins d'analyse, elle est détaillée encore par le type d'unité et la zone d'unité. Ce détail est vital pour valoriser l'expérience du candidat par rapport à un type donné d'unité et/ou de zone.

Quelques informations seront calculées et intégrées à la dimension candidat lors du chargement des données à savoir :

Aptitude au séjour sud (Apt_Sud) : cette information est nécessaire pour répondre au besoin d'information sur l'accomplissement ou non d'un certain parcours obligatoire dans la carrière, en l'occurrence le sud. Cette information sera déduite à partir du SI-RH.

Aptitude Physique, Psychologique ou familiale : ces informations peuvent être des contraintes quant à l'octroi d'un poste. Un exemple concret est le fait d'absence de prise en charge d'un cas médical ou social donné dans la zone d'implantation du poste en question. Ces données sont extraites à partir de la base de données du SI-RH. Donc les caractéristiques du candidat sont:

Désignation	Détails
ID_Candidat	Clé artificielle
Matricule	Le matricule du candidat, clé d'affaire
Nom	Nom du candidat
Prenom	Prenom du candidat
DateNaissance	Date de naissance du candidat
Sexe	Sexe du candidat
Nb_Mariage	Nombre de mariages du candidat
Nb_Enfant	Nombre d'enfant du candidat
Age_Max_Enf	Age de l'enfant aîné du candidat
DateRecrute	Date du recrute du candidat
Lib_Grade	Le grade actuel du candidat
Categ_Grade	Catégorie du grade actuel du candidat
Date_ChPos	Date de Changement de position du candidat
Lib_Fonction	Fonction actuelle occupée par le candidat

Lib_DiplômeMil	Plus haut diplôme militaire détenu par le candidat
Lib_SpécialitéMil	Spécialité militaire du candidat
Année_Dip_Mil	Année d'obtention du diplôme militaire
Lib_DiplômeCiv	Plus haut diplôme civil détenu par le candidat
Lib_SpécialitéCiv	Spécialité civile du candidat
Année_Dip_Civil	Année d'obtention du diplôme civil
Date_LastMut	Date de la dernière mutation du candidat
Date_Nom_Grd	Date de nomination dans le grade actuel du candidat
Lib_Unité	L'unité d'appartenance actuelle
Code_Wilaya	Wilaya de l'unité actuelle
Code_Region	Région de l'unité actuelle
LIB_TypeUnité	Type de l'unité actuelle du candidat
LIB_ZoneUnité	Zone de l'unité actuelle du candidat
DateFinContrat	Date de fin du contrat avec le candidat
DateDerContrat	Date du dernier contrat avec le candidat
Durée_Engage	Durée d'engagement du candidat
DateAdmi	Date admission du candidat
- Apt_Physique	Aptitude physique du candidat
- Apt_Psycho	Aptitude Psychologique du candidat
- Apt_Sud	Aptitude du candidat pour le service au Sud du pays
- Apt_Famille	Aptitude socio-familiale du candidat

Tableau V-2 Tableau descriptif de la dimension « Candidat »

(2) La Dimension Domaine de Compétence

Comme évoqué précédemment, l'évaluation sera mesurée autour de critères énumérés par rapport aux différents volets régissant l'organisme cible. Ces critères sont organisés en socle comportant deux cent quatre-vingt-quatre (284) domaines de compétences catégorisés en vingt-deux (22) compétences. Approuvé et validé par les services et structures concernés, ce socle de compétences est prêt pour être intégré à l'aspect appréciation et évaluation du système de production SI-RH.

1)	Sécurité publique
2)	Ordre public
3)	Sécurité routière
4)	Police judiciaire
...	
13)	Les qualités personnelles (Savoir-être)
14)	Télématique
...	
22)	Intendance

Figure V-7 Aperçu de la liste des 22 compétences recensées

N°	Compétence	Domaine de compétence
01	Sécurité publique	1) Renseignement
		2) Intervention lors de plans de secours, catastrophes
		3) Planification et conduite des opérations
		4) Police administrative
		...
		12) Technique de patrouille

Tableau V-3 Aperçu des domaines (non ordonnés) inhérents à une compétence donnée

(3) La Dimension Epreuve

L'évaluation sur épreuve est effectuée selon la liste des matières et disciplines nommées au niveau du service correspondant. Ces matières sont catégorisées selon des types plus génériques.

TD_Epreuve
<<PK>> ID_Epreuve
- LIB_Epreuve
- LIB_TypeEpreuve

Figure V-8 La Dimension Epreuve du volet «Evaluation directe»

V.4.2.3. Les mesurables :

Le schéma décisionnel relatif à cette partie comporte une table de faits qui englobe comme mesure l'évaluation (la note) et le rang. Ces indicateurs seront capitalisés par candidat, domaine de compétence, épreuve, et date d'évaluation.

Puisque il s'agit de deux types d'évaluation, décernée par le chef hiérarchique et attribuée suite à une épreuve, une ligne de table de fait correspondra à un seul type qui pourra être exclusivement une occurrence du domaine de compétence ou d'épreuve. L'autre type étant à une valeur neutre «**non applicable**» pour cette même ligne.

La note d'évaluation par rapport à un domaine du socle émane de la grille d'évaluation relative à chaque candidat dans le système de production. Cette grille est alimentée par les chefs hiérarchiques durant la période de service. Ces même grilles d'évaluation sont déduites à partir dudit socle selon la structure/l'unité du candidat. La notation est basée sur une échelle à six valeurs : (1/6 : Insuffisant, 2/6 : Médiocre, 3/6 : Bien, 4/6 : Très Bien, 5/6 : Excellent et 6/6 : Elite).

Les deux indicateurs envisagés pour l'évaluation sur épreuve sont la note obtenue et le rang par épreuve. L'observation de ces deux mesures est justifiée par le fait de leur complémentarité pour dénoter une évaluation plus correcte (voir illustration plus bas). En effet, une note élevée dans une épreuve donnée ne signifie pas forcément supériorité et maîtrise absolues sans la positionner par rapport aux notes du reste des concurrents, car les épreuves ne sont pas toutes du même **niveau de difficulté**.

Illustration : si on fait une analyse par rapport à la **difficulté des épreuves** on aura :

Cas de Difficultés	Choix du candidat à base de la note
Epreuves 1&2 modérées	Mêmes niveaux de difficulté des épreuves, la meilleure note l'emporte. Abstraction de la difficulté, la note suffit pour choisir.
Epreuves 1&2 difficiles	
Epreuve 1 difficile Epreuve 2 modérée	Si la note liée à l'épreuve la plus difficile est supérieure alors elle l'emporte, et donc la note suffit pour choisir.
Epreuve 1 modérée Epreuve 2 difficile	Sinon faire face à un problème de choix : favoriser la note Vs le niveau élevé et donc la meilleure note n'est pas forcément le bon choix, la note ne suffit pas.

Tableau V-4 Illustration d'indicateurs pour l'évaluation sur épreuve par rapport aux niveaux de difficulté

Du tableau, on constate qu'on peut immédiatement faire le choix d'après la note seulement dans le cas où on le niveau de difficulté est le même pour toutes les épreuves. Puisque ce cas n'est pas toujours vrai, le choix est tributaire du niveau de difficulté des épreuves (d'après le tableau). En effet, on doit se fier à un témoin plus objectif mais pratique pour dénoter cette difficulté. Ce témoin est le rang.

Il faut rappeler que pour l'homogénéité des mesures, l'évaluation est absolue, elle est exprimée en fraction (entre 0 et 1). Et puisque elle est proportionnelle au choix, tandis que le rang l'est en inverse, le rapport : **Eval = Note/Rang** désigne un indicateur plus précis quant à l'évaluation sur épreuve.

V.4.2.4. Le modèle en étoile de l'évaluation de base :

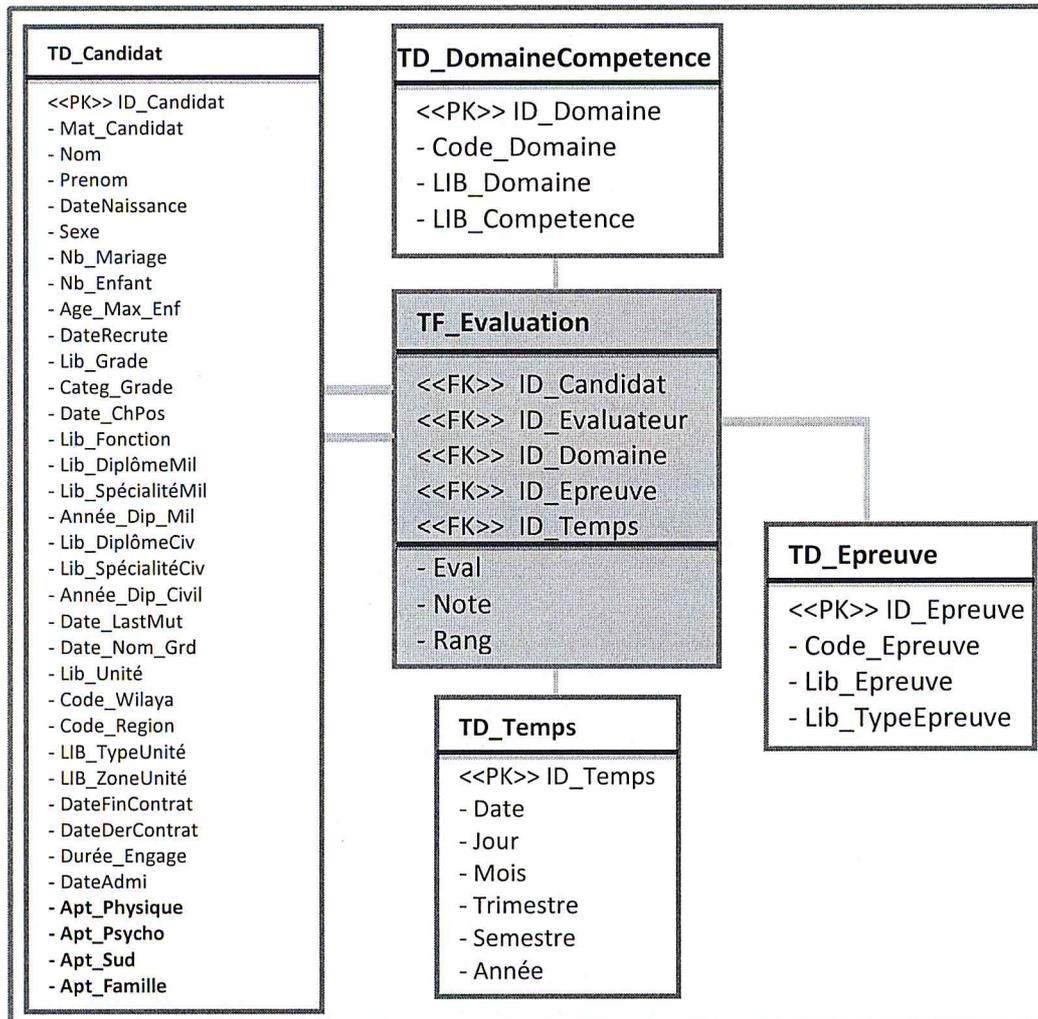


Figure V-9 Le modèle en étoile d'évaluation explicite des compétences

Lecture

Lecture abstraite : Le candidat **c1** avec sa description selon la table de la figure ci-dessus, s'est fait décerné, pour la date **dt1**, la valeur **Eval1 (fait1)**, pour le domaine de compétence particulier **d1**. Ce même candidat a obtenu, en date du **dt2**, la note **Note1** et a eu le classement **Rang1 (fait2)**, à l'épreuve **e1**.

Exemple concret : Le candidat **Lieutenant-colonel BOURAHLA Smail**, s'est fait attribué pour l'année **2017** la valeur **6(Elite)**, pour le domaine particulier de compétence **Techniques d'enquêtes judiciaires**. Ce même candidat a obtenu en date du **01/06/2010** un **14/20** et a été classé **5^{ème}**, à l'épreuve du **MRT(Méthodologie du Raisonnement Tactique)**.

V.4.2.5. Les agrégats

Dans ce qui suit on s'est inspiré de la démarche décrite par *C. Adamson* dans son livre «Mastering the Data Warehouse Aggregates, Solution for Star Schema Performance».

(1) Enumérer les agrégats potentiels/utiles à partir d'une étoile détaillée

Pour détecter les agrégats potentiels et choisir ceux à implémenter dans le Data Warehouse. Il est nécessaire de bien décrire chaque agrégat.

Le tableau suivant décrit, d'une manière simple et efficace, les agrégats potentiels du modèle dimensionnel de base du volet évaluation explicite :

Dimension	Agrégats potentiels	Nb agrégats possibles
Temps	Mois, trimestre, semestre, année	4
Candidat	Lib_Grade, Lib_Fonction, Lib_DiplômeMil, Lib_SpécialitéMil, Lib_Unité, Wilaya, Region, LIB_TypeUnité, LIB_ZoneUnité	9
Domaine de compétence	LIB_Domaine, LIB_Compétence	2
Epreuve	LIB_Epreuve, LIB_TypeEpreuve	2

Tableau V-5 Liste des agrégats potentiels pour le volet «Evaluation explicite»

Les agrégats potentiels ne sont pas forcément tous utiles, soit par le nombre de lignes agrégées ou par les informations fournies.

Si on élimine les agrégats non pertinents à partir des agrégats potentiels, on aura :

Dimension	Agrégats pertinents	Nb agrégats possibles
Temps	semestre, année	2
Domaine de compétence	LIB_Domaine, LIB_Compétence	2
Epreuve	LIB_Epreuve, LIB_TypeEpreuve	2

Tableau V-6 Descriptif des agrégats utiles du volet «Evaluation explicite»

(2) Construire le modèle agrégé

Pour la construction du modèle agrégé, on prend en considération les dimensions dérivées commune entre les différents modèles.

A partir du tableau précédent nous choisissons les agrégats qui nous semblent les plus pertinents et susceptibles de faire l'objet d'accès fréquents. Nous arrêtons la liste des modèles d'agrégats suivants :

- Evaluation **annuelle** par **candidat** par **compétence**.
- Evaluation **annuelle** par **candidat** par **domaine** de compétence.

- Evaluation **semestrielle** par **candidat** par **épreuve**.
- Evaluation **semestrielle** par **candidat** par **type** d'épreuve.
- Evaluation **annuelle** par **candidat** par **épreuve**.
- Evaluation **annuelle** par **candidat** par **type** d'épreuve.

Les agrégats sont conçus, en général, comme des modèles dimensionnels. La modélisation des agrégats se fait grâce aux principes de la modélisation dimensionnelle.

V.4.3 Volet : Quantification de l'activité du candidat

Les mesures du volet précédent (évaluation explicite pour un candidat) sont consolidées par des indicateurs chiffrés issus de l'activité opérationnelle (savoir-faire). Ces derniers sont calculés sur la base des contributions dans les différents services exécutés.

Les informations de ce volet proviennent, normalement, des différents systèmes de productions dédiés à toute l'activité administrative, technico-logistique et opérationnelle.

Néanmoins, pour cadrer notre étude dans les limites des systèmes de production existants et fonctionnels, nous allons nous intéresser aux informations en rapport avec l'activité opérationnelle seulement, notamment les affaires liées aux infractions dans le domaine de la police judiciaire. Comme ces infractions appartiennent à un éventail important de catégories (chapitre et titre), il est possible de faire, entre autres, extraire une spécialisation (profiling) des candidats afin de répondre à un contexte global ou à un besoin du commandement pour l'attribution d'un poste donnée ou encore pour la désignation à une formation ou à une mission.

Dans une implémentation avec un entrepôt de données, le schéma décisionnel relatif à cette partie du système pivote autour d'une table de faits qui englobe comme mesures les indicateurs d'activité cités précédemment. Ces indicateurs seront déduits des valeurs liés au candidat, type de service ou d'affaire, lieu du service et date du service. Ce qui a généré les tables de dimensions correspondantes (voir figure ci-après).

Pour plus d'objectivité dans l'évaluation, nous avons proposé de mettre en place aussi la dimension lieu du service pour prendre en compte la disparité et l'hétérogénéité des zones d'activités.

Désignation	Détails
LIB_LieuService	Niveau hiérarchique de la commune où l'affaire a eu lieu.
LIB_WilayaSvc	Niveau hiérarchique de la wilaya
LIB_RegionSvc	Niveau hiérarchique de région

Tableau V-8 Descriptif de la dimension « TD_LieuService »

V.4.3.3. Les mesurables :

A ce niveau, Il s'agit encore d'énumérer tous les indicateurs inhérents aux différents types d'affaires pour délimiter les aspects mesurables afin d'accomplir l'évaluation.

Pour cela, nous avons exposé les indicateurs suivants :

- Nombre d'infractions relevées ;
- Nombre de personnes impliquées : dénote l'ampleur de l'affaire et la masse de travail (auditions, etc.) ;
- Nombre de personnes arrêtées : mesure l'efficacité des travaux exécutés;
- Nombre de personnes écrouées : mesure l'aboutissement des travaux exécutés et la coordination avec les autorités judiciaires;
- Valeur restituée : conversion en montant des marchandises et biens saisis ;
- Nombre de décès ;
- Nombre de blessés ;
- Estimation du dégât matériel ;
- Nombre de participants : mesure l'apport dans l'exécution du service;
- Nombre de preuves collectées : mesure l'effort fourni dans l'observation, la fouille et la collecte des pièces à conviction;
- Nombre de pièces financières : quantifie l'expertise financière faite ;
- Nombre d'expertises abouties : mesure l'efficacité et l'aboutissement des réquisitions d'expertise générées;
- Contribution dans l'exécution du service ou l'élucidation de l'affaire ...;

Ces mesures sont le résultat des agrégations faites sur les informations reliées à une affaire menée dans le cadre de l'activité quotidienne sur le terrain.

Ces indicateurs seront capitalisés par candidat, Infraction, Lieu de service, et date de l'affaire.

V.4.3.4. Le modèle en étoile de l'évaluation par l'activité

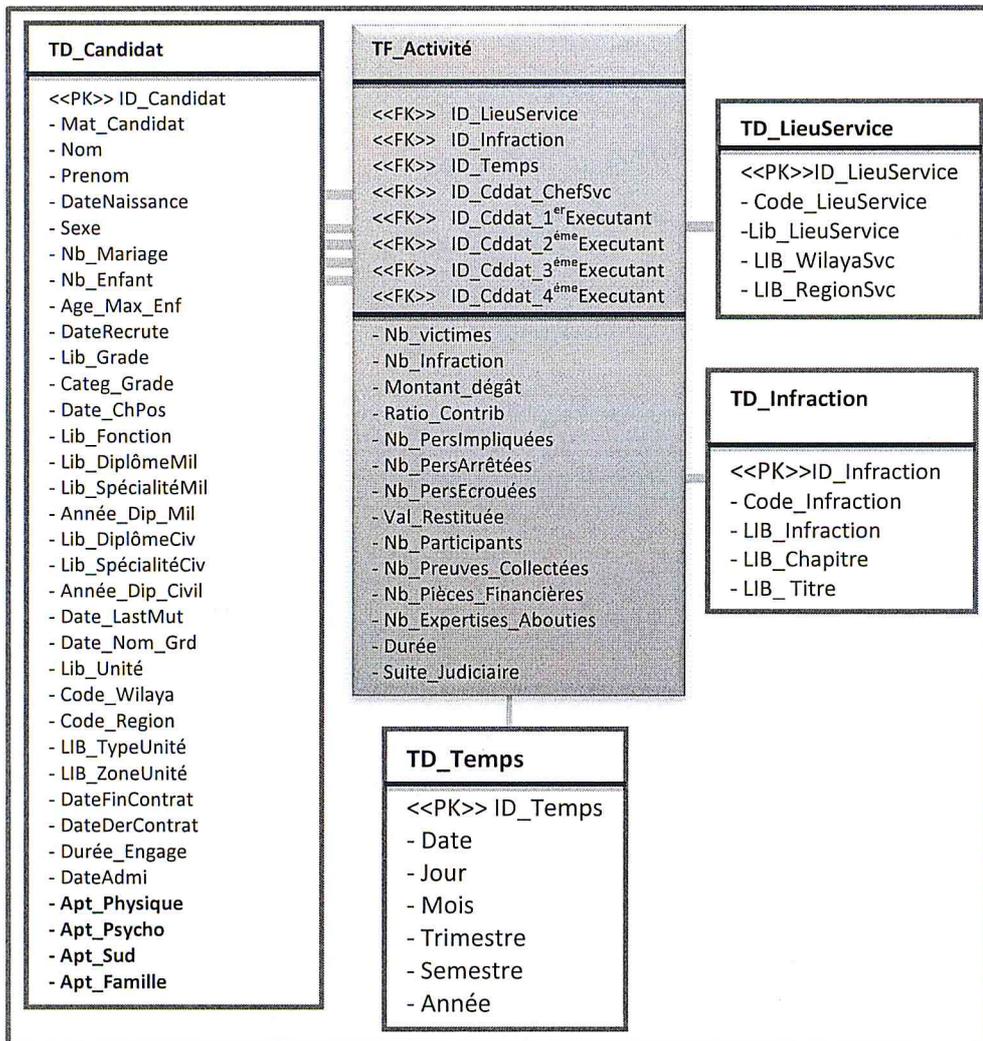


Figure V-12 Schéma décisionnel de l'activité par candidat, lieu de service, infraction et date

Lecture :

Lecture abstraite : Le candidat **c1** avec sa description selon la table de figure ci-dessus, a réalisé le jour **dt1** au niveau de la zone **z1** une affaire traitant l'infraction **i1**, le résultat était de : **Nb_Victimes1, Nb_PersImpliquées1, Nb_Personne_arrêtée1, Val_Restituée1,**

Exemple de valeurs : Le candidat **Lieutenant-colonel BOURAHLA Smail** avec sa description selon la table de figure ci-dessus, a réalisé le **12/01/2017** au niveau de la zone **Debdeb** une affaire traitant l'infraction **délits de contrebande**, le résultat était de : **03 personnes impliquées dont 02 arrêtées, et 5 000 000 DA** comme valeur restituée,

V.4.3.5. Les agrégats

(1) Enumérer les agrégats potentiels/utiles à partir d'une étoile détaillée

Le tableau suivant décrit les agrégats potentiels du volet évaluation par activité :

Dimension	Agrégats potentiels	Nb agrégats possibles
Temps	Mois, trimestre, semestre, année	4
Candidat	Lib_Grade, Lib_Fonction, Lib_DiplômeMil, Lib_SpécialitéMil, Lib_Unité, Wilaya, Region, LIB_TypeUnité, LIB_ZoneUnité	9
Infraction	LIB_Infraction, LIB_Chapitre, LIB_Titre	3
Lieu de service	LIB_LieuService, LIB_WilayaSvc, LIB_RegionSvc	3

Tableau V-9 Liste des agrégats potentiels pour le volet «évaluation par activité»

Les agrégats potentiels ne sont pas forcément tous utiles, soit par le nombre de lignes agrégées ou par les informations fournies.

Afin de réduire la liste des agrégats, on élimine les agrégats inutiles à partir des agrégats potentiels.

Dimension	Agrégats potentiels	Nb agrégats possibles
Temps	semestre, année	2
Infraction	LIB_Infraction, LIB_Titre	2
Lieu de service	LIB_WilayaSvc,	1

Tableau V-10 Tableau descriptif des agrégats utiles du volet «évaluation par activité»

(2) Construire le modèle agrégé

Pour la construction du modèle agrégé, on prend en considération les dimensions dérivées commune entre les différents modèles.

A partir du tableau précédent nous choisissons les agrégats qui nous semblent les plus pertinents et susceptibles de faire l'objet d'accès fréquents. Nous arrêtons la liste des modèles d'agrégats suivants :

- Evaluation **semestrielle** par **candidat** par **infraction** par **wilaya** ;
- Evaluation **annuelle** par **candidat** par **infraction** par **wilaya** ;
- Evaluation **semestrielle** par **candidat** par **titre d'infraction** par **wilaya** ;
- Evaluation **annuelle** par **candidat** par **titre d'infraction** par **wilaya** ;

La zone d'entreposage constitue la zone exploitable par les utilisateurs. La modélisation de cette zone se fait grâce à la modélisation dimensionnelle. Cette manière de représenter les données offre aux utilisateurs des modèles intuitifs et compréhensibles permettant de naviguer et de manipuler les données, détaillées ou agrégées, afin de satisfaire leurs besoins en analyse.

V.4.4 Le modèle en constellation de l'évaluation des candidats

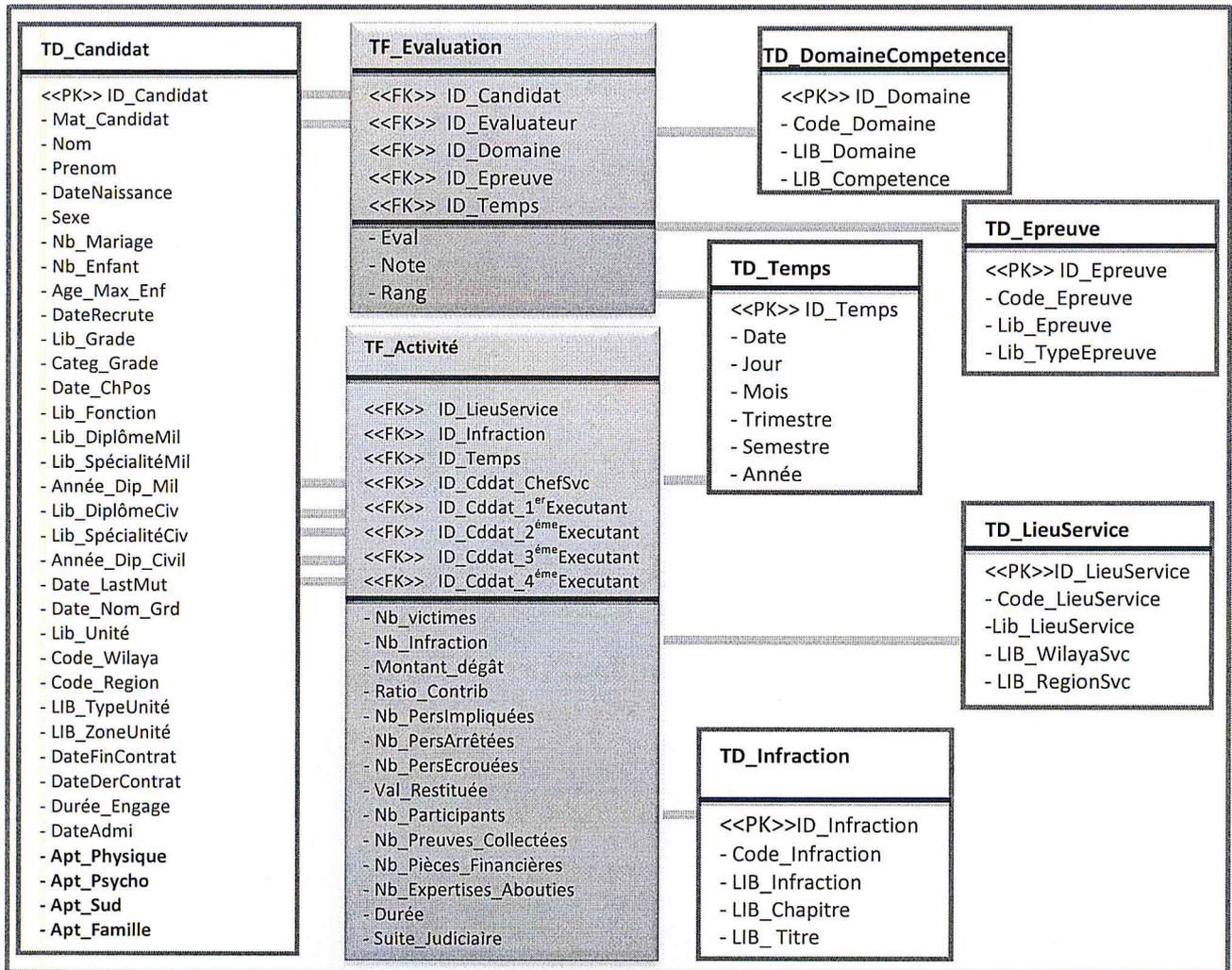


Figure V-13 Schéma décisionnel de l'évaluation globale des candidats

La finalisation de la conception d'une étoile de l'entrepôt, nous permet de passer à la construction de la zone d'alimentation. Cette zone d'alimentation constitue l'objet de la prochaine section.

V.5. Conception de la zone d'alimentation

L'alimentation du Data Warehouse, est une étape des plus importantes dans un projet Data Warehouse, elle représente 80% de la charge de travail, selon Kimball. Cette étape a pour objectif d'assurer l'acheminement des données des systèmes sources jusqu'à l'entrepôt de données, en passant par les différentes phases de nettoyage et de transformations nécessaires.

Ce processus comporte trois phases principales :

- Phase d'extraction des données en accédant aux différentes sources de données.

- Phase de transformation (changement de type de données . . .).
- Phase de chargement de données dans les Data Marts et l'entrepôt de données.

A Noter que ces différentes phases seront détaillées dans la partie suivante.

Dans la zone de préparation « *Staging area* » les données sont extraites à partir des sources de données, transformées et préparées pour le chargement final. Au niveau du serveur il est procédé à l'affectation de clés artificielles et à quelques transformations nécessaires avant le chargement final dans la zone d'entreposage.

V.5.1. Processus de chargement

Deux types de tables dans l'entrepôt de données « faits, dimensions » doivent être distingués. Chaque type de table diffère dans les informations qu'il contient, et donc adoption de deux processus de chargement.

V.5.1.1. Processus de chargement de dimension

Les tables de dimension constituent le contexte de la table de faits. Elles représentent le point d'entrée au Data Warehouse. Une dimension est généralement constituée : d'une clé artificielle, d'une clé naturelle et des attributs.

Le processus de chargement de dimension doit, outre charger des données, assurer :

- **La gestion des clés artificielles:** et mise en correspondance avec les naturelles.
- **La gestion de l'évolution de dimension :** gérer les changements que subissent les dimensions. Nous optons pour la méthode d'« **écrasement** » : consiste à mettre à jour l'attribut subissant un changement.

V.5.1.2. Processus de chargement des tables de faits

L'extraction des faits se fait avec les clés naturelles utilisées dans les systèmes sources. L'étape qui précède le chargement de la table des faits consiste à remplacer les clés naturelles par les clés artificielles. La substitution peut se faire directement par le biais des tables de dimension.

V.5.1.3. Processus de chargement particulier

Dans un entrepôt de données il y a des tables particulières, soit : la table de la dimension temps et les tables d'agrégats, nécessitent un traitement à part.

V.5.1.4. Processus de chargement de la dimension temps

Contrairement aux autres dimensions, la dimension temps contient uniquement des dates qui ne sont pas forcément extraites à partir des systèmes sources. Cette dimension doit contenir, en effet, toutes les dates qui coïncident, ou coïncideront, avec un fait donné. Elle

participe à toutes les étoiles et assure l'historisation. Dès lors, il est préférable de construire un calendrier.

V.5.1.5. Processus de construction d'agrégats

Après le chargement d'une étoile, les tables d'agrégats doivent être chargées par le biais de l'ETL et à partir des données détaillées. En plus du calcul des agrégats et de leur insertion, des mises à jour fréquentes de ces tables sont indispensables.

V.6. Contraintes du modèle

En plus des évaluations de critères, d'épreuves et d'indicateurs issus de l'activité opérationnelle, les compétences induites par le diplôme ainsi que les autres contraintes sont un autre type de paramètres à associer pour affiner davantage l'opération d'évaluation. Ces contraintes sont des conditions et des prérequis. Ils ne sont pas, généralement, assujettis à l'évaluation.

V.7. Modélisation du système de contexte

Comme il ne peut exister une liste universelle de compétences utilisables dans tous les secteurs d'activité, il est donc primordial d'élaborer des listes de compétences pour un poste ou un ensemble de poste. En effet, il s'agit d'en identifier les compétences requises.

Comment élaborer une liste des compétences d'un poste ou d'un emploi ?

Sachant qu'aucun poste n'a un contenu fixe dans le temps, le souci de flexibilité est donc prioritaire dans le contexte actuel. A cet effet, un mécanisme dynamique de sélection des prérequis pour un poste donné est à introduire.

Aussi il s'avère nécessaire d'exprimer cet aspect dynamique par l'introduction de la notion de pondération. Il s'agit d'appliquer des coefficients pour augmenter ou diminuer l'impact d'un critère par rapport aux autres dans le but de répondre aux différentes situations et profils à dégager.

Pour se faire, nous proposons de constituer des assemblages de critères, en concordance avec la fiche de poste en question (Voir figure V-14). Ces critères ne sont rien d'autres que les occurrences des tables de dimensions, de notre système décisionnel provoqué précédemment, à savoir domaines de compétence, épreuves, infractions et lieu de service.

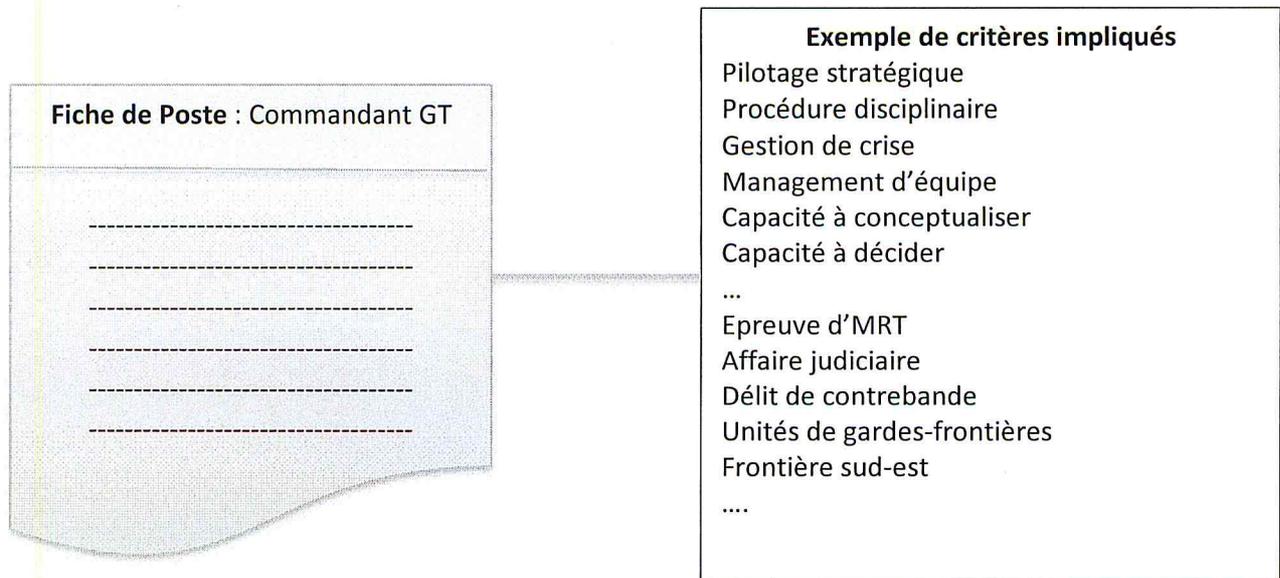


Figure V-14 Assemblage de critères par fiche de poste

Le résultat de cet assemblage est une matrice **M**, voir ci-après, qui associe à chaque candidat ces mesures par rapport aux critères impliqués pour le poste en question.

Dans le cas où plusieurs mesures se présentent, une matrice sera générée pour chacune d'elles. Ainsi pour le schéma décisionnel de quantification de l'activité, nous aurons une matrice de valeurs pour chacun des indicateurs à considérer.

Pour les domaines de compétence et les épreuves assujettis à l'évaluation pour un poste donné, le système décisionnel fait l'agrégation par rapport à une durée donnée, étendue pour l'ensemble des candidats et génère un noyau de données nécessaire (voir aperçu ci-après) pour l'analyse partielle: évaluation explicite des compétences.

Autrement dit, soient c_i un candidat et d_j un domaine de compétence. **Eval(i,j)** est alors l'évaluation de ce candidat par rapport à ce domaine pour une date donnée. L'agrégation de cette évaluation **Eval'(i,j)** par rapport à une période de service pour ce même candidat par rapport au même domaine, dans le cas le plus trivial¹⁵, donne la formule suivante de la moyenne des p évaluations :

$$\text{Eval}'(i,j) = \frac{\sum_{k=1}^p \text{Eval}_k(i,j)}{p}$$

¹⁵ Dans un système à pondération maximisée, les occurrences auront des poids différents et donc la moyenne sera à coefficients.

Pour une population de **n** candidat(s) évaluée(s) **p** fois pour **m** domaine(s) de on aura¹⁶

Candidats	...	c_i	...	c_n
Domaines				
d_j		$\frac{1}{p} \sum_{k=1}^p Eval_k(i, j)$		
d_m				$\frac{1}{p} \sum_{k=1}^p Eval_k(n, m)$

Tableau V-11 Évaluation des candidats pour les m domaines durant une période de service

De même, pour l'évaluation à base d'épreuve, si on travailler avec la mesure intégrant la note et le rang i.e. **Eval = Note/Rang**, on aura une illustration similaire à la précédente avec des épreuves à la place des domaines de compétence.

Par contre, si on préfère séparer l'agrégation des notes de celle des rangs, nous aurons l'illustration suivante. Soient c_i un candidat et e_j une épreuve. **Note(i,j)** et **Rang(i,j)** sont, respectivement, la note et le classement de ce candidat par rapport à une épreuve particulière. L'agrégation de ces mesures **Note'(i,j)** et **Rang'(i,j)** pour les p épreuves de même matière passées par ce candidat dans une durée donnée, dans le cas le plus trivial, donne les formules ci-après. **Eval'(i,j)** étant équivalente à la somme des notes divisée par la somme des rangs.

$$Note'(i,j) = \frac{\sum_{k=1}^p Note_k(i,j)}{p} \quad \text{et} \quad Rang'(i,j) = \frac{\sum_{k=1}^p Rang_k(i,j)}{p} \quad \text{avec} \quad Eval'(i,j) = \frac{Note'(i,j)}{Rang'(i,j)}$$

Candidats	...	c_i	...	c_n
Epreuves				
e_j		$\frac{1}{p} \sum_{k=1}^p Note_k(i, j)$ et $\frac{1}{p} \sum_{k=1}^p Rang_k(i, j)$		
e_m				$\frac{1}{p} \sum_{k=1}^p Note_k(n, m)$ et $\frac{1}{p} \sum_{k=1}^p Rang_k(n, m)$

Tableau V-12 L'évaluation des candidats pour les épreuves couvrant une période

¹⁶ Les valeurs sont explicitées par domaine de compétence en préparation de l'introduction des pondérations par domaine.

Pour l'activité soumise à la quantification pour un poste donné, le système décisionnel fait l'agrégation par rapport à une durée donnée, étendue pour l'ensemble des candidats et génère un noyau de données nécessaire (voir aperçu ci-après) pour l'analyse partielle : évaluation à base de l'activité.

Autrement dit, soient c_i un candidat, f_j un type d'infraction. **Mesure(i,j)** est alors un des indicateurs de ce candidat par rapport à cette infraction. L'agrégation de cette mesure **Mesure'(i,j)** par rapport au même lieu de service et à la même classification d'infraction ayant été effectués par ce candidat durant d'une période peut être assimilée au cumul dont la formule est la suivante :

$$\text{Mesure}'(i,j) = \sum_{k=1}^p \text{Mesure}_k(i,j)$$

Pour une population de n candidat(s), on aura la quantification suivante des p affaires menées classées selon les m types d'infraction :

Candidats				
Infractions	...	c_i	...	c_n
f_j		$\sum_{k=1}^p \text{Mesure}_k(i,j)$		
f_m				$\sum_{k=1}^p \text{Mesure}_k(n,m)$

Tableau V-13 Quantification de l'activité des candidats par type d'infraction

Récapitulatif :

Critères \rightarrow *Candidats* $\vec{c}_1 \quad \vec{c}_2 \quad \dots \quad \vec{c}_n$

$$\begin{matrix} \text{critère}_1 \\ \text{critère}_2 \\ \vdots \\ \text{critère}_m \end{matrix} \rightarrow \left(\begin{array}{c|c|c|c} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & & x_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1m} & x_{2m} & & x_{nm} \end{array} \right) \text{ d'où } M = \left(\begin{array}{c|c|c|c} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & & x_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1m} & x_{2m} & & x_{nm} \end{array} \right)$$

Avec : x_{ij} = évaluation du candidat i par rapport au critère j .

Le sous ensemble de critères avec leurs pondérations associées modélisant le contexte imposé varie non seulement d'un poste à un autre, mais aussi dans le temps. De ce fait, il est nécessaire d'ajouter une couche de persistance. Cette dernière constitue une sorte de référentiel de contextes ou de fiches de poste pour une éventuelle réutilisation, traçabilité, etc.

A ce stade deux cas de figure peuvent se présenter :

1^{er} cas : le résultat sera un vecteur V constitué de valeurs calculées sur la base des évaluations pondérées pour chaque candidat comme illustré ci-après.

<i>Critères</i> → <i>Pondérations</i>	\vec{P}	<u><i>Candidats</i></u>			
		\vec{c}_1	\vec{c}_2	...	\vec{c}_n
critère ₁ → pondération ₁	<i>donc: V = (p₁p₂...p_m) * </i>	$\begin{pmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & & x_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1m} & x_{2m} & & x_{nm} \end{pmatrix}$			
critère ₂ → pondération ₂					
⋮					
critère _m → pondération _m					

$$\text{d'où : } V = \left(\sum_{j=1}^m p_j * x_{1j} \mid \sum_{j=1}^m p_j * x_{2j} \mid \dots \mid \sum_{j=1}^m p_j * x_{nj} \right)$$

$V = (r_1 \mid r_2 \mid \dots \mid r_n)$ r_i : nouvelle évaluation après pondération, pour le candidat i .

Tel que : $r_i = \sum_{j=1}^m p_j * x_{ij}$ où : p_j = la pondération introduite pour le critère j .

En fin, le résultat est le vecteur V' engendré par le tri des colonnes (C_i) du vecteur V .

2^{ème} cas : le résultat restera une matrice d'évaluations mais cette fois multipliées par les pondérations correspondantes comme ci-après.

$$M' = \left(\begin{array}{c|c|c|c} x_{11} * p_1 & x_{21} * p_1 & \dots & x_{n1} * p_1 \\ x_{12} * p_2 & x_{22} * p_2 & & x_{n2} * p_2 \\ \vdots & \vdots & & \vdots \\ x_{1m} * p_m & x_{2m} * p_m & & x_{nm} * p_m \end{array} \right) \text{ ou bien } M' = \left(\begin{array}{c|c|c|c} \acute{x}_{11} & \acute{x}_{21} & \dots & \acute{x}_{n1} \\ \acute{x}_{12} & \acute{x}_{22} & & \acute{x}_{n2} \\ \vdots & \vdots & & \vdots \\ \acute{x}_{1m} & \acute{x}_{2m} & & \acute{x}_{nm} \end{array} \right)$$

avec $\acute{x}_{ij} = x_{ij} * p_j$

M' est un ensemble de candidats (en colonnes) avec des valeurs x'_{ij} à maximiser. Ce dernier résultat a l'avantage d'être exploitable via les techniques de data mining.

Pour plus de souplesse, il est conseillé de garder les résultats partiels séparés (une matrice pour chaque type : $M_{\text{évaluation}}$ et $M_{\text{activité}}$) pour un affichage détaillé ou pour en appliquer peut être un autre niveau de pondération (par volet d'évaluation).

V.7.1. Choix du Relationnel/SQL comme infrastructure de développement

Afin de valider notre proposition, nous avons besoin de cette infrastructure qui est capable de représenter et exploiter le modèle relationnel, les tables et les relations avec un langage dédié qui est le SQL. Ce choix permet de persister les contextes et leurs pondérations ainsi que de les manipuler via le langage SQL. Les deux critères fondamentaux ayant motivés notre choix de cette infrastructure pour la mise en œuvre de notre approche sont :

- (i) La nécessité de gestion d'un grand volume de données. En effet, avec l'émergence des nouvelles technologies de l'information et de la communication et surtout grâce aux nouveaux centres de données, une masse considérable de données est actuellement disponible et doit être prise en charge par les différents systèmes de gestion de données. De ce fait notre choix s'est orienté vers les ED relationnels qui permettent de répondre à cette exigence ;
- (ii) Comme nous voulons étendre le modèle noyau de l'ED proposé précédemment avec notre modèle de contexte, l'infrastructure doit permettre l'extension de la manipulation via SQL.

Ainsi, pour concrétiser notre idée, nous empruntons la méthode d'implémentation sous forme d'un schéma relationnel.

V.7.2. Persistance du modèle de contexte dans l'ED (Extension)

Dans cette section, nous détaillons comment le modèle de contexte proposé peut être utilisé pour étendre le modèle décisionnel afin de prendre en charge le contexte. Pour cela, nous utilisons le langage SQL afin de manipuler les différentes tables et étendre le modèle décisionnel avec le contexte.

En effet, nous introduisons la notion de **Table de Pondération**. Celle-ci est préfixée conventionnellement par "TP_". Chacune des tables de pondération est raccordée à une table de dimension désignée pour être pondérée, (voir figure V-16). Un enregistrement d'une table de pondération TP_Dim_i devra associer une valeur de pondération **p** à une instance **critère** de la dimension TD_Dim_i auquel il est lié.

Pour chaque mesure Mes_j de la table de fait : **TF_Fait**, la pondération globale associée sera calculée à partir des pondérations relatives aux occurrences des dimensions impliquées (décrivant le présent fait). Une des solutions consiste en la somme des différentes pondérations.

Le même enregistrement de la table de pondération TP_Dim_i est aussi fils d'un enregistrement père qui représente la fiche de poste (le contexte). Ainsi nous introduisons une autre notion qui est la **Table de Contexte**. Préfixée conventionnellement "TC_", cette table constitue le support pour le référentiel des fiches de poste.



Figure V-15 Modèle décisionnel étendue avec le modèle du contexte

Nous parlerons désormais de contextes qui, autour des mêmes évaluations, associe des rendus et des lectures contextuelles différents pour les décideurs et ce en jouant sur les pondérations correspondantes.

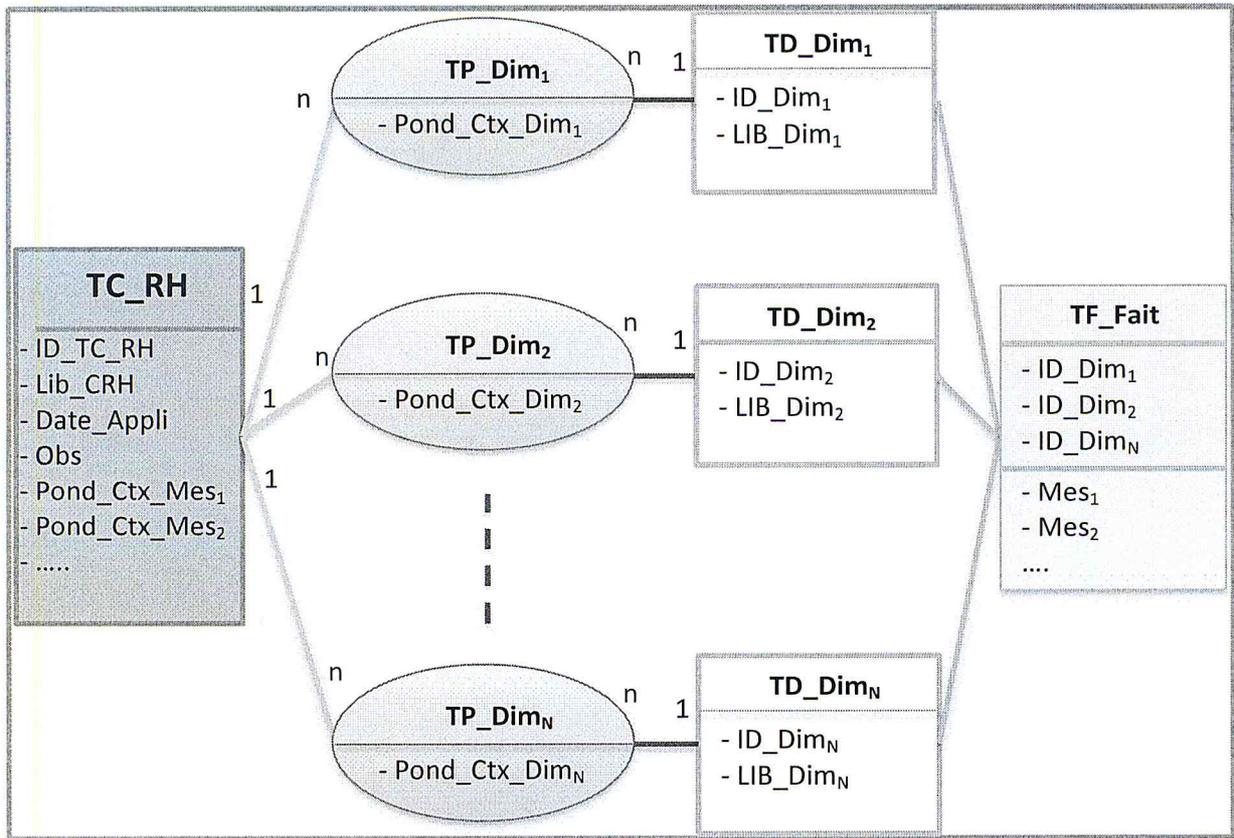


Figure V-16 Modèle décisionnel étendue avec le modèle conceptuel du contexte

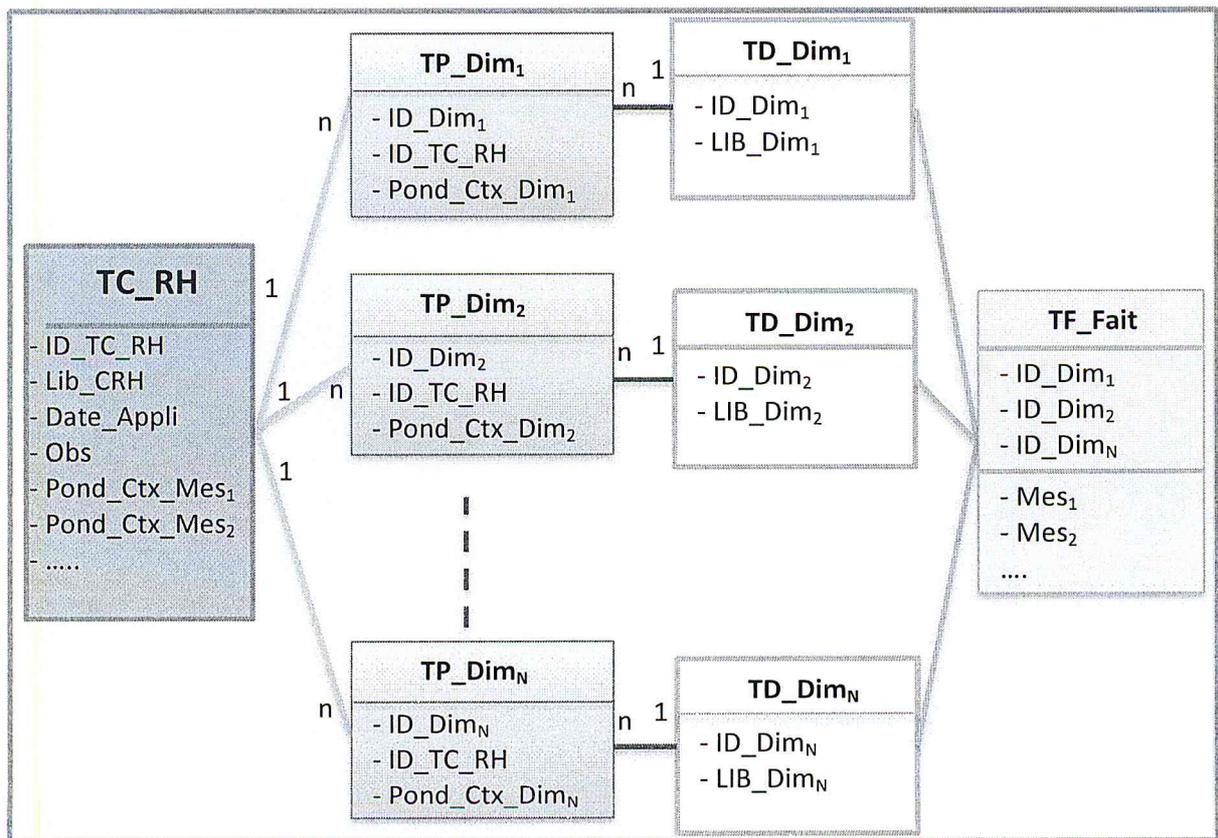


Figure V-17 Schéma du modèle de pondération et référentiel de fiches de postes (contextes)

V.8. Conclusion

Dans ce chapitre, nous avons conçu graduellement notre modèle dynamique d'aide à la décision pour la gestion des compétences. Ceci à travers un système décisionnel basé sur un entrepôt de données. En effet, nous avons commencé par concevoir ledit entrepôt et la zone d'alimentation, en passant par la conception des cubes.

Ensuite, nous avons proposé une approche pour la prise en charge du contexte dans la conception des systèmes d'entreposage de données afin qu'elle puisse gérer et expliciter les informations contextuelles dont peuvent dépendre ses données. Nous avons d'abord intégré, dans le modèle ED, le modèle de contexte proposé. Concrètement, ceci a été réalisé en créant l'ensemble des éléments constituant notre modèle de contexte et en mettant en œuvre son lien avec le modèle décisionnel. De cette manière, nous avons offert la possibilité de stocker n'importe quelle information contextuelle et de l'associer avec les critères correspondants.

Nous rappelons que cette approche est valable sur n'importe quel modèle décisionnel respectant une sémantique similaire.

Chapitre VI Réalisation & Déploiement

Pour la réalisation et la mise en place de la solution préconisée, il a été nécessaire de faire appel à un certain nombre d'outils et de mettre en place des environnements pour héberger l'exécution du système.

VI.1. Environnement de développement

Pour anticiper tout conflit potentiel de compatibilité entre les différentes sources de données, nous décrivons, dans ce qui suit, la plateforme **technique et fonctionnelle** déployée.

1. Serveur de bases de données sources

- 1 Serveur Windows 2008.

2. L'environnement de manipulation :

- **Matériel**

- Micro-ordinateur:

- Processeur Intel® Xeon® CPU E5-1650 v3 @ 3.50 GHZ ;
- RAM de 8,00 GO ;
- Disque dur de Capacité de 01 TB.
- Système d'exploitation Windows 8.1 Professionnel (2013);

- Serveur :

- Processeur Intel® Xeon® CPU E5-2670 v3 @ 2.50 GHZ 2.50 GHZ (2 Processeurs);
- RAM de 32,0 GO ;
- Type de système : 64-bit SE et Processeur ;
- Disque dur de Capacité de 03 TB ;
- SE : Windows Serveur 2012 r2 Standard et Serveur 2008 r2 Datacenter.

- **Logiciel**

- Oracle 10gR2 pour la sauvegarde des données du Data Warehouse ;
- Oracle Warehouse Builder pour le chargement et l'administration du DW ;
- BIEE 11gR2 ;
- Oracle Data Miner.

VI.2. Architecture technique de la solution

La figure suivante illustre la structure et l'architecture technique de la solution proposée :

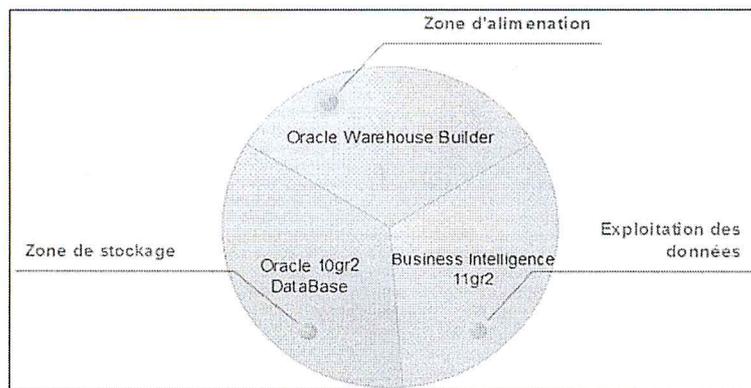


Figure VI-1 Architecture technique de la solution

VI.3. Les outils technologiques choisies

VI.3.1. Oracle Data Base 10g r2

Nous optons pour le SGBD Oracle pour ses performances avérées et sa stabilité. De plus la Gendarmerie Nationale l'utilise d'ores et déjà notamment sa version 10gR2 pour l'hébergement de ces Systèmes d'Information opérationnels.

VI.3.2. Oracle Warehouse Builder

Oracle Warehouse Builder (OWB) est un outil unique et complet pour tous les aspects d'intégration des données [62]. En effet, il exploite Oracle Database pour transformer des données en informations de haute qualité. Il fournit outre la qualité et l'audit des données, la modélisation relationnelle et dimensionnelle de manière entièrement intégrée.

OWB permet de créer des entrepôts de données, de migrer des données à partir de systèmes existants, de consolider des données provenant de sources disparates, de les nettoyer et de les transformer pour fournir des informations de qualité, en bref, il permet la gestion complète du cycle de vie des données et des métadonnées de l'organisation.

VI.3.2.1. Avantages d'Oracle Warehouse Builder

- Accès à une large variété de sources de données ;
- Possibilité de catégoriser, transformer et nettoyer les données ;
- Possibilité d'implémenter différentes modélisations pour diverses applications ;
- Audit et traçabilité de toutes les actions dans l'entrepôt de données.

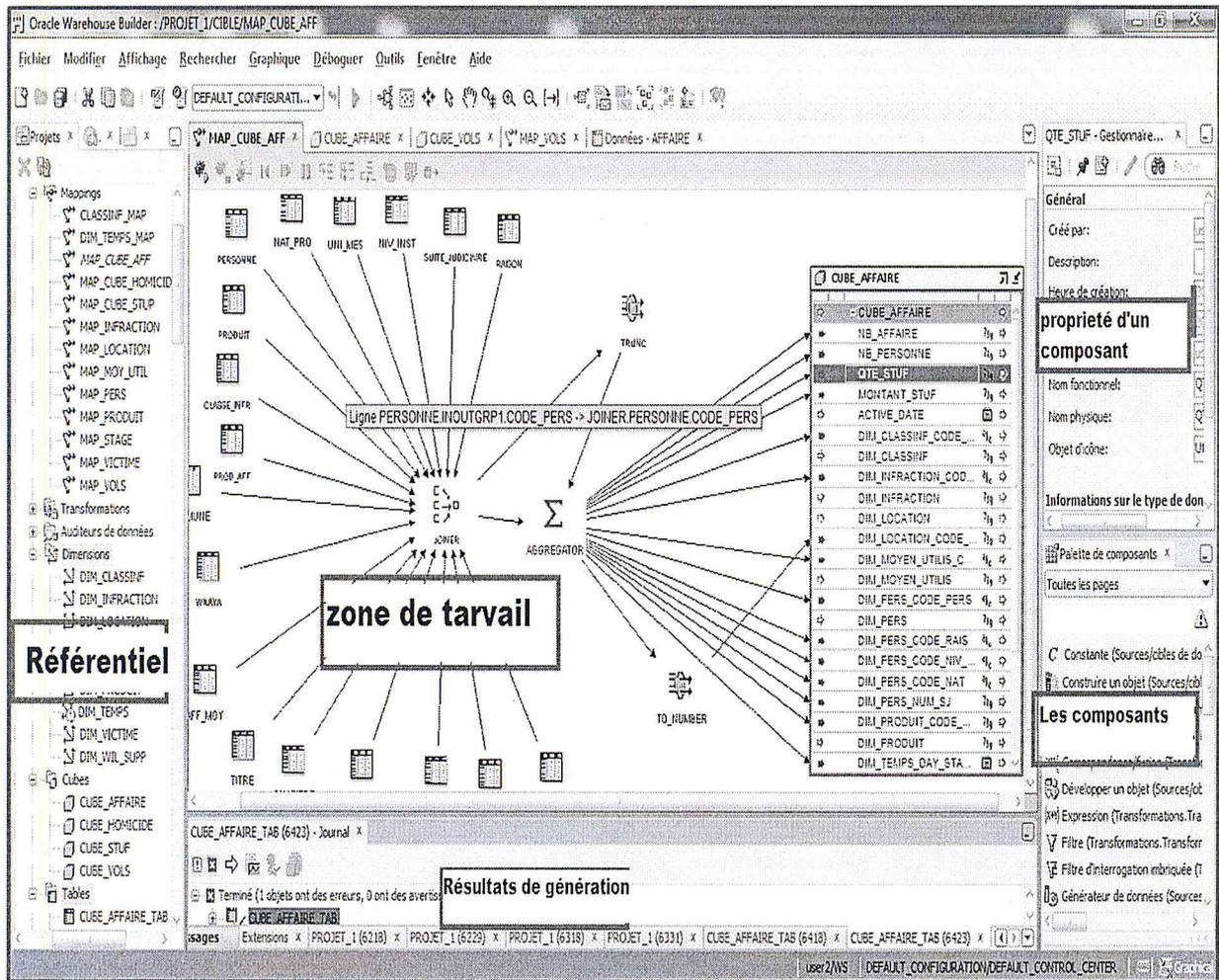


Figure VI-2 L'interface d'OWB

VI.3.2.2. Architecture et composants de Warehouse Builder

- **Design Center** : interface utilisateur pour concevoir, gérer, planifier et déployer des processus ETL pour le déplacement et la transformation de données.
- Toutes les métadonnées associées au travail effectué dans le Design Center sont stockées dans le dépôt Warehouse Builder.
- Le référentiel est hébergé sur une base de données Oracle et vous pouvez utiliser le **Repository Browser** pour faire rapport sur les métadonnées dans le dépôt. Aussi hébergé sur une base de données Oracle est le **Target Schema** auquel Warehouse Builder charge les données résultant des processus ETL que vous exécutez dans le **Control Center Service**.

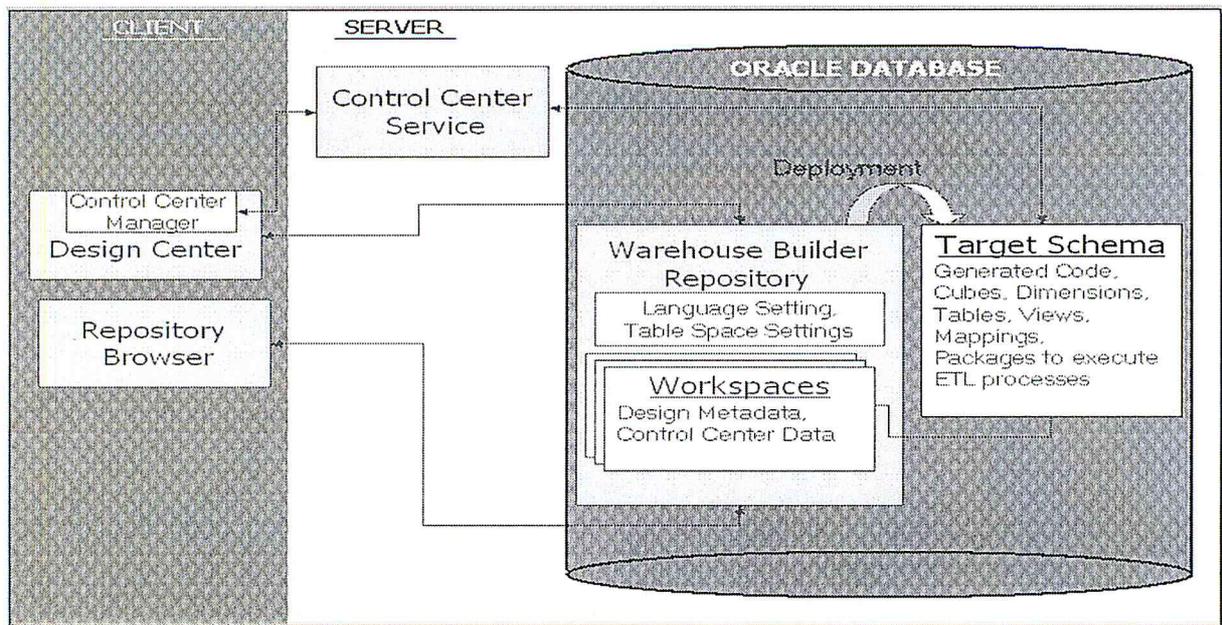


Figure VI-3 L'interaction des principaux composants du logiciel OWB [63]

VI.3.3. Oracle Business Intelligence

Oracle Business Intelligence Enterprise Edition 11g (OBIEE) est une plate-forme complète d'intelligence d'affaires et d'analyse qui offre une gamme complète de fonctionnalités y compris les tableaux de bord interactifs, les notifications et les alertes, les rapports d'organisation et financiers, la gestion des tableaux de bord et de la stratégie, l'invocation des processus métier et plus encore.

OBIEE 11g repose sur une base technologique éprouvée et moderne qui prend en charge les charges de travail et les déploiements les plus complexes, tout en fournissant des informations propices aux utilisateurs d'une organisation à un coût réduit.

Oracle BI inclut dans sa partie administration, une base référentielle (référentiel) dans lequel sont importées les métas data de l'entrepôt de donnée, les connexions et les jointures. Ensuite la couche présentation inclut les outils *Publisher*, *Answers* et *Dashboard*.

VI.3.3.1. Oracle BI Publisher

Est la solution de reporting pour autoriser, gérer et livrer tous ses rapports et documents plus facilement et plus rapidement que les outils de reporting traditionnels. Utilisez le navigateur Web ou des outils de bureau familiers pour créer tout, à partir de documents faisant face à un client parfait pour les pixels et des rapports de gestion interactifs contre pratiquement n'importe quelle source de données. Voir les rapports en ligne ou les planifier et

fournir des dizaines de milliers de documents par heure avec un impact minimal sur les systèmes transactionnels.

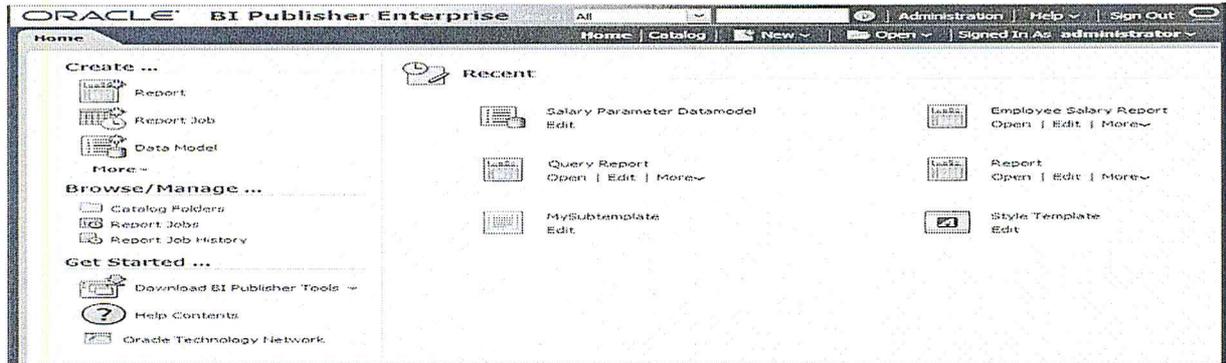


Figure VI-4 Interface de BI Publisher

VI.3.3.2. BI Answers

Oracle BI Answers est un outil de la suite Oracle BI, utilisé essentiellement dans la création des analyses métier, qui répondent aux questions des gestionnaires. Ces analyses seront intégrées dans le tableau de bord (Dashboard).

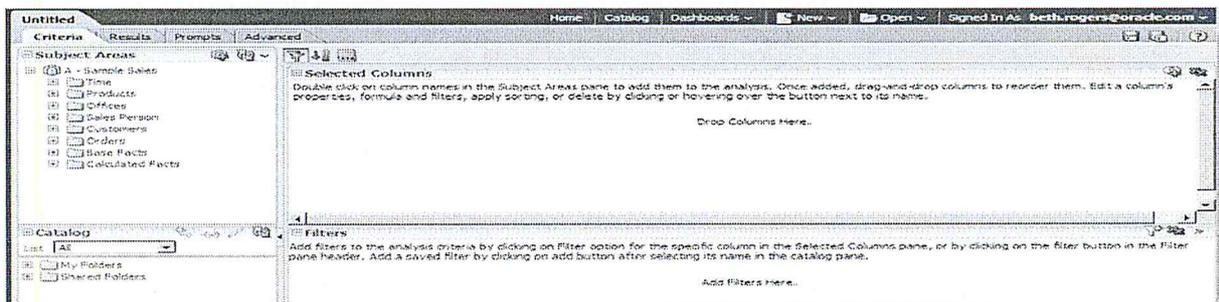


Figure VI-5 Interface de BI Answers

VI.3.3.3. BI Dashboard

Oracle BI Dashboard est l'outil graphique qui permet l'intégration de tous les composants créés par le publisher (rapport), answers (analyses) et mapviewer (représentation sur carte). Il permet de donner un accès direct à l'utilisateur de créer, personnaliser et éditer ses pages et contenu du tableau de bord.

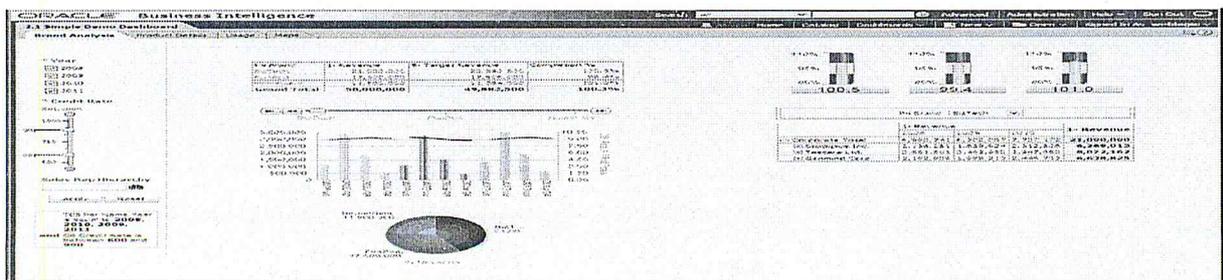


Figure VI-6 BI Dashboard

VI.4. Implémentation des DataMarts

Dans ce chapitre, nous allons présenter l'implémentation des deux (02) Data Marts que nous avons conçu dans le chapitre V à travers la création des dimensions et le déploiement des cubes.

Nous passerons, ensuite, à l'implantation du Business Intelligence pour l'affichage des analyses, graphes, filtrages et affichage des candidats selon les critères étudiés.

VI.4.1.1. DataMart évaluation directe (explicite)

La première étape de création du Data mart Evaluation directe est la création des dimensions, leurs mapping avec les sources de données, et leurs déploiement et chargement.

Les étapes suivantes sont la création du cube de donnée et la définition de ses mesures, son mapping, son déploiement et son chargement.

Pour la création des dimensions, nous avons choisi de présenter, à titre illustratif, le processus de création, mapping et déploiement d'une seule dimension qui est l'INFRACTION relative au deuxième Data Mart (exemple plus parlant). Le processus pour le reste des dimensions est effectué d'une manière similaire. Ainsi, nous nous trouvons contraints de présenter la création, le mapping et le déploiement du cube relatif au deuxième Data Mart.

VI.4.1.2. DataMart évaluation à base d'activité (implicite)

(1) Création de la dimension « INFRACTION » :

La dimension infraction est composée des niveaux infraction, chapitre et titre :

On entre le nom de la dimension, en respectant une règle unifiée pour nommer les dimensions de l'entrepôt de données :

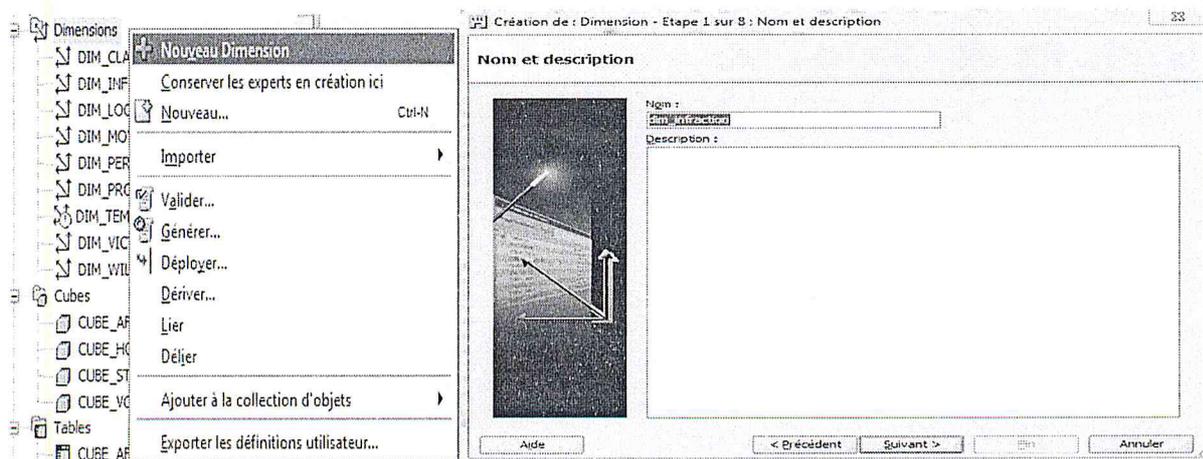


Figure VI-7 L'ajout d'une nouvelle dimension

On entre les informations relatives à la dimension (champs de la dimension).

A noter que pour chaque dimension, nous aurons une clé primaire de la dimension qui sera généré automatiquement de type *surrogate*, (ID dans notre cas), puis viennent les champs de type business qui sont les clés étrangères des sources de données (application métier) comme suit :

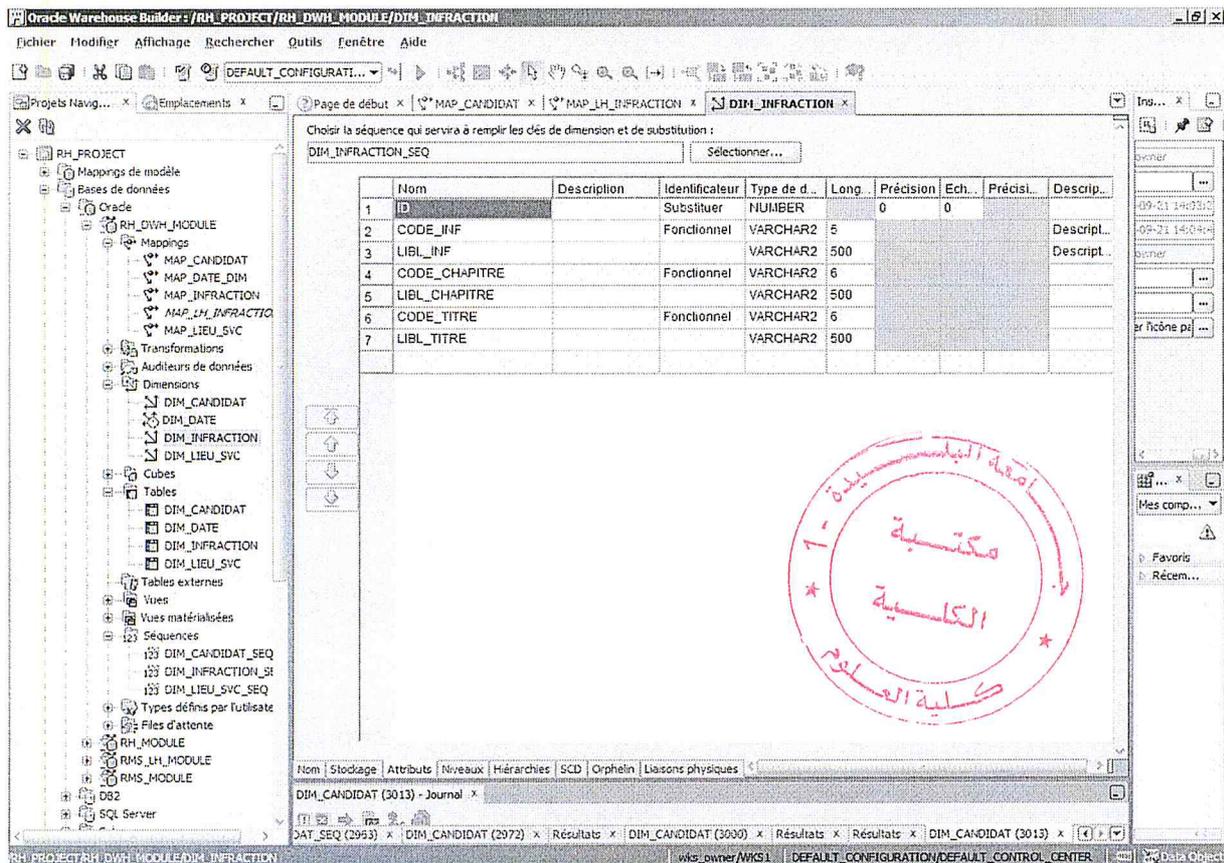


Figure VI-8 Spécification des attributs de la dimension

Lors de l'utilisation de l'assistant de création des dimensions de OWB, celui-ci crée une seule hiérarchie (par défaut), dans lequel on spécifie les différents niveaux.

Le cas contraire, si on veut créer plusieurs hiérarchies, il faudra passer par le mode manuel.

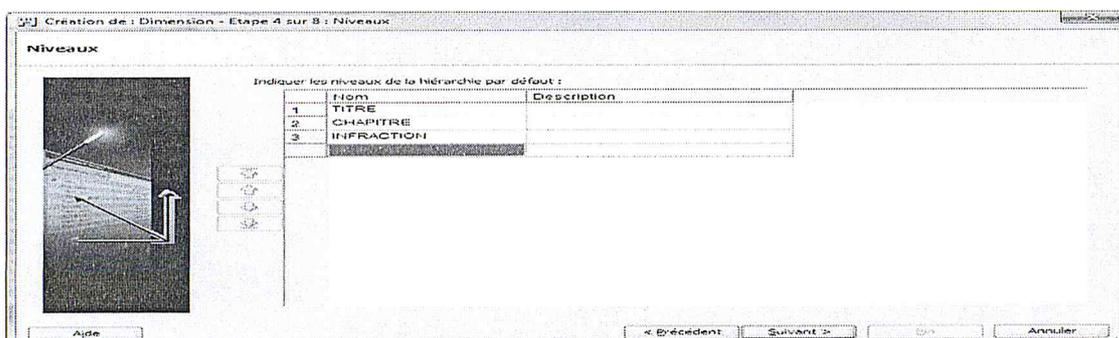


Figure VI-9 Spécification des niveaux

Ensuite, nous aurons à sélectionner les attributs de chaque niveau comme suit :

A l'issue de la création de la dimension, on procède à la validation de la dimension, afin de vérifier la dimension, si elle ne contient aucune erreur de création.

Puis on procède à la génération de la dimension comme suit :

Niveau	Description
1	TITRE
2	CHAPITRE
3	INFRACTION

Niveau	Nom d'attribut de dimen...	Applica...	Nom d'attribut de niveau	Description	Vale
1	ID	<input checked="" type="checkbox"/>	ID		-1
2	CODE_INF	<input type="checkbox"/>			
3	LIBL_INF	<input type="checkbox"/>			
4	CODE_CHAPITRE	<input checked="" type="checkbox"/>	CODE_CHAPITRE		
5	LIBL_CHAPITRE	<input checked="" type="checkbox"/>	LIBL_CHAPITRE		
6	CODE_TITRE	<input type="checkbox"/>			
7	LIBL_TITRE	<input type="checkbox"/>			

Figure VI-10 Spécification des attributs des niveaux

La création du reste des dimensions se fait de manière similaire. Nous allons uniquement présenter dans les sections suivantes le mapping et le déploiement de cette dimension, et nous présenterons le résultat du déploiement.

(2) Mapping et déploiement des dimensions

Après la création de la dimension, il faudra procéder au mapping de la dimension avec les sources de données correspondantes dans la base de données source (exemple : Table infraction, chapitre et titre).

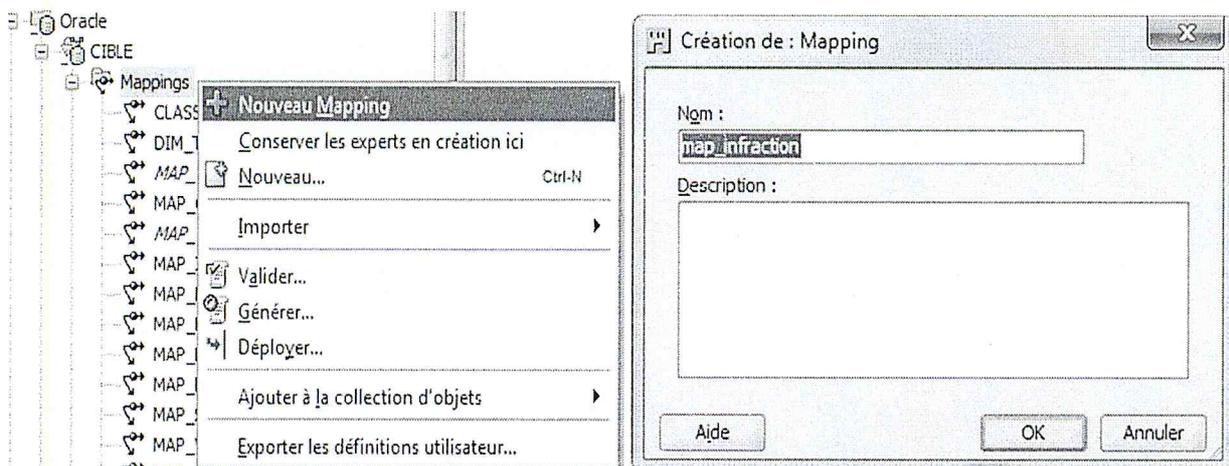


Figure VI-11 L'ajout d'un nouveau mapping

On choisit d'abord nos données source qui alimenteront la dimension.

Ensuite, on utilise un opérateur *JOINER* pour regrouper les données, avec spécification des conditions de jointure entre les sources de données.

Une fois les sources de données sont réunies dans le *JOINER*, on n'oublie pas de spécifier la clause de jointure des sources de données :

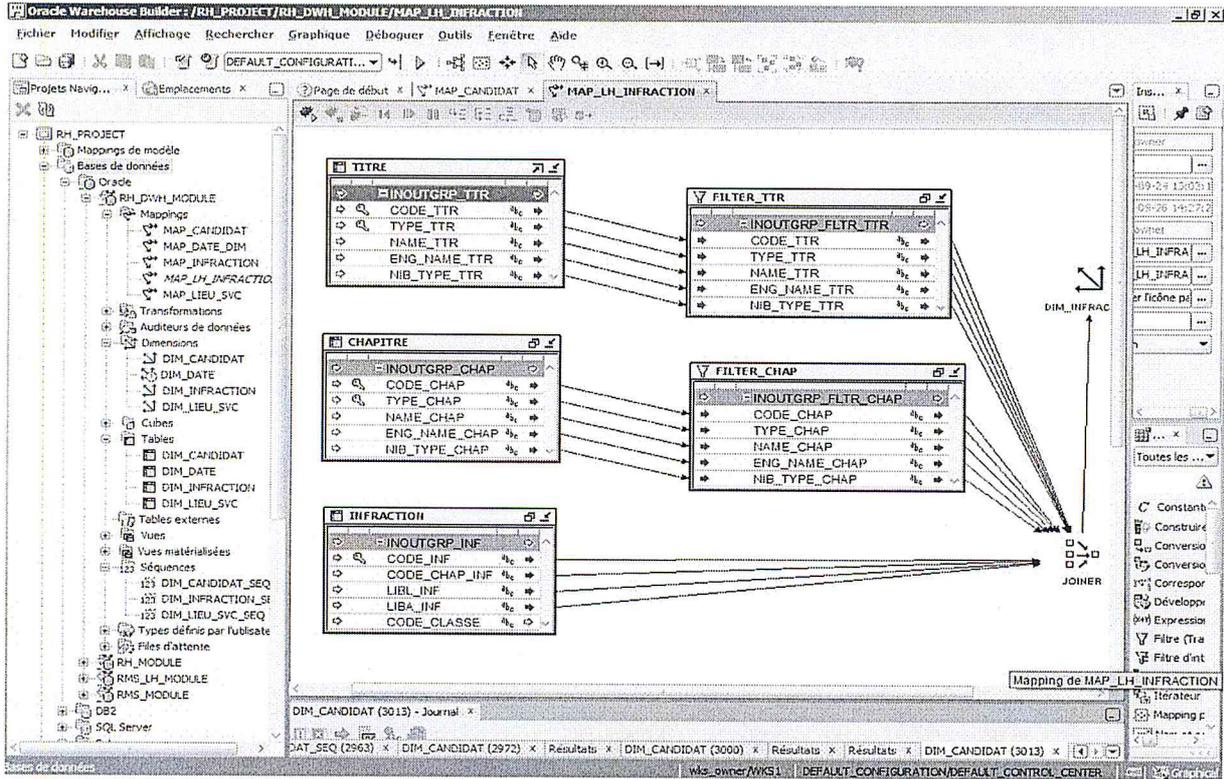


Figure VI-12 La sélection des tables sources

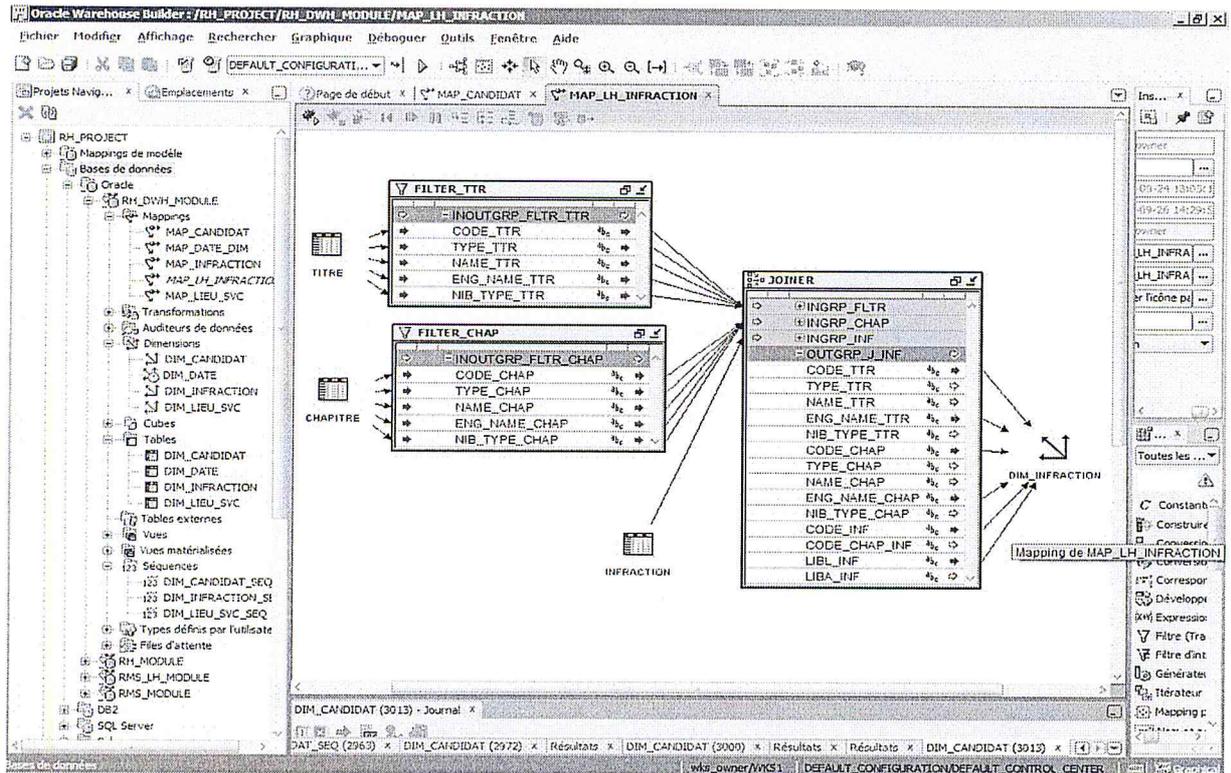


Figure VI-13 Regroupement des données avec l'opérateur JOINER

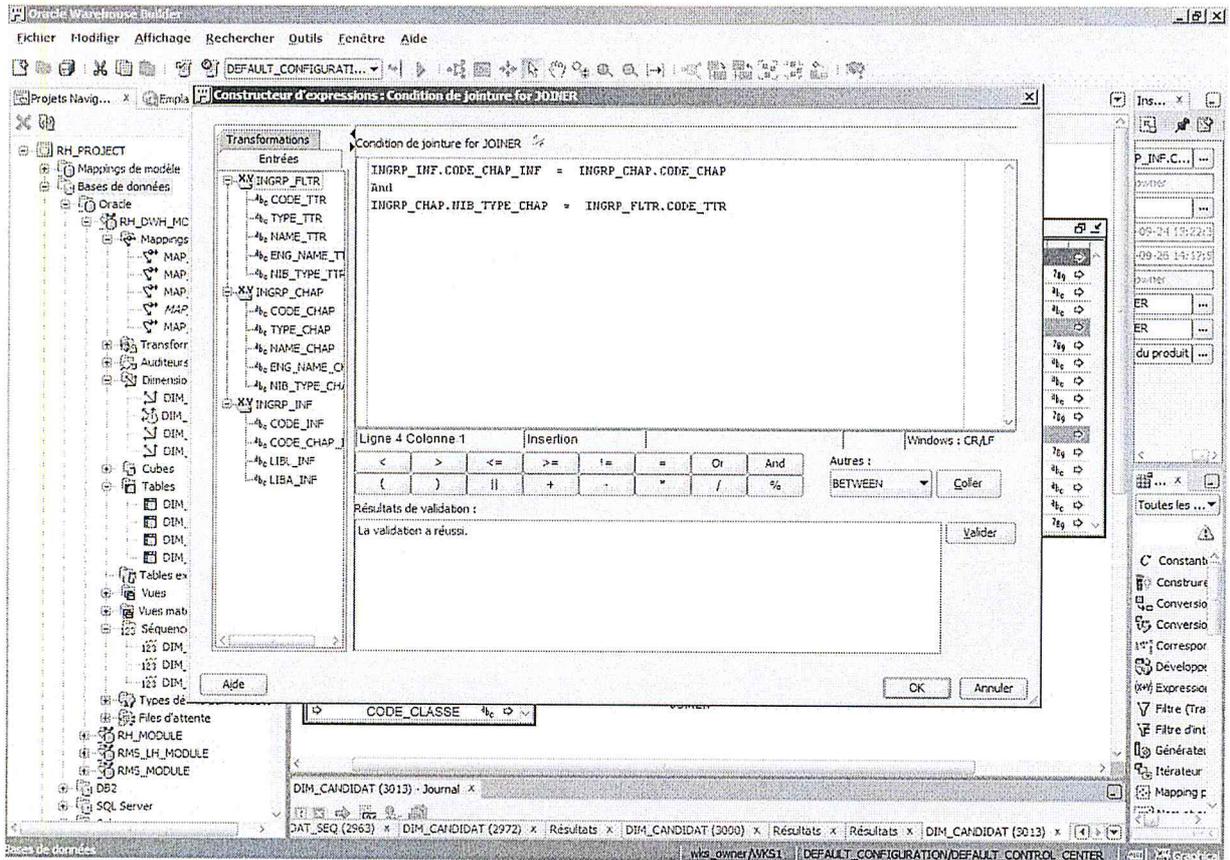


Figure VI-14 Spécification de la condition de jointure

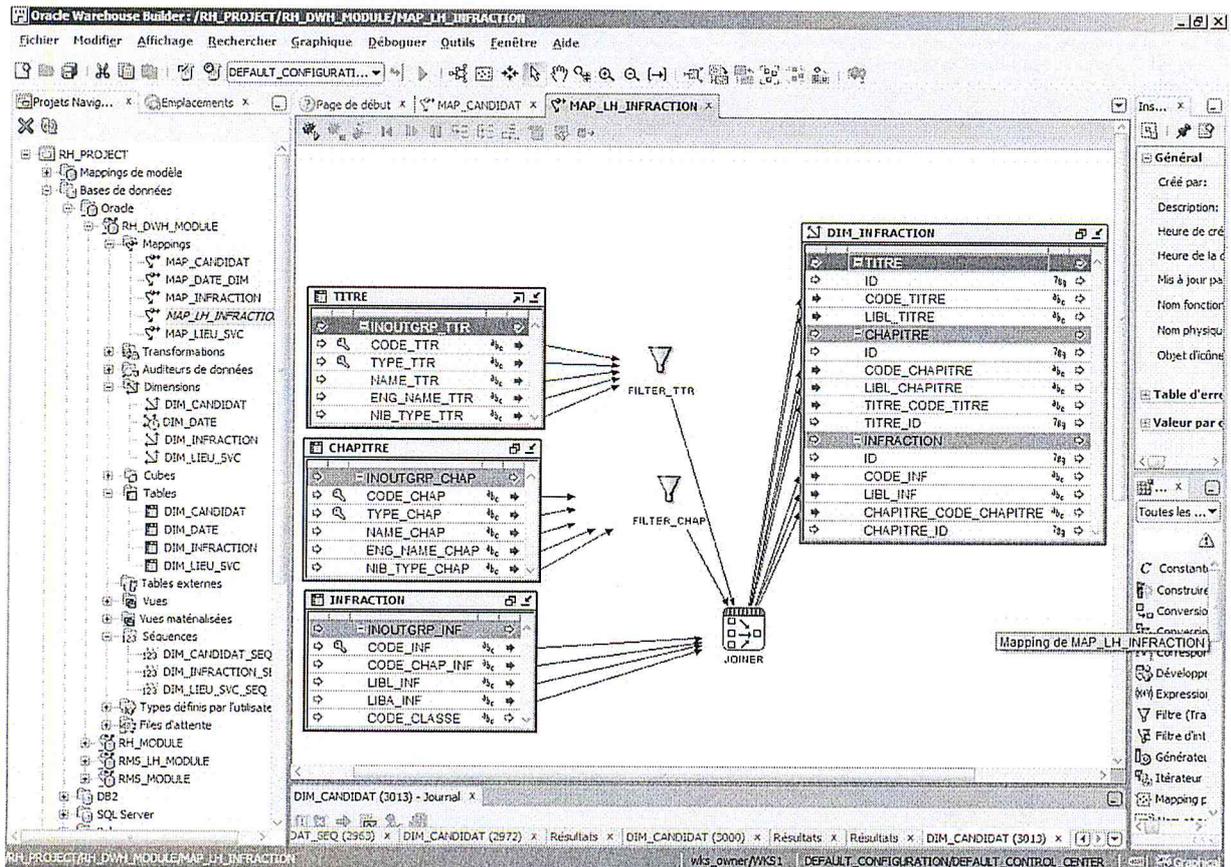


Figure VI-15 Liaison entre JOINER et « DIM_INFRACTION »

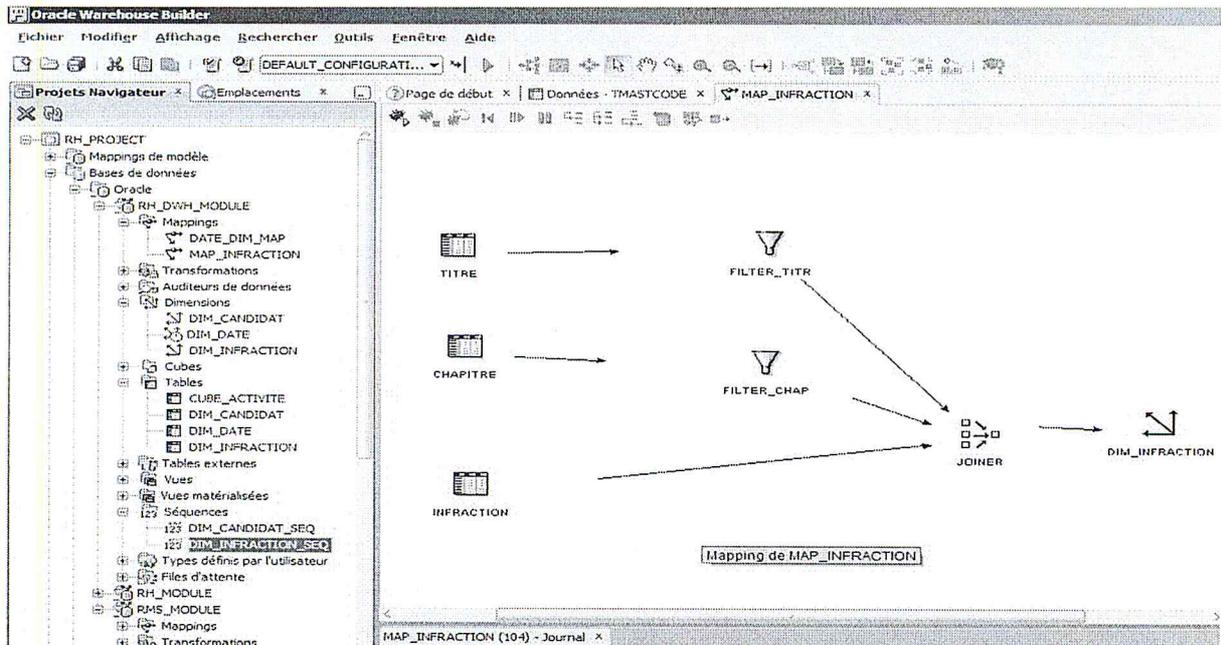


Figure VI-16 Schéma final du mapping INFRACTION

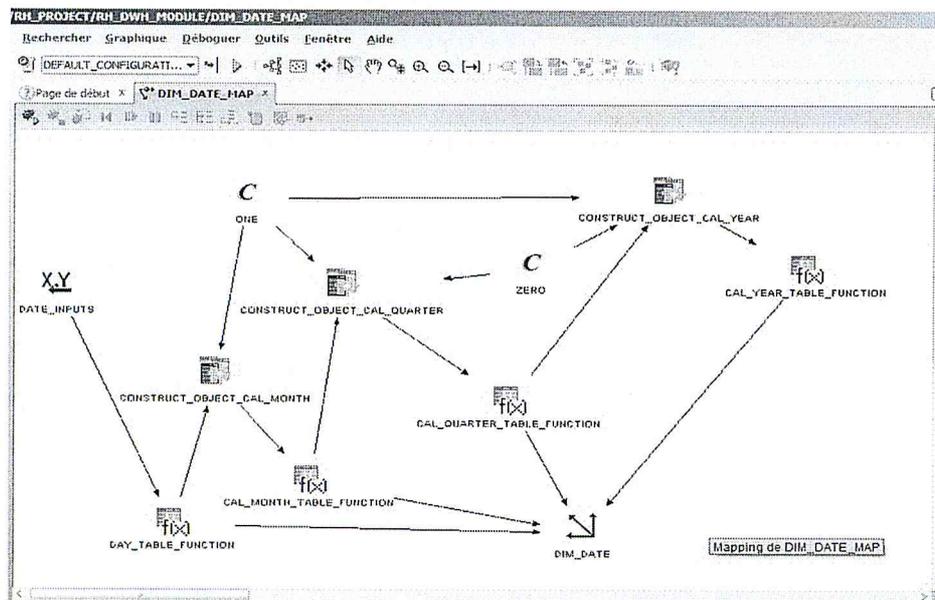


Figure VI-17 Schéma final du mapping de la dimension Date

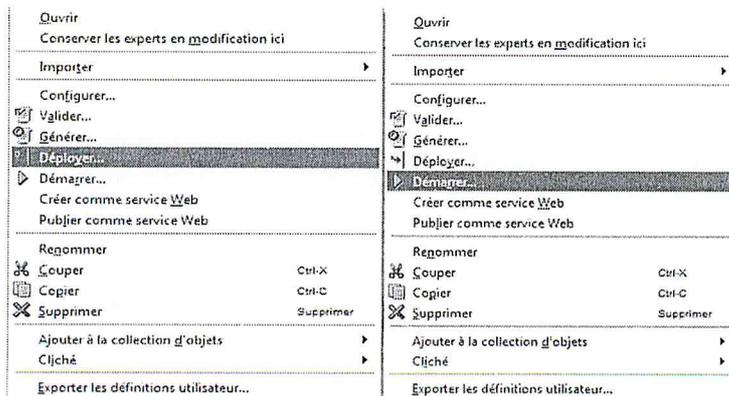


Figure VI-18 Déploiement et démarrage du « MAP_INFRACTION »

Afin de vérifier le résultat du mapping, on peut afficher les données correspondantes à la dimension :

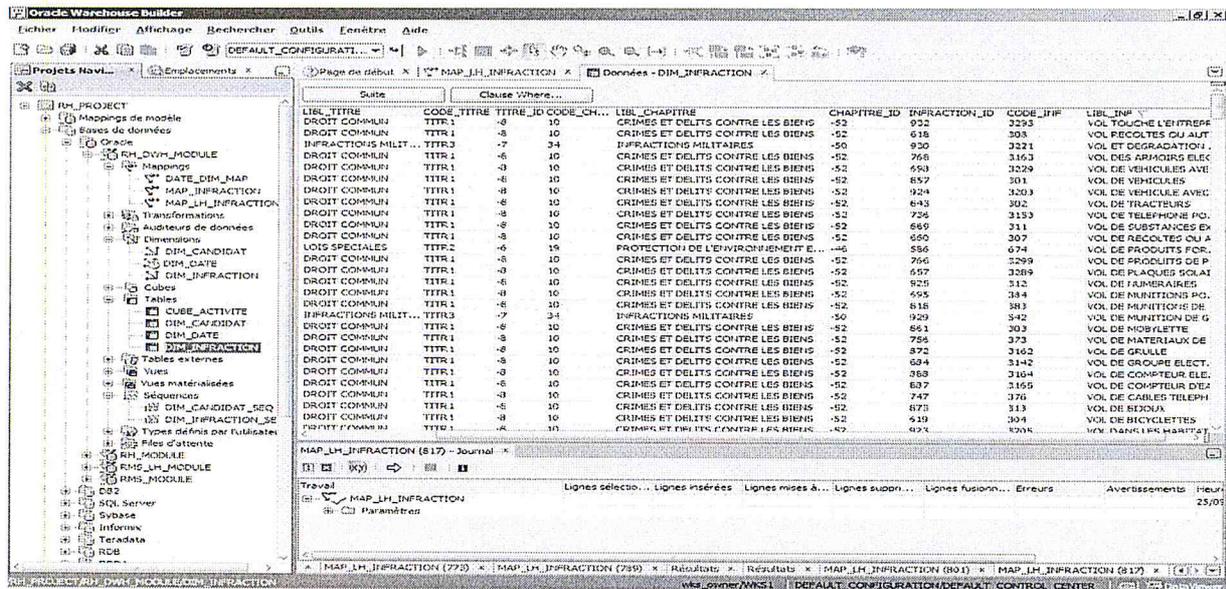


Figure VI-19 Les données de la dimension « INFRACTION »

(3) Création, mapping et déploiement du cube

Il faut juste veiller à sélectionner les dimensions du cube de données, puis introduire les mesures voulues, et procéder directement au mapping du cube comme illustré dans la figure.

A noter que lors du mapping du cube de données, on procède globalement comme lors du mapping de la dimension, sauf qu'il faut utiliser l'opérateur AGGREGATOR, afin d'agréger les données du cube, et de le regrouper par des clauses group by, du fait que notre objectif, et de regrouper les données du cube selon les mesures définies lors de sa création.

A noter également que l'on doit utiliser l'opérateur TRUNC pour transformer l'attribut Date de notre source de donnée, et le mapper avec le champ "DIM-TEMPS-DAY-STARTDATE", pour des raisons de compatibilité avec la dimension temps créée auparavant.

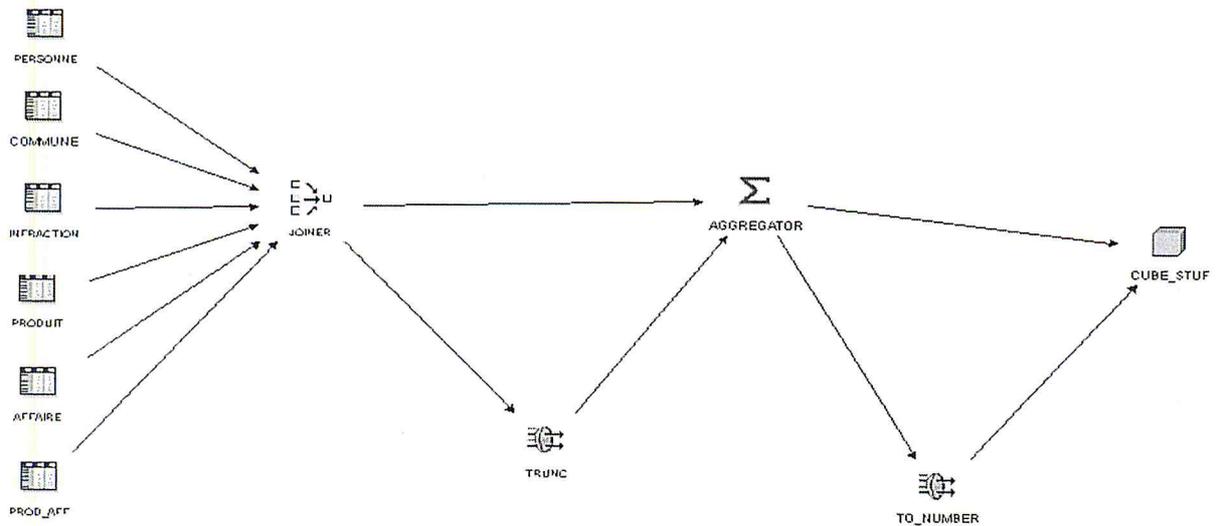


Figure VI-23 Schéma final du cube

(4) Validation, création et déploiement des cubes

Pour se faire nous lançons le « *Centre de contrôle* » du Warehouse Builder, après nous ouvrons l'onglet cible_location2 pour sélectionner pour chaque cube trois composants (un mapping, une table et le cube lui-même). Nous sélectionnons l'action : créer dans l'onglet action de déploiement comme la montre la figure ci-dessous. Enfin, nous lançons leurs créations, cette action prend beaucoup de temps du fait du nombre élevé d'enregistrements.

Objet	Statut de conce.	Action de déploy.	Déployé	Statut de déplo...	Emplacement	Module
MAP_CUBE_HOMICIDES	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE
MAP_CUBE_STUF	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE
MAP_VOLS	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE
CUBE_HOMICIDE	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE
CUBE_STUF	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE
CUBE_VOLS	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE
CUBE_HOMICIDE_TAB	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE
CUBE_STUF_TAB	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE
CUBE_VOLS_TAB	Nouveau	Créer	Non déployé	Non déployé	CIBLE_LOC...	CIBLE

Travail	ID	Statut	Travaux terminés	Propriétaire
PROJET_1	5525	✓	*	USER2
PROJET_1	5522	✓	*	USER2

Figure VI-24 Déploiement des cubes

Dans les figures suivantes nous affichons les données des Data Marts, obtenues après le déploiement.

Les colonnes encadrées en vert représentent les données des indicateurs définis pour les différents DataMarts.

Remarque :

- La colonne de l'indicateur n'est pas vide, nous y retournerons lors de l'implémentation du BI.
- Nous avons aussi créé un autre Data Mart « *FAIT_RH* » englobant tous les types d'évaluations. Il ressemble aux autres DataMarts en termes de dimensions et de création.

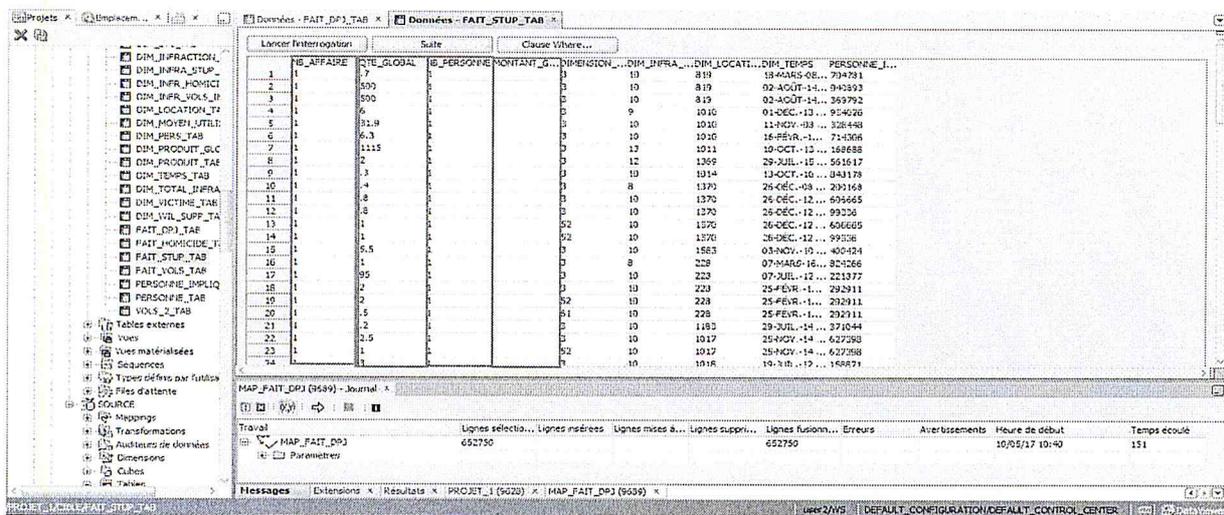


Figure VI-25 Affichage des données du datamart

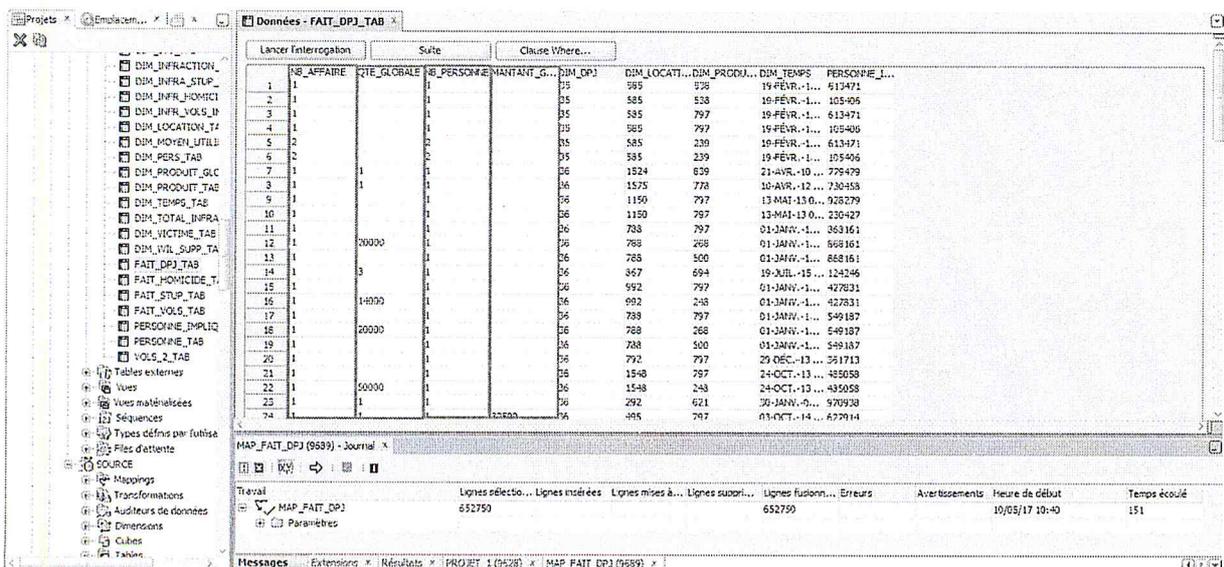


Figure VI-26 Affichage des données du datamart global

VI.5. Déroulement de l'étude de cas

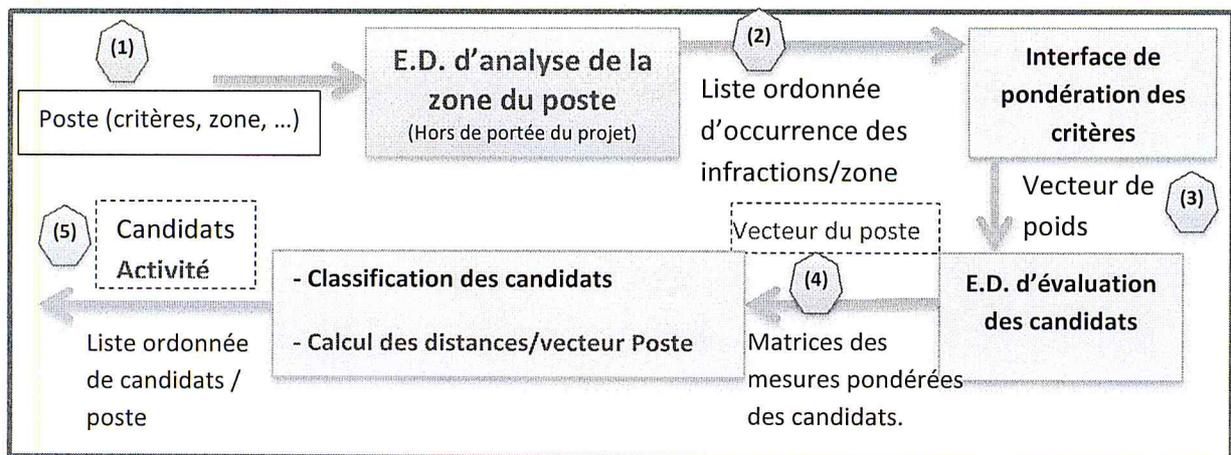
Pour valider notre approche qui consiste à créer un ED d'évaluation des candidats, stocker et exploiter les contextes (fiches de poste), nous reconsidérons le cas d'étude présenté précédemment où on souhaite interroger le système pour avoir la liste des candidats selon les trois scénarios posés. Pour cela nous allons dérouler le processus schématisé dans le chapitre de modélisation de la solution comme suit :

Phase I : Présélection des données relatives aux différentes dimensions par rapport au choix du poste ciblé. Il s'agit de définir des contraintes où sont fixées les conditions d'accès au poste en question

- Procéder à une présélection des candidats vis-à-vis du grade, sexe, diplôme, spécialité, etc, en tenant compte du décalage de la mise à jour des systèmes de productions pour la portion des candidats en phases de transition par rapport au grade (propositions au grade supérieur) et au diplôme (en cours de stage), etc.

Phase II : définition des critères pour lesquelles les candidats seront évalués et classés par ordre de mérite pour l'obtention du poste en question.

- Cibler les domaines de compétence évalués et les épreuves passées selon le poste ;
- Récupérer les évaluations correspondantes ;
- Cibler les type d'affaires judiciaires concernées par le scénario du poste en cours;
- Récupérer les mesures des différents indicateurs ;
- Appliquer les pondérations à ces mesures ;
- Réduction à une même échelle de valeurs puis combinaison des résultats.



Soit le descriptif du poste suivant :

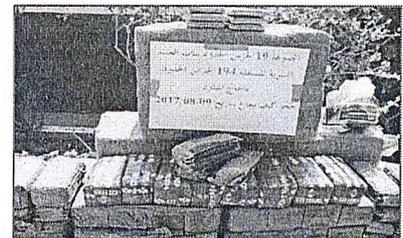
- a. Poste : **Commandant de Groupement Territorial** ;
 - b. Domaine d'emploi : **Opérationnel** ;
 - c. Famille fonctionnelle : **Sécurité publique** ;
 - d. Dominante : **Police Judiciaire** ;
-
- e. Grade : **Lieutenant-colonel/Colonel** ;
 - f. Diplôme : **Etat-major** ;
 - g. Expérience professionnelle : **Avoir servi en unité territoriale ou de recherche** ;
 - h. Age minimum : **40 ans** ;
 - i. Aptitude physique : **Apte** ;

(1) Du descriptif du poste, on en tire dans un premier temps le domaine qui est **la PJ**, l'échelon hiérarchique du poste qui est du **niveau wilaya**, ainsi que **la zone (les trois contextes)**. Ces données seront exploitées moyennant les mécanismes décisionnels existants (hors du cadre du présent projet) pour déduire les caractéristiques réelles et actuelles de la zone en question.

(2) Soit le bilan de la Police judiciaire (La Criminalité régnante) pour les trois scénarios :

Contexte du poste n°1 : Zone frontalière Nord-ouest

- Trafic de stupéfiants¹⁷ ;
- Contrebande ;
- Immigration clandestine ;
- Commerce illicite du tabac.



→ Sélection et pondération des infractions correspondantes, à savoir :

¹⁷ Image récupérée du site : « http://www.mdn.dz/site_principal/index.php?L=fr » consulté le:17/09/2017.

- Armes et explosifs → Poids = 5
- Trafic de stupéfiants → Poids = 4
- Délits de contrebande → Poids = 3
- Etrangers (immigration) → Poids = 3
- Atteintes à l'économie nationale → Poids = 1

Contexte n°2 : Zone haut plateaux (traitement par lot de wilayas semblables)

- Vol de cheptel¹⁸ ;
- Vol de véhicule ;
- Agression des personnes suivie de vol;
- Commerce illicite des boissons alcoolisées.

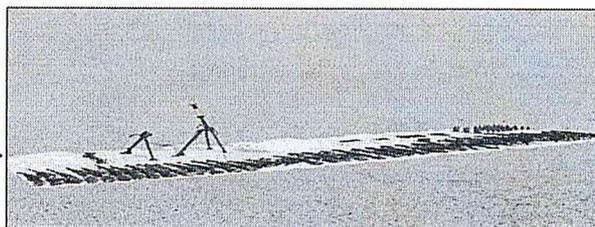


→ Sélection et pondération des infractions correspondantes, à savoir :

- Crimes et délits contre les personnes → Poids = 6
- Crimes et délits contre les biens → Poids = 5
- Délits de boissons → Poids = 2
- Règlementation des prix et fct des marches → Poids = 2
- Autres atteintes à l'économie nationale → Poids = 1
- Autres infractions aux lois spéciales → Poids = 1

Contexte n°3 : Zone frontalière Sud Est

- Terrorisme¹⁹ ;
- Trafic d'armes ;
- Contrebande ;
- Immigration clandestine.
- Vol de véhicule ;



→ Sélection et pondération des infractions correspondantes, à savoir :

- Crimes et délits contre la défense nationale → Poids = 8
- Crimes et délits contre la sécurité publique → Poids = 7
- Armes et explosifs → Poids = 6
- Crimes et délits contre les personnes → Poids = 5
- Autres atteintes à l'économie nationale → Poids = 3
- Délits de contrebande → Poids = 2
- Etrangers (immigration) → Poids = 2
- Crimes et délits contre les biens → Poids = 1

(3) Les vecteurs des poids pour les trois scénarios : \vec{P}_x

Contexte 1 : Infractions considérées $(i_1 \ i_2 \ i_3 \ i_4 \ i_5)$

$$\vec{P}_{\text{Frontalière-NW}} = (5 \ 4 \ 3 \ 3 \ 1)$$

¹⁸ Adresse de l'image : http://fibladi.dz/media/k2/items/src/33_actualite2-83457.jpg consulté le:17/09/2017.

¹⁹ Image récupérée du site : « http://www.mdn.dz/site_principal/index.php?L=fr » consulté le:17/09/2017.

Contexte 2 : Infractions considérées $(i'_1 \ i'_2 \ i'_3 \ i'_4 \ i'_5 \ i'_6)$

$$\vec{P}_{\text{Hauts plateaux}} = \begin{pmatrix} 6 & 5 & 2 & 2 & 1 & 1 \end{pmatrix}$$

Contexte 3 : Infractions considérées $(i''_1 \ i''_2 \ i''_3 \ i''_4 \ i''_5 \ i''_6 \ i''_7 \ i''_8)$

$$\vec{P}_{\text{Frontalière-SE}} = \begin{pmatrix} 8 & 7 & 6 & 5 & 3 & 2 & 2 & 1 \end{pmatrix}$$

Ces trois vecteurs seront persistés et rattachés aux trois contextes dans la base de données du Contexte suivant notre modèle proposé dans le chapitre de modélisation de la solution.

Une fois l'information contextuelle stockée dans la BDD, nous pouvons utiliser des jointures ou des fonctions pour l'interroger pour la génération des nouvelles évaluations (évaluations combinées avec les pondérations correspondantes).

(4) De l'entrepôt de données, sont générées les matrices suivantes des mesures x_{ij} des candidats :

(i) Matrice d'évaluation explicite basée sur les Compétences et les épreuves. Il est à noter que seulement une sélection de critères du socle est considérée conformément au poste en cours. Il en est de même pour les épreuves.

Les compétences considérées pour le poste en question : (valeurs entre [1/6, 6/6])

Savoir

- 1- Management et commandement - **Pilotage stratégique**
- 2- Gestion administrative du personnel - **Notation (aspects juridiques et procéduraux)**
- 3- Gestion administrative du personnel - **Procédure disciplinaire**

Savoir faire

- 4- **Communication interne**
- 5- **Communication externe en situation de crise**
- 6- Compétences transversales - **Travail en réseau/partenariat**
- 7- Compétences transversales - **Inspection/contrôle**
- 8- Compétences transversales - **Gestion de crise**
- 9- Sécurité publique - **Renseignement**
- 10- Sécurité publique - **Planification et conduite des opérations**
- 11- Management et commandement - **Management d'équipe**

Savoir-être (Les qualités personnelles)

12- Capacité à conceptualiser

13- Discrétion et confidentialité

14- Ouverture d'esprit et curiosité

15- Sens de l'analyse

16- Capacité à décider

17- Sens de la synthèse

Les Epreuves considérées pour le poste en question : (Valeurs entre [0/20, 20/20]).

- 1- Epreuve de police judiciaire ;
- 2- Epreuve MRT ;
- 3- Renseignement ;
- 4- Management ;
- 5- Expression ;
- 6- Conception, Analyse et Prise de décision.

\vec{c}_i est le vecteur des évaluations du candidat i par rapport aux 23 (17 + 6) critères et épreuves suscités.

$$M = \begin{pmatrix} \vec{c}_1 & \vec{c}_2 & \dots & \vec{c}_n \\ \begin{array}{c|c|c|c} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & & x_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1m} & x_{2m} & & x_{nm} \end{array} \end{pmatrix} \text{ avec } m = 23 \text{ et } n = \text{nombre de candidats.}$$

Cette matrice étant fixe pour les trois scénarios.

Donc soit $\vec{V}_{\text{Eval}_{\text{Explicite}}} = \left(\frac{1}{m} \sum_{i=1}^m x_{1i} \quad \frac{1}{m} \sum_{i=1}^m x_{2i} \quad \dots \quad \frac{1}{m} \sum_{i=1}^m x_{ni} \right)$

(ii) Matrices d'évaluation à base de l'activité issues de l'ED.

Pour simplifier la démonstration, on se contente d'une seule mesure à savoir le **nombre d'infractions** relevées. Il y aura autant de matrice que de mesure.

$$M_{\text{Frontalière-NW}} = \begin{pmatrix} \begin{array}{c|c|c|c} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & & x_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{15} & x_{25} & & x_{n5} \end{array} \end{pmatrix} \text{ avec } n = \text{nombre de candidats.}$$

donc: $\vec{V}_{\text{Frontalière-NW}}^{\text{Eval_Activité}} = \vec{P}_{\text{Frontalière-NW}} * M_{\text{Frontalière-NW}}$

$$\vec{V}_{\text{Frontalière-NW}}^{\text{Eval_Activité}} = (5 \quad 4 \quad 3 \quad 3 \quad 1) * \begin{pmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & & x_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{15} & x_{25} & & x_{n5} \end{pmatrix}$$

d'où : $\vec{V}_{\text{Frontalière-NW}}^{\text{Eval_Activité}} = (5x_{11} + 4x_{12} + 3x_{13} + 3x_{14} + 1x_{15} \quad \dots \quad \dots \quad \dots)$

De la même manière nous obtiendrons les évaluations pour les autres contextes.

d'où : $\vec{V}_{\text{Hauts plateaux}}^{\text{Eval_Activité}} = (6x_{11} + 5x_{12} + 2x_{13} + 2x_{14} + 1x_{15} + 1x_{16} \quad \dots \quad \dots \quad \dots)$

d'où : $\vec{V}_{\text{Frontalière-SE}}^{\text{Eval_Activité}} = (8x_{11} + 7x_{12} + 6x_{13} + 5x_{14} + 3x_{15} + 2x_{16} + 2x_{17} + 1x_{18} \quad \dots \quad \dots \quad \dots)$

(5) Donc le résultat final sera la combinaison des deux vecteurs $\vec{V}_{\text{Eval_Explicite}}$ issu de (i). et $\vec{V}_{\text{Contexte}}^{\text{Eval_Activité}}$ issu de (ii).

Il faut signaler que les valeurs du 2^{ème} vecteur doivent être homogénéisées en considérant les valeurs maximales ; Valeur = (mesure_i / mesure_{Max}).

VI.6. Implémentation du BI

Cette section est l'étape où nous exploitons les données des cubes créés précédemment. Pour se faire nous avons utilisé le logiciel *Business Intelligence 11gr2* pour construire une sorte de tableau de bord.

IV.5.1. Création de référentiel

Avec l'outil administration du BI nous avons créé un nouveau référentiel et lui attribué la source des données. Ensuite, nous avons importé les différents cubes et leurs dimensions.

L'outil administration est constitué de trois couches :

- **La couche physique** : contient des informations sur les sources de données physiques auxquelles Oracle BI Server soumet des requêtes. La manière la plus courante de peupler la couche physique est en important les métadonnées des bases de données et des autres sources de données.

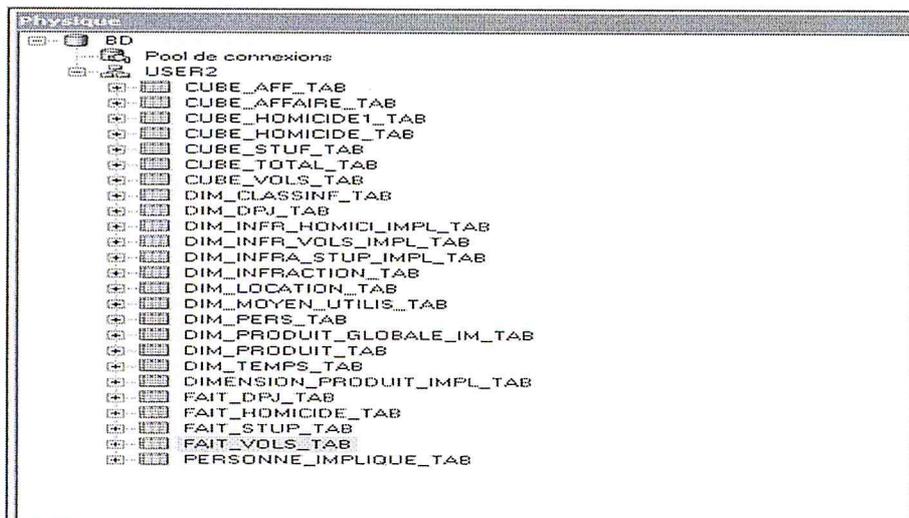


Figure VI-27 La couche physique

- **La couche modèle de gestion et correspondance (ou Logique)** : spécifie la correspondance entre les modèles métiers et les schémas de la couche physique. C'est là que les schémas physiques sont simplifiés pour constituer la base pour la vue des utilisateurs sur les données.
- **Le modèle présentation** : contient un ou plusieurs objets du modèle logique.

1/ Spécification du type d'agrégation pour les différents indicateurs :

Cette étape vient après la création des différents modèles logiques à partir des tables du model physique. La figure suivante montre l'indication du type d'agrégation pour l'indicateur « NB_AFFAIRE ».

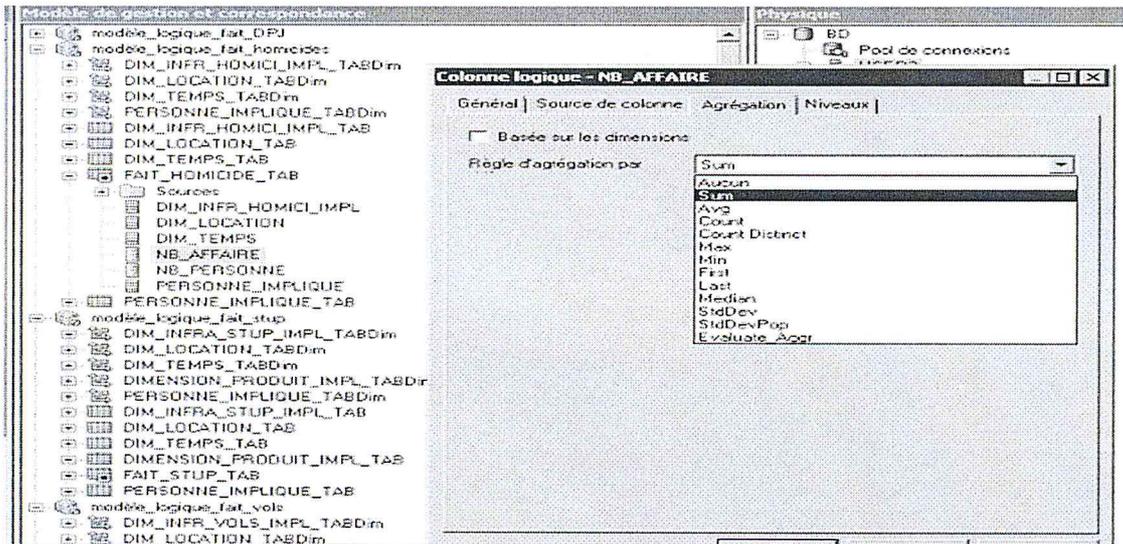


Figure VI-28 Indication du type d'agrégation

2/ Création des dimensions :

Pour chaque dimension on va créer sa propre dimension logique, la hiérarchie de la dimension est basée sur les niveaux.

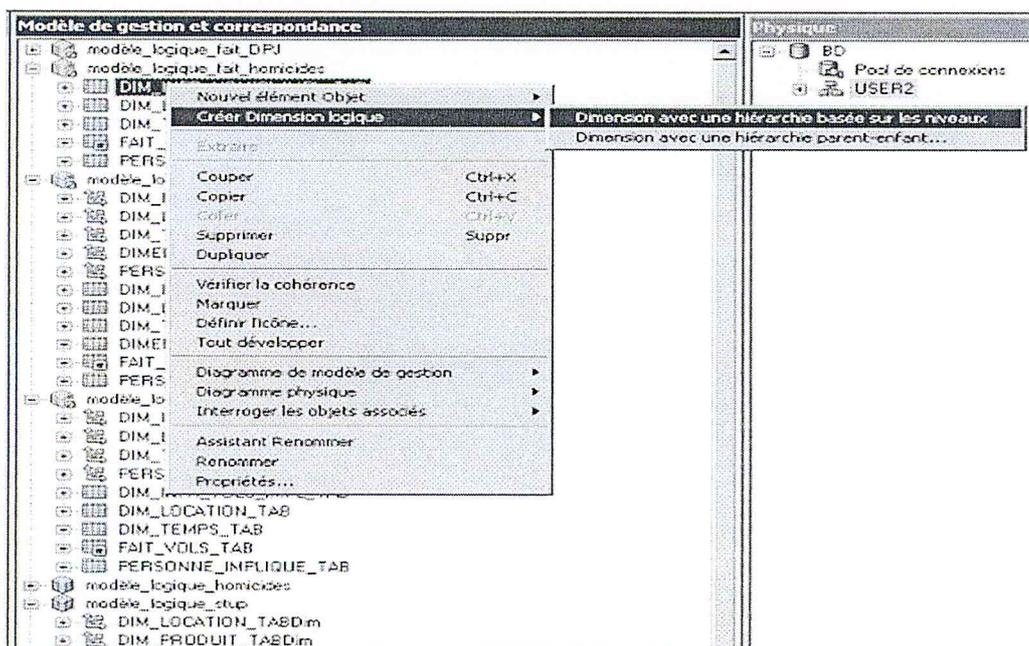


Figure VI-29 Création d'une dimension logique

3/ Création des niveaux de la dimension logique :

Dans cette étape, pour chaque dimension logique on crée ses niveaux avec la spécification du nombre de lignes pour chaque niveau, voir la figure suivante :

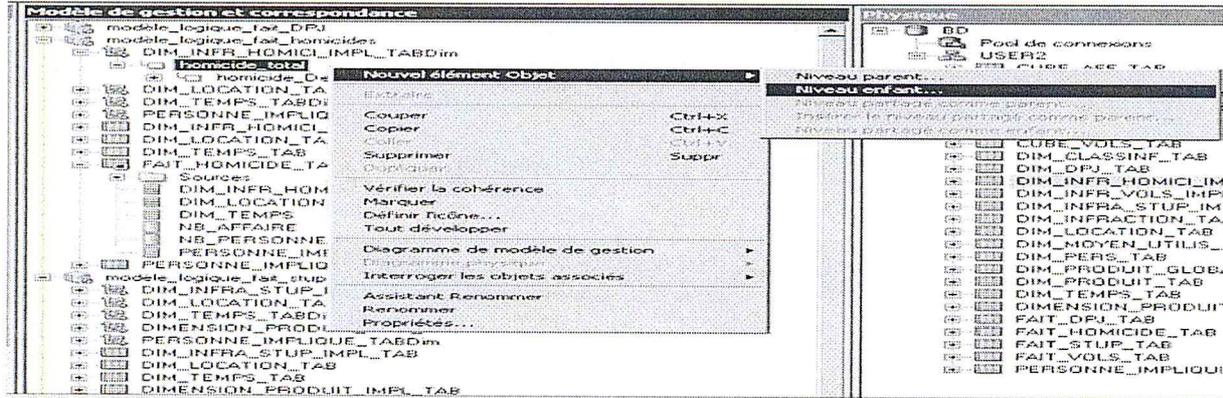


Figure VI-30 Création d'un niveau

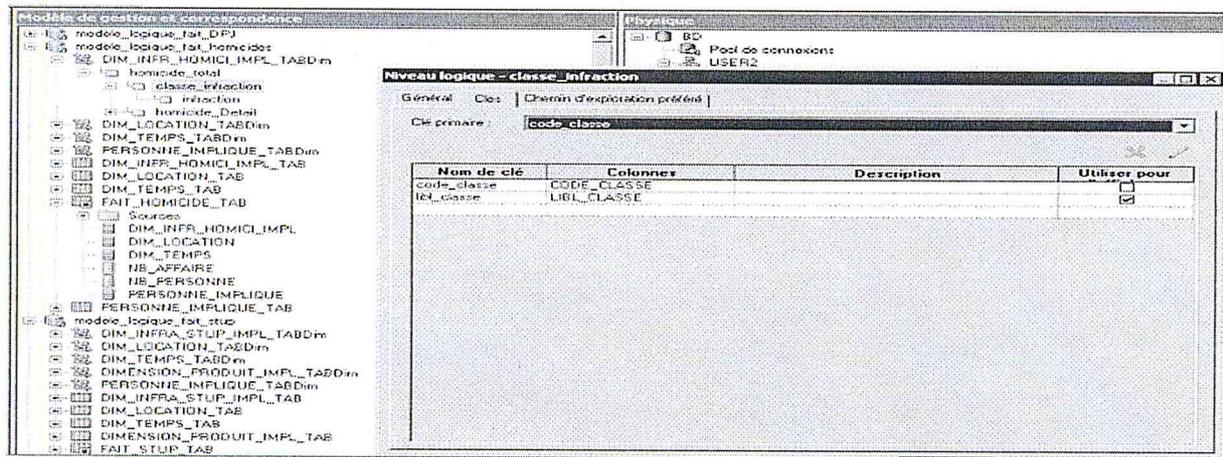


Figure VI-31 L'introduction des informations sur le niveau

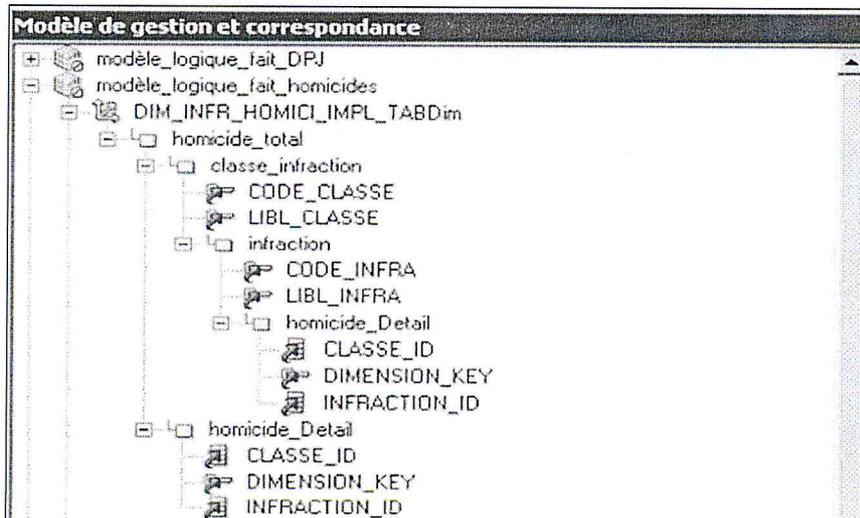


Figure VI-32 Dimension logique infraction finale

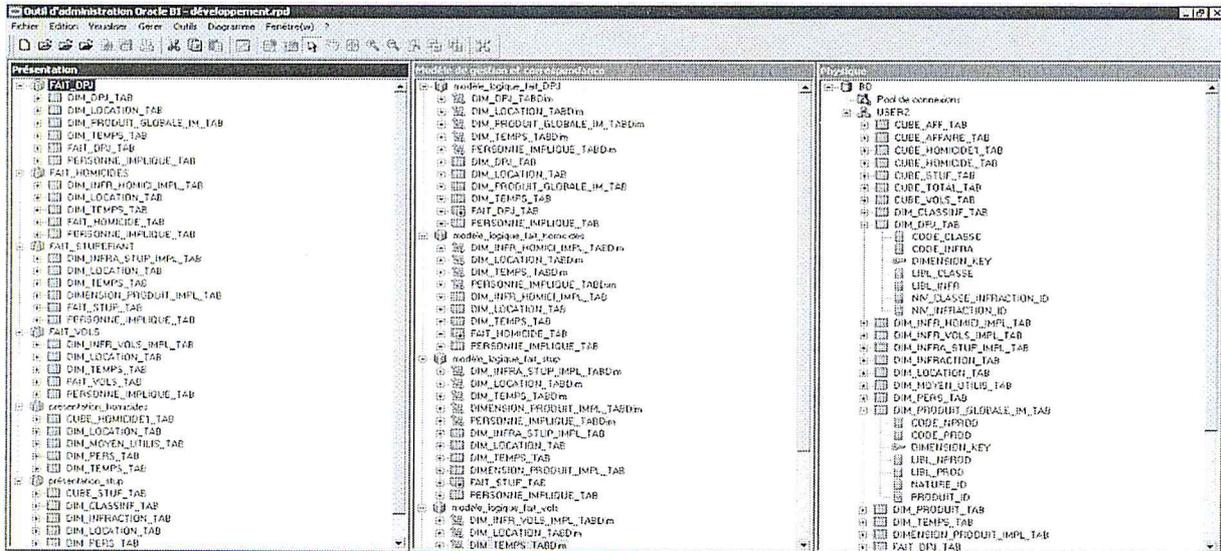


Figure VI-33 Administration final

IV.5.2. Le chargement de référentiel dans le BI

Une fois les données des différents DataMarts sont chargées dans le référentiel, nous allons charger le référentiel créé dans le BI (voir figure ci-dessous) :

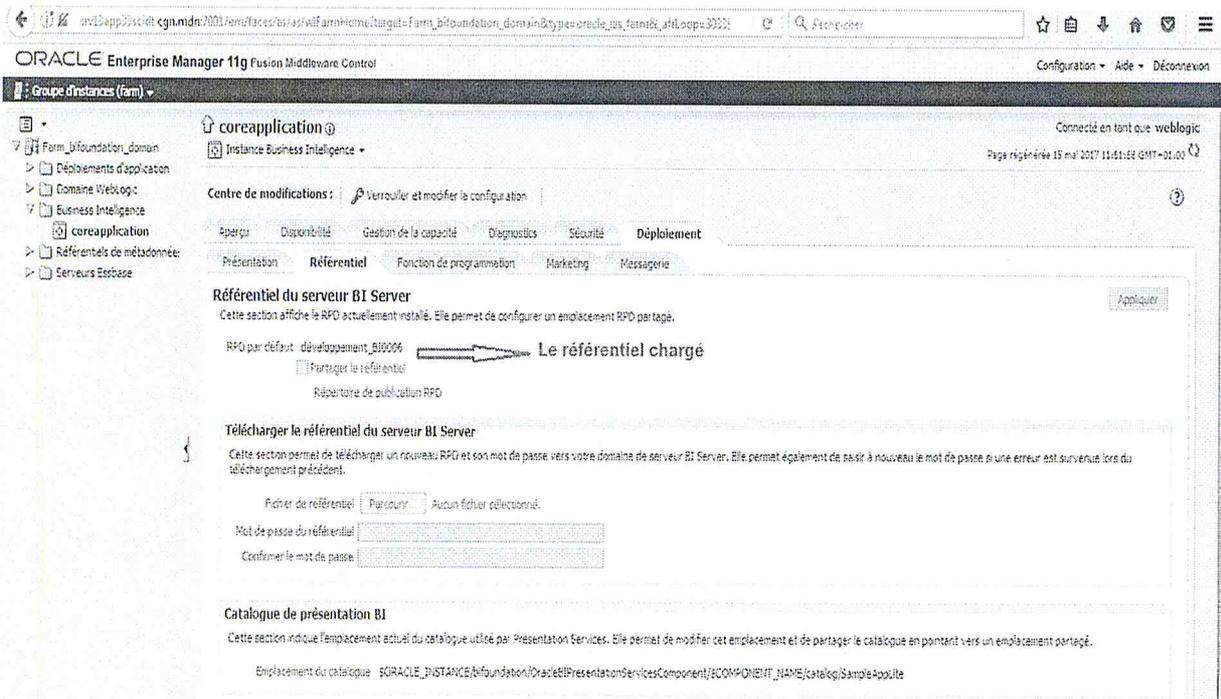


Figure VI-34 Le montage du référentiel

Après le chargement du référentiel il faut redémarrer tous les services du BI (voir figure ci-dessous) :

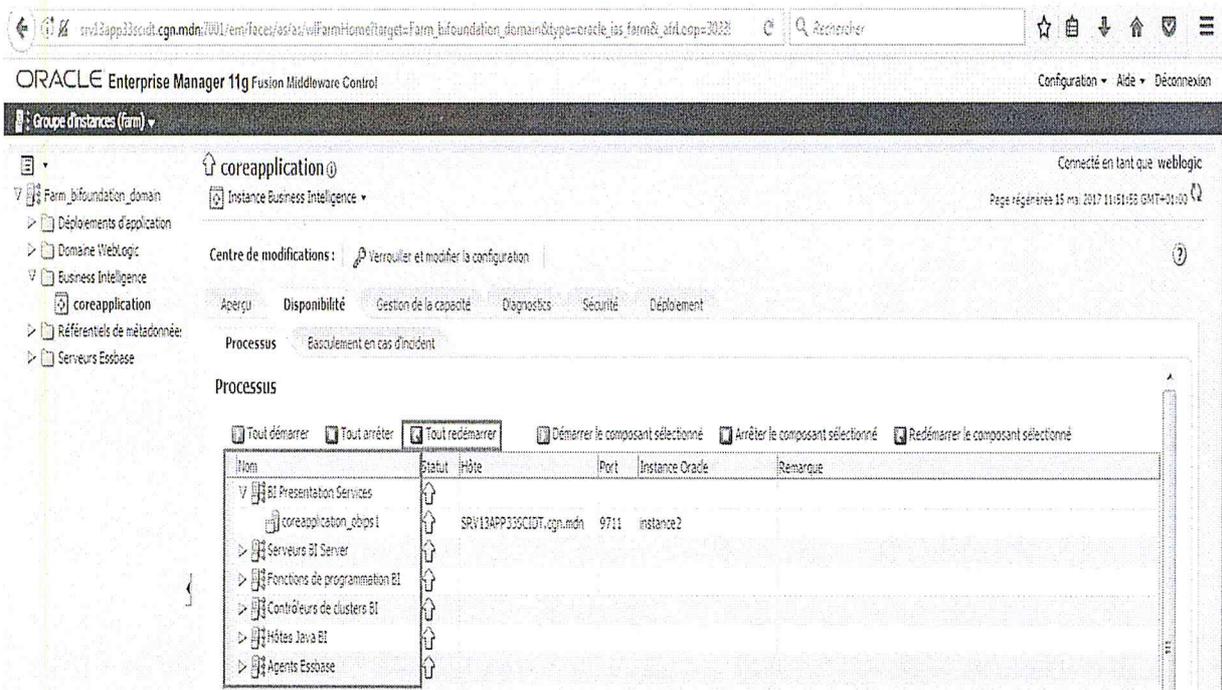


Figure VI-35 Le redémarrage des services du BI

VI.7. Conclusion

Dans ce chapitre nous avons exploité les outils de travail adoptés en l'occurrence : Oracle 10g et Oracle Warehouse Builder comme ETL pour les données conventionnelles. Ensuite nous avons validé notre approche de contextualisation des critères en présentant le déroulement du processus suivant la démarche préconisée en considérant les scénarios du cas d'étude décrits précédemment. Enfin, un aperçu a été donné sur le BI admin et OBIEE pour la création des référentiels et pour enrichir l'affichage des statistiques.

C ONCLUSION G ENERALE

L'évaluation des compétences reste une notion difficilement maîtrisable. En effet, plusieurs outils existent, mais tendent à mettre en avant des compétences plutôt générales, sans pousser l'analyse en tenant compte des caractéristiques de chaque individu. Aucun outil ne permet d'analyser la personne dans son intégralité et donc de mettre en œuvre le potentiel de chacun. Or, ceci reste une étape importante pour les organisations qui recherchent de plus en plus l'efficacité.

Par ailleurs, les entrepôts de données sont nés d'un besoin utilisateur qui n'était pas satisfait par les Systèmes de Gestion de Bases de Données traditionnels. Souvent en lien étroit avec le monde de la recherche, les constructeurs ont intégré dans leurs outils des techniques pour répondre à ces nouveaux besoins, exploiter les données à disposition de l'organisation afin de leur donner de la valeur ajoutée.

La DRH a exprimé le besoin de disposer d'un outil décisionnel lui permettant de gérer les compétences par rapports à des indicateurs clés de performance dont les évaluations explicites et le résultat du traitement des affaires judiciaires (activité), et de les présenter sous forme synthétique et graphique au commandement.

Afin de mener à bien ce travail, nous avons commencé par introduire et définir les différents concepts utilisés par l'approche des systèmes décisionnels. Ensuite, nous avons présenté les différentes phases suivies pour mettre en œuvre un Data Warehouse en partant de plusieurs bases de données opérationnelles.

L'étape suivante a été consacrée à l'étude de l'existant décisionnel au sein des différents services impliqués et la description des différents besoins. La modélisation de la zone de stockage des données s'est faite grâce aux principes de la modélisation dimensionnelle. Cette modélisation offre une vision claire et une compréhension intuitive des modèles proposés. Nous avons, de ce fait, proposé des modèles en étoile afin d'améliorer les performances du futur système.

La partie alimentation de l'entrepôt de données a été la plus fastidieuse et consommatrice en temps, elle nous a permis de vérifier le principe selon lequel il est nécessaire d'y consacrer plus de 70% du temps de réalisation d'un entrepôt de données. Cette étape nous a permis aussi de concevoir et de réaliser, grâce à l'outil Oracle Warehouse Builder, l'extraction, la transformation et le chargement des données.

Intégrer le contexte dans ce cycle de vie complexifie davantage le processus de construction. En effet, passer d'une vision mono-contexte à une vision multi-contextuelle permet de rendre les entrepôts de données sensibles au contexte.

La dernière phase était bien sûr l'exploitation de cet entrepôt de données pour générer des états compréhensibles dotés souvent par des graphes et tableaux susceptibles d'aider les responsables pour prendre des décisions d'une manière stratégique.

Afin de permettre aux décideurs d'interroger et d'exploiter les données stockées dans l'entrepôt de données, nous avons utilisé la plateforme Oracle Business Intelligence, qui offre une palette complète et intégrée d'outils d'analyse et de reporting des données agrégées directement à partir des cubes multidimensionnels conçu.

Notre projet de gestion de compétences n'est qu'une contribution à un plus grand projet de mise en place d'un système global d'aide à la décision au sein de la GN. Notre système présente toujours des imperfections, particulièrement en raison du fait qu'il ne couvre pas l'ensemble des activités ni les autres volets technico-logistique et administratif régissant l'institution. De ce fait, nous proposons, en perspectives, l'intégration, entre autres, de :

- La base de données des activités terroristes ;
- L'application de suivi des missions de police et de renseignement aux frontières ;
- Le système de gestion et de suivi des événements ;
- L'application de suivi des missions de sécurité publique (troubles à l'ordre publique et le maintien d'ordre, attroupement ... etc.).
- La sécurité routière ;
- La lutte contre la cybercriminalité ;
- Les missions de police militaire et administrative ;Etc.

Aussi, nous proposons de considérer dorénavant l'intégration dans la conception de toute solution informatique des mentions relatives à l'auteur et à la qualité d'exécution pour tout type d'activité et de tâche. Ceci permettra d'enrichir davantage les sources de données de l'activité et d'en constituer d'autres pour couvrir les autres volets de l'organisation.

REFERENCES BIBLIOGRAPHIQUE

- [1] Thèse TE0449, 1993, Construction de la compétence.
- [2] <https://www.morganmckinley.fr/fr/article/business-intelligence-un-outil-de-pilotage-rh-indispensable>
- [3] H. Dresner ; « BI : Making the Data Make Sens » ; Gartner Group 2001.
- [4] Zaraté, P. (2013) Outils pour la décision coopérative. Paris: Lavoisier. ISBN 2746295253, 9782746295254
- [5] Le Moigne J.L., « *La théorie du système général, théorie de la modélisation* », P.U.F., 1977.
- [6] J.F. Goglin ; « La Construction du Datawarehouse: du Datamart au Dataweb » ; Hermes 1998.
- [7] Le projet CETISME (2002) Economic Intelligence: A Guide for Beginners and Practitioners. Communautés européennes. I.S.B.N.: 84-451-2350-5
- [8] <http://perso.univ-lyon1.fr/haytham.elghazel/BI/presentation.html> (consulté le 12/05/2017).
- [9] Panos Vassiliadis, Alkis Simitis, and Eftychia Baikousi. A taxonomy of ETL activities. In Proceedings of the ACM twelfth international workshop on Data warehousing and OLAP, pages 25–32. ACM, 2009.
- [10] Alkis Simitis, Panos Vassiliadis, Spiros Skiadopoulos, and Timos Sellis. Data warehouse refreshment. 2007.
- [11] <http://www.piloter.org/> (consulté le 13/05/2017).

- [12] <https://www.coursehero.com> (consulté le 16/05/2017).
- [13] <https://www.leblogdudirigeant.com/> (consulté le 14/05/2017).
- [14] Fernandez, A. 2002. *La création et la publication de rapport d'activité : Le reporting, qu'est-ce que c'est ?* Récupéré Mai 2017 depuis <http://www.piloter.org/businessintelligence/reporting.htm>
- [15] Edgar. F. Codd ; « *Providing OLAP (On-Line Analytical Processing) to User-Analysts : an IT mandate.* » ; Technical report ; E.F. Codd & Associates; 1993.
- [16] Ruysen, P. 2010. *Data Mining et Business Intelligence*. Récupéré Mai 2017 depuis <http://dev.af83.com/2010/03/24/data-mining-et-business-intelligence.html>
- [17] Elmasri et Navathe. *Fundamentals of Database Systems* (6th Edition).
- [18] W. H. Inmon; « *Building the Data Warehouse Third Edition*». Wiley Computer Publishing 2002.
- [19] Mémoire de fin d'études, ESI, 2010 : « *Conception et réalisation d'un Data Warehouse pour la mise en place d'un système décisionnel* » ; Filali Abderrahmane, Kedjnane Sofiane.
- [20] Lecompte, S and T, Boulanger. 2008. *XML par la pratique: bases indispensables, concepts et cas pratiques*. France : Editions ENI. ISBN 978-2-746-4644-3.
- [21] Ponniah, P. 2001. *Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals*. The University of Michigan: Wiley. ISBN 0-471-22162-7.
- [22] R. Kimball et M. Ross ; « *Entrepôts de Données : Guide Pratique de Modélisation Dimensionnelle 2ème édition* ». Vuibert 2002.
- [23] Ralph Kimball and Margy Ross. *The data warehouse toolkit: the complete guide to dimensional modeling*. John Wiley & Sons, 2011.
- [24] Matteo Golfarelli and Stefano Rizzi. *Data warehouse design: Modern principles and methodologies*. McGraw-Hill, Inc., 2009.
- [25] Robert Bruckner, Beate List, and Josef Scheifer. *Developing requirements for data warehouse systems with use cases*. AMCIS 2001 Proceedings, page 66, 2001.

- [26] Martin Glinz. On non-functional requirements. In 15th IEEE International Requirements Engineering Conference (RE 2007), pages 21–26. IEEE, 2007.
- [27] R. Kimball et J. Caserta ; « The Data warehouse ETL Toolkit » ;Wiley Publisshing, INC 2004.
- [28] <http://grim.developpez.com/cours/businessintelligence/concepts/conception-datawarehouse/> (consulté le 20/05/2017).
- [29] R. Kimball; « The data warehouse »; 2005.
- [30] Edgar F Codd, Sharon B Codd, and Clynch T Salley. Providing olap (on-line analytical processing) to useranalysts: An it mandate. Codd and Date, 32, 1993.
- [31] Didier Nakache; « Data Warehouse et Data Mining »; Conservatoire National des Arts et Métiers de Lille; Version 1.1; 15 juin 1998.
- [32] http://blerubrus.free.fr/cnam/ueeng111/solap_onehtml/index.html(consulté le 18/05/2017).
- [33] <http://www.emaze.com> consulté le 05/06/2017.
- [34] Lamri Chouder; « Entrepôt Distribué de Données » ; Thèse de Magistère Option : SI ; Institut National de Formation en Informatique (ESI ex. I.N.I) 2007.
- [35] https://www.tutorialspoint.com/dwh/dwh_olap.htm (consulté le 20/07/2017).
- [36] Hong-Tai Chou and David J DeWitt. An evaluation of buffer management strategies for relational database systems. *Algorithmica*, 1(1-4):311–336, 1986.
- [37] Stefano Ceri, Mauro Negri, and Giuseppe Pelagatti. Horizontal data partitioning in database design. In *Proceedings of the 1982 ACM SIGMOD international conference on Management of data*, pages 128–136. ACM, 1982.
- [38] L. Candillier ; "Contextualisation, visualisation et évaluation en apprentissage non supervisé", thèse de Doctorat, Université de Charles de Gaulle, Lille 3, France, 2006.
- [39] G. Cabanes ; "Classification non supervisée à deux niveaux guidée par le voisinage et la densité", thèse de Doctorat, Université de Paris 13, France, 2010.
- [40] P. Bekhin ; "Survey of clustering data mining techniques", Rapport technique, Accrue Software, 2002.
- [41] A. Meghaoui; "Exploitation des arbres fréquents de dépendance pour la présentation et la classification automatiques de textes", thèse de Maitre ès sciences (M.Sc.), Université de Sherbrooke, Canada, 2008.

- [42] J. Han, et M. Kamber ; "Data Mining : Concepts and Techniques.", Morgan Kaufmann Publishers, 2001.
- [43] C. Ramdan; "Le clustering des données : une nouvelle approche évolutionnaire quantique", thèse de Magistère, Université Mentouri de Constantine, Algérie, 2006.
- [44] E. W. Forgy ; "Cluster analysis of multivariate data: efficiency versus interpretability of classifications", in *Biometrics*, Vol. 21, pp. 768-769, 1965.
- [45] J. B. MacQueen; "Some Methods for classification and Analysis of Multivariate Observations. ", In the Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–297, 1967.
- [46] J. C. Bezdek, R. Ehrlich, et W. Full ; "FCM: The fuzzy c-means clustering algorithm", *Computers & Geosciences*, Vol. 10, N°. 2-3, pp.191-203, 1984.
- [47] R. Krishnapuram, et J. M. Keller ; "The possibilistic C-means algorithm: insights and recommendations", *IEEE Transactions on Fuzzy Systems*, Vol. 4, N°. 3, pp. 385-393, 1996.
- [48] L. J. Heyer, S. Kruglyak¹, et S. Yooseph; "Exploring Expression Data: Identification and Analysis of Coexpressed Genes", *Genome Research*, Vol. 9, pp. 1106-1115, 1999.
- [49] G. Forestier; "Connaissances et clustering collaboratif d'objets complexes multisources", thèse de Doctorat, Université de Strasbourg, France, 2010.
- [50] S. T. Hadjitodoro et L. I. Kuncheva; "Selecting diversifying heuristics for cluster ensembles". Chapitre, "Multiple Classifier Systems", LNCS 4472, pp 200-209, 2007.
- [51] A. Fred, et A. Jain; "Combining multiple clustering using evidence accumulation ", *IEEE Transactions on Pattern Analysis and Machine Learning*, Vol. 27, N. 6, 2005.
- [52] J. Handl et J. Knowles; "Evolutionary Multiobjective Clustering". In the Proceedings of the Eighth International Conference on Parallel Problem Solving from Nature (PPSN VIII), pp 1081-1091. LNCS 3242, 2004.
- [53] J. Handl et J. Knowles; "Exploiting the trade-off -- the benefits of multiple objectives in data clustering". In the Proceedings of the Third International Conference on Evolutionary Multi-Criterion Optimization, pp 547-560, LNCS 3410, 2005.
- [54] J. Handl and J. Knowles; "Improvements to the Scalability of Multiobjective Clustering". In the Proceedings of the Congress on Evolutionary Computation, Vol 3, pp. 2372-2379, 2005.
- [55] J. Handl et J. Knowles; "Multiobjective Clustering Around Medoids". In the Proceedings of the Congress on Evolutionary Computation (CEC 2005) Copyright IEEE, Vol.1, pp 632-639, 2005.

- [56] K. Faceli, A. de Carvalho, et M. de Souto; "Multi-objective Clustering Ensemble". In the Proceedings of the Sixth International Conference on Hybrid Intelligent Systems (HIS'06), pp. 51-51, 2006.
- [57] M. H. C. Law, A. P. Topchy et A. K. Jain; "Multiobjective Data Clustering", In the proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp 424-430, 2004.
- [58] W. Pedrycz; "Collaborative Fuzzy Clustering". Pattern Recognition Letters, Vol. 23, pp 1675-1686, 2002.
- [59] N. Grozavu, et Y. BENNANI; "Classification collaborative non supervisée", In proceings of Conférence francophone sur l'apprentissage automatique CAap'10, 2010.
- [60] A. B. S. Serapião, G. S. Corrêa, F. B. Gonçalves, et V. O. Carvalho; "Combining K-Means and K-Harmonic with Fish School Search Algorithm for data clustering task on graphics processing units", Elsevier Applied Soft Computing, Vol. 41, pp 290-304, 2016.
- [61] T. E. Sweeney, A. C. Chen, et O. Gavaert; "Combined Mapping of Multiple clUsteriNg ALgorithms (COMMUNAL): A Robust Method for Selection of Cluster Number, K", Scientific Reports 5, Article number: 16971, 2015.
- [62] [Robert Griesemer, 2009] «Oracle Warehouse Builder 11g Getting Started-Packt Publishing (2009) ».
- [63] http://www.oracle.com/webfolder/technetwork/tutorials/obe/db/11g/r1/owb/owb11g_update_getting_started_intro/lesson1/less1_start.htm (consulté le 25/05/2017).

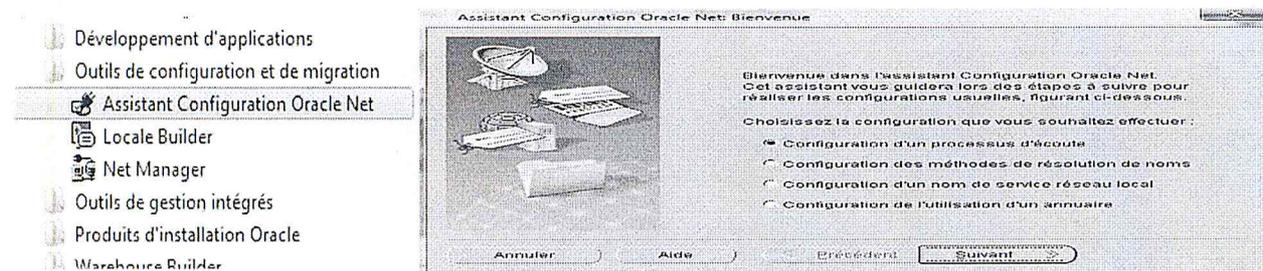
Annexe A Installation & Configuration du OWB

(Oracle Warehouse Builder 11gR2)

1. Installation du logiciel Oracle Database

D'abord nous téléchargeons le fichier d'installation, pour la version appropriée à partir de site officiel oracle <http://www.oracle.com>. Ensuite, nous conduisons l'installation complète mais qui ne comprend pas automatiquement les composantes du client *Warehouse Builder*.

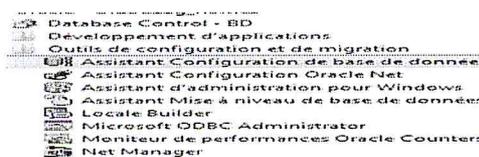
2. Configuration de listener



Configuration du listener.

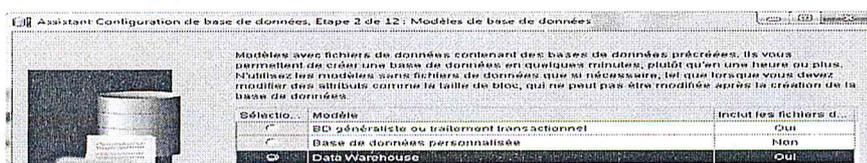
Dans l'assistant de configuration Oracle Net configurer un processus d'écoute. Nommer le et préciser le type de Protocol « **TCP** » et le port « **1521** ».

3. Création de la base de données



Lancement de l'assistant de création de la base de données.

Dans l'assistant de configuration de base de données suivre l'installation et préciser le modèle : **Data Warehouse**



Configuration et création de la base de données.

4. Configuration d'Oracle Warehouse Builder et son référentiel

- ❑ S'assurer d'avoir un compte avec les privilèges SYSDBA.
- ❑ Créer un schéma OWBSYS, puis y installer les objets du référentiel. Les scripts en ligne de commande sont disponibles dans le répertoire « C:\OWB112\owb\UnifiedRepos\ » :

- clean_owbsys.sql
- cat_owb.sql
- reset_owbcc_home.sql
- remote_owb_install.sql

5. Configuration du référentiel et des espaces de travail

Créé et déverrouillé, le schéma OWBSYS est prêt pour héberger les objets d'OWB. Nous utilisons l'application Assistant de référentiel pour configurer l'entrepôt, créer un espace de travail et créer les objets requis pour l'exécution d'OWB.

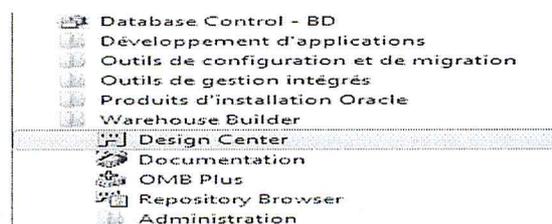


Lancement de l'assistant de création du référentiel.



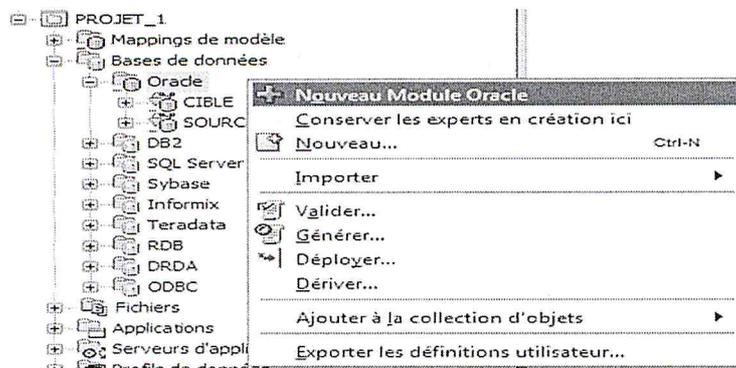
Création du référentiel.

Lancement du Design Center (voir figure ci-dessous).

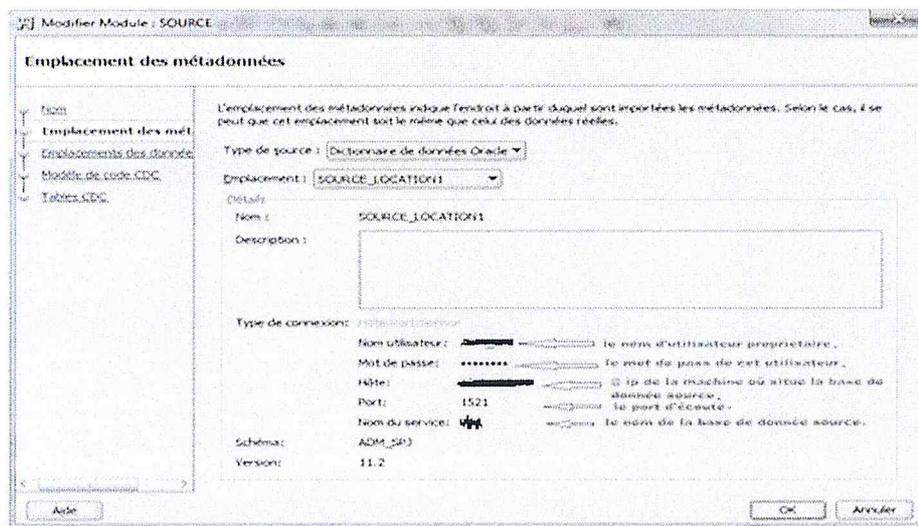


Lancement du Design Center.

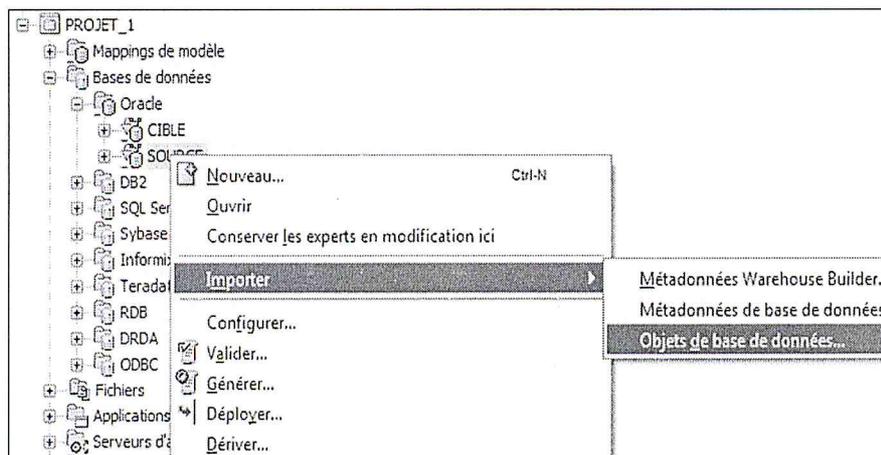
6. Création d'un module source et d'un module cible pour les données



Création d'un nouveau module.



Spécification des informations nécessaire à la création de ce module.



L'importation des objets de la base de données source.

Annexe **B** Installation & Configuration du OBIEE

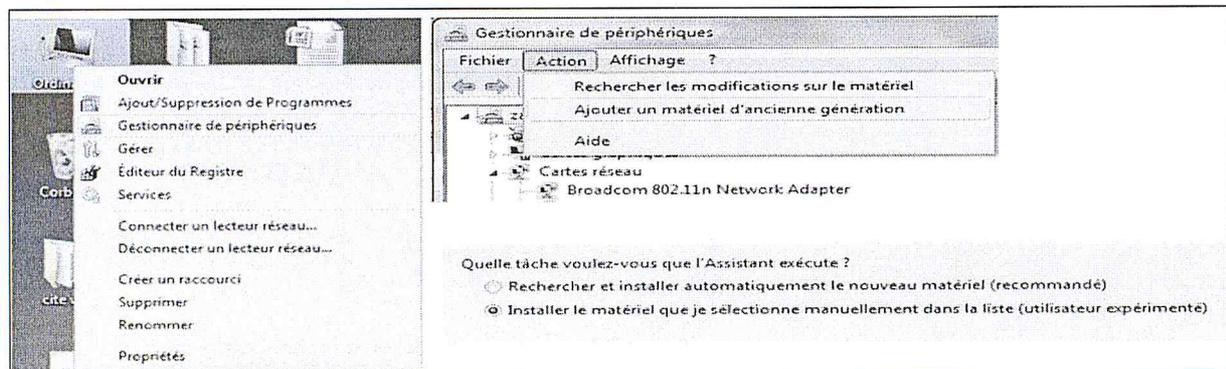
1. Téléchargement des fichiers d'installation du OBIEE + RCU depuis le site oracle

obi_WINDOWSX64_11.1.1.9.0_1of5.zip	31/03/2016 07:19	WinRAR ZIP archive	1 928 102 Ko
obi_WINDOWSX64_11.1.1.9.0_2of5.zip	31/03/2016 17:03	WinRAR ZIP archive	1 883 593 Ko
obi_WINDOWSX64_11.1.1.9.0_3of5.zip	31/03/2016 07:02	WinRAR ZIP archive	1 782 952 Ko
obi_WINDOWSX64_11.1.1.9.0_4of5.zip	30/03/2016 11:20	WinRAR ZIP archive	698 797 Ko
obi_WINDOWSX64_11.1.1.9.0_5of5.zip	31/03/2016 05:15	WinRAR ZIP archive	1 636 643 Ko
rcuHome.zip	30/03/2016 12:25	WinRAR ZIP archive	390 503 Ko

Les fichiers nécessaires pour l'installation du OBIEE.

2. Install Loopback Adapter

Dans *Device manager* on clique sur nom de l'ordinateur et on sélectionne *ajouter un matériel d'ancienne génération* comme indiqué dans la capture d'écran ci-dessous :

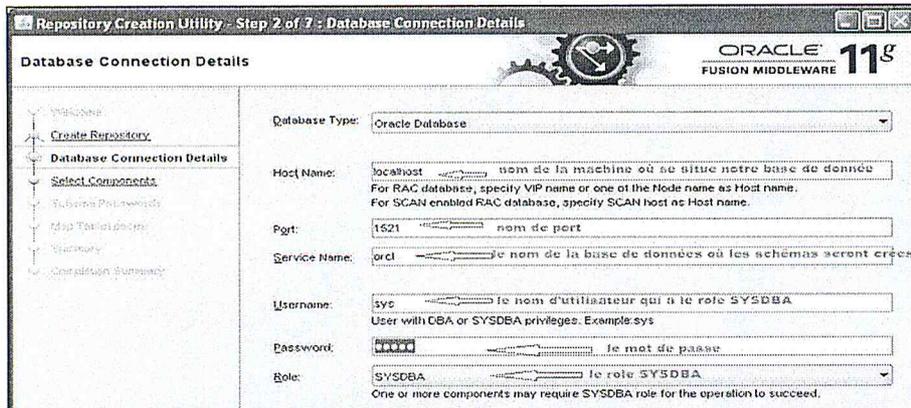


Etapes d'installation de carte de bouclage.

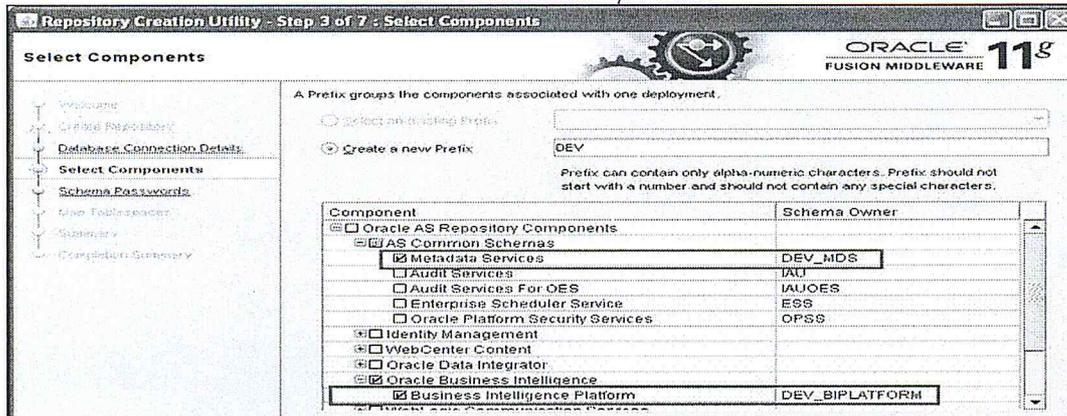
3. Installation du RCU (Schémas de base de données créé par RCU)

Avant d'installer Oracle Business Intelligence, on utilisera RCU pour créer les schémas de Metadata Services (*MDS*) et BI Platform dans la base de données. Pour les installations utiliser le compte SYS car le privilège DBA et le rôle SYSDBA y sont requis.

Avec la console, exécuter le fichier *rcu.bat* dans le dossier *rcuHome*.

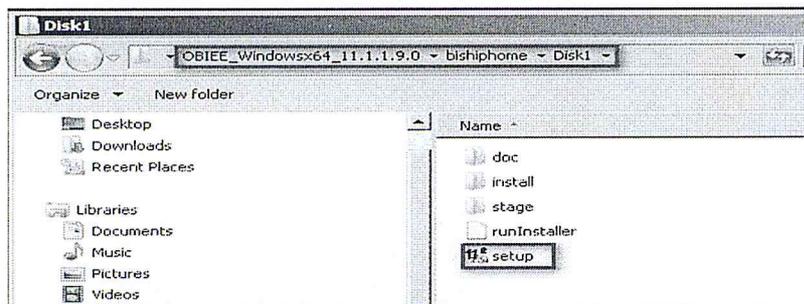


Installation des Schémas par RCU.

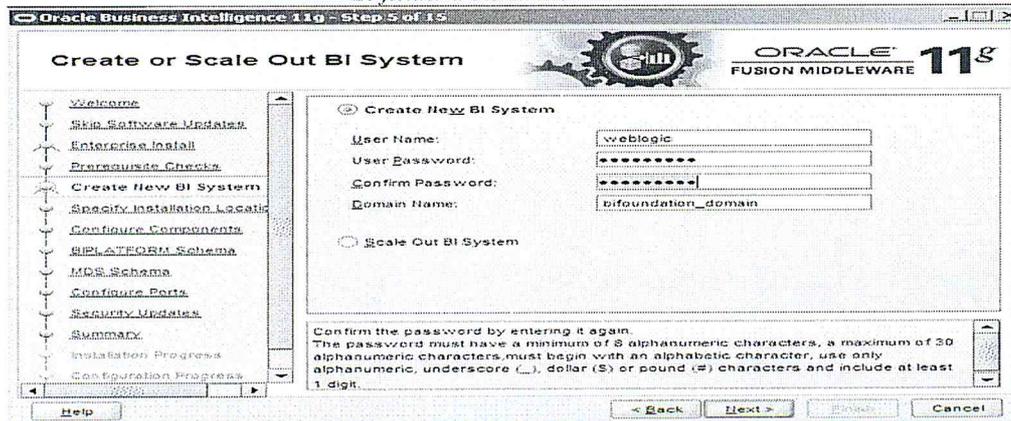


Spécification des schémas MDS et BIPLATFORM pour l'installer.

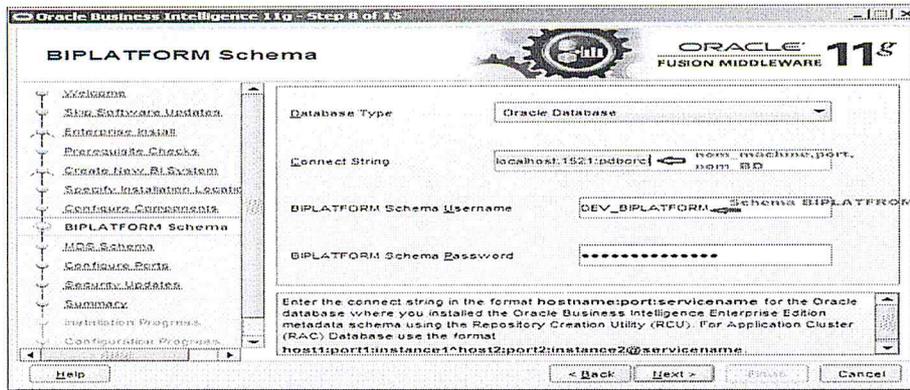
4. Install & Configure OBIEE 11g



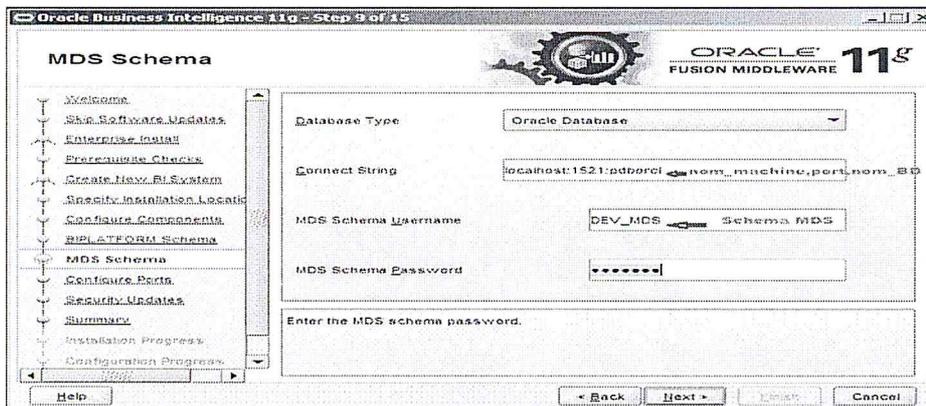
Le fichier d'installation du BIEE.



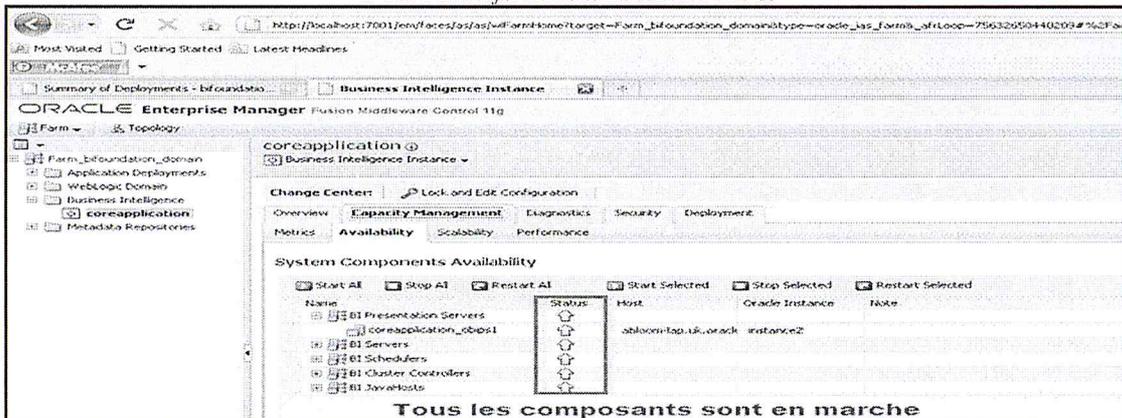
Lancement d'installation du BIEE.



L'entre des informations du BIPLATFORM.



L'entre des informations du schéma MDS.



Tous les composants sont en marche

La disponibilité des différents des composants

```

C:\Windows\system32\cmd.exe
C:\>cd OBIEE11G
C:\OBIEE11G>cd instances
C:\OBIEE11G\instances>cd instance1
C:\OBIEE11G\instances\instance1>cd bin
C:\OBIEE11G\instances\instance1\bin>opmnctl startall
opmnctl startall: starting opmn and all managed processes...
C:\OBIEE11G\instances\instance1\bin>opmnctl status
Processes in Instance: instance1
ias-component      process-type      pid      status
-----
coreapplication_obics1  OracleBIClusterCo~  7916    Alive
coreapplication_obisch1 OracleBIScheduler~  9132    Alive
coreapplication_obijh1  OracleBIJavaHostC~  7592    Alive
coreapplication_obips1  OracleBIPresentat~  7980    Alive
coreapplication_obis1   OracleBIServerCom~  9184    Alive
  
```

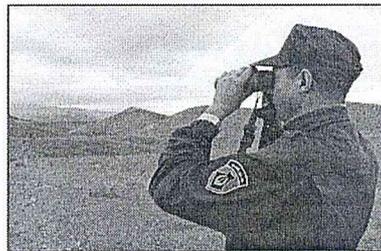
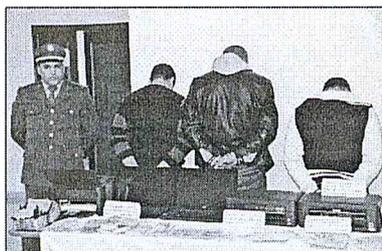
Lancement des services du BIEE.

Annexe **C** Bilan PJ. 1^{er} Semestre 2015 : « Lutte contre le crime et le crime organisé »²⁰

- 6588 Enquêtes et arrestation de 7884 personnes impliquées.
- Arrestation de 579 contrebandiers et saisie de près de 1,3 millions de litres de carburant.
- Saisie de 43,6 t de drogues, 82 k de cocaïne et plus 82 000 comprimés de psychotropes.
- Saisie de plus de 289 t de différents produits alimentaires.

Dans le cadre de la lutte contre la criminalité organisée les unités GN ont traité pas moins de **6 588 affaires** ayant conduit à l'arrestation de **7 884** personnes impliquées, dont **4 283** ont été écrouées.

Les affaires de la **contrebande** et le **trafic de stupéfiants** occupent la première place avec un taux de plus de 31% pour chacune, suivie des affaires de **trafic d'armes et de munitions** avec 892 affaires traitées, soit un taux de 13,54% et 2,23% pour le **trafic de véhicules** avec 147 affaires traitées.



1. Contrebande :

A signaler que, le plus grand nombre d'affaires liées à la contrebande a été enregistré à la frontière Ouest avec plus de 800 affaires traitées, suivie de la frontière Est avec plus de 500 affaires et de l'extrême sud avec plus de 100 affaires.

L'ensemble des affaires de contrebande traitées a conduit à l'arrestation de 579 individus impliqués et la saisie de: 1 298 875 litres de carburant et 136 328 paquets de cigarettes et plus de 289 tonnes de différents produits alimentaires, ainsi que la récupération de 3 445 têtes de cheptel. Notons que la contrebande du carburant représente plus de 70% des affaires de contrebande traitées.

²⁰ Récupérée du site : « http://www.mdn.dz/site_principal/index.php?L=fr » consulté le:17/09/2017.

2. Drogues et psychotropes :

Cependant, le plus grand nombre d'affaires liées au trafic de stupéfiants a été enregistré au niveau du centre (770), l'est (573) et l'ouest (543) du pays.

Durant le premier semestre de l'année en cours la saisie d'une importante quantité de drogue et de psychotropes dans différentes prises qui ont permis de saisir 20619 Kg de kif traité au niveau des frontières, particulièrement celles du Ouest et du Sud-ouest du pays, en plus de 22031 Kg saisi à l'intérieur du pays en provenance de l'Ouest et autres 993 Kg jetée par la mer, d'un total égal à 43643 Kg.

En matière de trafic de psychotropes une quantité estimée à 82476 comprimés a été saisi au cours de la même période, ainsi que la saisie de 82828 grammes de cocaïne à l'ouest et au centre du pays.

3. Le trafic d'armes et de munitions:

892 affaires liées au trafic d'armes et de munitions ont été traitées durant les six premiers mois de l'année 2015, lesquelles ont conduit à l'arrestation de 889 individus et la saisie de 109 armes à feu.

Le plus grand nombre d'affaires traitées a été enregistré à l'Est du pays avec 322 affaires, suivi du centre du pays avec 310 affaires et l'Ouest du pays avec 204 affaires.

4. Le trafic de véhicule:

Dans ce cadre, 147 affaires ont été traitées ayant conduit à l'arrestation de 130 individus impliqués et la saisie de 116 véhicules de tous types.

Le plus grand nombre d'affaires a été traité à l'Ouest du pays avec 71 affaires, suivi de l'Est du pays avec 38 affaires et le Centre avec 23 affaires.

