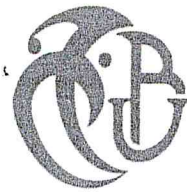
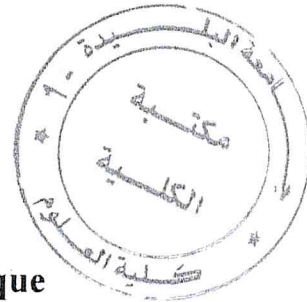


Université Saad DAHLAB de Blida



Faculté des Sciences

Département d'Informatique



Mémoire présenter par :

M<sup>elle</sup> MEHDAOUI Hanane

M<sup>elle</sup> BOUTHELDJA Ouassila

En vue d'obtenir le diplôme de Master

**Domaine:** MI

**Filière:** Informatique

**Spécialité:** Informatique

**Option:** Ingénierie de logiciel

Sujet :

**Spécification des besoins du décideur pour  
l'aide à la découverte des connaissances  
utiles dans une base de données**

Soutenue le:

devant le jury composé de :

M<sup>me</sup> MADANI Amina

M<sup>me</sup> ZAHRA Ratiba

M<sup>lle</sup> GUENDOZ Amina

M<sup>me</sup> BOUMAHDI Fatima

M<sup>me</sup> ZERF Nadjat

Présidente

Examinatrice

Examinatrice

Promotrice

Co-promotrice

Promotion : 2012/2013

MA-004-206-1

# Résumé

## Résumé

---

### Résumé

Aujourd'hui avec l'expansion des TIC, Les systèmes d'information donnent accès à un grand nombre de sources de données hétérogènes et distribuées. A fur et à mesure que les sources se multiplient et que le volume de données disponibles s'accroît, le décideur se voit confronté à une surcharge informationnelle dans laquelle il est difficile de distinguer l'information pertinente de l'information secondaire.

Il devient important de proposer une solution pour faciliter la prise de décision telle que l'extraction d'une connaissance adéquate guide à la bonne décision.

Notre travail a pour objectif d'intégrer au processus KDD, une phase de spécification des besoins du décideur afin de mettre à sa disposition uniquement les connaissances dont il a besoin en utilisant la méthode KADS.

**Mots clefs :** ECD, Aide à la décision, Ingénierie des besoins, méthodes d'ingénierie connaissance., la méthode KADS .

### Abstract

Today with the expansion of ICT, information systems provide access to a wide range of heterogeneous data sources and distributed. As the sources are multiplying and the volume of available data increases, the decision maker is confronted with an information overload in which it is difficult to distinguish relevant information from secondary information.

It became important to propose a solution for facilitate the decision making, suitable Knowledge extraction guide to the best decision

Our work aims to integrate KDD process, a requirements specification phase of the decision maker to make available only the knowledge they need with using KADS method.

**Keywords:** KDD, decision support, needs engineering, knowledge engineering methods, KADS method.

# ملخص

اليوم مع التوسع في تكنولوجيات المعلومات والاتصالات، تسمح نظم المعلومات الوصول إلى مجموعة واسعة من مصادر معلوماتية غير متجانسة وموزعة. وأمام تضاعف مصادر وحجم المعلومات المتاحة، يجد صانع القرار نفسه أمام الحمل الزائد للمعلومات أين يصعب التمييز بين المعلومات ذات الصلة من المعلومات الثانوية .

أصبح من الضروري اقتراح حل من أجل تسهيل اتخاذ القرار حيث أن استخراج المعارف الملائمة يقود إلى القرار الجيد.

عملنا يهدف إلى استخراج المعارف من القاعدة المعلوماتية و هذا بدمج نموذج لتحديد احتياجات صانع القرار و وضع في متناوله المعارف التي يحتاجها فقط.

كلمات البحث: استخراج المعارف من القاعدة المعلوماتية، متطلبات الهندسة، طرق هندسة المعارف، طريقة اقتناء المعلومات و دعم النماذج.

# Sommaire

# Sommaire

1. Introduction Générale.....	1
2. Problématique.....	1
3. Objectifs .....	2
4. Structure du mémoire.....	2

## Partie I : Etat de l'art

### CHAPITRE 01 : Aide à la Décision

1. Introduction.....	3
2. La décision .....	4
2.1. Les Décisions de régulation ou opérationnelles.....	5
2.2. Les Décisions d'adaptation ou tactiques .....	5
2.3. Les Décisions stratégiques .....	5
3. Le décideur .....	5
4. Prise de décision .....	6
4.1. Rationalité limitée .....	6
4.2. Modèle IDC .....	7
5. Outils et méthodes de l'aide à la décision.....	8
5.1. Outils de l'aide à la décision .....	9
5.2. Les Méthodes de l'aide à la décision .....	14
6. Conclusion .....	15

### CHAPITRE 02 :Extraction de Connaissances à partir de Données (ECD ou KDD)

1. Introduction.....	16
2. Découverte de connaissances « KnowledgeDiscovery(KD) » .....	16
3. Processus Extraction des connaissances a partir des données (KDD) .....	17
3.1. Sélection des données .....	19
3.2. Préparation des données .....	20
3.3. Data Mining .....	20
3.4. L'interprétation .....	21
4. Data Mining (fouille de données) .....	21
4.1. Définition .....	21
4.2. Les tache de Data Mining .....	22
4.2.1. La classification .....	22
4.2.2. L'estimation .....	23
4.2.3. La prédiction .....	23
4.2.4. Le groupement par similitude .....	23
4.2.5. L'analyse des clusters .....	23
4.2.6. La description.....	23
5. Conclusion.....	24

## **CHAPITRE 03 :l'ingénierie des besoins**

1. Introduction .....	25
2. Définition besoin ou exigence .....	26
3. L'Ingénierie des Besoins (IB) .....	27
3.1. Définition .....	28
3.2. LeconceptBut .....	29
3.3. Leconceptscénario .....	30
4. Leprocessusde l'IB .....	31
5. Lesapprochesdel'IB .....	32
5.1. L'IB pour GL.....	32
5.1.1. La Méthode KAOS.....	32
5.1.2. La MéthodeCREWSL'Ecritoire.....	33
5.2. L'IB pour l'aide à la décision .....	35
5.2.1. La MéthodeROMC .....	35
5.2.2.L'Approche Annoni (2007).....	37
6. Positionnement de notre recherche dans IB .....	38
7. Conclusion .....	39

## **CHAPITRE 04 :l'ingénierie de connaissances**

1. Introduction.....	40
2. Acquisition de connaissances .....	40
3. Ingénierie de connaissances.....	41
4. Les Méthodes d'ingénierie de connaissances.....	41
4.1. La méthode KADS.....	42
4.1.1. Cycle de vie de KADS .....	43
4.1.2. Les modèles de KADS.....	44
4.1.2.1. Le modèle de l'organisation.....	45
4.1.2.2. Le modèle d'application.....	45
4.1.2.3. Le modèle de tâches.....	45
4.1.2.4. Le modèle de coopération.....	46
4.1.2.5. Le Modèle d'Expertise.....	47
4.1.2.6. Le Modèle Conceptuel.....	48
4.2. La méthodeKOD.....	48
4.3. La méthode MACAO.....	49
5. Conclusion.....	51

## **PARTIE II : le modèle de découverte des connaissances utiles dans une base de données et l'implémentation**

### **CHAPITRE 5 : le modèle de découverte des connaissances utiles dans une base de données**

1. Introduction .....	52
2. Positionnement de la méthode KADS dans le processus KDD.....	52
3. Le modèle proposé .....	53
4. Modèle de spécification des besoins du décideur.....	54
4.1. Le modèle d'Organisation/ Application.....	54
4.2. Le modèle de tâche .....	57

4.3. Le modèle d'Expertise.....	59
4.3.1. La connaissance du domaine.....	59
4.3.1.1. Le schéma du domaine.....	60
4.3.1.2. La base de connaissances.....	61
4.3.2. La connaissance d'inférences.....	61
4.3.2.1. Les inférences et les rôles de connaissances.....	61
4.3.2.2. La structure d'inférences.....	62
4.3.3. La connaissance de la tâche.....	63
5. Etude de cas .....	69
6. Utilisation des résultats du modèle proposé par le processus KDD.....	68
7. Architecture générale du système .....	70
8. Conclusion.....	73

## **CHAPITRE 6 : IMPLEMENTATION**

1. Introduction.....	74
2. Environnement de développement Visual Studio 2008.....	74
3. Langage de programmation.....	74
4. L'interface d'application.....	75
5. Conclusion.....	77

<b>Conclusion générale .....</b>	<b>78</b>
----------------------------------	-----------

<b>Bibliographies.....</b>	<b>79</b>
----------------------------	-----------



# Liste des figures

## Liste des figures

Figure 1.1 : Processus de décision [Simon 60].....	7
Figure 1.2 : Classification des outils de l'aide à la décision [Ghomari 08].....	9
Figure 2.1 : Processus KDD [Fayyad 96].....	19
Figure 3.1: l'ingénierie des SI et l'ingénierie des besoins[Rolland 03] .....	26
Figure 3.2 : Besoins et Exigences [Essame 02].....	27
Figure 3.3 : Le processus de l'IB [Rolland 03].....	32
Figure 3.4 : le méta-modèle KAOS [Matoussi 2011].....	33
Figure 3.5 : La notion de fragment de besoin [Rolland 03].....	35
Figure 3.6: Déroulement de la méthode ROMC [Sprague 82].....	37
Figure 3.7 : Diagramme décisionnel d'Annoni[Annoni 07].....	38
Figure 4.1 : Modèles et espaces de développement de KADS [Dieng 90].....	43
Figure 4.2 : Interactions des modèles KADS.....	44
Figure 4.4 : Les différents modèles de KADS.....	44
Figure 4.4 : les interactions du modèles de coopération .....	46
Figure 4.5 : Le Modèle de KOD[Dieng 90].....	49
Figure 5.1: Le modèle de spécification des besoins du décideur.....	53
Figure 5.2 : Enchaînement des modèles de spécification/KDD.....	54
Figure 5.3 : Diagramme de classe du modèle Organisation /Application .....	55
Figure 5.4 : Diagramme du modèle de tâche.....	58
Figure 5.5 : Différents types de connaissances.....	59
Figure 5.6 : Exemple d'une relation.....	60
Figure 5.7 : Représentation graphique de l'inférence « spécifier ».....	62
Figure 5.8 : Diagramme du modèle de tâche pour identifier les critères de priorité.....	64
Figure 5.9 : Les Concepts du domaine distribution logement.....	66
Figure 5.10 : Structure d'inférence pour la tâche d'évaluation d'un critère.....	67
Figure 5.11 : diagramme de la tâche Création critère.....	68
Figure 5.12 : diagramme de la tâche Modification critère.....	69
Figure 5.13 : diagramme de la tâche Suppression critère.....	69
Figure 5. 14 : L'incorporation les connaissances du décideur au processus KDD.....	70
Figure 5.15 : Architecture générale du système.....	72
Figure 6.1 : Interface Authentification.....	75
Figure 6.2 : Interface Traitement décideur selon le modèle organisation /application.....	75
Figure 6.3 : Interface Traitement de requête décideur.....	76
Figure 6.4 : Interface sélection Critère du décideur.....	76
Figure 6.5 : Interface du résultat de la sélection critère.....	77
Figure 6.6 : Interface de classification et visualisation.....	77

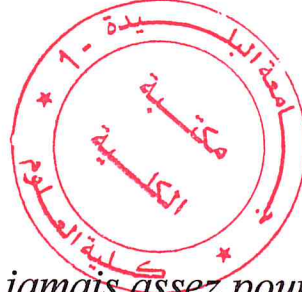
## Liste des tableaux

Tableau 2.1 : Liste des différentes étapes de processus d'ECD [Andrassyová 02]. .....	18
Tableau 5.1 : Description de structure d'inférence d'évaluation du critère.....	68

# **Remerciement**

## **Dédicace**

# Dédicace



*Je dédie ce travail*

*À mes parents que je ne remercierai jamais assez pour toute  
l'aide qu'ils m'ont prodiguée,*

*À mes frères et mes sœurs*

*À tous mes amis,*

*À toutes les personnes que je connais et que je n'ai pas citées*

*Wassila*

## *Dédicace*

*Que ce fut long et dure d'arriver à la fin et de réaliser ce travail qui n'a pas été facile à réaliser sans la présence de personnes chères à nous pour nous aider, je dédie mon travail tout d'abord aux membres de ma famille mon cher père ma chère mère mes frères mes sœurs et sans oublier ma deuxième famille BOUTHELDJA qui m'ont aidé et tout mes amis formidable surtout Dalila sans oublier toute personne qui m'aide de près ou de loin dans mon travail.*

*Hanane*

# Remerciement

*Avant toute chose, nous tenons à remercier le Dieu le tout puissant de nous avoir donné la volonté, la patience et le courage de poursuivre et d'achever ce travail dans les bonnes conditions.*

*En second lieu Nous tenons à remercier chaleureusement et affectueusement tous ceux et toutes celles qui de près ou de loin ont contribué à la réalisation de notre projet.*

*Nos vifs remerciements accompagnés de toute notre gratitude vont tout d'abord à notre Co-promotrice M<sup>me</sup> ZERF Nadjet pour nous avoir proposé ce sujet, pour les conseils qu'il n'a cessé de nous prodiguer, ses encouragements et surtout pour la confiance qu'il nous a accordée pour la réalisation de ce projet, Nous prions Allah de lui rendre grâce pour avoir fait de notre travail avec elle, un réel honneur et grand plaisir.*

*Nous tenons également à remercier Mme BOUMAHDI fatima et Mr R. CHALAL pour leurs conseils avisés, leurs idées riches et leurs encouragements, Nous prions Allah de leurs rendre grâce pour avoir aidé nous, un réel honneur et grand plaisir.*

*Nous remercions les membres du jury d'avoir consentis à évaluer notre travail.*

*Par ailleurs ; nous rendons un vibrant hommage à l'ensemble du corps professoral du département d'informatique de l'université Saâd DAHLAB de Blida surtout SALIMA et HAYATE qui ont contribué activement et vaillamment à notre formation.*

*A tous ceux et à toutes celles dont les noms n'apparaissent pas sur cette page, qu'ils demeurent convaincus, que nous ne les avons point oubliés et qu'ils soient assurés de notre profonde gratitude.*

*Merci*

# **Introduction Générale**



# INTRODUCTION GENERALE

---

## 1. Introduction Générale

La prise rapide de décision étant aujourd'hui un enjeu crucial pour la compétitivité de l'entreprise. En effet, l'environnement des décideurs est de plus en plus complexe et évolue rapidement.

L'aide à la décision [Roy 85] est devenue une discipline tout entière qui tente d'appliquer les théories de la décision et les nouvelles technologies pour aboutir à des outils et méthodes qui supportent un ou plusieurs décideurs dans leurs tâches de prise de décision.

Le processus KDD (KnowledgeDiscovery in Databases) est un outil d'aide à la décision et définit comme un processus non trivial, permettant l'identification, au sein des données, de patterns valides, nouveaux, potentiellement utiles et les plus compréhensibles possible [Fayyad 96].

Les connaissances extraites à partir des bases de données opérationnelles sont censées être potentiellement cruciales [Grundstein 03] ou utiles pour le décideur afin de favoriser une meilleure prise de décision.

Le domaine de l'ingénierie des besoins pour l'aide à la décision a connu, ces dernières années, l'apparition de quelques approches essayant de résoudre le problème d'analyse des besoins décisionnels stratégiques.

L'IC est un domaine qui s'intéresse à la conception de systèmes qui visent à traiter ou à aider des opérateurs à traiter des problèmes mal posés, faisant appel à des connaissances jusqu'ici non explicités ou non modélisées.

## 2. Problématique

Il est devenu difficile de prendre une décision, cette difficulté a tendance à l'accroissement de plusieurs facteurs :

- Notre problème principale est extraction des connaissances potentiellement utile par le décideur (les connaissances extraites n'exprime pas ses propre besoins)
- le volume de données est très grand.
- Le temps de réponse élevé.
- le coût des erreurs de décision est plus grand, en raison de la complexité et de l'importance des conséquences engendrées par une décision.

## 3. Objectif

Notre travail de recherche a pour objectif d'intégrer au processus KDD, un modèle de spécification des besoins du décideur afin de mettre à sa disposition uniquement les connaissances dont il a besoin.

Dans ce mémoire, nous allons montrer comment les méthodes d'IC qui s'intéressent aux décideurs peuvent aider à déterminer les connaissances utiles durant le processus KDD.

Donc nous nous essayons de pallier ce déficit qu'on trouve au niveau de la première phase du KDD (phase sélection).

## 4. Structure de mémoire

Ce mémoire est structuré de la manière suivante :

**Chapitre 1** : est consacré au rappel aux différents aspects de l'aide à la décision. La prise de décision est présentée en premier, le modèle IDC, suivi par un aperçu bref sur les outils et leurs méthodes de réalisation d'aide à la décision.

**Chapitre 2**: présente une description détaillée du processus d'extraction des connaissances à partir des bases de données (KDD). Il définit ce processus et détaille ses phases (sélection de données, prétraitement de données, fouille de données et validation des résultats).

**Chapitre 3** : pour montrer l'importance de la spécification des besoins, nous passerons en revue l'ingénierie des besoins, le processus d'ingénierie des besoins et les différentes approches et les concepts de l'ingénierie des besoins.

**Chapitre 4** : présente les méthodes d'ingénierie de connaissance qui s'intéressent au décideur comme la méthode KADS, MACAO et KOD, en détaillant la méthode « KADS » (Knowledge Acquisition and Design Support) [Dieng 90], Il présente cette méthode et décrit son cycle de vie (modèle organisation, application, tâche, expertise, coopération, fonctionnel, et à la conception détaillée) et une description brève de la méthode KOD et la méthode MACAO.

**Chapitre 5** : présente notre approche, à travers la description, des étapes du processus de spécifications des besoins du décideur et les modèles qui supportent le processus proposé.

**Chapitre 6** : présente notre test et validation pour notre modèle.

Le mémoire se termine par une conclusion générale qui dresse un bilan général du travail réalisé.

# **Partie I**

## **L'état de l'art**

# **Chapitre 01**

## **AIDE A LA DECISION**

## 1. Introduction

L'aide à la décision stratégique nécessite toujours une information importante sur l'environnement de l'entreprise (clients, marchés, concurrents, technologie disponible, environnement économique et physique...).

L'aide à la décision est définie comme étant << L'activité de celui qui, prenant appui sur des modèles clairement explicites mais non nécessairement complètement formalisés, aide à obtenir des éléments de réponse aux questions que se pose un intervenant dans un processus de décision, éléments concourant à éclairer la décision et notamment à prescrire ou simplement à favoriser, un comportement de nature à accroître la cohérence entre l'évolution du processus d'une part, les objectifs et le système de valeurs au service desquels cet intervenant se trouve place d'autre part >>[Roy 85].

Cette définition traduit le fait qu'on ne "résout" pas un problème, on aide le décideur à construire une représentation pertinente de la situation.

Le but de l'aide à la décision est de rendre le travail plus intelligent et d'opérer un changement dans les méthodes de travail personnelles (efficacité) et de l'entreprise (efficacité, réactivité, productivité, innovation).

Les attentes des managers, au regard des potentialités des Systèmes d'Aide à la Décision, annoncées dans les recherches sont triples [Carlsson 02]:

- ✓ les décideurs peuvent résoudre des problèmes complexes d'une manière plus efficace.
- ✓ les décideurs peuvent prendre des décisions plus rationnelles, sans pour autant utiliser des outils d'optimisation.
- ✓ les décideurs peuvent profiter pleinement de leurs capacités cognitives, les activités routinières étant à la charge du SIAD.

L'aide à la décision représente l'ensemble des moyens (modèles, méthodes, outils, concepts,...) mis à la disposition du décideur pour faciliter la prise de décision. Chaque outil informatique est doté de ses propres méthodes de réalisation et de mise en œuvre: méthodes KOD ou KADS pour les systèmes experts, Data Mining, KDD pour l'extraction des connaissances à partir des données, Les méthodes sont diversifiées, nombreuses, non reliées, utilisant des concepts ou formalismes différents selon les outils.

Dans ce premier chapitre, nous essayons de présenter les concepts clé qui tournent autour de l'aide à la décision. Nous commençons d'abord par le concept de base qui est la décision.

# CHAPITRE 01 : AIDE A LA DECISION

---

Nous présenterons le décideur, Nous verrons par la suite le processus de prise de décision et le modèle IDC et nous évoquerons les méthodes et les outils d'aide à la décision.

## 2. La décision

La décision est le fait d'un acteur (ou d'un ensemble plus ou moins cohérent d'acteurs) qui effectue un choix entre plusieurs solutions susceptibles de résoudre le problème ou la situation auxquels il est confronté.

On peut définir la décision comme étant un «*acte par lequel un ou des décideurs opèrent un choix entre plusieurs options permettant d'apporter une solution satisfaisante à un problème donné* ». Cette notion de décision a évolué dans le temps au fur et à mesure que ce sont transformés et complexifiés les procédures de prise de décision.

Au sens classique du terme on assimile la décision à l'acte par lequel un individu (disposant du pouvoir de décider) prend les mesures favorisant la création et la répartition des richesses dans une entreprise en s'appuyant sur un ensemble d'informations à sa disposition sur le marché. Dans son approche plus moderne, la prise de décision apparaît plutôt comme « un processus d'engagement progressif, connecté à d'autres, marqué par l'existence reconnue de plusieurs chemins pour parvenir au même et unique but » [Sfez 92]. Les évolutions du concept de décision est révélatrice d'un certain nombre d'évolutions dans la manière d'appréhender le processus de la prise de décision :

La décision n'est plus un acte unique et constant fondé sur la recherche du profit mais repose sur un ensemble successif de décisions de moindre portée.

La décision n'est plus fondée sur la recherche d'un seul objectif mais intègre un nombre plus important de variables. La décision intervient dans un contexte plus aléatoire dans le sens où la manière d'atteindre l'objectif poursuivi peut passer par différents types d'actions.

Ces évolutions sont compréhensibles car elles ne font que souligner les mutations du système productif : l'environnement de l'entreprise est devenu plus complexe, plus incertain aussi et la prise de décision ne repose plus sur un seul individu mais peut être partagée entre un nombre élevé d'acteurs agissant au sein de l'entreprise. Cette multiplication du nombre de décideurs reflète par ailleurs la diversité des décisions qui doivent être prises dans une entreprise.

Pour notre part, nous adoptons une vision systémique de la décision, c'est le processus de décision et son résultat.

Les décisions sont classées en fonction de leur horizon temporel (court, moyen ou long terme), de leur degré d'incertitude (avenir certain, incertain ou aléatoire), de leur champ

d'application (l'entreprise dans son ensemble, une fonction ou un service) et de leur degré de structuration d'un problème (pas forcément facilement identifiable, ce qui suppose une analyse préalable permettant de structurer le problème avant d'envisager des solutions possibles).

En fonction de ces caractéristiques de décision, il a pu être défini 3 niveaux de décisions, les décisions stratégiques, les décisions tactiques et les décisions opérationnelles (classées par ordre décroissant d'importance).

### **2.1. Les Décisions de régulation ou opérationnelles**

C'est des décisions normalisées et structurées. Elles s'appuient sur des informations internes portant sur le fonctionnement de l'entreprise qui constitue un environnement stable et prévisible.

### **2.2. Les Décisions d'adaptation ou tactiques**

Ces décisions portent sur les pratiques d'adaptation à des changements de l'environnement tels que l'entreprise ne peut utiliser ses processus habituels. Toutefois ces changements ne sont pas suffisamment influents pour remettre en cause profondément les savoir-faire, et ceux-ci permettront de trouver de nouveaux processus. Les décisions auront besoin d'aller chercher ces savoir-faire et de les mettre en rapport avec les informations dont elle dispose sur son environnement.

### **2.3. Les Décisions stratégiques**

Ces décisions recouvrent l'attitude proactive de l'entreprise ou l'attitude réactive face à des environnements incertains et non probabilisables. Ces décisions sont de type peu structuré et peu normalisé, et auront un besoin majeur d'informations sur l'environnement, des informations externes à l'entreprise.

## **3. Le décideur**

Il existe plusieurs définitions du décideur, nous ne retenons que celles qui concernent notre problématique.

[David 01] donnent une définition basée sur le rôle du décideur dans l'environnement interne et externe de l'entreprise : « *celui que est apte à identifier et à poser le problème à résoudre en terme d'enjeu, de risque ou de menace qui pèse sur l'entreprise* ».

Une autre définition donnée par [Schneider 00] repose également sur le rôle joué par le décideur dans son environnement. Ce dernier considère le décideur comme « un acteur social. Il doit être décrit et traité comme un système ouvert qui, activement, saisit de

## CHAPITRE 01 : AIDE A LA DECISION

---

l'information, qui agit. Il possède une certaine indépendance cognitive ». En se basant sur ces deux définitions nous considérons que le décideur, sans être le seul, est un observateur bien placé dans son environnement interne et externe, et c'est grâce à sa position qu'il est sensible à détecter les signaux de l'environnement [Bouaka et al 02].

### 4. Prise de décision

L'activité de prise de décision est définie comme l'ensemble des actions que le décideur effectue pour prendre sa décision et la mettre en œuvre (consultation, recherche d'informations) [Barabel 96]. Ces actions, qui constituent le processus de décision répondent aux objectifs de l'organisation en intégrant les contraintes inhérentes à la situation analysée (situation de décision). [Chalal 07].

Ce processus est déclenché par le stimulus (incitation à l'action) [Bouaka 03] [Berard 09] et s'étale jusqu'à l'instant où l'engagement est pris [Chalal 07].

Les décisions sont souvent prises sur la base d'intuitions et d'expériences passées [Kahnemann 82]. Ce type de stratégies ne peut s'appliquer qu'à des problèmes familiers, Lorsque nous sommes confrontés à des situations nouvelles, la tâche de prise de décision devient beaucoup plus difficile [Simon 82].

#### 4.1. Rationalité limitée

*Pourquoi est-il impossible d'avoir une rationalité « parfaite », c'est-à-dire une rationalité où les moyens sont en parfaite adéquation avec leurs fins et permettent de choisir l'action la plus efficace compte tenu des contraintes en ressources ?*

Les individus n'ont pas la capacité de traiter l'ensemble des informations en provenance de leur environnement : «Chaque organisme humain vit dans un environnement qui produit des millions de bits de nouvelle information chaque seconde, mais le goulot d'étranglement de l'appareil de perception n'admet certainement pas plus de 1000 bits par seconde et probablement moins» [Simon 59].

En effet L'homme est limité par la structure de sa mémoire à court terme et l'insuffisance de ses connaissances, la limitation des ressources,...etc.

Ainsi, lorsque des individus rencontrent des situations de décision, ils ne peuvent anticiper toutes les alternatives qui se présentent à eux ni ne disposent de toutes les informations, et se contenteront de solutions « satisfaisantes ». Ce que cherche le décideur c'est une solution satisfaisante compte tenu de ses fins. C'est le principe de rationalité limitée. Dans ce contexte, le décideur cherche les ou la premières solutions qui lui procurent un



# CHAPITRE 01 : AIDE A LA DECISION

résultat satisfaisant et l'affirmation que l'on peut organiser rationnellement le processus de recherche de cette décision.

## 4.2. Modèle IDC

Le Modèle IDC (Intelligence, Design, Choice) est décrit par Simon comme un processus cognitif qui est assorti d'une situation de traitement d'information. Cette dernière présente une structure générique en trois phases qui possèdent chacune ses propres buts et des mécanismes spécifiques de traitement d'information.

Soit le modèle IDC, qui se compose des trois phases suivantes : Intelligence (compréhension), Design (conception) et Choice (choix). Une quatrième phase appelée « Review ou implémentation » fut rajoutée ultérieurement au modèle (figure 1.1).

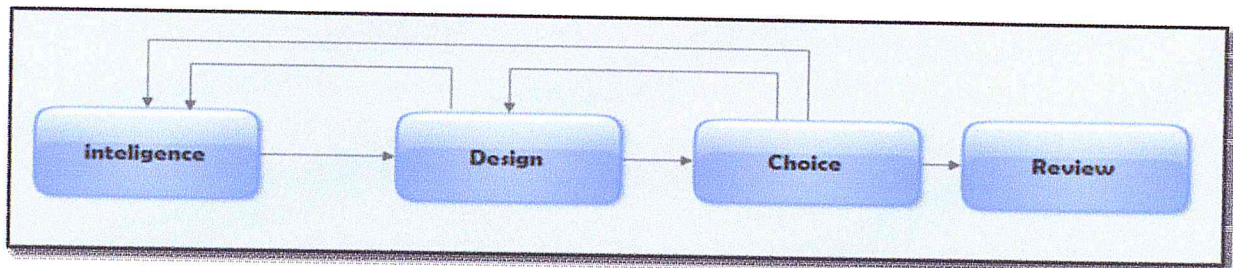


Figure 1.1 : *Processus de décision selon H.A. Simon [Simon 60].*

Ces phases ne se déroulent pas de façon séquentielle, elles font l'objet de retours en arrière, d'ajustements, de corrections, de reprises de résultats antérieurs,... : etc.

• La phase « *Intelligence* » comprend plusieurs tâches :

- ✚ Définir le problème à résoudre.
- ✚ Rechercher les informations pertinentes selon les questions que se pose le décideur. La phase implique ainsi une classification de ces informations. En effet, dans certains problèmes, il est parfois difficile de trouver les informations pertinentes ; or ce sont elles qui sont à l'origine du processus de décision et leur choix est crucial. La recherche d'information est une phase longue et difficile. Les informations pertinentes sont le plus souvent disséminées, peu accessibles, hétérogènes, non structurées Elles influencent fortement les autres phases puisque les autres choix en découlent.
- ✚ Rassembler les informations.
- ✚ Organiser les informations dans une représentation du problème utilisable par la mémoire du décideur.

## CHAPITRE 01 : AIDE A LA DECISION

---

- ✚ Feedback (ou retour en arrière): toute activité peut être interrompue pour reprendre des activités antérieures (principe de recherche) : compréhension de la situation de décision, recherche de nouvelles informations,....

- **La phase «*Design*» ou Conception**

Consiste à construire et étudier des solutions, à partir et en utilisant les informations sélectionnées. Durant cette phase, différents ensembles d'actions possibles sont analysés. Le modèle de la situation de décision est construit et conceptualise le problème et ses abstractions sous une forme quantitative et/ou qualitative. Si nécessaire, des hypothèses simplificatrices seront faites afin de rendre le problème plus facile à comprendre et à appréhender. Il faudra trouver alors un compromis entre une représentation la plus proche possible de la réalité et une représentation simplifiée permettant une résolution du problème. Mais cette simplification ne doit pas déformer la représentativité du modèle. Si un problème est trop simplifié, et ceci à un niveau trop élevé, il sera certes plus facile de trouver des solutions, mais ceci se fera au détriment de la réalité sensée être décrite.

- **La phase «*Choice*» ou Choix**

Consiste à sélectionner une des solutions élaborées, celle-ci devenant le résultat de la décision. Elle se base sur des indicateurs agrégés, qualitatifs ou quantitatifs. Elle Permet d'effectuer la comparaison des différentes alternatives en fonction de leurs conséquences supposées.

Les alternatives sont alors ordonnées suivant une échelle de préférences et mènent à un choix rationnel de la meilleure d'entre elles.

- **La phase d'évaluation «*review*» :**

Une évaluation des alternatives par le décideur, en accord avec ses objectifs, est possible. Cette évaluation permet de recommander la solution la plus appropriée au modèle.

### 5. Outils et méthodes de l'aide à la décision

L'aide à la décision utilise des techniques et des méthodologies issues du domaine des mathématiques appliquées telles que l'optimisation, les statistiques, la théorie de la décision ainsi que des théories de domaines moins formelles telles que l'analyse des organisations et les sciences cognitives.

L'aide à la décision représente l'ensemble des moyens (modèles, méthodes, outils, concepts,...) mis à la disposition du décideur pour faciliter la prise de décision.

## 5.1. Outils de l'aide à la décision

De nombreuses techniques permettent au décideur de prendre certaines décisions. Les outils qu'il devra mettre en œuvre, dépendent du problème initial et de la connaissance plus ou moins précise du décideur.

Les outils de l'aide à la décision se caractérisent à la fois par leur continuité et par leur variété. Aucun outil informatique de type aide à la décision n'a été abandonné. De plus en plus d'outils émergent grâce à l'évolution des TIC (Technologies de l'Information et de la Communication).

La variété d'outils d'aide à la décision pose cependant problème, que ce soit pour :

- mesurer l'impact réel.
- le choix des outils adaptés à une situation déterminée.
- adopter les méthodes de mise en place adéquates.

Selon [Ghomari 08] les outils de l'aide à la décision sont classés selon qu'ils soient destinés à l'usage individuel, à un groupe ou à l'organisation dans son ensemble.

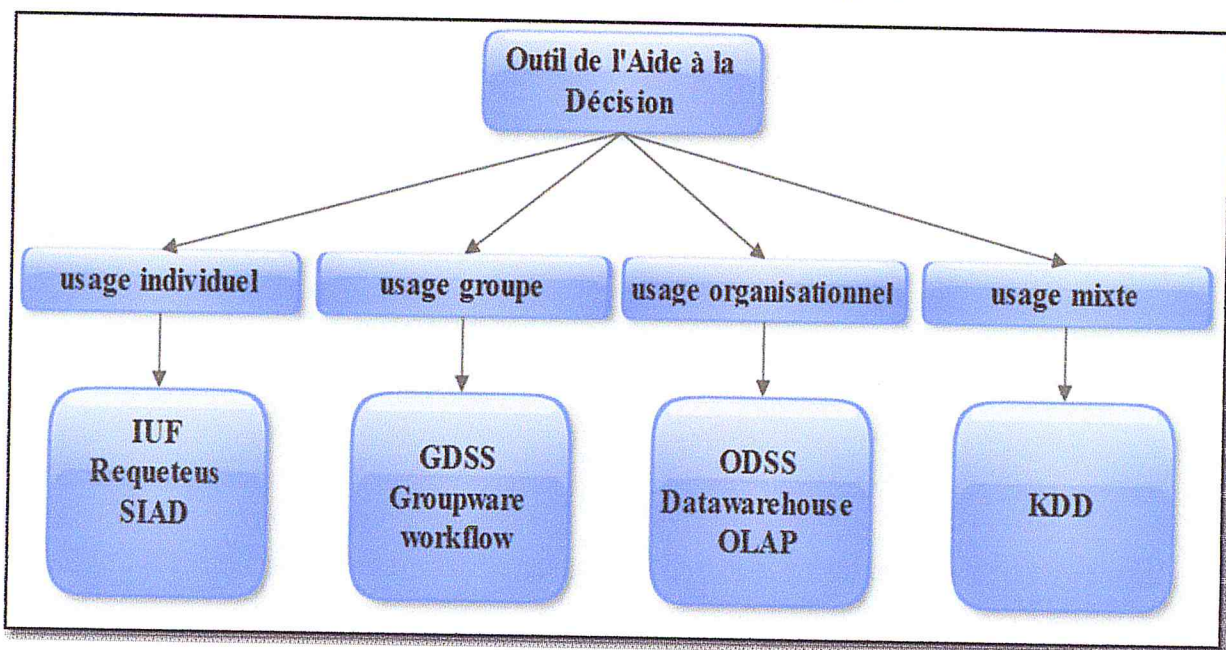


Figure 1.2 : Classification des outils de l'aide à la décision centrée usages [Ghomari 08].

### 5.1.1. Outils de l'aide à la décision à usage individuel

#### 5.1.1.1. IUF (Informatique de l'Utilisateur Final)

Elle regroupe l'ensemble des outils dont a besoin un décideur durant son travail. Ils constituent les supports pour l'aide à sa décision individuelle. Ils permettent :

- la réalisation de documents, de présentations, de graphiques... etc.

## CHAPITRE 01 : AIDE A LA DECISION

---

- la collecte des données nécessaires à sa prise de décision, par l'extraction de données internes ou externes, la constitution de bases de données, ... etc
- l'organisation personnelle de son travail (gestion de projet, modèles de documents, etc.).
- la construction de SIAD personnels grâce à un tableur.

L'apport évident de ce type d'outils est l'aide personnelle à la formalisation des processus cognitifs du décideur ; grâce à ses efforts de présentation synthétique et agrégée le décideur pourra exprimer ses modèles mentaux. Cette connaissance que le décideur, ou tout autre acteur de l'entreprise, crée grâce à ce type d'instruments de travail personnel, constitue des connaissances déclaratives : formalisées et stockées, elles peuvent s'intégrer dans un processus de gestion de connaissances ou tout au moins constituer un élément d'analyse durant la phase d'ingénierie de la connaissance.

### 5.1.1.2. Requêteur

Un requêteur permet à l'utilisateur final d'accéder aux données de l'entreprise de manière autonome, dans un langage proche de celui de son métier. Ces outils peuvent nécessiter la connaissance de la structure de la base que l'on interroge.

Les outils d'interrogation (requêteurs) sont ainsi les moyens par excellence pour les échanges de données. Ces derniers sont non procéduraux : ils n'exigent pas la description de la manière d'extraire les données de la base. Il suffit à l'utilisateur de spécifier ce qu'il recherche ou ce qu'il désire modifier. Ce dernier point implique une autonomie des données, d'où la nécessité de la connaissance du métier. Plus la requête est ciblée, sophistiquée et moins elle est utilisable par d'autres.

### 5.1.1.3. SIAD (Systèmes Interactifs d'Aide à la décision)

A. P. Keen et M. Scott Morton [Keen 78] ont pris en compte la dimension cognitive du décideur en proposant alors la définition suivante: *«Les SIAD réunissent les ressources intellectuelles des individus avec les potentialités des ordinateurs dans le but d'améliorer les décisions prises ».*

Dans l'expression du SIAD (Système Interactif d'Aide à la Décision), en anglais DSS (Décision Support System), on trouve :

- Le mot « *système* » qui se réfère à un système de traitement de l'information ayant une architecture plus ou moins complexe.

## CHAPITRE 01 : AIDE A LA DECISION

---

- Le mot « *interactif* » qui traduit l'existence d'un dialogue coopératif entre l'utilisateur et le système. P. Lévine souligne en effet que ce qui distingue un SIAD d'un autre système informatique, c'est son interactivité (tout ou une partie du contrôle dans la conduite du processus de décision est laissé à l'utilisateur). [Lévine 89].
- Le terme « *aide à la décision* » indique que la priorité est donnée à l'utilisateur. C'est lui qui exerce le contrôle dans le processus de décision et prend les décisions ultimes. Le système ne fait que l'aider dans sa démarche.

Courbon J.C lui définit un SIAD comme « *un système homme-machine qui, à travers un dialogue, permet à un décideur d'amplifier son raisonnement dans l'identification et la résolution de problèmes mal structurés* » [Courbon 99].

Enfin c'est dans le domaine des problèmes « mal structurés » que le SIAD trouve sa raison d'être. H.Simon identifie un problème mal structuré comme « *une situation pour laquelle le système ne possède pas de procédures spécifiques de résolution du problème rencontré, et où l'on doit se rabattre sur une capacité générale à générer une action intelligente, adaptative et centrée sur le problème* ».

### 5.1.2. Outils de l'aide à la décision à usage groupe

#### 5.1.2.1. GDSS (Group Dècision Support Systems)

Ils consistent en un ensemble de technologies qui soutiennent les activités effectuées par des décideurs organisés en un groupe. Ils favorisent un effort concerté et coordonné dans le groupe vers l'achèvement de tâches collectives de prise de décision.

Le rôle principal d'un GDSS est d'augmenter l'efficacité de groupes de décision en incluant la génération d'idées, le partage d'informations, l'analyse de la décision, l'évaluation ou le vote des options alternatives et l'élaboration du rapport de la session de prise de décision et ce de façon interactive entre les membres du groupe et l'ordinateur [Huber 84].

#### 5.1.2.2. Groupware

Les outils de Groupware aident les individus à travailler ensemble et ce en évitant les contraintes de lieu et de temps [Saadoun 96].

## CHAPITRE 01 : AIDE A LA DECISION

---

Le groupware est une technologie qui recouvre des domaines aussi vastes que la coopération, l'interaction homme-machine et l'interaction interpersonnelle via des techniques numériques.

Ceci avec l'objectif d'améliorer et de transformer les entreprises de manière radicale. En effet, l'adoption des outils de groupware est basée sur la communication et le partage du savoir, ce qui nécessite une organisation sociale aplatie, démocratique et coopérative.

Dans cette optique, ces technologies ne peuvent pas être utilisées de manière optimale avec des modèles de management de type autoritaire, hiérarchique et bureaucratique.

### 5.1.2.3. Workflow

La WPMC (WorkFlow Management Coalition), organisme international de normalisation des opérateurs de ce secteur définit le workflow comme *«l'automatisation de tout ou partie d'un processus d'entreprise, au cours duquel l'information circule d'une tâche à l'autre, c'est à dire d'un participant à l'autre, pour une action en fonction d'un ensemble de règles de gestion»* [WPMC 99].

Souvent la technologie Workflow est considérée comme une application du Groupware. La technologie Groupware permet aux individus de travailler ensemble, par la communication, la collaboration et la coordination. On y inclut des activités, aussi variées que les forums électroniques, l'édition conjointe et le Workflow. Ces activités se distinguent par leur structuration (plus ou moins forte) et leur orientation (données ou processus). Dans ce contexte, le Workflow est une technologie qui supporte la coordination d'activités structurées et qui est centrée autour des processus.

### 5.1.3. Outils de l'aide à la décision à usage organisationnel

#### 5.1.3.1. ODSS (Organizational Decision Support System)

King et Star [King 90], le concept d'ODSS renvoie à *«l'application de technologies informatiques et de communications dans le but d'améliorer le processus de prise de décision collective dans l'entreprise»*.

Alquier insiste sur le fait qu'un ODSS est *« un système focalisé sur les besoins de prise de décision stratégique affectant le management de processus transverses de l'entreprise. »*[Alquier 2001].

Les Fonctionnalités/caractéristiques d'un ODSS [Turban 01] :

## CHAPITRE 01 : AIDE A LA DECISION

---

OLAP est une technologie créée en 1993 par E. Codd, pour répondre aux besoins d'analyse et de reportions des applications de gestion. Ces bases de données relationnelles étaient nées du modèle OLTP (*On Line Transactionnal Processing*). Codd a poursuivi dans la même direction en définissant pour les bases multidimensionnelles le modèle OLAP « *On Line Analytical Processing* ».

### 5.1.4. Outils d'aide à la décision à usage mixte

#### 5.1.4.1. KDD

La définition la plus souvent rencontrée au sujet du concept de KDD est celle de [Fayyad 96] « l'acquisition de connaissances nouvelles, intelligibles et potentiellement utiles à partir de faits cachés au sein de grandes quantités de données ». Il s'agit donc d'une discipline qui s'intéresse à l'exploration informatisée de grandes quantités de données avec l'objectif d'en extraire des régularités ou relations, inconnues avant même que ne débute l'analyse de ces données. En effet le KDD est une discipline émergente, construite par l'assemblage de diverses disciplines : « ... statistiques, intelligence artificielle, apprentissage automatique, reconnaissance de formes, base de données, visualisation des données et linguistique » [Crie 01].

Le KDD est un processus semi-automatique et itératif, constitué de plusieurs étapes allant de la sélection et la préparation des données jusqu'à l'interprétation des résultats, en passant par la phase de recherche des connaissances « le data mining ». En effet le datamining se réfère à une étape particulière dans le processus KDD pour l'application des algorithmes spécifiques pour extraction de motifs à partir des données.

### 5.2. Les Méthodes de l'aide à la décision :

Les méthodes de l'aide à la décision sont nombreuses, diversifiées, non reliées, utilisant des concepts et des formalismes différents selon les outils. Seule la modélisation des données reste un point à peu près partagé.

Il n'existe pas de panorama complet des méthodes, incluant définition des besoins, choix d'un type d'outil et réalisation.

D'après [Ghomari 08], La conception de système d'aide à la décision peut être abordée selon différentes approches, chacune de ces approches privilégiant un aspect particulier de l'environnement informatique multi-utilisateurs [Touati 95]:

## CHAPITRE 01 : AIDE A LA DECISION

---

- ✓ *Conception orientée génie logiciel* : Cette approche a pour objet de définir les méthodes d'analyse et de conception permettant d'identifier les différents composants des SIAD.
- ✓ *Conception orientée espace d'information partagé* : cette approche est basée sur les systèmes de bases de données multi-utilisateurs. La coopération est, ici, fondée sur l'accès aux informations communes. Ces informations peuvent être utilisées comme moyen de perception des autres.

### 6. Conclusion

Nous avons vu dans ce chapitre l'aide à la décision, qui est un domaine en soit, fait interagir le décideur qui émerge dans une situation décisionnelle et qui est appelé à prendre une décision en utilisant un outil d'aide à la décision qui ne fournisse pas la décision la plus optimale mais qui l'aide à trancher pour la décision qui convient au mieux au cas qui se pose.

Le décideur est un acteur qui joue un rôle au sein de l'organisation dans un des niveaux hiérarchiques ; stratégique, tactique ou opérationnel. En face d'un problème décisionnel, et un panorama des principaux outils d'aide à la décision. Parmi ces outils nous nous intéressons, dans notre travail de recherche, au processus KDD (Knowledge Discovery from Databases).

Nous allons présenter dans le chapitre suivant en détaille le processus KDD.



# **Chapitre 02**

## **KDD**

### 1. Introduction

Les faibles coûts des machines en termes de stockage et de puissance de calcul et le développement des techniques informatiques ont encouragé les entreprises et les organisations à accumuler de grandes masses de données qui sont souvent sous exploitées alors qu'elles peuvent renfermer des connaissances stratégiques que des experts peuvent ignorer.

Ces réservoirs de données représentent une importante mine d'informations que les entreprises doivent exploiter et explorer pour découvrir des informations pertinentes et utiles à des fins de prédiction et de prise de décisions. Les mécanismes de Data Mining (DM) ou encore de Knowledge Discovery in Databases (KDD) sont alors mis en place pour répondre à ce besoin à travers un nouveau domaine de recherche qui se situe au carrefour de plusieurs disciplines à savoir : les bases de données, l'apprentissage automatique et les statistiques.

Le problème de recherche ou de découverte de relations, de liens et de règles générales à partir d'instances de données est devenu un problème très populaire et fait en général référence aux machines d'apprentissage (learning machine) [Kodratoff 93] et aux techniques de découverte de Dépendances Fonctionnelles [Mannila 94]. La différence essentielle entre ces deux méthodes, est le domaine d'application. L'apprentissage (avec par exemple les arbres de décision et les réseaux de neurones) classe les exemples et offre ainsi un moyen très important pour faire de la prédiction et de la prise de décision, alors que les techniques d'extraction de dépendances fonctionnelles infèrent des dépendances fonctionnelles valides ou non valides, certaines ou incertaines à partir de relations et offre ainsi un moyen efficace d'aide dans la conception du schéma de bases de données ou encore peuvent être utilisées dans le domaine du reverse engineering.

### 2. Découverte de connaissances « Knowledge Discovery(KD) »

La Découverte de Connaissances réfère au processus qui permet de découvrir de la connaissance utile à partir de données pré organisées à la sortie du Data Mining, et de s'assurer qu'elle est utile, c'est à dire que c'est bien une connaissance nouvelle. Elle ajoute à la fouille de données des étapes d'incorporation de connaissances antérieures, d'interprétation des résultats,...

La découverte de connaissances peut s'appuyer sur des techniques anciennes comme l'analyse de données (analyse factorielle des correspondances, analyse en composantes principales), ou des techniques plus récentes comme l'exploration multidimensionnelle ou les EIS (Executive information system).

En effet, les outils ont des fonctions évoluées (calculer des agrégats et des ruptures parmi de dimensions multiples, présentations graphiques,...) qui supportent des présentations plus appropriées pour la découverte ou la mise en valeur de connaissances.

### 3. Extraction des connaissances a partir des données (ECD ou KDD)

[Fayyad 96] a défini le processus ECD (Knowledge Discovery in database KDD) par: "Le processus ECD est un processus non trivial d'identification de modèles valides, nouveaux, potentiellement utiles, compréhensibles à partir d'une base de données". En analysant les termes utilisés dans cette définition, nous obtenons une définition plus pertinente:

- **Modèle** : c'est une représentation d'un phénomène réel, le plus souvent constitués d'objets mathématiques comme par exemple des ensembles de données, des tables, des matrices, des fonctions, des relations, des listes de règles, des systèmes d'équations, des arbres, des graphes, des hypergraphes, des réseaux, des opérateurs fonctionnels linéaires et non-linéaires, etc.
- **Processus** : c'est l'exécution d'un ensemble de plusieurs tâches qui peut être exécuté d'une façon itérative.
- **Processus Non-trivial** : c'est un processus qui exécute des tâches dans un ordre spécifique
- **valide** : un modèle valide est un modèle appliqué à une base de test doit renvoyer un degré de certitude.

Le KDD s'inscrit dans une démarche d'aide à la décision : les données disponibles sont d'abord transformées en une information dont on va extraire une connaissance intelligible, utile, intéressante dans le contexte (apport du datamining), qui favorisera une meilleure décision.

Il y a plusieurs définitions du processus d'ECD. Le nombre d'étapes défini dans un processus d'ECD varie de 4 à 9 suivant la définition donnée.

Dans le tableau 2.1, défini par Eva Andrassyová [Andrassyová 02], nous présentons une liste représentative des processus d'ECD avec leurs différentes phases. Ce tableau est organisé de façon à retrouver sur une même ligne des tâches similaires.

## CHAPITRE 02 : KDD

Simoudis 96	Mannila 97	Fayyad 96	Brachman & Anand 96
	Compréhension du domaine	Apprentissage du Domaine d'application	<b>Découverte des actions à réaliser</b>
<b>Sélection des données</b>		Création d'un ensemble de données cibles	<b>Découverte des données</b>
		Nettoyage et prétraitement des données	<b>Nettoyage des données</b>
<b>Transformation des données</b>	Préparation des Données	Réduction et projection de données	
		Sélection des fonctions de data mining	<b>Développement de modèles</b>
<b>Data Mining (Fouille de données)</b>	Découverte de modèles	Sélection des algorithmes de data mining Data mining	<b>Analyse des données</b>
<b>Interprétation des Résultats</b>	post-traitement des modèles	Interprétation	<b>Génération de rendement</b>
	<b>Utilisation des résultats</b>	<b>Exploitation des connaissances découvertes</b>	

**Tableau 2.1 :** Liste des différentes étapes de processus d'ECD [Andrassyová 02]

Le modèle du processus d'extraction des connaissances que nous avons retenu se décompose en plusieurs phases : la sélection des données, la préparation des données, le data mining et enfin l'interprétation et l'évaluation des résultats (voir Figure 2.1). Le choix de ces quatre phases provient de la synthèse des processus ECD du Tableau 2.1 :

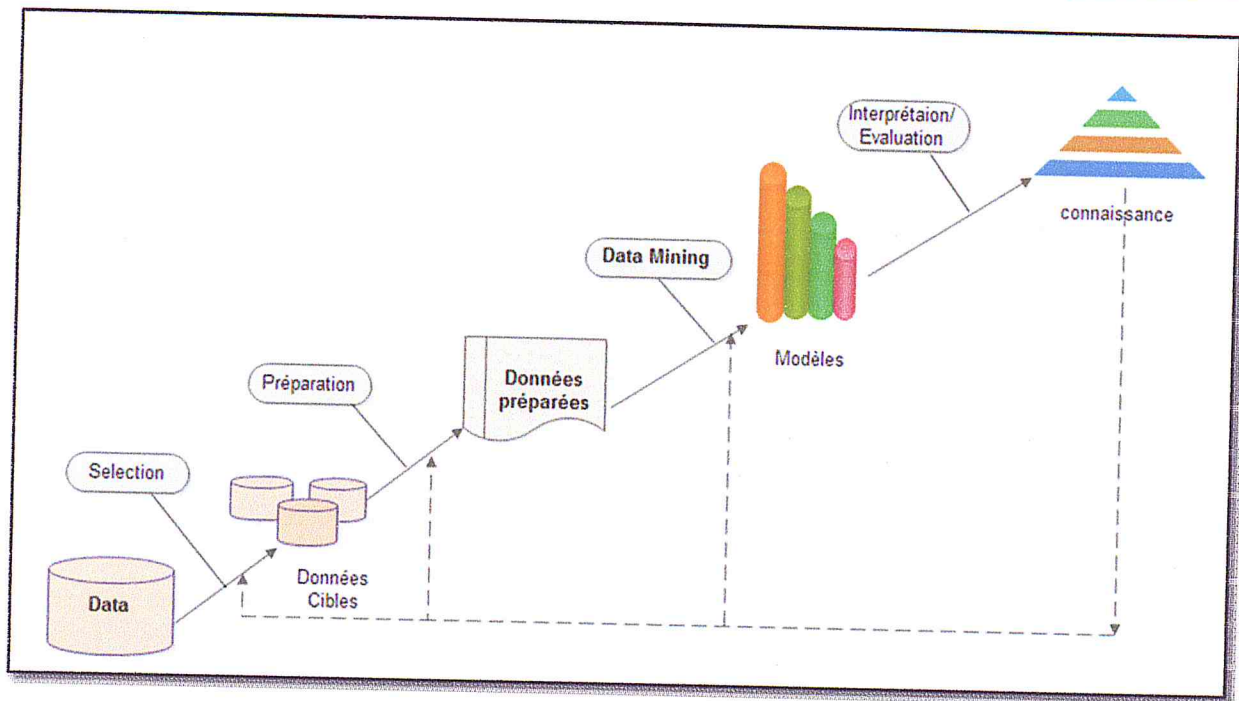


Figure 2.1 : Processus KDD [Fayyad 96].

### 3.1. Sélection des données

Cette phase ne se limite pas à la seule sélection des données qui vont être exploitées par le processus KDD. Elle comprend également l'analyse du problème à résoudre, ce qui permet d'en déduire le ou les types de données qui sont exploitées, ainsi que les méthodes qui pourraient être utilisées pour résoudre ce problème.

Un système KDD idéal est un système qui nécessite l'intervention d'aucune entité, c'est-à-dire un système automatisé qui va extraire de nouvelles connaissances à partir de grandes bases de données mises à sa disposition sans l'intervention de l'utilisateur. Actuellement, ce type de système présente de nombreux inconvénients. Le premier de ceux-ci est la perte de temps et de ressources nécessaires à l'exploitation de l'ensemble des données disponible au système. [ZERF 10]

Parmi les conséquences de ces inconvénients, on peut citer :

- Les recherches lancées par le système peuvent toucher divers domaines ou thèmes qui n'ont aucun rapport avec l'objectif défini par l'utilisateur.
- Le système peut fournir des connaissances qui ne présentent aucun intérêt ou sont incompréhensibles pour l'utilisateur.
- L'utilisateur submergé de nouvelles connaissances, ne peut distinguer des connaissances proposées celles qui lui sont réellement intéressantes.

Ceci implique que l'utilisateur (décideur) doit avoir la possibilité de communiquer avec le système afin d'orienter la recherche selon ses objectifs. Pour faciliter la communication de l'utilisateur (décideur) avec le système KDD, un ensemble de primitives (data mining primitives) a été conçu. Ces primitives incluent :

- ✓ Spécification des données
- ✓ Spécification du type de connaissances à extraire
- ✓ Spécification des connaissances préalables
- ✓ Spécification de la mesure
- ✓ Représentation de la connaissance extraite

Pour pallier ce déficit, nous proposons notre modèle de spécification des besoins du décideur (détaille chapitre 05).

### 3.2. Préparation des données

Les données à analyser par les méthodes de data mining sont parfois incomplètes, inconsistantes, erronées, incompatibles entre elles, inadaptées ou encombrantes [Jiawei 00].

Ces types de données sont courants et se retrouvent régulièrement dans les bases de données et d'entrepôts de données.

Dans cette phase, plusieurs procédures sont nécessaires et chacune d'entre-elles a des tâches bien précises dans le traitement des données.

- La procédure de nettoyage des données permet de compléter les données manquantes et de régulariser les données erronées et inconsistantes.
- La procédure de transformation permet de modeler les données sous une forme exploitable par les méthodes de data mining.
- La procédure de réduction des données permet de réduire la taille des données tout en gardant leur intégrité.

### 3.3. Data Mining

C'est le cœur du processus de KDD. Il s'agit à ce niveau de trouver des connaissances à partir des données. Tout le travail consiste à appliquer des méthodes intelligentes dans le but d'extraire cette connaissance. Il est possible de définir la qualité d'un modèle en fonction de critères comme les performances obtenus, la fiabilité, la compréhensibilité, la rapidité de construction et d'utilisation et enfin l'évolutivité.

Tout le problème du data mining réside dans le choix de la méthode adéquate à un problème donné. Il est possible de combiner plusieurs méthodes pour essayer d'obtenir une solution optimale globale.

Les méthodes de data mining qui sont les plus couramment utilisées dans les systèmes KDD sont les méthodes de type classification, régression, structuration et association

- ✚ Méthodes de classification et de structuration (algorithme des k-moyennes (k-means, algorithme du plus proche voisin, SVM),
- ✚ Méthodes d'explication et de prédiction (arbre de décision, réseaux de neurones, réseaux bayésiens, règles d'associations),
- ✚ Méthodes de visualisation et de description.

### 3.4. L'interprétation

La validation complète d'un système à base de connaissances consisterait à s'assurer de l'adéquation entre la connaissance modélisée dans la base de connaissances et la connaissance de l'expert. Deux types de techniques de validation peuvent être envisagés :

- ✚ La validation par l'examen des résultats obtenus à partir de la base de connaissances.
- ✚ La validation par l'étude de la cohérence de la base de connaissances.

Pour certains domaines d'application (le diagnostic médical, par exemple), le modèle présenté doit être compréhensible. Une première validation doit être effectuée par un expert qui juge la compréhensibilité du modèle. Cette validation peut être, éventuellement, accompagnée par une technique statistique.

## 4. Data Mining (fouille de données)

Le data mining est le noyau du processus KDD il couvre plusieurs domaines, l'analyse de données, les bases de données, l'apprentissage, les statistiques, les systèmes à base de règles. Il dispose d'outils performants afin de structurer et d'extraire des connaissances : la classification, la segmentation, la recherche de règle d'association, ... etc.

### 4.1. Définition

La fouille de données correspond à un besoin d'accès intelligent à l'information contenue dans les bases de données. Le terme de Data Mining signifie littéralement forage de données. Comme dans tout forage, son but est de pouvoir extraire un élément : la connaissance. Ces concepts s'appuient sur le constat qu'il existe au sein de chaque entreprise

des informations cachées dans le gisement de données. Ils permettent, grâce à un certain nombre de techniques spécifiques, de faire apparaître des connaissances.

Nous appellerons Data Mining l'ensemble des techniques qui permettent de transformer les données en connaissances.

### 4.2. Les tâches de Data Mining

Beaucoup de problèmes intellectuels, économiques ou même commerciaux peuvent être exprimés en termes des six tâches suivantes :

- La classification
- L'estimation
- La prédiction
- Le groupement par similitude
- L'analyse des clusters
- La description

Les trois premières tâches sont des exemples du Data Mining supervisé dont le but est d'utiliser les données disponibles pour créer un modèle décrivant une variable particulière prise comme but en termes de ces données. Le groupement par similitude et l'analyse des clusters sont des tâches non-supervisées ou le but est d'établir un certain rapport entre toutes les variable.

La description appartient à ces deux catégories de tâche, elle est vue comme une tâche supervisée et non-supervisée en même temps.

#### 4.2.1. La classification

« *La classification consiste à examiner des caractéristiques d'un élément nouvellement présenté afin de l'affecter à une classe d'un ensemble prédéfini.* » [Berry 97]

Dans le cadre informatique, les éléments sont représentés par un enregistrement et le résultat de la classification viendra alimenter un champ supplémentaire.

La classification permet de créer des classes d'individus (terme à prendre dans son acception statistique).

Celles-ci sont discrètes : homme / femme, oui / non, rouge / vert / bleu, ...

Les techniques les plus appropriées à la classification sont :

- ✚ Les arbres de décision,
- ✚ Le raisonnement basé sur la mémoire,
- ✚ Eventuellement l'analyse des liens.



# **Chapitre 03**

## **L'INGENIERIE DES BESOINS**

## 1. Introduction

Un grand nombre d'études a montré que les échecs dans la mise en œuvre et l'utilisation des systèmes informatiques sont dus à une mauvaise compréhension des besoins auxquels ces systèmes tentent de répondre pour cela ils on basé sur la phase d'analyse.

En informatique cette phase a prouvé son importance .En 1976 déjà, [Bell et Tayer 76] ont fait observer que l'inadéquation des fonctionnalités du système aux besoins des usagers, l'incohérence, l'incomplétude et l'ambiguïté des documents de besoins avaient une influence majeure sur la qualité du système final.

Ces auteurs concluent que « *the requirements for a system do not arise naturally, instead they need to be engineered and have continuing review and revision* ».

L'IB est concernée par l'identification de buts assignés au système envisagé et l'opérationnalisation de ces buts en contraintes et exigences imposées au système et aux agents qui en assureront le fonctionnement [Rolland 03]. Elle consiste aussi à vérifier tout au long du développement que les besoins sont respectés. Sans perdre de vue qu'elle tienne aussi compte de l'évolution des besoins non seulement durant le développement du système mais aussi durant tout son cycle de vie.

Les erreurs d'analyse sont les plus nombreuses, les plus tenaces et les plus dangereuses pour la suite du développement [Heisel 99]. Elles sont aussi les plus coûteuses à corriger : une erreur d'analyse détectée en phase de conception coûte cinq fois plus chère que si elle était découverte pendant la phase d'analyse, ce rapport passant à deux cent fois si la détection a lieu en phase d'exploitation. Ces résultats, issus d'enquêtes récentes sur les réussites et les échecs des projets informatiques, ont été publiés par le [Standish Group 99].

L'enquête européenne menée auprès de 3800 organisations dans 17 pays différents [European Software Institute 96] a conclu dans le même sens : les principaux problèmes sont liés à la spécification des besoins (>50%) et à la documentation des besoins (50%) [Rolland 03].

En conséquence apparait clairement à travers ces faits et chiffres, l'importance de l'ingénierie des besoins.

Pour Rolland, le développement des SI est assuré par deux processus complémentaires [Fredj 07] :

- ✓ *L'ingénierie des besoins* : qui consiste à acquérir et représenter les connaissances dont les utilisateurs ont besoin. Ceci est traduit par l'élaboration d'un schéma conceptuel du SI.
- ✓ *L'ingénierie des systèmes* : est le processus par lequel les besoins du SI sont

transformés en une solution logicielle fiable. Elle consiste à implémenter le schéma conceptuel du SI sur un système technologique en utilisant les technologies de l'information.

C'est le schéma conceptuel qui articule la transformation des besoins en une solution logicielle

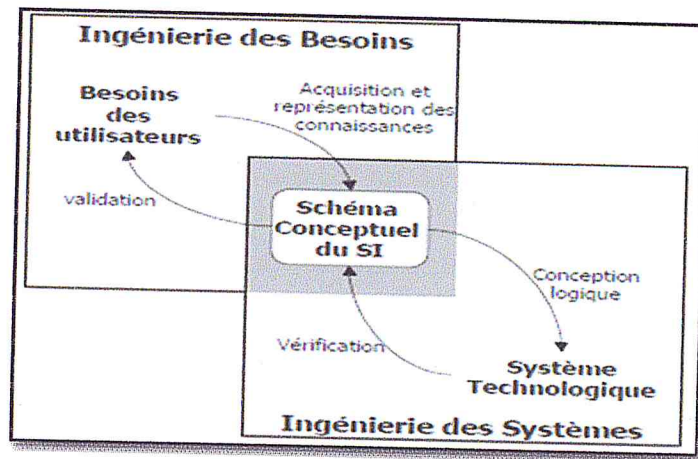


Figure 3.1 : l'ingénierie des SI et l'ingénierie des besoins [Rolland 03].

## 2. Définition besoin ou exigence ?

Pour traduire le terme anglais « Requirement », nous pouvons recenser plusieurs traductions différentes dans la littérature de « Requirement Engineering » :

- ✚ « Requirement » est traduit par le terme « besoin » comme dans [Boukrara&Chalal 13], [Ginting 00]
- ✚ « Requirement » est traduit par le terme « exigence » comme dans [Gam 08], [Gnaho&Semmak 10].
- ✚ « Requirement » est traduit par les deux termes sans faire de différence entre eux comme pour [Delotte 06].
- ✚ Dans [Konaté 09], les deux termes sont utilisés différemment. Où le terme « besoin » est utilisé de point de vue utilisateur, pour exprimer ce qu'il veut ou ce qu'il attend du système. Et le terme « exigence » est utilisé de point de vue concepteur ou développeur du système, pour exprimer ce que le système doit assurer pour répondre aux besoins de l'utilisateur. Ainsi, Une exigence est un besoin qui est techniquement satisfaisable ou dont la solution peut être implémentée. Les exigences formalisent l'expression des besoins.
- ✚ Dans la méthode KAOS, une exigence est un but placé sous la responsabilité d'un agent du système, tandis qu'un but placé sous la responsabilité d'un agent de

l'environnement du système est appelé une attente (expectation) [Matoussi 11]. Cependant, le passage des besoins aux exigences est une phase très critique ce qui fait de l'IB un processus assez complexe à exécuter.

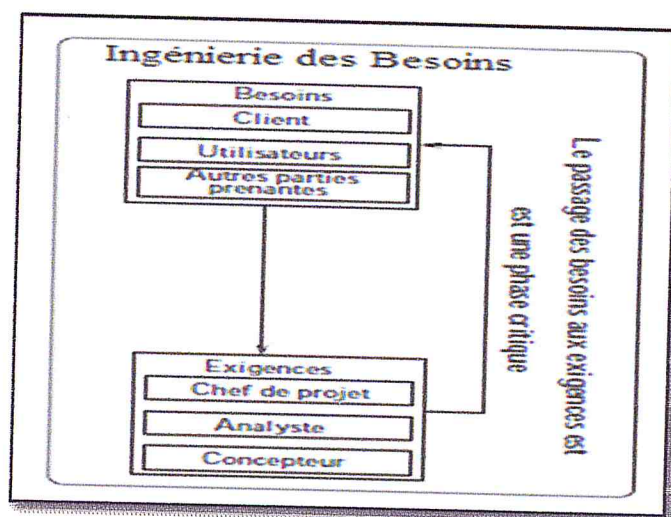


Figure 3.2 : Besoins et Exigences [Essame 02].

### 3. L'Ingénierie des Besoins (IB)

Dans l'industrie, un client, auquel on veut fournir un produit ou un service, peut avoir dans certains cas des exigences explicites (cahier des charges, contrat, ...) qui sont exprimées clairement et qui facilitent, par conséquent, la satisfaction de ses besoins en lui fournissant le produit ou le service en question tel qu'il l'a désiré. Mais ce n'est pas toujours le cas.

Le client exprime ses besoins à sa façon et naturellement, dans son propre langage. Et ainsi ces besoins sont souvent implicites qu'il faudra simplifier. Plusieurs situations problématiques peuvent être envisagées, par exemple:

- ✚ Le client exprime ce qu'il veut, mais pas ce qu'il a vraiment besoin.
- ✚ Les besoins sont mal exprimés ou incomplets.
- ✚ Le client et les ingénieurs des besoins (ou concepteurs des systèmes) conservent leurs propres langages qui ne se comprennent pas forcément. Et par conséquent, Les choses sont exprimées différemment...etc.

Le risque principal de cette incompréhension est de fournir des produits ou des services qui ne correspondent pas aux exigences réelles du client [Gam 08]. D'où l'intérêt de l'ingénierie des besoins (IB) qui veille pour garantir la bonne correspondance entre les produits et services fournis par les organisations et les besoins réels des clients.

### 3.1. Définition

Il existe plusieurs définitions de l'ingénierie des besoins, pour cela nous adoptons les deux définitions :

✚ Selon Colette Rolland : « L'Ingénierie des Besoins (IB) est concernée par l'identification de buts assignés au système envisagé et l'opérationnalisation de ces buts en contraintes et exigences imposées au système et aux agents qui en assureront le fonctionnement. L'IB peut être vue comme le processus qui permet de transformer une idée floue en spécification précise des besoins servant de support à la spécification du système et de ses interfaces avec l'environnement. Le processus d'IB inclue donc des activités de découverte, de spécification, de négociation et de validation des besoins. Au cours de l'IB, les besoins à l'égard du système sont découverts, négociés, validés, spécifiés et répertoriés dans des documents de spécification des besoins. Produire un document des besoins de qualité est difficile et crucial. De nombreuses enquêtes confirment en effet l'importance de l'IB dans la réussite des projets de systèmes d'information » [Rolland 03].

✚ L'IEEE (Institute of Electronic and Electrical Engineers)[IEEE 90], dans sa définition standard considère que « l'ingénierie du besoin est le processus d'analyse des besoins des utilisateurs permettant d'étudier, d'affiner, et enfin de parvenir à une définition des besoins du système matériel ou logiciel ». Elle doit déterminer *pourquoi* le système informatique est utile, les *fonctionnalités* dont il doit accomplir et les *spécifications* à base desquelles il sera construit.

Ces définitions mettent l'accent sur des éléments très importants à savoir:

- ❖ Les objectifs du monde réel ou de l'utilisateur. qui sont à l'origine du développement du système informatique. ils représentent le pourquoi du système.
- ❖ Les spécifications qui traduisent les besoins des utilisateurs, la définition de ce que les développeurs doivent construire et serviront pour la vérification du système produit.
- ❖ L'évolution des besoins du monde réel au cours du temps et la répercussion de leur évolution sur le système informatique.

De là nous retenons que l'IB est une discipline qui est venue pour combler sur le plan méthodologique, plus particulièrement la phase analyse des besoins des utilisateurs comme phase initiale du cycle de vie d'un projet.

C'est le processus d'analyse des besoins des utilisateurs permettant d'étudier, d'affiner, et enfin de parvenir à une définition des spécifications du système informatique et veiller à leur alignement conformément à l'évolution des besoins du monde réel.

L'IB se centre sur le **POURQUOI** pour découvrir les besoins correspondants à la mission et aux objectifs de l'organisation. Si l'on veut éviter de développer des systèmes techniquement parfaits mais inutilisés parce qu'inadaptés aux besoins réels de leurs utilisateurs, il faut se donner les moyens de comprendre à quoi le système va servir dans son contexte organisationnel. C'est la mission de l'IB [Rolland 03].

### 3.2. Le concept But

C'est un concept qui est abordé d'une façon abondante dans la littérature portant sur l'Ingénierie des besoins. « *Un but correspond à un objectif que le système doit réaliser à travers la coopération des acteurs qui interviennent dans le futur système informatique* » [Lamsweerde 00].

Anton atteste que « les buts sont des objectifs de haut niveau économique, organisationnel ou système. Ils capturent le pourquoi du système et guide les décisions à différents niveaux dans l'entreprise » [Anton96], alors que pour Jackson un but exprime des intentions [Jackson 95].

Généralement les buts sont formulés au cours des premières phases du processus de l'IB [Darimont98][Dubois98] vu le rôle qu'ils jouent dans ce processus.

#### L'importance d'étude les buts :

- ❖ Les nouvelles méthodes de l'ingénierie des besoins produisent le document des spécifications techniques du futur système. Ainsi donc les buts permettent de prescrire clairement ce que doit réaliser le système informatique avec les spécifications. Ils définissent aussi le raisonnement portant sur leur opérationnalisation et Le graphe de raffinement des buts fournit les liens de *traçabilité* à partir des buts stratégiques de haut niveau aux besoins opérationnels de bas niveau [Lapouchnian05].
- ❖ Ils fournissent un critère précis pour la complétude suffisante pour la spécification des besoins. La spécification est complète si tous les buts peuvent

être prouvés à être réalisés à partir de la spécification et les propriétés connues du domaine, et Ils fournissent un critère précis des *besoins pertinents*. Un besoin est pertinent si sa spécification est utilisée au moins dans une preuve de but [YUE 87].

- ❖ La modélisation des buts offre un mécanisme naturel pour la *structuration des documents* complexes [Lamsweerde 2001] .
- ❖ Les buts permettent la détection et la gestion des *conflits* entre les besoins [Lamsweerde 96].

Plusieurs classifications des buts ont été proposées dans la littérature

### ✚ Fonctionnels et non fonctionnels

Les buts fonctionnels expriment les fonctionnalités et les services du système. Les buts non-fonctionnels sont liés aux qualités attendues du comportement du système dans son environnement en termes de performances, de sécurité, de couts ou d'adaptabilité.

### ✚ But soft et hard

Pour un but soft, la satisfaction ne peut pas être établie dans un sens clair [Mylopoulos 92] et ils sont utiles pour comparer les alternatives de raffinement des buts et choisir le meilleur qui contribue. Pour un but hard, la satisfaction peut être établie par des techniques de vérification [Dardene 93].

### ✚ Classification selon le comportement temporel

Les buts réalisés génèrent les comportements du système dans la mesure où ils nécessitent une certaine propriété de la cible pour être satisfaits dans certains états futurs.

A propos de la modélisation des buts, [Lamsweerde 01] atteste que les approches de modélisation semi formelles comme les diagrammes entité-relation, les diagrammes d'état, les flux de données, etc. ne sont pas très appropriées pour l'ingénierie des besoins, elles couvrent juste les aspects fonctionnels. Des approches de modélisation plus formelles basées sur des fondements théoriques comme la logique temporelle à temps réel sont proposées pour la modélisation des buts.

### 3.3. Le concept scénario

Un scénario est une séquence d'interactions entre le système envisagé et son

environnement énoncée dans le contexte restreint d'un propos particulier. Il décrit une 'histoire', c'est un écrit concret de la façon dont un acteur imagine d'interagir avec le système.

Les scénarios ont différents contenus. Ils peuvent décrire : Des activités, actions ou événements actuels ou futurs, des objets mis en jeu par le système, le cadre de travail actuel ou imaginé des acteurs, des contraintes de qualité attendues du système, des informations sur l'organisation tels que sa structure, ses départements, ses groupes, ses agents etc., les intervenants dans le processus de l'IB, leurs caractéristiques, leurs vues et leurs aspirations [Nardi 92].

Les scénarios sont exprimés dans différentes notations :

- ❖ Les scénarios informels sont décrits à l'aide du langage naturel [Rolland 97].
- ❖ Les scénarios semi-formels utilisent des notations structurées comme des tableaux [Potts 94] ou des scripts [Rubin 92].
- ❖ Les scénarios formels sont décrits à l'aide de langages basés sur des grammaires régulières [Harel 87] ou des diagrammes d'états [Cockburn 95]. Ils peuvent être utilisés pour simuler le fonctionnement futur du système et juger des réactions des utilisateurs.

Notons pour terminer que de nombreux auteurs suggèrent de combiner buts et scénarios dans le processus d'ingénierie des besoins [Rolland 03].

#### 4. Le processus de l'IB

Jusqu'à ce jour il n'y a pas de définition précise sur le processus de l'IB. En effet plusieurs auteurs ont travaillé dessus et distinguent presque tous quatre grandes activités principales à savoir l'explicitation, la négociation, la spécification et la validation.

Par exemple [Rolland 03] a elle aussi définit le processus de l'IB. Le processus couvre de multiples activités qui sont inter-reliées. Il s'agit, selon la Figure 3. 3 de :

- ✚ L'extraction ou découverte des besoins.
- ✚ La spécification des besoins dans une forme précise et non ambiguë.
- ✚ La validation des besoins.
- ✚ La négociation conduisant au choix des besoins qui seront implémentés.



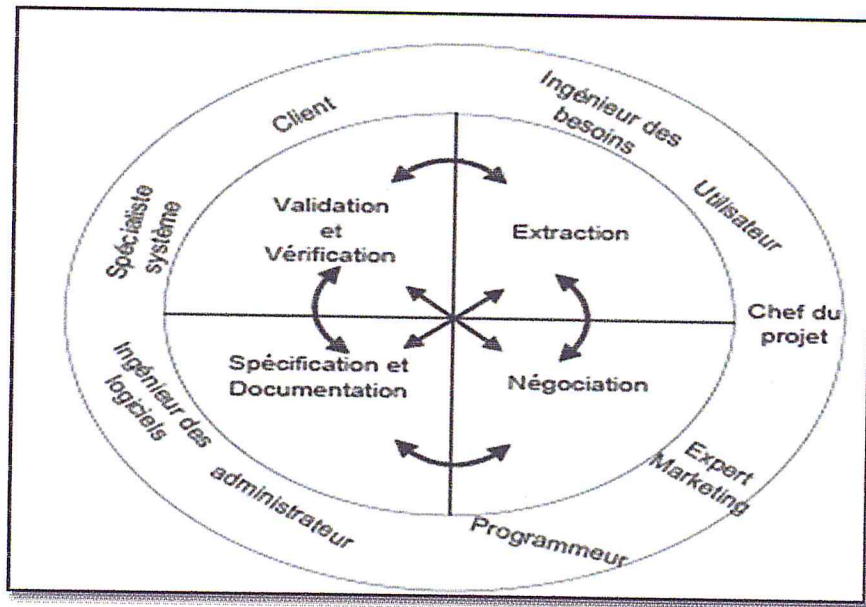


Figure 3.3 : Le processus de l'IB selon C.Rolland[Rolland 03].

A ces quatre activités principales d'un cycle d'IB s'ajoute l'activité d'évolution des besoins permettant de corriger des erreurs et/ou de prendre en compte de nouveaux objectifs.

## 5. Les approches de l'IB

### 5.1. L'IB pour GL

Nous présentons dans ce qui suit deux exemples de méthodes l'IB (la méthode KAOS et la méthode CREWS). Ces deux méthodes sont d'après [Gam et al 2006] très répandues dans le monde de l'IB.

#### 5.1.1. La Méthode KAOS

KAOS (Knowledge Acquisition in autOmedated Specification) ou « acquisition automatique des connaissances sur les spécifications » est une méthode pour l'IB orientée buts, résulte des travaux du professeur Axel Van Lamsweerde en 1991 dans l'université de Louvain [Gnahou&semmak 10].

Un but KAOS définit un objectif que le système et son environnement doivent réaliser grâce à la coopération des de différents agents (matériel, logiciel ou humain).

#### Objectifs :

- Identifier et raffiner les buts progressivement jusqu'à obtenir des exigences qui sont assignables à des agents.
- Identifier des objets et des actions progressivement à partir des buts.

- Assigner les exigences aux agents.

**Principes :**

Cette approche est composée de :

1. Un langage de spécification qui utilise la logique temporelle et il est basé sur les concepts tels que l'objet, l'action, l'agent, le but, la contrainte etc.
2. Une méthode d'élaboration pour la transformation des buts des intervenants en exigences pour le système logiciel en se basant sur les questions classiques comment et pourquoi.
3. Une base de connaissance méta-niveau utilisée pour orientée les décisions pendant le processus d'élaboration. Elle contient
  - Une classification des buts.
  - Les règles pour assurer la cohérence et l'exhaustivité des besoins.
  - Les tactiques et heuristiques pour conduire l'élaboration et la sélection des différents buts.

Elle repose sur quatre modèles à savoir le modèle des buts, le modèle des responsabilités, le modèle des opérations et le modèle objet.

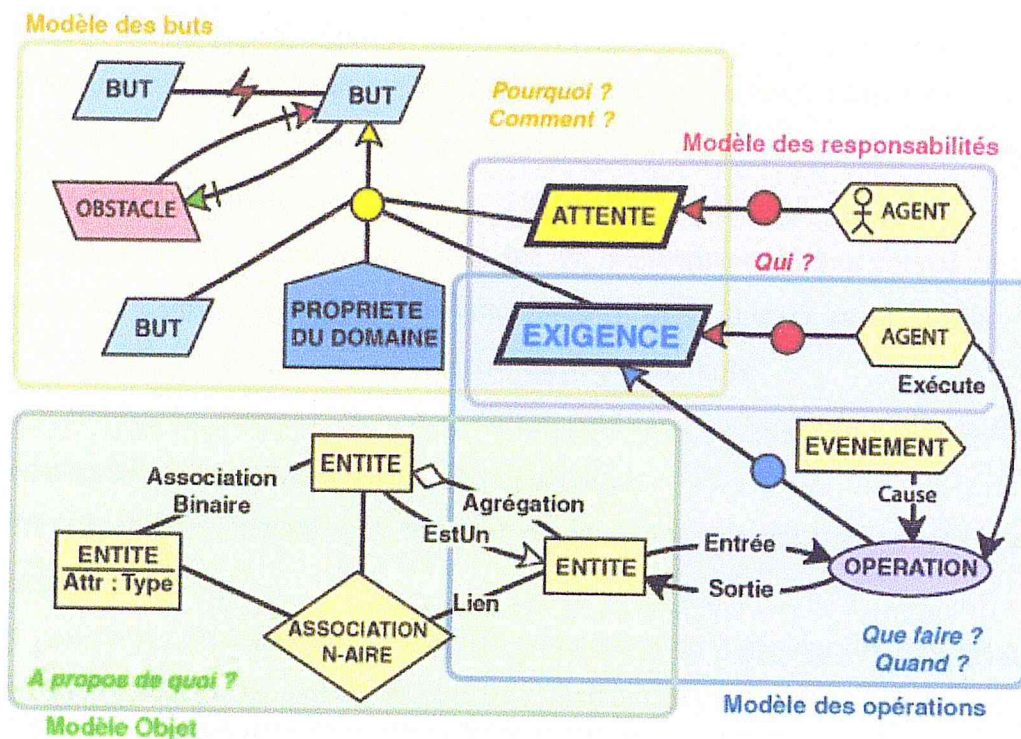


Figure 3.4 : le méta-modèle KAOS [Matoussi 2011].

Par nature, un but est 'intentionnel' tandis qu'un scénario est 'opérationnel'. Le couplage but scénario permet donc de concrétiser le but au moyen d'un scénario.

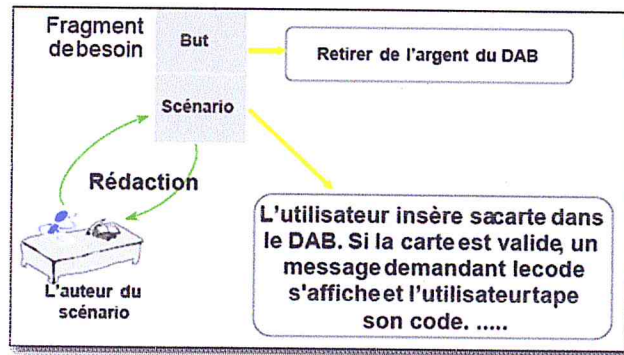


Figure 3.5 : La notion de fragment de besoin [Rolland 03]

## 5.2. L'IB pour l'aide à la décision

Pour un décideur, le processus de prise de décision commence par la prise en compte d'un problème décisionnel, que l'on peut traduire comme son besoin décisionnel. La résolution d'un besoin décisionnel comporte l'identification de ses besoins informationnels [Afolabi & Thiery 05].

Les besoins informationnels des décideurs sont fournis par des systèmes d'information d'aide à la décision. La conception et le développement de ces systèmes doivent faire face aux problèmes de multiplicité et la complexité des informations et des connaissances sur l'organisation et son environnement [Gam 08].

Le problème de l'aide à la décision se pose plus persistant au niveau stratégique. Les décideurs stratégiques représentent les intérêts généraux de l'organisation, et leurs besoins stratégiques sont de nouveaux besoins à évaluer pour réussir un projet décisionnel.

### 5.2.1. Les méthodes de l'IB pour l'aide à la décision

Pour explorer comment les besoins décisionnels sont modélisés et analysés, nous présentons dans cette section trois approches proposées dans le domaine de l'IB pour l'aide à la décision.

#### 5.2.1.1. La Méthode ROMC

Développée par Sprague et Carlson [Sprague 82], la méthode ROMC (Représentations, Opérations, Mémoire-aides, Contrôle-aides) est une démarche de prototypage pour la conception des SIAD individuels. Son objectif est la conception des interfaces utilisateur fournissant les capacités et caractéristiques nécessaires d'un SIAD

individuel. ROMC s'articule autour de quatre entités.

Les deux premières pour supporter la phase de raisonnement ou d'analyse, et les deux dernières pour l'utilisation de SIAD (Figure 3.6):

- ✚ **Représentation** : l'activité de prise de décision se déroule dans un contexte de conceptualisation de l'information utilisée dans l'activité. La représentation des informations ou de concepts peut être sous forme de (graphes, icônes, cartes, fenêtres graphiques, organigrammes, histogrammes, courbes,...).La conceptualisation permet de transmettre les informations à l'utilisateur de l'interface de SIAD sous forme de différentes représentations. Et c'est à l'utilisateur (décideur, analyste,...) d'évaluer et d'appropriier les présentations choisies.
- ✚ **Opérations** : Les opérations sont les tâches spécifiques que l'utilisateur peut effectuer sur un SIAD. Elles permettent de manipuler les informations fournies par l'interface utilisateur. Exemples d'opérations: collecter des données, identifier les objectifs, générer des rapports, ajouter des alternatives, ...etc. Une opération peut être utilisée dans plusieurs activités. Les paramètres des opérations sont donnés par des représentations. Par exemples : un point sélectionné sur une carte peut être lié à une valeur d'un document ou une requête, une liste d'employés prioritaires peut faire l'entrée d'une planification de personnel, ...etc.
- ✚ **Mémoires** : Les aides mémoire sont extrêmement utiles dans le processus de décision et doivent être prévus dans les SIAD. Ils peuvent être sous forme de bases de données, vues, déclencheurs automatiques, bloc-notes, ....etc.
- ✚ **Contrôles** (les aides contrôles): ce sont les mécanismes utilisés pour contrôler le fonctionnement du système (menus, fenêtres, messages d'erreurs, commentaires d'aide, procédures, ...). Les décideurs s'attendent à exercer un contrôle direct sur le système, ils doivent être capables d'analyser les entrées et d'interpréter les sorties.

Phases de développement d'un logiciel				
Etapes du SIAD	Compréhension du problème	Conception des solutions possibles	Choix, réalisation de la solution	Installation et tests
<b>R</b>	Objectifs, schémas, figures, documents	Schémas conceptuels de données, dictionnaire des données Choix des outils	Modèles logique et technique des données	Saisie des données Test Manuel
<b>O</b>	Traitements de données Description de calculs	États et flux de données, formalisation des calculs Choix des outils	Réalisation - de la base de modèles - des modèles	Manipulation sur les données et les opérations Manuel (suite)
<b>Aide-Mémoire</b>	Aides, Contraintes d'intégrité et de sécurité	Formalisation	Réalisation	Test Fin des manuels
<b>Contrôle</b>	Type de dialogue	Menus, écrans	Réalisation interface	Test Manuel à compléter

Figure 3. 6: Déroulement de la méthode ROMC [Sprague 82].

### 5.2.1.2. L'Approche Annoni (2007)

L'approche d'Annoni [Annoni2007] est une approche pour l'analyse et la conception des systèmes d'information décisionnels SID (entrepôts de données). Elle répond sur la question : « comment définir une méthode de développement rapide de SID fiables au regard des besoins de tous ses acteurs ? ».

Dans cette approche, l'auteur distingue deux aspects : l'aspect statique (les données) et l'aspect dynamique (les besoins décisionnels) du SID. Elle spécifie les besoins à base de modélisation multidimensionnelle des données, et à partir de cette formalisation, un processus automatique guide le choix de l'architecture du SID.

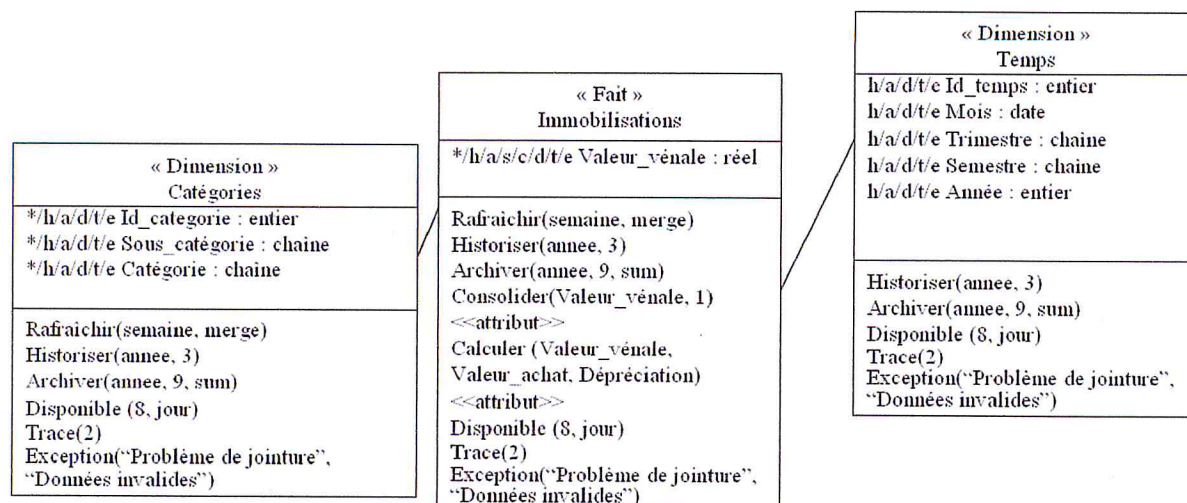


Figure 3.7 : Diagramme décisionnel d'Annoni[Annoni 07]

Le diagramme décisionnel ne contient que deux types de classes afin de faciliter la compréhension des décideurs. La définition de ces deux types de classes vise à représenter les deux concepts reconnus par l'ensemble de chercheurs et des industriels.

Les mesures des faits et les paramètres des dimensions sont représentés par les attributs des classes multidimensionnelles. Les traitements associés aux faits et aux dimensions sont représentés par des opérations. La modélisation orientée objet nous permet d'une part, de traiter séparément les données et les traitements bien qu'étant dans le même modèle et d'autre part, de favoriser la réutilisation.

Pour modéliser les données appelées « besoins système », l'approche Annoni utilise le modèle entité-association.

L'analyse des besoins dans l'approche Annoni est guidée par l'application d'un ensemble de règles (règles de transformation, règles de sélection, règles de structuration, règles syntaxiques et règles de fusion). Ces règles sont appliquées sur les modèles des besoins pour une confrontation des besoins des acteurs afin de modéliser un système d'information décisionnel fiable par rapport aux besoins utilisateurs et aux sources.

L'objectif de la phase d'analyse est d'évaluer la dynamique du système d'information décisionnel (SID) relative à la composante technique et à la composante décision.

## 6. Positionnement de notre recherche dans IB

D'après [Yu 97], « la nature complexe des problèmes du domaine nécessitent des outils qui aident au mieux les analystes. Un volume important de connaissances est dégagé durant cette phase, portant sur les objectifs, l'environnement du système, les alternatives de mise en œuvre et l'implication des acteurs. Il est donc important de

préservé et maintenir ces connaissances sur le développement du système et prendre en charge les changements durant tout le cycle de vie du système ».

Il a été démontré que les premières phases du cycle de vie de tout système sont cruciales pour assurer le succès de son développement, déploiement et évolution. Raison pour laquelle beaucoup de recherche sur l'IB ont pris comme point de départ les besoins des utilisateurs portant sur le futur système. Ces besoins initiaux sont généralement ambigus, incomplets, inconscients et dans la plus part du temps informels.

Notre travail tente de clarifier, et faciliter la compréhension des besoins du décideur dans processus d'aide à la décision KDD.

### 7. Conclusion

Ce parcours dans L'IB a mis en évidence : la richesse de cette discipline en termes de concepts et de techniques, son importance dans la conduite de projet informatique prouvée par les enquêtes évoquées dans les sections précédentes, et l'intérêt qui lui est accordé par la communauté des chercheurs surtout durant les deux dernières décennies. Les besoins initiaux des utilisateurs sont généralement ambigus, incomplets, inconscients et dans la plus part du temps informels.

Pour cela L'IB doit poser la question *POURQUOI* développer un système et aider les parties prenantes du projet à y répondre, le rôle de l'IB est ensuite de déterminer les fonctionnalités que le système doit mettre en œuvre pour aider à la satisfaction de ces buts et identifier les contraintes qui restreignent la mise en œuvre de ces fonctions. Ces buts, fonctions et contraintes, constituent les 'besoins' qui doivent ultérieurement être convertis en une spécification précise permettant le développement du système, et l'évolution des besoins au cours du temps et la répercussion de leur évolution sur la spécification du système doit aussi être prise en compte.

Dans le chapitre suivant, nous allons voir les différentes méthodes de l'IC qui se base sur le décideur.

# **Chapitre 04**

## **L'INGENIERIE DE CONNAISSANCE**



## 1. Introduction

Les processus décisionnels créent et utilisent des connaissances. Ces connaissances vont nécessiter un management des connaissances, donc au préalable un travail d'ingénierie des connaissances (IC).

L'IC est un domaine qui s'intéresse à la conception de systèmes qui visent à traiter ou à aider des opérateurs à traiter des problèmes mal posés, faisant appel à des connaissances jusqu'ici non explicités ou non modélisées [Charlet 02].

Historiquement, la formalisation des connaissances a d'abord été étudiée en Intelligence Artificielle (IA), dans le cadre du développement des Systèmes à Base de Connaissances (SBC) et des Systèmes Experts (SE). A ce moment là, on distingue, la méthode MACAO [Aussenac 89], la méthode KOD [Vogel 88], la méthode KADS [Breuker 95].

Suite aux travaux de [Newel 82] sur la modélisation conceptuelle du monde (knowledge level), l'ingénierie de connaissance s'impose comme une discipline et propose de modéliser l'expertise recueillie une fois le problème défini et avant la conception du SBC.

Pour l'histoire, la communauté des BD a introduit dès 1975 la notion de modélisation du monde réel. par la suite, c'est l'idée même de ce qu'est un « système intelligent » qui a évolué : initialement localisée dans le seul SE, « l'intelligence » se déplace vers le couple « système-utilisateur » (ou humain centered design).

L'IC a pris en compte le groupe d'acteurs concernés par la mise en place d'un S.I., et pas seulement l'expert humain qui serait la référence en matière de « bonne résolution de problème ». C'est ainsi que l'IC s'intéressa de plus en plus à la mise à disposition de connaissances auprès des utilisateurs. a cet effet, plusieurs méthodes dites d'IC sont apparues: La méthode REX [Eichenbaum 94], la démarche CORPUS [Grunstein 95], la méthode MUSIC [Alquier 93], la méthode MKSM [Ermine 96] ainsi que quelques expériences sur des cas ponctuels tel que le modèle DOLMEN [Simon 97] ou le modèle de Maret [Maret 95], MEREX [Corbel 97], CYGMA [Bourne 97], et bien d'autres encore.

Pour la suite, nous présenterons, les méthodes qui s'intéressent au décideur, puis nous allons détailler la méthode KADS.

## 2. Acquisition de connaissances

L'acquisition des connaissances est une des difficultés majeures de la conception des systèmes experts". Combien de centaines d'articles commencent par une telle introduction.

L'acquisition de connaissances se définit comme « le transfert et la transformation d'une expertise d'une source de connaissances à un programme ».

L'acquisition de connaissances n'est plus donc perçue comme un processus de transformation de connaissances sous forme d'un programme compréhensible par la machine, mais comme un processus de construction d'un modèle qui tient compte de la subjectivité et du caractère relatif et implicite des connaissances [Studer 98].

L'acquisition de connaissances à un but pratique, elle fait partie de l'ingénierie de la connaissance.

### 3. Ingénierie de connaissances

L'IC (Ingénierie de Connaissances) est défini comme « *l'étude des concepts, méthodes et techniques permettant de modéliser et/ou d'acquérir les connaissances pour des systèmes réalisant ou aidant des humains à réaliser des tâches se formalisant a priori ou pas* » [Charlet 00].

On ne parle plus alors de l'extraction des connaissances des têtes des experts et de leur formalisation à l'aide de règles de production, mais de la modélisation des connaissances expertes selon différents points de vue [Charlet 00].

### 4. Les Méthodes d'ingénierie de connaissances

Les méthodes d'ingénierie de connaissances ont été regroupées en trois types d'approches : approche ascendante, approche descendante et approche mixte.

✚ Approche ascendante (Bottom up) dirigée par les données (data driven) : se base sur une étape d'élicitation des connaissances des experts (entretiens, analyse de documents, analyse des tâches, etc.) suivie d'une étape de conceptualisation. La méthode KOD (Knowledge Oriented Design) [Vogel 1989], MACAO [Aussenac 1989] et la méthode MIKE (Model-based and Incremental Knowledge Engineering) [Angele 1996] sont des exemples typiques de cette famille.

✚ Approche descendante dirigée par les modèles : Une méthode descendante consiste à identifier en premier lieu le type de problème auquel nous sommes confrontés et de réduire à son minimum une analyse non guidée. Cette démarche est basée sur l'adaptation de modèles génériques existants. Une fois le modèle sélectionné, il sert de guide à l'acquisition. Nous présentons dans le détail la méthodologie CommonKADS (Knowledge Acquisition & Documentation Structuring) [Wielienga 92].

✚ Approche mixte : adoptant une mixture entre les deux approches précédentes.

## CHAPITRE 04 : L'INGENIERIE DE CONNAISSANCE

---

Selon [Ghomari 08] Les méthodes d'IC sont classées, en se basant sur la dichotomie :

- ✚ Méthodes orientées acquisition de connaissance des individus (KOD, MIKE, MACAO).
- ✚ Méthodes orientées acquisition de connaissances de l'organisation (MKSM, MUSIC).

Nous adoptons pour notre part, une troisième décomposition, qui regroupe les méthodes d'IC selon qu'elles soient intéressées au décideur (KOD, KADS, MACAO) ou à la situation décisionnelle (MKSM, MUSIC) [Chalal 07].

Pour notre projet de recherche, nous basons sur les méthodes qui s'intéressent au décideur, et plus précisément la méthode KADS car elle propose un cadre de modélisation au niveau des connaissances. De plus KADS est devenue un standard dans le développement de SBC. Dans la méthode KADS, la modélisation se fait grâce à la construction d'une succession de modèles. Ces modèles correspondent à plusieurs perspectives de la résolution du problème dans son environnement organisationnel et son contexte d'application.

Ainsi, il nous sera possible de situer notre objectif dans le contexte adéquat pour l'application, pour une modélisation optimale.

Dans les méthodes de modélisation des SBC, il s'agit d'extraire les savoirs d'experts, savoirs structurés, stables et constitués, extraits à partir de leurs actes de discours, la conception étant faite de manière descendante.

Nous présentons la méthode KADS qui s'intéresse au décideur qui nous permet de mieux positionner nos travaux.

### 4.1. La méthode KADS

La méthode KADS (Knowledge Acquisition and Design Support) [Breuker 94], [Breuker 95], [Anjewierden 87], [Dieng 90], est issue du projet ESPRIT 1, KADS n°1098 du programme de recherches européen ESPRIT (1985-1990).

L'objectif de KADS (Knowledge Acquisition and Design Support) (projet n°1098) est d'assister le transfert de la technologie des SBC vers le domaine commercial en proposant un guide méthodologique pour le processus de développement de ces systèmes.

Dans la suite des idées développées dans KADS, il y eut quatre projets ESPRIT 2 (KADS II n° 5248, VITAL n°, ACKnowledge n°2576 et VALID n°2148). Des méthodes et des techniques ont été développées pour couvrir tous les aspects de l'analyse et de la conception d'un SBC. Ce processus est décomposé dans KADS en la construction successive de plusieurs modèles.

### 4.1.1. Cycle de vie de KADS

KADS propose un cycle de vie basé sur les méthodes de développement de logiciels. Une analyse complète des données précède la conception et l'implantation du SE, ce qui diffère donc du prototypage rapide ou des approches incrémentales.

Les résultats de la phase d'analyse (contraintes externes, Modèle Conceptuel des objets du domaine et des méthodes de résolution) servent d'entrée à la phase de conception où ils sont transformés en spécification de l'architecture fonctionnelle du SBC.

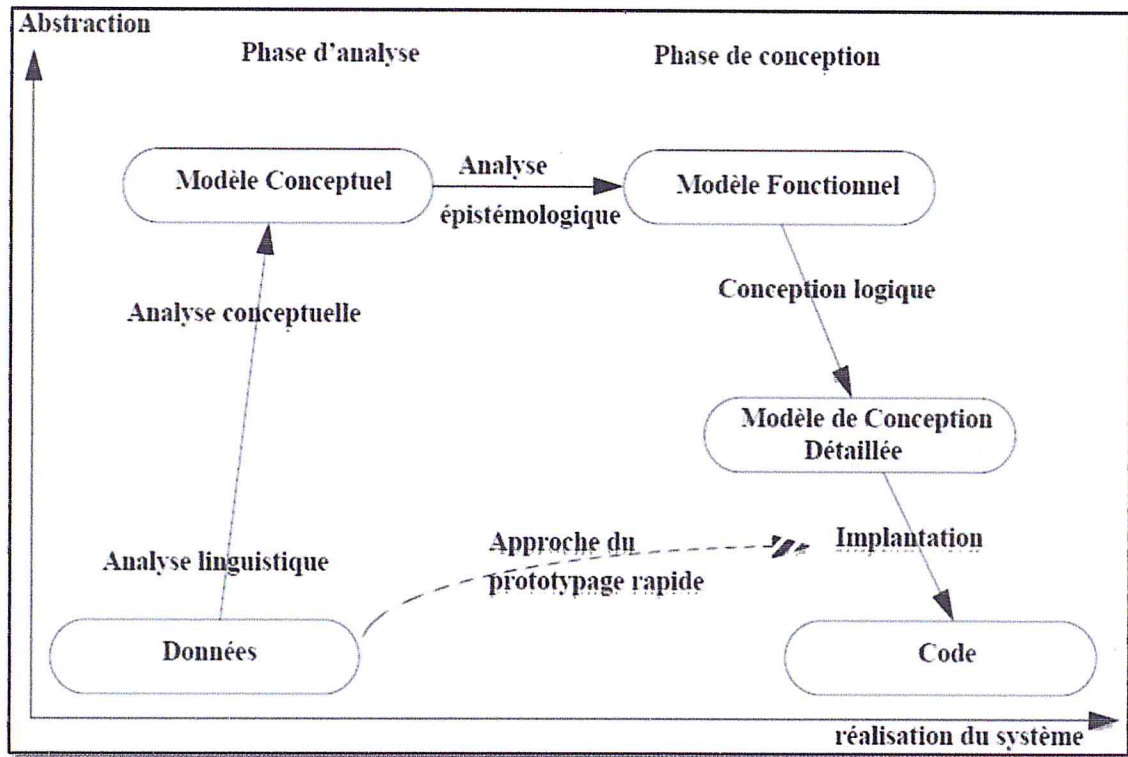


Figure 4.1 : Modèles et espaces de développement de KADS [Dieng 90].

Ce cycle de vie est constitué d'un ensemble d'étapes modèles consécutives. Il a été enrichi au fil des différentes versions de la méthode, et peut pour un projet donné, être adapté pour détailler ou accélérer l'analyse.

KADS repose sur l'hypothèse que les phases d'analyse et de conception doivent être séparées : il n'y a pas de retour sur la phase de conceptualisation une fois celle-ci achevée, mais le raffinement de modèles exige l'alternance de recueils et d'analyses de ceux-ci.

Le cogniticien peut donc extraire les connaissances utiles de l'expert indépendamment des conditions d'implantation et vision par vision il existe plusieurs sous-modèles au Modèle Conceptuel L'information recueillie dans le Modèle Conceptuel est ainsi moins biaisée est structurée et d'un haut d'abstraction.

## CHAPITRE 04 : L'INGENIERIE DE CONNAISSANCE

Les différents modèles proposés dans KADS s'enchaînent chronologiquement et sont en interaction entre eux comme le décrit la figure ci-dessous.

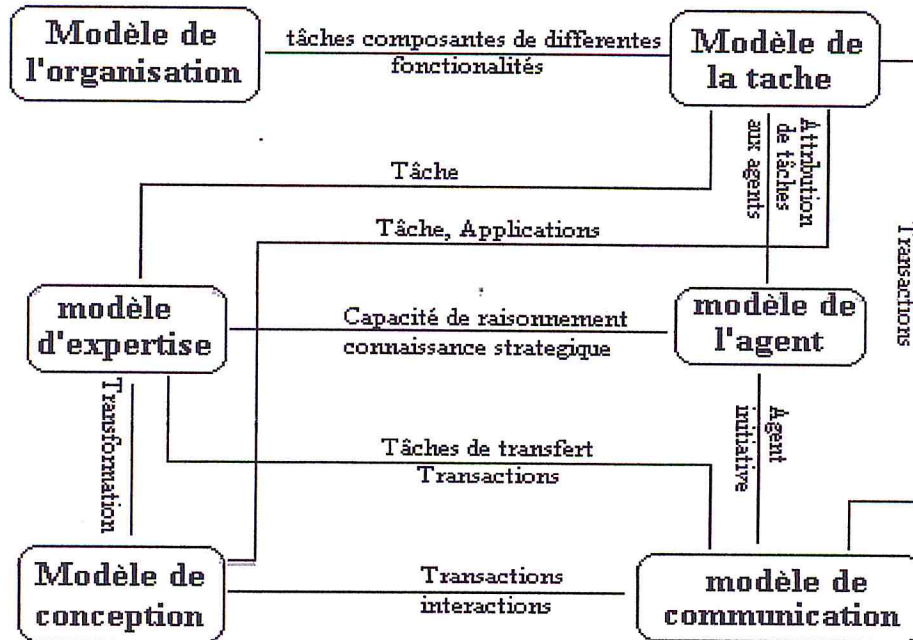
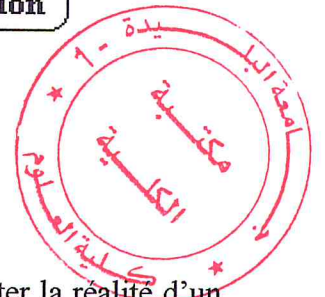


Figure 4.2 : Interactions des modèles KADS.



### 4.1.2. Les modèles de KADS

Les différents modèles proposés dans KADS permettent de représenter la réalité d'un problème de capitalisation de connaissances en facilitant la compréhension du problème même.

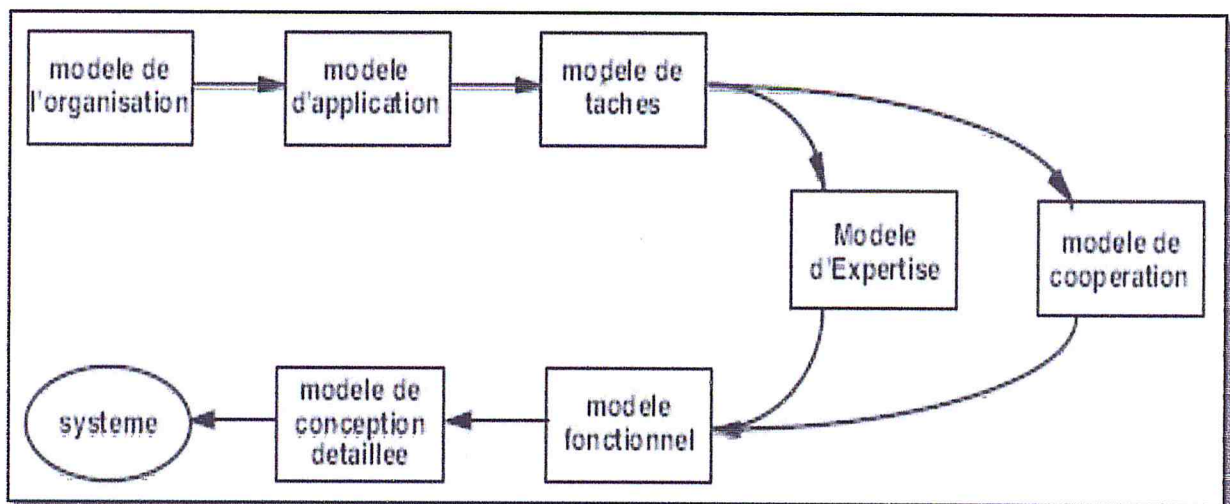


Figure 4.3 : Les différents modèles de KADS.

### 4.1.2.1. Le modèle de l'organisation

Décrit l'environnement du SBC ainsi que les interactions (au niveau humain ou social) avec les utilisateurs. Doivent donc être spécifiés: l'objectif du manager, le rôle du ou des experts, les types d'utilisateurs, leurs besoins et utilisations du SBC.

Ceci est important pour permettre de bien cibler l'utilité et la fonction de l'outil, d'assurer son adéquation aux besoins des utilisateurs, de s'assurer de sa faisabilité

### 4.1.2.2. Le modèle d'application

Définit quel est le problème à résoudre, l'utilité d'un SBC, sa fonction et ses contraintes techniques (langage, système d'exploitation, performance ...).

### 4.1.2.3. Le modèle de tâches

Spécifie comment la fonction est décomposée en tâches et sous tâches, indique leurs entrées-sorties et les agents internes ou externes qui les exécutera.

La création d'un modèle de tâches passe par plusieurs étapes: la première étape est l'identification de la tâche réelle globale (équivalent à définir l'objet du système à base de connaissances à développer).

Cette tâche principale est ensuite décomposée en sous tâches hiérarchisées pour lesquelles il faut:

- Identifier les interdépendances (transferts d'objets).
- Déterminer les spécifications de leur communication.
- Attribuer les sous tâches aux agents en essayant de limiter les dépendances dans l'exécution des tâches et de tenir compte des compétences des agents.

[Greef 92] donne certains conseils pour la décomposition de la tâche globale en sous-tâches et pour la distribution : limitations des dépendances, exploitation des compétences, etc. Il préconise pour tester ces deux actions, l'utiliser de la technique du Magicien d'Oz (où l'expert joue le rôle du futur SBC en communiquant avec l'utilisateur via un terminal, mais toutefois en langage naturel). Il semble intéressant d'utiliser les relevés de ces expériences pour spécifier la communication associée à chaque tâche de transfert et les stratégies explicatives à sélectionner.

Notons cependant que la distribution des sous-tâches entre le système et l'utilisateur peut parfois ne pas être totale (i.e. maintien de tâches partagées mais où les rôles de chacun sont cependant clairs), et ce pour au moins trois raisons :

## CHAPITRE 04 : L'INGENIERIE DE CONNAISSANCE

✓ Il ne peut être déterminé de manière fixe qui va faire quoi. L'allocation doit donc être dynamique, après **négociation** entre l'utilisateur et le système. Ce dernier doit alors avoir une liste d'alternatives pour choisir.

✓ La sous-tâche peut devoir être effectuée en parallèle par le système et l'utilisateur.

✓ Dans la sous-tâche, l'un des deux agents donne à l'autre des instructions et ce dernier les exécute

Le modèle de tâche ainsi formé sert d'entrée à l'analyse de l'expertise d'une part, et raffiné pour inclure les aspects contrôle et communication d'autre part... etc.

Enfin, il faut assurer la synchronisation entre ces tâches et le Modèle d'Expertise.

### 4.1.2.4. Le modèle de coopération

Le modèle de coopération n'est dans le KADS standard qu'une simple annexe du modèle de tâches, définissant sommairement les tâches de transfert : pour chacune, il précise qui, de l'utilisateur ou du système, a l'initiative de l'échange, et quel est le type d'«ingrédients» transférés : «information», «connaissance», «talent». [Greef 92].

KADS définit quatre types de transferts comme le décrit la figure ci-dessous:

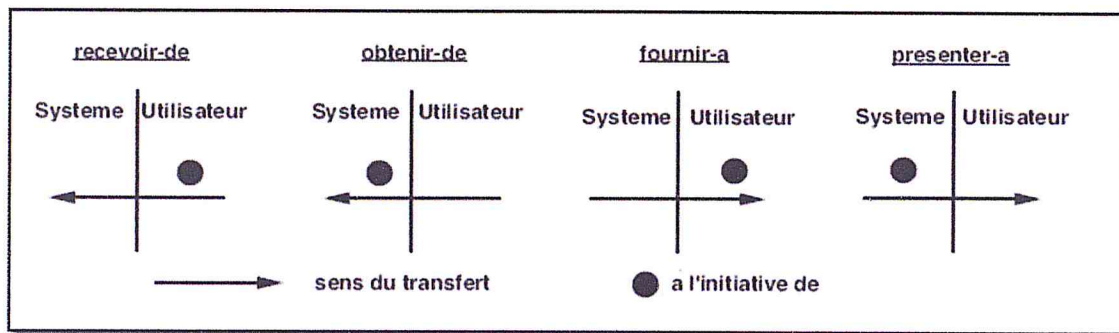


Figure 4.4 : les interactions modèles de coopération

- ✚ **Recevoir de**: le système qui n'a pas l'initiative reçoit de l'utilisateur un flot sans qu'il puisse le contrôler,
- ✚ **Obtenir de**: le système qui a l'initiative demande et obtient de l'utilisateur un flot sur lequel il a le contrôle,
- ✚ **Fournir à**: le système qui n'a pas l'initiative répond à une demande de l'utilisateur en fournissant un flot sur lequel l'utilisateur a le contrôle,
- ✚ **Présenter à**: le système qui a l'initiative fournit à l'utilisateur un flot sur lequel il garde le contrôle.

## CHAPITRE 04 : L'INGENIERIE DE CONNAISSANCE

Un cinquième type de tâche de transfert, «**négociier**», peut être utile pour transférer des informations sur (le type de) la coopération ou sur la résolution de problème. Exemples : sélection d'une distribution de tâches, correction d'une information donnée préalablement.

Trois types majeurs d'ingrédients peuvent être distingués : «l'information», la «connaissance» et le «talent» :

➤ **L'information** réfère à des états spécifique du monde réel ou conceptuel (il y a donc beaucoup d'analogie avec les méta-classes des structures d'inférences). Typiquement :

- des données, comme une valeur de variables ou une description de structure au sens KADS.
- des états de problèmes.
- des états internes ou mentaux : intention, évaluation, histoire ou état actuel de la résolution, processus de communication, on retrouve particulièrement des méta-classes de même nom en entrées ou sorties des structures d'inférences.

➤ **La connaissance** n'est pas spécifique d'une situation mais générique. Elle peut être utilisée en tant qu'ingrédient pour des buts d'explication ou d'enseignement.

➤ **Le talent** est transféré quand l'objectif est d'instruire l'autre agent sur la façon d'exécuter une sous-tâche (idem pour la connaissance). Ceci arrive par exemple lorsque les compétences de l'utilisateur sont moins grandes que prévues.

### 4.1.2.5. Le Modèle d'Expertise

Utilise le modèle de tâches et décrit le comportement et le type de connaissances nécessaires au système pour la résolution du problème, indépendamment de toutes contraintes d'implantation.

Il s'agit donc d'un modèle de la connaissance de l'expert, subjectif car il s'agit d'une vision particulière du monde réel, mais ce n'est pas un modèle cognitif de l'expert car il est orienté par ce que peut et doit faire le futur SBC.

Ce modèle joue donc un peu le rôle de spécifications fonctionnelles dans les développements classiques: c'est «un cadre sémantique partagé par des utilisateurs d'un programme et ses développeurs qui leur permettent de communiquer» [Aussenac 89] ou encore «le modèle qu'a en tête le concepteur quand il va implanter le SBC» [Brunet 91].

Préalablement à la conception et l'implémentation. Dans ce modèle, les connaissances sont structurées en quatre niveaux :

✚ **Le niveau domaine:** il décrit la connaissance propre au domaine, indépendamment



de toute utilisation. Il rassemble les concepts du domaine (savoir générique), les relations entre les concepts et les relations entre les propriétés du concept. Ce niveau peut être vu comme une ontologie du domaine.

✚ **Le niveau inférence:** ce niveau mentionne les inférences (au sens dérivation de connaissances à partir d'autres connaissances) à effectuer sur le domaine.

✚ **Le niveau tâche:** ce niveau explicite comment contrôler les inférences élémentaires pour atteindre un but (un état espéré).

Les tâches sont décomposées en trois types de sous tâches qui sont: les primitives (une inférence élémentaire), les tâches composées (autorisent la récursivité) et les tâches de transfert.

✚ **Le niveau stratégie:** détermine les buts à poursuivre pour résoudre un problème et les stratégies de réaction en cas d'échec.

### 4.1.2.6. Le Modèle Conceptuel

Est la réunion des deux derniers modèles et décrit donc de façon abstraite les objets, comportements et opérations nécessaires au futur SBC. Par extension, il désigne toutes les activités menant à ce modèle (analyse linguistique puis conceptuelle).

Certains générateurs de S.E. interprète partiellement le Modèle Conceptuel pour certains environnements ou type de tâches : Model-K [Karbach 91], ZDEST-2 [Tong 88].

- Une phase de conception plus "classique" et plus courte que celle de l'analyse (ex: 1 à 2 mois pour un prototype ayant demandé 6 mois d'analyse [Brunet 91]) comprenant :

- ✚ **Le modèle fonctionnel** qui correspond à l'architecture fonctionnelle du futur SBC et s'obtient par une analyse épistémologique du Modèle Conceptuel, là peuvent intervenir les termes du formalisme de représentation des connaissances du langage d'implantation.

- ✚ **Le modèle de conception détaillée**, après une analyse logique effectuée sur le modèle fonctionnel.

## 4.2. La méthode KOD (Knowledge Oriented Design)

C'est est une méthode d'aide à l'extraction et à la modélisation des connaissances qui a été développée par [Vogel 89]. Elle consiste à élaborer trois modèles :

➤ **Le modèle pratique** : permet d'identifier les éléments de l'expertise. Ce modèle permet d'identifier et d'extraire plusieurs types de manifestations verbales recueillies en entretien avec l'expert : (1) les *taxèmes* : objets du monde physique, (2) les *actèmes*:

## CHAPITRE 04 : L'INGENIERIE DE CONNAISSANCE

des descriptions de changement d'état causé par un sujet sur un destinataire à l'aide de ressources et (3) les *schémèmes* : des schémas d'interprétation (déduction, induction) qui se manifestent par des contraintes et des inférences.

- **Le modèle cognitif** : permet de structurer et valider les connaissances acquises. L'expert utilise : (1) des *taxinomies* : représentations hiérarchiques des entités du monde physique, (2) les *actinomies* : structuration des schémas mentaux de l'expert, permettant de répondre à des situations déjà connues, et (3) les *schémas d'interprétation* : des aides à l'interprétation du réel. Ce sont des ensembles de règles d'inférence associés correspondant aux stéréotypes de raisonnement de l'expert, qui se déclenchent selon les observations du domaine.
- **Le modèle informatique** : permet la mise en forme des connaissances dans un programme informatique. Il est constitué d'objets, de méthodes et de règles.

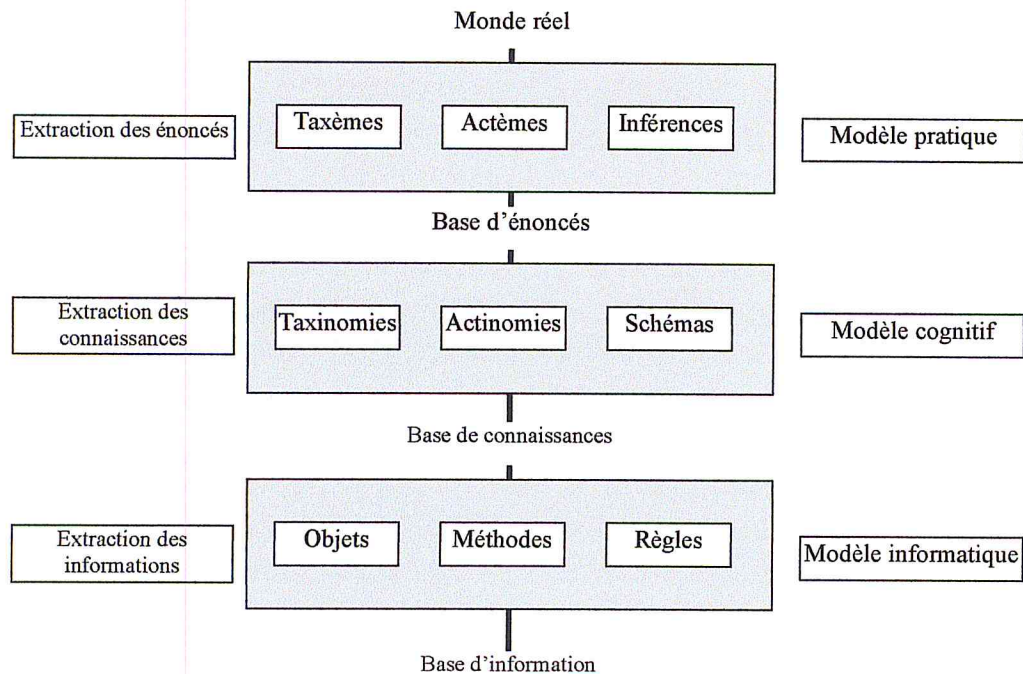


Figure 4.5 : Le Modèle de KOD [Dieng 90].

### 4.3. La méthode MACAO

S'adressant plutôt au cognicien, mais permettant aussi à l'expert d'exprimer directement ses connaissances. Selon [Aussenac 92], MACAO est à la fois une méthode complète de développement d'un SBC, et un outil supportant cette méthode.

Elle fait référence à un modèle cognitif de l'expertise pour définir un formalisme de représentation médiatrice des connaissances. Ce formalisme constitue un intermédiaire entre

## CHAPITRE 04 : L'INGENIERIE DE CONNAISSANCE

---

l'expression verbale de l'expert et la forme codée des connaissances, sa représentation externe doit être accessible à l'expert.

Le modèle cognitif a permis également de sélectionner un ensemble de techniques de recueil de connaissances. La méthode suppose une coopération active entre l'expert et le cogniticien. Elle intègre des techniques ergonomiques pour l'analyse de l'activité de l'expert et du futur utilisateur, et accorde une grande importance à l'activité comme moyen d'explicitation de l'expertise.

La méthode se décompose en quatre étapes [Aussenac 05]:

- A. Identification de l'expertise :** cette étape est basée sur une succession de dialogues expert-cogniticien et sur des observations de l'expert au travail. On peut ainsi spécifier l'environnement de l'utilisateur, ainsi que les principales caractéristiques de l'expertise.
- B. Recueil des protocoles de résolution de problèmes :** L'expert est invité à résoudre divers problèmes des catégories identifiées dans la première étape, il explique ce qu'il fait en raisonnement en fournissant toute information supplémentaire demandée par le cogniticien. Ces dialogues sont enregistrés. L'expert note dans le logiciel tous les objets qu'il manipule de façon à constituer le glossaire du problème. Puis au cours d'entretien centré, il définit les concepts associés à chaque objet figurant dans le glossaire. Il note ces concepts dans le logiciel tout en justifiant oralement leur utilisation.
- C. Conceptualisation, analyse et formalisation de la connaissance :** le cogniticien, guidé par le logiciel MACAO, analyse les informations précédentes en ce basant sur le modèle cognitif. Il formalise les concepts et relations en terme d'unités fonctionnelles et de déterminants (entités de base du modèle cognitif). Pour ce faire, il confronte les problèmes d'une même catégorie et ceux de catégories différentes. Il identifie des relations statique ou dynamique entre unités fonctionnelles. MACAO prévient le cogniticien en cas de formalisation inconsistante et peut suggérer de regrouper certaines unités fonctionnelles en procédures. Grâce à cette analyse, on obtient plusieurs graphes de connaissance, correspondant à chaque problème (graphes instanciés) ou décrivant chaque catégorie de problème (graphes génériques).
- D. Validation interactive des connaissances par l'expert :** La validation a lieu directement par l'expert qui, lors d'entretien centré, peut corriger le graphe générale de connaissances et explorer la hiérarchie du graphe. MACAO lui offre un éditeur interactif pour afficher ou mettre à jour la connaissance avec vérification automatique de la consistance.

### 5. Conclusion

Ce chapitre a été consacré aux méthodes d'IC, et focalisé surtaux à l'étude de la méthode KADS (Knowledge Acquisition and Design Support).

Nous avons vu que La méthode KADS est une méthode orientée modèles. Elle offre la possibilité de choisir un ou plusieurs modèles, étant donné que ses modèles sont indépendants.

Les utilisateurs de KADS se restreignent aux plusieurs étapes successives allant du général au particulier : modèle d'organisation, modèle des tâches, modèle application, modèle de coopération, modèle d'expertise et modèle conceptuelle.

Pour notre cas, nous voulons utiliser cette méthode dans le sens de définir les besoins d'un décideur de façon bien précise, en respectant les étapes et le principe de la méthode KADS.

Dans le chapitre suivant, nous allons voir comment utiliser la méthode KADS pour Spécifier les besoins du décideur avant l'utilisation du processus KDD.

# **Partie II**

**Le modèle de découverte des  
connaissances utiles dans une  
base de données**

**Et**

**Implémentation**

# **Chapitre 05**

## **LE MODELE DE DECOUVERTE DES CONNAISSANCES UTILES DANS UNE BASE DE DONNES**

## 1. Introduction

Le développement du volume de données dans systèmes d'information est devenue une réalité à la quelle chaque entreprise doit faire face, notamment, elle doit permettre à ses responsables « managers » de découvrir les informations pertinentes à fin de prendre les meilleurs décisions dans les plus brefs délais.

Avec la mondialisation des échanges et à la concurrence les dirigeants des entreprises ont besoin d'avoir une vision claire de leur environnement cette vision est basée sur l'utilisation des outils faciles, qui protègent toujours le système opérationnel.

A fin de pouvoir prendre les décisions les plus efficaces chaque entreprise désire d'exploiter intelligemment ses données et avoir plus d'informations que ses concurrents.

Parmi ces outils, nous nous intéressons du processus KDD (Knowledge Discovery from Databases) qui est considéré comme l'un des outils d'agrégation les plus intéressants.

Ce dernier, s'est inscrit dans une démarche d'aide à la décision : les données disponibles sont d'abord transformées en une information dont on va extraire une connaissance intelligible, utile, intéressante qui favorisera une meilleure décision. L'une des caractéristiques de ce processus qui est itératif (plusieurs étapes) et interactif, l'utilisateur ou l'analyste est impliqué à chaque étape dans ce processus. Tandis que l'analyste ne peut pas découvrir ce que le décideur à besoin de savoir quel que soit le soin avec lequel sera étudié un système.

Le décideur essaiera de rassembler le maximum des connaissances à travers le processus KDD. Pour prendre une bonne décision .Néanmoins, les décideurs n'utilisent que des connaissances nécessaires et censées d'êtres utiliser. C'est pour cela nous essaierons d'incorporer la méthode KADS, pour extraire uniquement les connaissances utiles pour le décideur.

## 2. Positionnement de la méthode KADS dans le processus KDD

Le processus KDD est une suite d'étapes allant de la sélection de données jusqu'a l'interprétation des résultats. Les connaissances découvertes doivent être intéressantes, parce qu'elles sont destinées au décideur pour qu'il puisse prendre la bonne décision.

La première phase du processus KDD est l'acquisition des données. Cette phase vise à cibler, de façon grossière, l'espace des données qui va être exploré.

Dans ce cas, le problème étudié par l'analyste doit être bien délimité, pour qu'il choisisse que les données nécessaires. Alors il ne se lance pas dans le processus KDD sans

## CHAPITRE 05 : LE MODELE DE DECOUVERTE DES CONNAISSANCES UTILES DANS UNE BASE DE DONNES

avoir une certaine idée des objectifs du décideur, le processus KDD est efficace lorsque le problème est bien posé.

Pour comprendre et préciser les objectifs du décideur, nous proposons d'ajouter une étape au processus KDD. Cette dernière précède la phase de sélection. Elle vise à spécifier les besoins du décideur.

### 3. Le modèle proposé

Pour répondre à notre besoin on propose un modèle nommé (processus spécification des besoins du décideur) Figure 5.1, qui incorpore la méthode KADS avec le processus KDD dont le but de consolider ce processus. Ce modèle est composé de deux processus :

- ✓ Processus de spécification des besoins du décideur avec l'utilisation de la méthode KADS.
- ✓ Processus KDD.

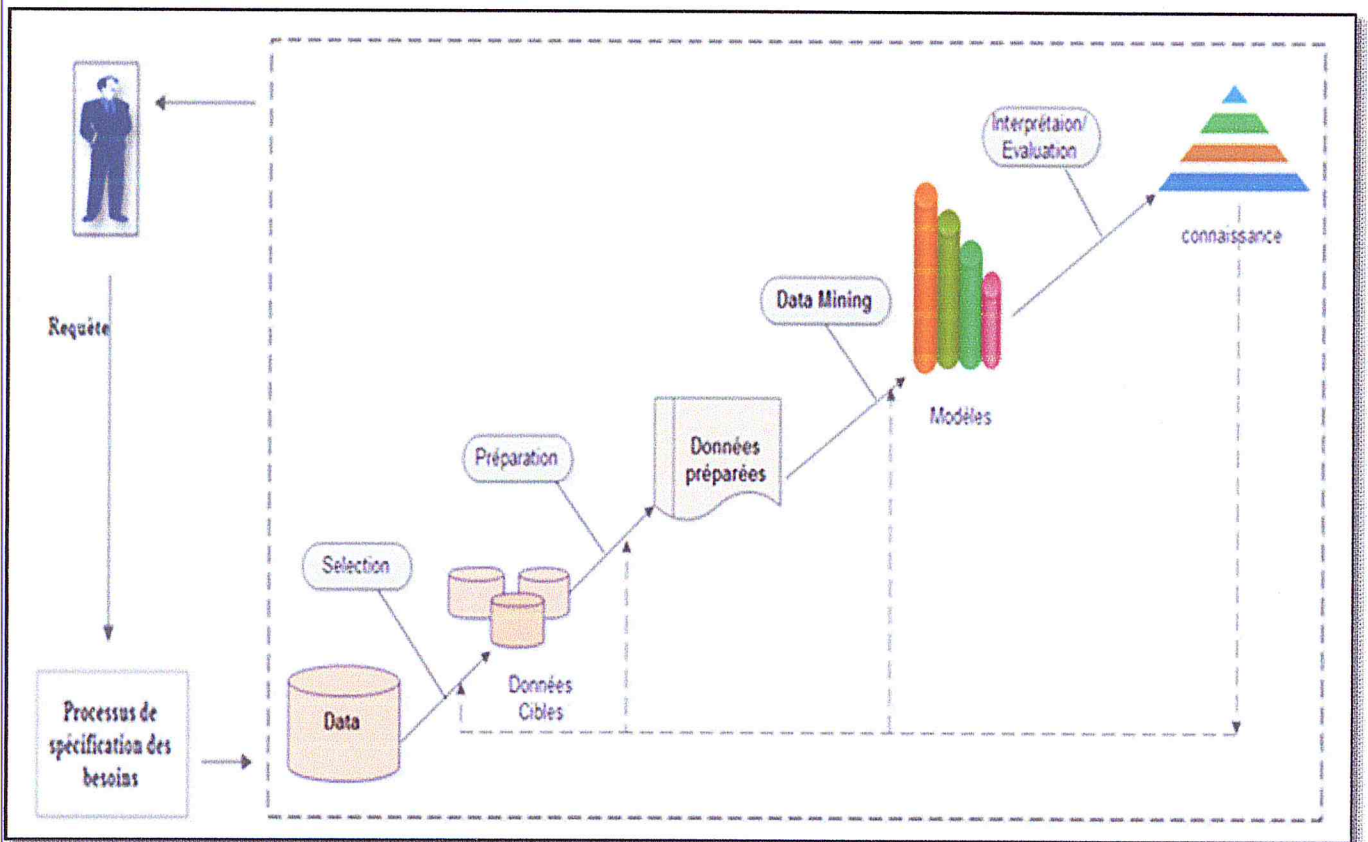


Figure 5.1: Le modèle de spécification des besoins du décideur



## 4. Modèle de spécification des besoins du décideur

Pour extraire ces connaissances utiles à partir des données volumineuses traitées par le processus KDD nous ferons appel à la méthode d'ingénierie des connaissances au décideur.

Nous proposons un modèle pour la spécification des besoins du décideur. Ce modèle est basé sur la méthode KADS [Knowledge Acquisition and Design Support], ce modèle nous aide à modéliser les connaissances dont le décideur a besoin. Cette modélisation est exprimée dans la partie spécification des besoins du décideur du modèle proposé, par un processus qui repose sur les modèles dans l'ordre (figure 5.2) :

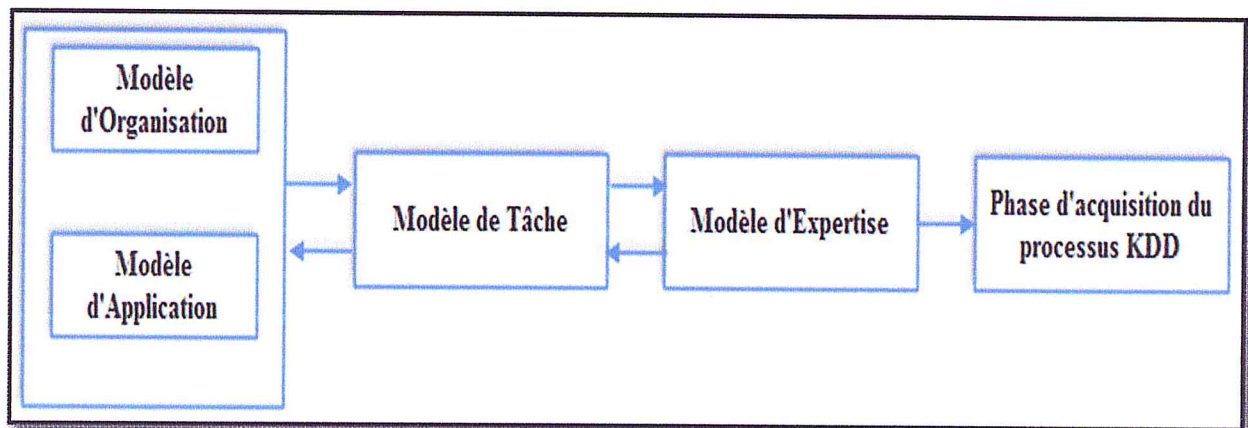


Figure 5.2 : Enchaînement des modèles de spécification/ KDD.

### 4.1. Le modèle d'Organisation/ Application

Ce modèle est l'intégration du modèle d'organisation et le modèle d'application, il définit aussi quel est le problème à résoudre, et décrit la requête lancée par le décideur, ainsi que les contraintes de l'objectif décisionnel.

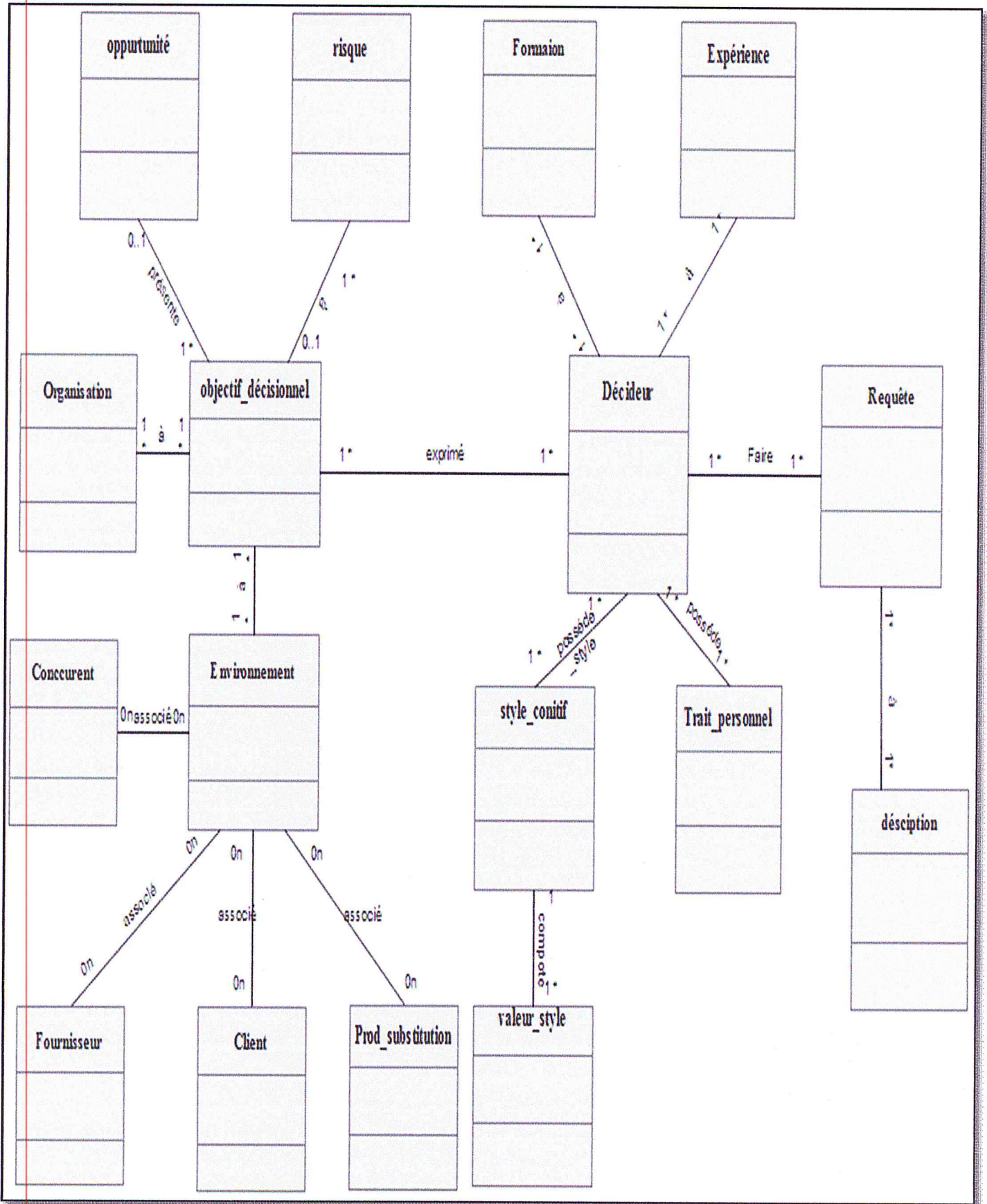


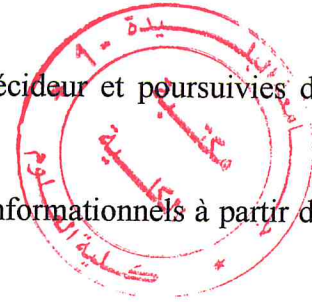
Figure 5.3 : diagramme de classe du modèle Organisation /Application .

## CHAPITRE 05 : LE MODELE DE DECOUVERTE DES CONNAISSANCES UTILES DANS UNE BASE DE DONNES

---

### ✚ Description des différentes classes du modèle d'Organisation/Application :

- **Décideur** : une personne physique, qui est devant un problème décisionnel et veut le résoudre.
- **expérience** : exprimée au nombre d'années d'ancienneté. L'expérience peut influencer sur le style cognitif du décideur et donc sur sa façon de percevoir et de résoudre le problème.
- **formation** : exprime les différentes formations du décideur et poursuivies durant son parcours professionnel.
- **Requête** : ce que veut le décideur comme besoins informationnels à partir d'une base de données.
- **Description** : explication détaillée de la requête collectée au près du décideur via des interviews dans le but de comprendre le vrai objectif de la requête.
- **trait personnalité** : C'est l'ensemble de comportement adapté par un individu dans les différentes situations
- **style cognitif** : « le style cognitif a une influence sur la façon dont les individus scrutent leur environnement pour recueillir de l'information, sur la façon dont ils organisent et interprètent cette information et sur la façon dont ils intègrent leurs interprétations dans les modèles mentaux qui guident leur actions » [Hayes 98]. Dans notre cas, le style cognitif est défini comme la façon propre à chaque décideur de comprendre l'information perçue à travers les différentes modalités qui sont à sa disposition, par exemple : sous quel format le décideur veut voir les résultats de recherche d'information (rapports, graphiques, ... etc.).
- **valeur style** : les différentes réponses pour les questions posées concernant le style cognitif du décideur.
- **Objectif décisionnel** : un objectif décisionnel est la décision à prendre par rapport à une situation décisionnelle, aussi c'est une expression particulière et contextuelle de la stratégie de l'entreprise et des priorités actuelles du décideur.
- **Risque** : les différents risques qu'il faut éviter lors de prise d'une décision et les impacts négatifs d'un objectif décisionnel.



- **Opportunité** : les différentes occasions qu'il faut prendre en considération lors de prise une décision et les impacts positifs d'un objectif décisionnel.
- **Organisation** : organisation supporte la coordination du travail entre les acteurs telle que ces acteurs qui vont mettre en œuvre les actions qui découlent des décisions.
- **Environnement** : environnement permet de circonscrire la relation de l'entreprise avec le monde qui l'entoure, Il a pour but de repérer les facteurs que le décideur considère comme sensibles Nous retenons comme paramètres de l'environnement ceux basés sur les cinq forces de Porter à savoir les *concurrents*, les *fournisseurs*, les *clients* et *produits de substitution*.

### 4.2. Le modèle de tâche

Le modèle des tâches permet de décrire les tâches et leurs répartitions, les entrées et les sorties, les conditions préalables et les critères d'exécution, en tant que ressources et compétences nécessaires pour le système. Le modèle des tâches sert donc à décrire les tâches qui seront exécutées durant la phase de spécification des besoins du décideur par le futur système lors de la réalisation d'une fonction précise de l'organisation.

Une tâche se décrit par ses entrées requises, ses sorties désirées (atteindre un but à valeur ajoutée), les ressources qu'elle utilise, les connaissances et les compétences qu'elle requiert pour être exécutée et celles qu'elle génère. La description d'une tâche est indépendante des entités qui les réalisent. En effet à chaque tâche, il sera possible de rattacher une entité désignée par le terme agent, permettant de la réaliser.

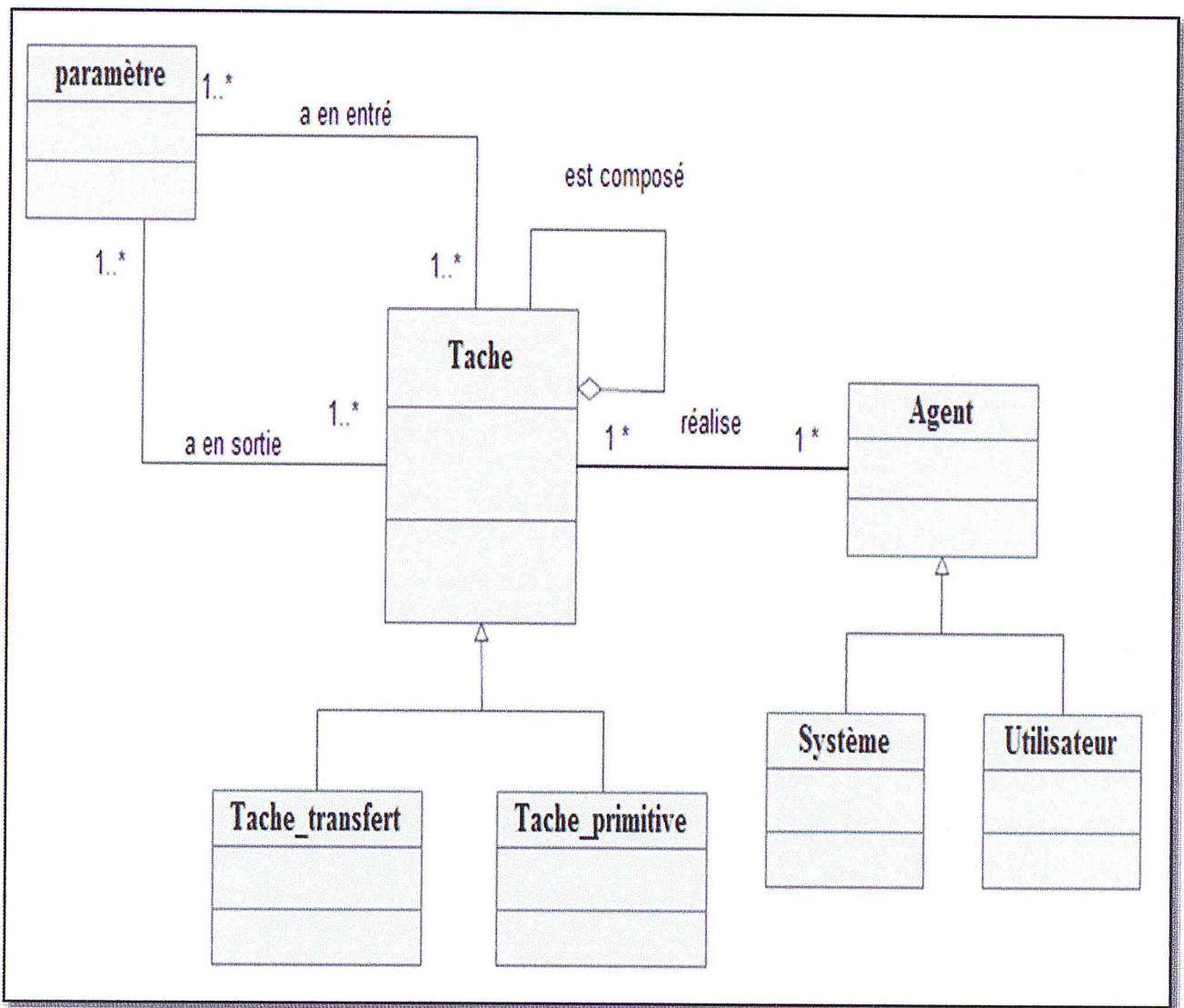


Figure 5.4 : Diagramme du modèle de tâche.

✚ Description des différentes classes du modèle de tâche :

- **Tache** : une tâche est une unité de calcul en résolution des problèmes, décrit les actions à effectuer pour attendre les buts, une tâche peut être simple ou composé des sous (tâche) la récursivité est ainsi possible.
- **Tache\_transfert** : exigeant une interaction avec un agent externe, elles sont spécifiées dans le modèle de coopération.
- **Tache\_primitive** : sont les inférences élémentaire indécomposable.
- **Agent** : est le responsable de la réalisation d'une tâche, il peut être utilisateur (analyste, décideur) ou système.
- **Paramètre** : sont les paramètres utiliser comme des entrés et des paramètres obtenu comme résultat lors d'exécution d'une tâche.

*Remarque:* Un paramètre de sortie d'une tâche peut être un paramètre d'entrée d'une autre tâche.

### 4.3. Le modèle d'Expertise :

Le modèle d'expertise c'est la partie la plus spécifique à un système à base de connaissances. Il décrit le comportement et le type de connaissances nécessaires au système ainsi que les compétences en résolution de problèmes des différents agents.

Comme nous l'avons déjà présenté dans le chapitre précédent, quatre types de connaissances sont à identifier et formaliser dans KADS : *la connaissance du domaine*, *la connaissance d'inférences* et *la connaissance de la tâche* et *la connaissance stratégie*.

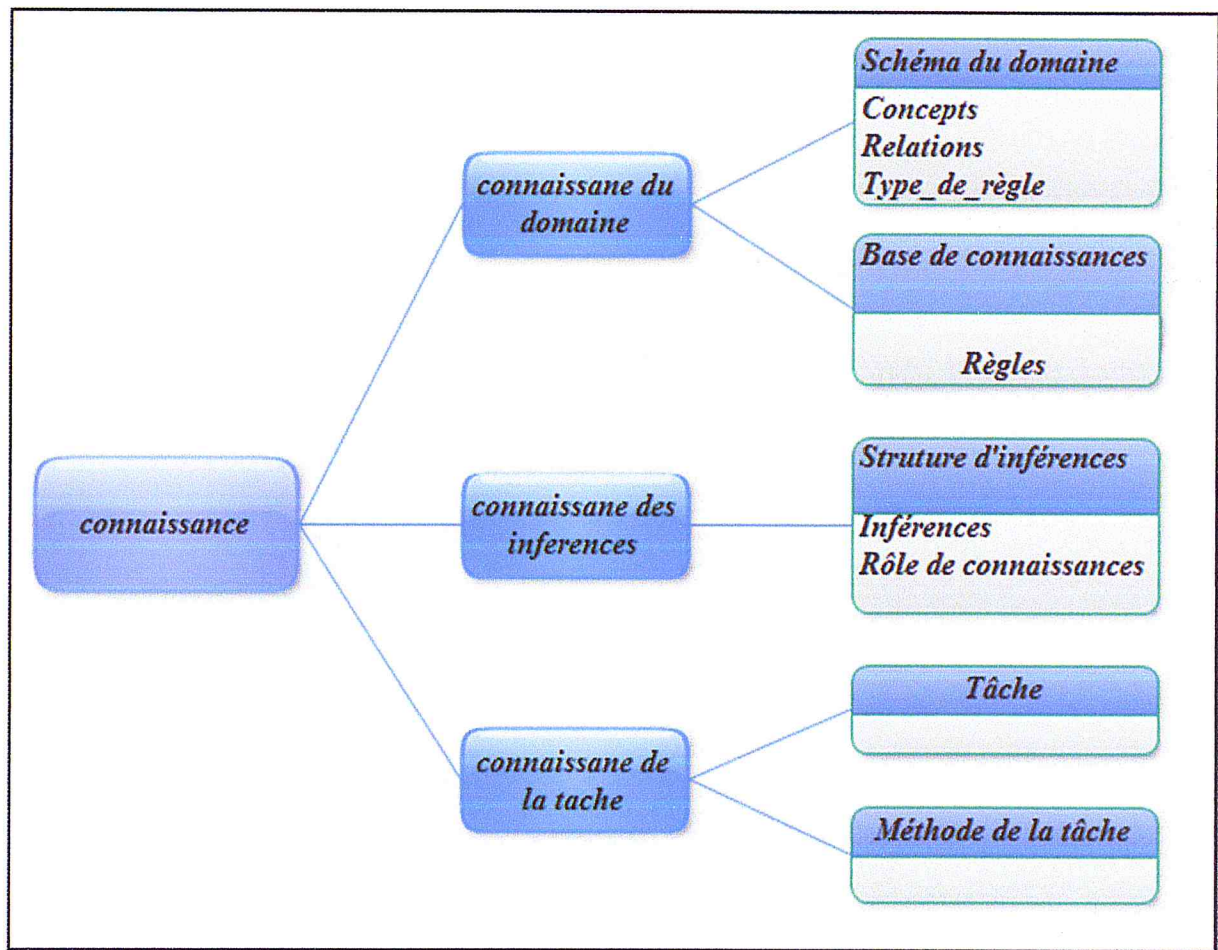


Figure 5.5 : Différents types de connaissances.

#### 4.3.1. La connaissance du domaine :

La connaissance du domaine décrit l'information statique du domaine d'application. Elle est habituellement constituée de deux ingrédients: le schéma du domaine (modèle des

## CHAPITRE 05 : LE MODELE DE DECOUVERTE DES CONNAISSANCES UTILES DANS UNE BASE DE DONNES

données) et la base de connaissances (les instances des types spécifiés au schéma du domaine). Il s'agit ici d'identifier les concepts et les relations, les types-de-règle (rule type) et de décrire la base de connaissances (BC) pour chaque relation et type-de-règle.

### 4.3.1.1. Le schéma du domaine :

Les éléments d'un schéma de domaine sont les concepts, les relations, et les types-de règle.

#### ✚ Les concepts

Un concept décrit un ensemble d'objets ou d'instances qui existent dans le domaine d'application et qui ont des caractéristiques similaires. Il correspond à la notion de classe dans la modélisation orientée-objet, à la différence que les fonctions ne sont pas incluses dans la définition d'une classe. Nous utilisons la notation UML [Ericksson 1998] pour décrire graphiquement les éléments du schéma du domaine.

Un concept a des attributs qui peuvent prendre des valeurs. Pour chaque attribut, on définit un type-de-valeur.

#### ✚ Les relations :

Les liens entre les différents concepts sont définis par une construction de type relation. Une relation peut elle-même avoir des attributs. Un exemple d'une relation est présenté à la figure ci-dessous. Dans ce cas, on voit qu'un Demandeur peut produire un ou plusieurs Demande , alors qu'un Demande est produit par un seul Demandeur, et que la relation "fait" n'a aucune attribut .

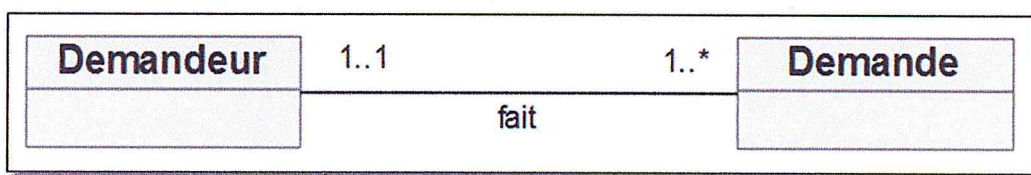


Figure 5.6: Exemple d'une relation.

#### ✚ Les types-de-règle

Les types-de-règle, qui différencient le schéma du domaine d'un modèle de données traditionnel, nous permettent de représenter les dépendances entre différents concepts. Ils nous permettent d'exprimer des relations logiques entre les concepts à un niveau schématique abstrait, c.-à-d. sans avoir à énumérer toutes les instances. Ils correspondent habituellement à des expressions concernant les valeurs des attributs d'un concept. Un type-de règle ressemble à une relation dont les arguments sont des antécédents et des conséquents. Ces arguments ne

## CHAPITRE 05 : LE MODELE DE DECOUVERTE DES CONNAISSANCES UTILES DANS UNE BASE DE DONNES

---

sont pas des instances de concept, mais plutôt des expressions à propos de ces concepts. Ils ne décrivent pas toujours une implication, mais peuvent aussi représenter des relations heuristiques entre les expressions du domaine. Les antécédents et les conséquents d'un type-de règle sont un concept ou une relation.

Dans notre travail, nous avons utilisés la notation UML, Donc les types-de-règle sont : généralisation, agrégation, composition... etc.

### 4.3.1.2. La base de connaissances :

Le deuxième élément est la base de connaissances. Celle-ci contient des instances de ces types de connaissance. Une base de connaissances est typique de la modélisation des connaissances; elle contient les instances sur lesquelles le raisonnement devra s'effectuer. La séparation du schéma du domaine et de la base de connaissances implique que l'acquisition de connaissances se fait en deux étapes. La première consiste à définir les types de connaissances et la deuxième à obtenir les instances de ces connaissances avec des va-et-vient entre les deux.

Une base de connaissances contient les instanciations des relations et de type-de-règle.

### 4.3.2. La connaissance d'inférences :

La connaissance d'inférences, deuxième type de connaissances dans KADS, décrit les inférences de base pour résoudre un problème donné en utilisant les connaissances du domaine.

Deux éléments constituent la connaissance d'inférences : *les inférences* et *les rôles de connaissance*. [Schreiber et al 94]

#### 4.3.2.1. Les inférences et les rôles de connaissances :

- **Définition Inférences :**

Elles forment le plus bas niveau des unités de traitement de l'information nécessaires dans un processus de raisonnement. Elles opèrent sur des données d'entrée, et ont la capacité de produire d'autres informations en sortie.

- **Définition Rôles de connaissance :**

Ils forment les entrées et les sorties d'une inférence. On les appelle rôles de connaissance car ils représentent les rôles que jouent les concepts du domaine par rapport aux inférences.



## CHAPITRE 05 : LE MODELE DE DECOUVERTE DES CONNAISSANCES UTILES DANS UNE BASE DE DONNES

Un rôle peut faire référence à un ou plusieurs concepts du domaine. De plus, les noms des rôles doivent être suffisamment abstraits pour favoriser leur réutilisation dans d'autres applications.

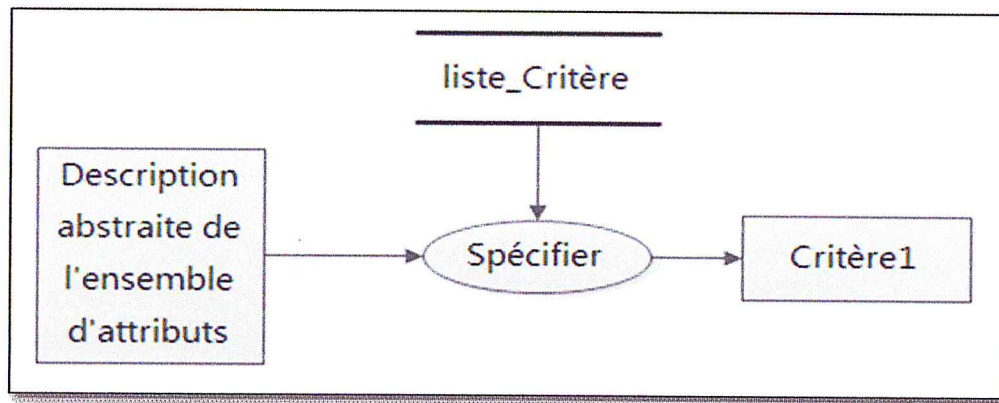


Figure 5.7 : Représentation graphique de l'inférence « spécifier ».

Dans la Figure 5.7 « Spécifier » est une inférence alors que « Description abstraite de l'ensemble d'attributs » et « Critère1 » sont des rôles de connaissances.

On distingue deux types de rôles de connaissance : les *rôles dynamiques* et les *rôles statiques*. Les rôles dynamiques sont les entrées et les sorties (durant le temps d'exécution) des inférences.

Les rôles statiques sont des ressources stables qui représentent des connaissances du domaine utilisées pour exécuter une inférence. Ils spécifient la connaissance du domaine dont on a besoin pour réaliser l'inférence.

### 4.3.2.2. La structure d'inférences

L'ensemble des inférences et des rôles de connaissances ainsi que les relations entre eux sont organisés dans une *structure d'inférence*

#### Définition Structure d'inférence

*Elle fournit un mécanisme d'abstraction sur les détails du processus de raisonnement. Elle constitue une trace du raisonnement dans la résolution de la tâche.*

Les conventions graphiques dans une structure d'inférence sont les suivantes : Des *ovales* représentent les inférences. Des *rectangles* représentent les rôles dynamiques de la connaissance. Des *traits épais horizontaux* indiquent les rôles statiques. Des *flèches* relient les rôles et les inférences.

### 4.3.3. La connaissance de la tâche

La connaissance de la tâche, troisième type de connaissances dans KADS, permet de définir l'application (la tâche), ses objectifs et les moyens pour les réaliser. Les tâches peuvent être composées en sous-tâches, qui peuvent être composées de tâches terminales indécomposables (inférences). Les connaissances de la tâche spécifient aussi le contrôle (ordre d'exécution, composition, etc.) autour des sous-tâches et inférences qui la composent. Dans notre cas, la tâche globale est *l'identification des critères de sélection*.

La méthode de la tâche représente la façon dont la tâche est effectuée. Les inférences sont les éléments de raisonnement utilisés pour la méthode pour résoudre la tâche. Une même inférence peut être utilisée par plusieurs méthodes et une même méthode peut utiliser plusieurs inférences. Il est possible, de voir à travers la manière dont les inférences sont appliquées, une stratégie d'exécution de la tâche réalisée.

## 5. Etude de cas

Pour faire une validation dans le monde réel (instanciation du modèle), nous avons pris comme exemple l'agence AADL (Agence nationale de l'Amélioration et du Développement du logement), L'instanciation des classes dans ce cas est faite comme suit :

### 5.1. Le modèle Organisation /Application

- **Décideur** : le directeur de l'AADL (Agence nationale de l'Amélioration et du Développement du logement).
  - **expérience** : plus de dix ans au domaine logement.
  - **formation** : plusieurs formations concernant le domaine logement.
  - **style cognitif** : rationnel.
- **Requête** : Sélectionner les critères de priorité de distribution des logements
- **Objectif décisionnel** : commencer un nouveau projet AADL (un nouveau site).
  - **Risque** : le nombre des demandeurs ne correspond pas au nombre des logements c.-à-d. le nombre des logements est supérieur du nombre des demandeurs.
  - **Opportunité** : l'amélioration et le développement du projet AADL

- **Organisation :** L'organisation garante la coordination du travail entre les acteurs comme les bureaux d'études, direction des finances et de comptabilité qui font plusieurs opérations comme l'élaboration des différentes études du projet selon les conditions qui sont prises en considération.
- **Environnement :**
  - **Client :** sont les habitants qui font une demande.
  - **Concurrent :** sont les agences qui offrent la même activité.
  - **Fournisseurs :** les agences ou l'état qui peuvent participer au projet.

### 5.2. Le modèle de tâche :

Dans notre cas, la tâche globale (requête) comporte trois tâches principales:

- ✓ Créer un critère,
- ✓ Modifier un critère
- ✓ Supprimer un critère.

Chacune de ces tâches se décompose en sous-tâches. La figure 5.8 présente la décomposition de ces tâches.

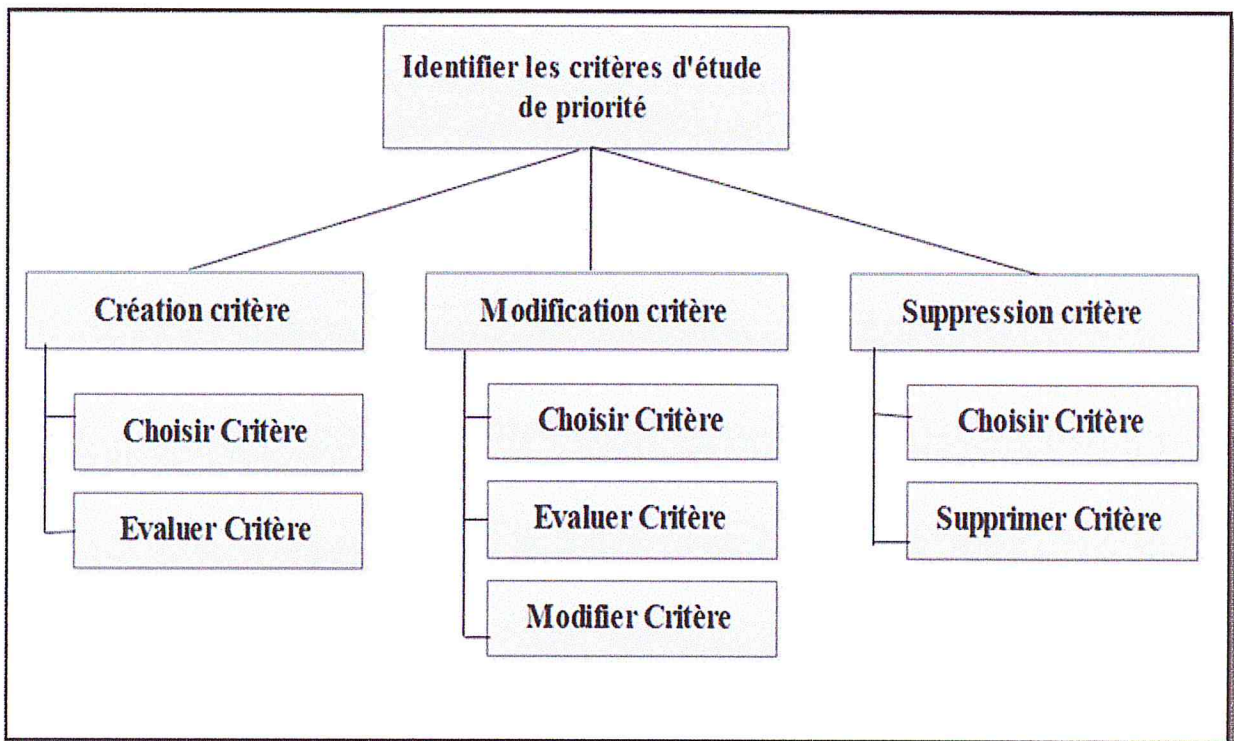


Figure 5.8 : Diagramme du modèle de tâche pour identifier les critères d'étude de priorité.

### ✚ Description des décompositions des tâches :

- la tâche de création d'un critère comprend une première sous-tâche « Choisir Critère ». Ensuite une deuxième sous-tâche, « évaluer critère», consiste à évaluer la valeur du critère à partir de la liste des critères.
- La tâche de modification d'un critère permet d'éditer les critères de sélection Elle se décompose en trois sous-tâches : la tâche « Choisir critère» qui consiste à identifier le critère que nous voulons modifier et la tâche « évaluer critère», consiste à évaluer la valeur du critère, et la tâche « Modifier critère» qui permet de changer les critères.
- La tâche de suppression d'un critère permet d'éliminer un critère. Elle se décompose en deux sous-tâches : la tâche « Choisir critère » qui consiste à identifier les critères que le décideur veut effacer et la tâche « Supprimer critère» qui consiste à effacer critère sélectionner pour la suppression.

### 5.3. Le modèle Expertise

#### 5.3.1. Connaissance du domaine

La figure 5.9 présente les différents concepts du domaine (schéma du domaine distribution logement).

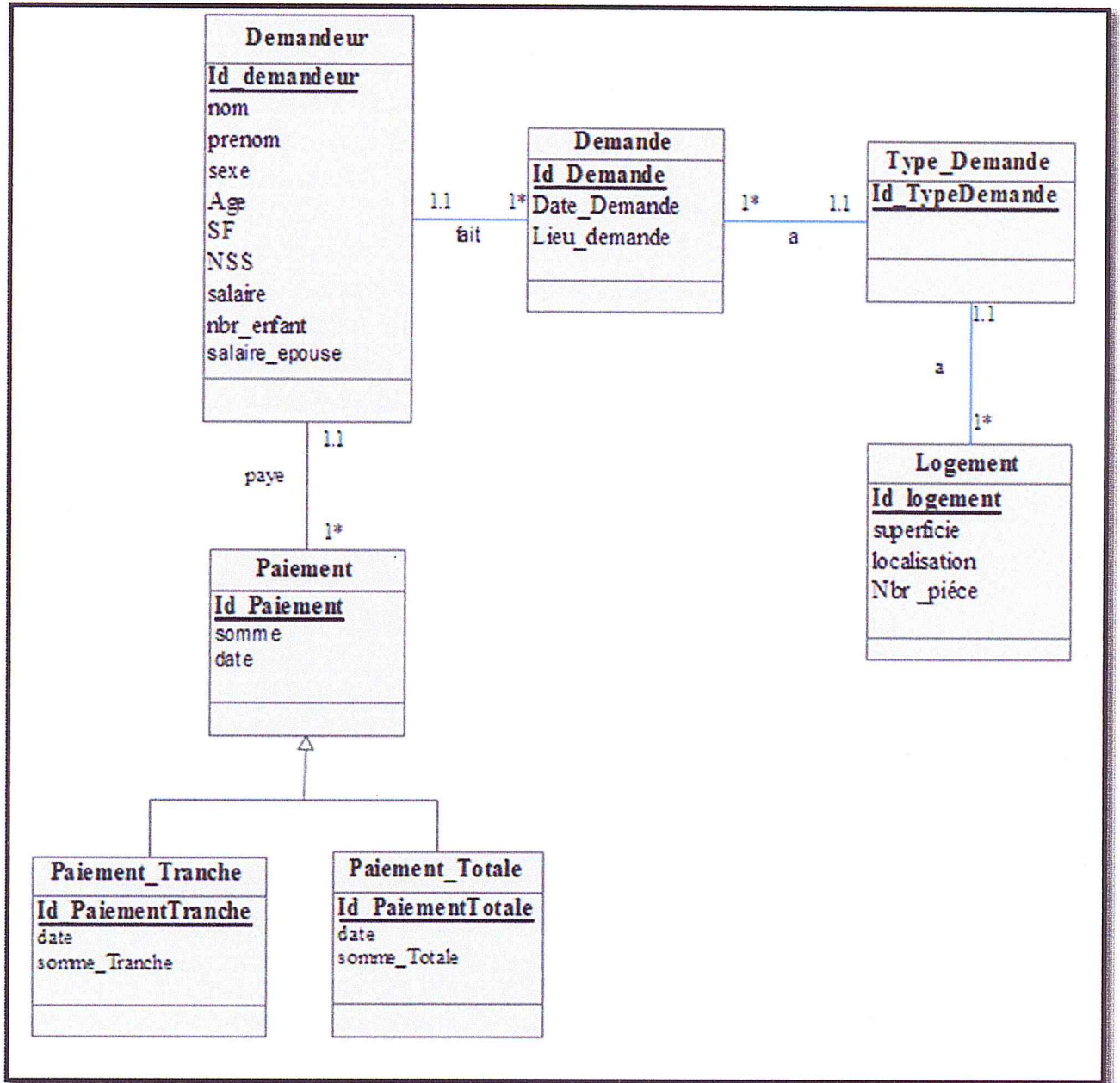


Figure 5.9 : Les Concepts du domaine distribution logement.

5.3.2. Connaissance d'inférence

Nous avons choisi de présenter dans notre cas la structure d'inférence relative à la sélection des attributs d'étude.

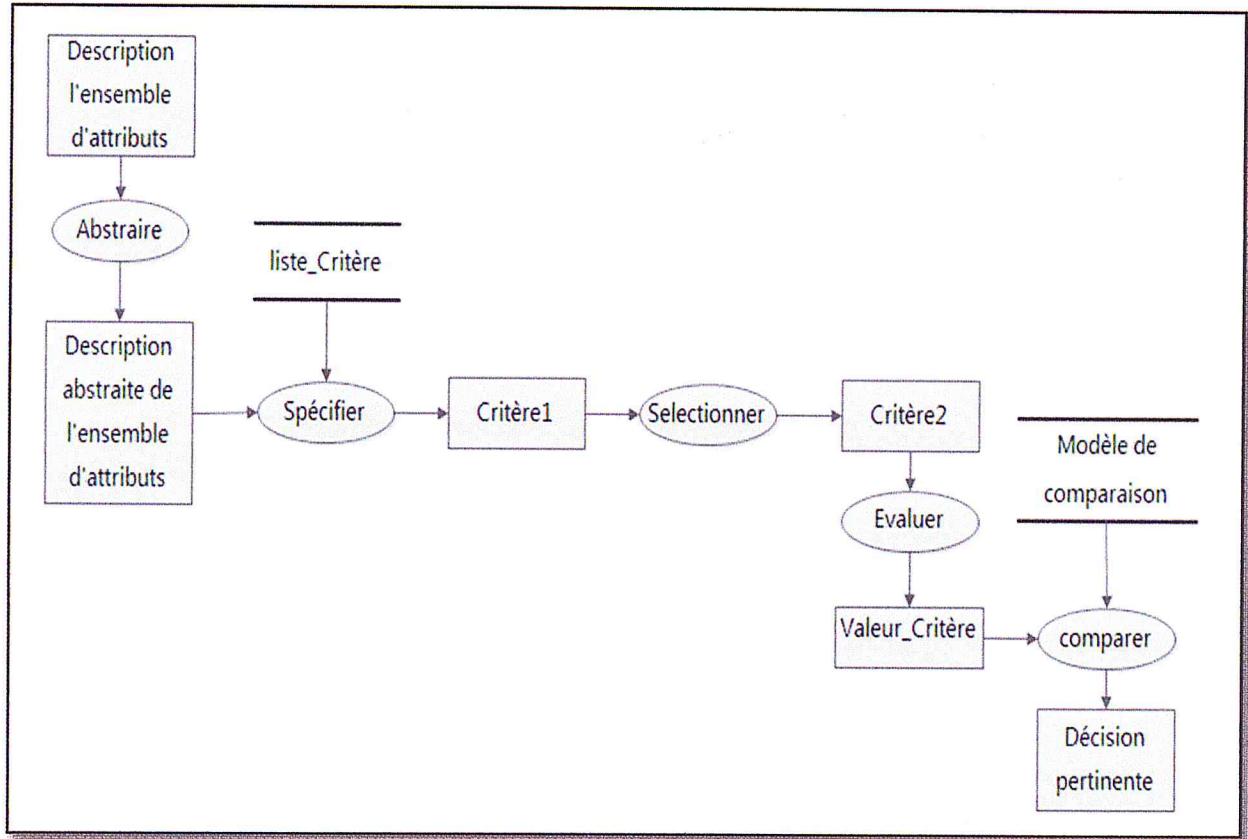


Figure 5.10 : Structure d'inférence pour la tâche d'évaluation d'un critère.

Et le tableau suivant présente la description du structure d'inférence d'évaluation d'un critère

# CHAPITRE 05 : LE MODELE DE DECOUVERTE DES CONNAISSANCES UTILES DANS UNE BASE DE DONNES

Inférence	Input	Output	R_Statique	Description
Abstraire	Description_attribut	Description_abstraite_attribut	/	l'entrée est un attribut. la sortie est composée du même attribut mais dont la description est enrichie avec de nouvelles variables permettant de juger la pertinence de l'attribut.
Spécifier	Description_abstraite_attribut	Critère1	Liste_critères	Il s'agit de rechercher les critères qui peuvent être utilisés pour l'évaluation des attributs. Les mêmes connaissances statiques d'abstraction sont utilisées
Sélectionner	Critère1	Critère2	/	Il s'agit de sélectionner l'un ou plusieurs de critères pour l'évaluation
Evaluer	Critère2	Valeur_critère;	/	déterminer la valeur du critère (ou critère) choisi pour un cas donné. Les contraintes issues du Besoin interviennent dans l'évaluation.
Comparer	Valeur_Critère	Décision_pertinence	Modèle de comparaison	en compare les résultats des évaluations selon les différents critères choisi, décider si l'attribut est pertinent. Si oui, il est candidat pour être choisi comme attribut d'étude. Sinon, l'attribut est rejeté

Tableau 5.1 : Description de structure d'inférence d'évaluation du critère.

### 5.3.3. Connaissance de tache

La figure 5.11 ci-dessous présente la décomposition de la tâche de création d'un critère.

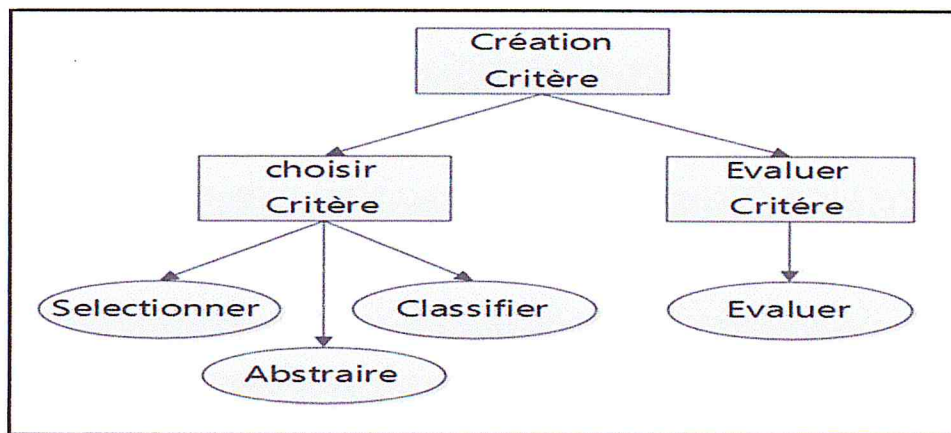


Figure 5.11 : diagramme de la tâche Création critère.

La figure 5.12 ci-dessous présente la décomposition de la tâche de modification d'un critère.

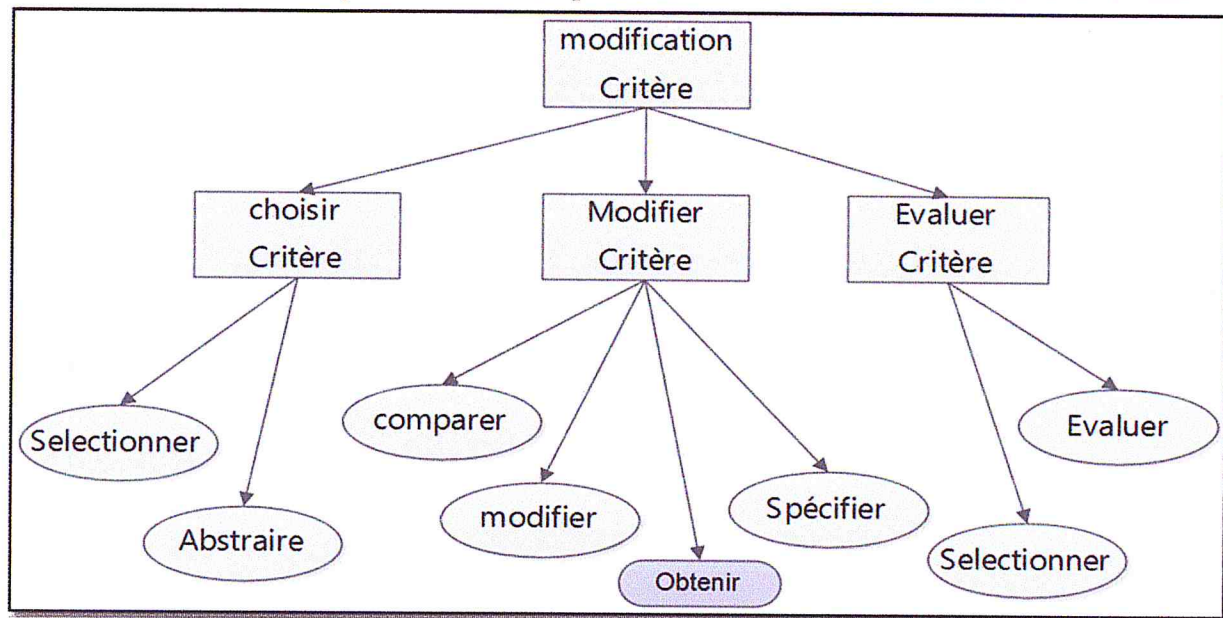


Figure 5.12 : *diagramme de la tâche Modification critère.*

La figure 5.13 ci-dessous présente la décomposition de la tâche de suppression d'un critère.

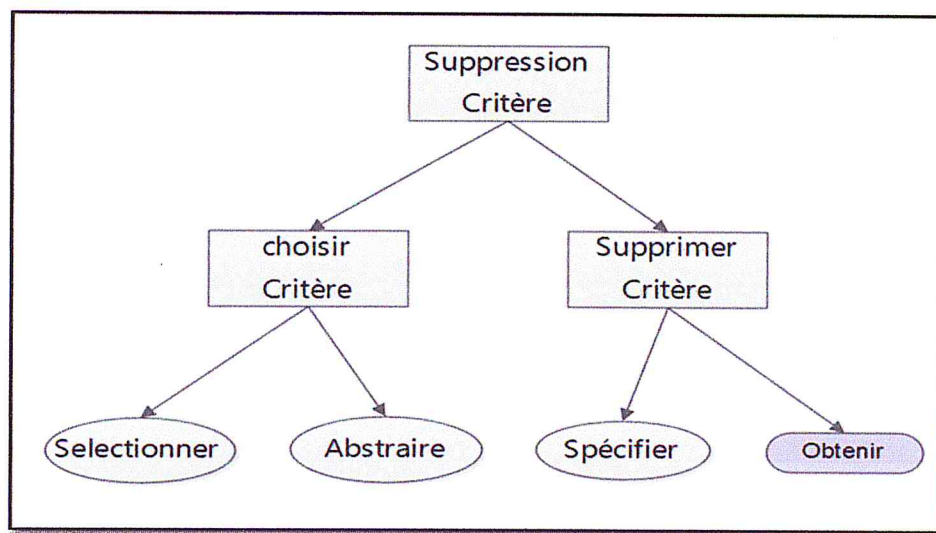


Figure 5.13 : *diagramme de la tâche Suppression critère.*

## 6. Utilisation des résultats du modèle proposé par le processus KDD

Notre contribution est composée de deux parties important :

- ✚ Modèle proposé en utilisons la méthode KADS en a inspiré les trois modèles : modèle organisation /application, modèle tache et le modèle d'expertise.
- ✚ Processus KDD qui compose de plusieurs phases : sélection des données, préparation des données, Datamining, et l'interprétation et visualisation.



## CHAPITRE 05 : LE MODELE DE DECOUVERTE DES CONNAISSANCES UTILES DANS UNE BASE DE DONNES

Notre objectif est la spécification des besoins du décideur et pour atteindre notre objectif on basé sur la phase sélection et pallié le déficit qui existe au niveau de cette phase comme la figure ci-dessous montre :

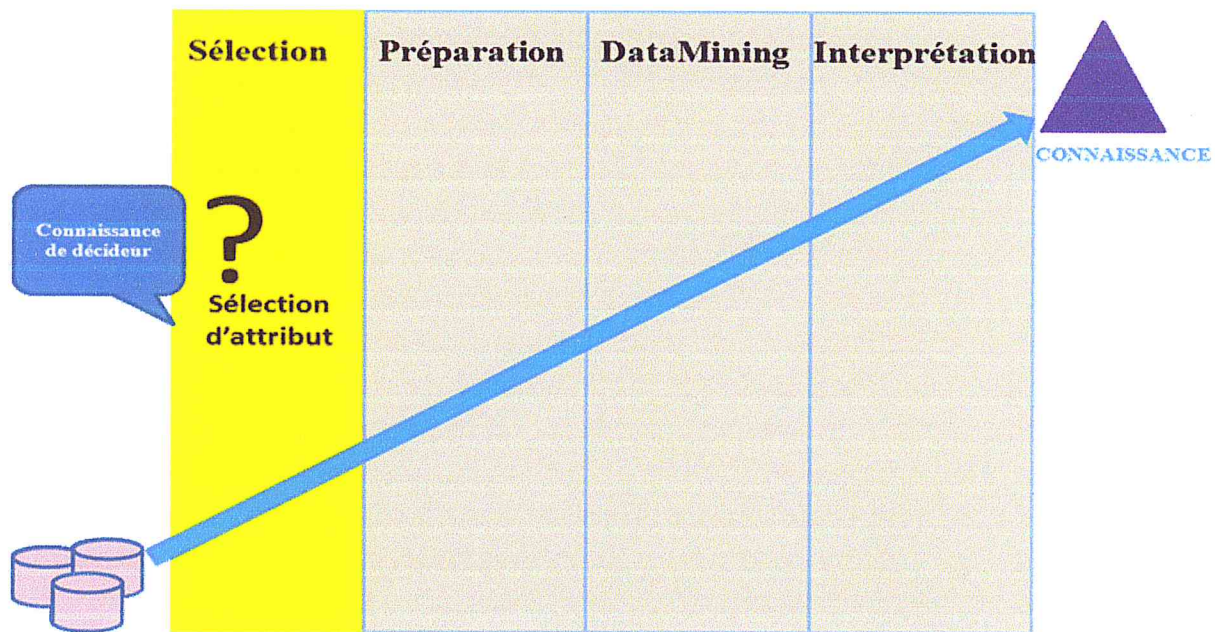


Figure 5. 14 : L'incorporation des connaissances du décideur au processus KDD.

Dans notre cas, il y a plusieurs décideurs ou chaque décideur a sa particularité et sa façon de prise de décision.

Pour la prise de décision, le décideur prend en considération des critères, par exemple dans notre cas de savoir la priorité de distribution des logements.

Il existe des décideurs qui s'intéressent au critère « *salaires* », « *Age* » et « *la situation familiale* » des demandeurs ;

Alors qu'il existe d'autres décideurs qui s'intéressent au critère « *Age* », « *sexe* », « *la situation familiale* » et même pour « *le nombre d'enfant du demandeur* » d'où vient la différence de décision.

Sachant que chaque critère soumet à plusieurs contraintes, par exemple le critère « *salaires* » soit (faible entre  $[0, 40000]$ , moyen entre  $[40000, 60000]$  et élevé entre  $[60000, 90000]$ ) et le critère « *situation familiale* » soit (célibataire, marié, divorcés, veuf).

Après la sélection des critères, on commence le processus KDD en traitant que les critères sélectionnés par le décideur en passant par la phase sélection qui devient rapide et

facile et bien spécifier ,ensuite la phase préparation de donné en arrivant à la phase Datamining qui est le cœur du processus KDD en utilise un algorithme de classification supervisé (SVM Machines à vecteurs supports) telle que les SVM peuvent être utilisés pour résoudre des problèmes de discrimination, c'est-à-dire décider à quelle classe appartient un échantillon, ou de régression, c'est-à-dire prédire la valeur numérique d'une variable. La résolution de ces deux problèmes passe par la construction d'une fonction  $h$  qui à un vecteur d'entrée  $x$  fait correspondre une sortie  $y : y = h(x)$ .

Et pour la phase interprétation et visualisation est selon le style cognitif du décideur (Graphe, rapport,... etc.)

Afin de savoir qui parmi les demandeurs a la priorité selon les différents critères sélectionnés par le décideur.

### **7. Architecture générale du système :**

La principale contribution de notre travail est la spécification des besoins du décideur pour extraire les connaissances utiles dans le processus KDD, il est basé sur la sélection des données pertinentes dans la partie d'acquisition de données.

Pour cela, avant le lancement du processus KDD nous avons ajoutés un modèle basé sur la méthode KADS, comporte les modèles suivant : modèle de l'organisation /application, modèle de tâche, modèle d'expertise.

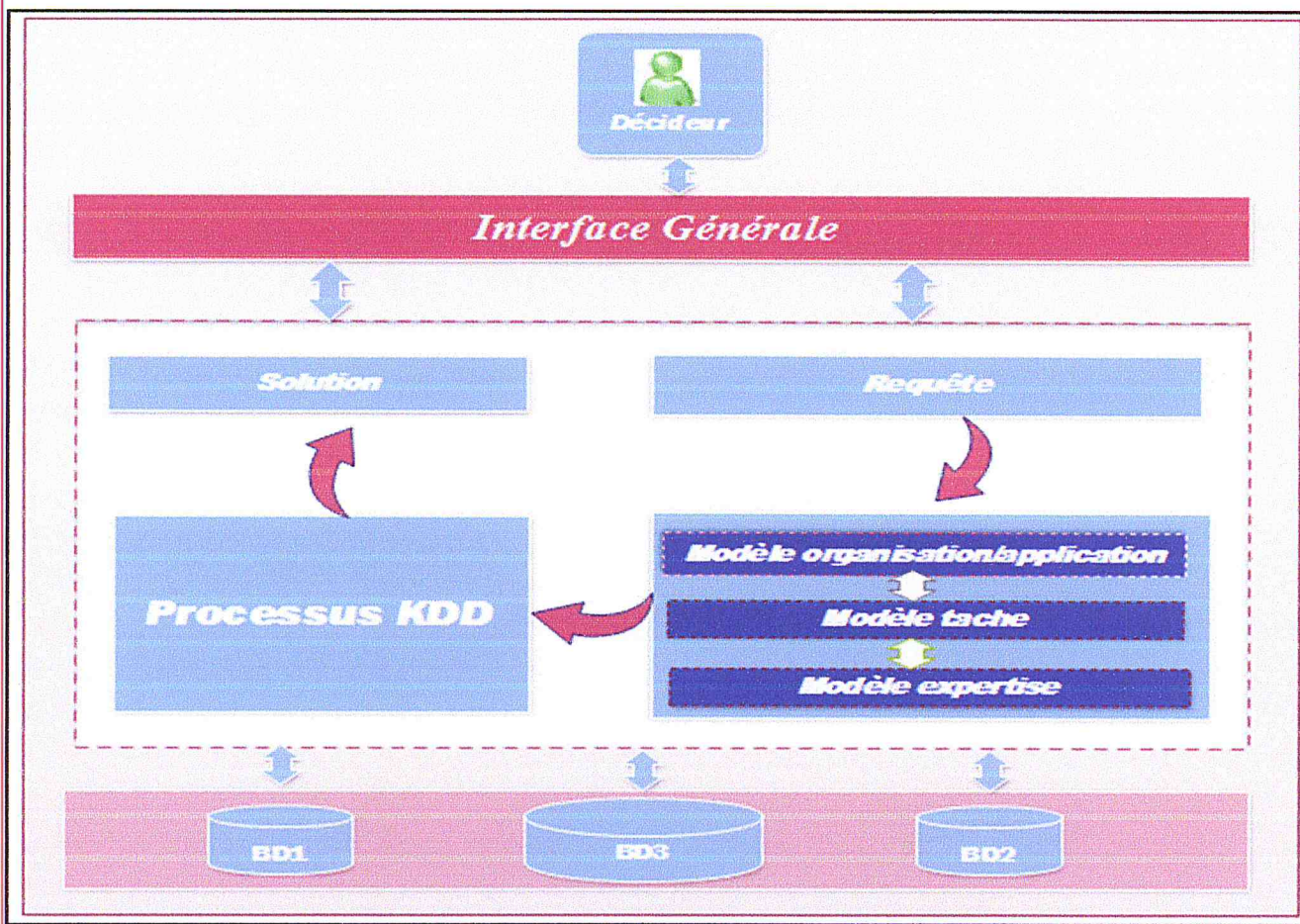


Figure 5.15 : Architecture générale du système.

Certains modèles seront plus détaillés que d'autres en fonction de leur importance dans notre projet sachant que nous avons basé sur le modèle d'expertise.

Dans le schéma ci-dessus nous représentons les principaux modules de notre architecture dans lesquels chaque module a sa propre fonctionnalité comme suit :

L'architecture est composée de cinq (05) modules (module requête, module solution, module du modèle proposé, module processus KDD, module base de données) :

- **Module requête** : Ce module permet de spécifier les requêtes (besoins) d'information du décideur sous forme d'une requête.
- **Module solution** : ce module permet au système de visualiser les résultats de la requête du décideur.
- **Module de modèle proposé** : ce module comprend le modèle organisation/application, le modèle de tâche et le modèle expertise pour décrire :
  - Le profil du décideur selon le modèle organisation /application,
  - La décomposition de tâche (requête) du décideur selon le modèle de tâche,

- La résolution du problème par le système selon le modèle d'expertise.
- **Module processus KDD** : Ce module permet de lancer le processus KDD pour l'extraction des connaissances à partir d'une base de données.
- **Module base de données** : qui regroupe trois types de base de données,
  - BD1 : stocke les différentes requêtes du décideur,
  - BD2 : stocke les données des modèles d'acquisition,
  - BD3 : stocke les données exploitées par le processus KDD.

### 8. Conclusion

Notre contribution est un modèle de spécification des besoins du décideur, ou nous avons montré l'incorporation de la méthode KADS (qui est utilisée au début dans l'IC pour extraire les connaissances de l'expert pour la construction des SBC) avec le processus KDD.

Dans notre travail, la méthode KADS est utilisée pour extraire les connaissances du décideur pour spécifier ses besoins, en prend en considération son profil, ses objectifs, l'environnement requête, sous requête ainsi que les paramètres utilisé dans la réalisation d'une tâche toutes ces données seront exploitées pour aider le processus KDD à extraire les connaissances utiles à partir de bases de données et facilite la prise de décision.

# **Chapitre 06**

## **IMPLEMENTATION**

## 1. Introduction

Dans ce chapitre, nous allons introduire les aspects techniques liés à l'implémentation de notre projet, en expliquant les outils utilisés dans l'élaboration du système, nous avons utilisé le langage C#, ce langage est orienté objet et il est destiné au développement de aspect professionnel.

Plus loin dans ce chapitre, nous allons décrire les buts désirés de l'application et aussi les différentes interfaces qui la compose et les différents résultats qui sont générés.

Les bases de données sont gérées par des logiciels spécialisés appelés système de gestion de base de données .Il existe plusieurs SGBD: Access, SQL server, MYSQL, Oracle.

Nous avons utilisé une base de donnée Access car elle offre un ensemble d'outils permettant de : Saisir, de mettre à jour, de manipuler, d'interroger et d'imprimer des données.

## 2. Environnement de développement Visual Studio 2008

Visual Studio 2008 est un environnement de programmation riche en outils comportant toutes les fonctionnalités nécessaires pour créer des projets C# de toute taille. Vous pouvez même créer des projets qui combinent harmonieusement des modules compilés utilisant des langages de programmation différents.

## 3. Langage de programmation

C# est un langage récent. Il a été disponible en versions beta successives depuis l'année 2000 avant d'être officiellement disponible en février 2002 en même temps que la plate-forme .NET 1.0 de Microsoft à laquelle il est lié. C# ne peut fonctionner qu'avec cet environnement d'exécution. Celui-ci rend disponible aux programmes qui s'exécutent en son sein un ensemble très important de classes.

En première approximation, on peut dire que la plate-forme .NET est un environnement d'exécution analogue à une machine virtuelle Java. On peut noter cependant deux différences

- Java s'exécute sur différents OS (windows, unix, macintosh) depuis ses débuts. En 2002, la plate-forme .NET ne s'exécutait que sur les machines Windows. Depuis quelques années le projet Mono [<http://www.mono-project.com>] permet d'utiliser la plate-forme .NET sur des OS tels qu'Unix et Linux. La version actuelle de Mono (février 2008) supporte .NET 1.1 et des éléments de .NET 2.0.
- La plate-forme .NET permet l'exécution de programmes écrits en différents langages. Il suffit que le compilateur de ceux-ci sache produire du code IL (Intermediate

## CHAPITRE 06 : IMPLEMENTATION

Language), code exécuté par la machine virtuelle .NET. Toutes les classes de .NET sont disponibles aux langages compatibles .NET ce qui tend à gommer les différences entre langages dans la mesure où les programmes utilisent largement ces classes. Le choix d'un langage .NET devient affaire de goût plus que de performances.

### 4. Interface :

#### 4.1. Authentification :

Le décideur qui accède au système, il faut passer par l'interface d'authentification pour introduire son login et son mot de passe.

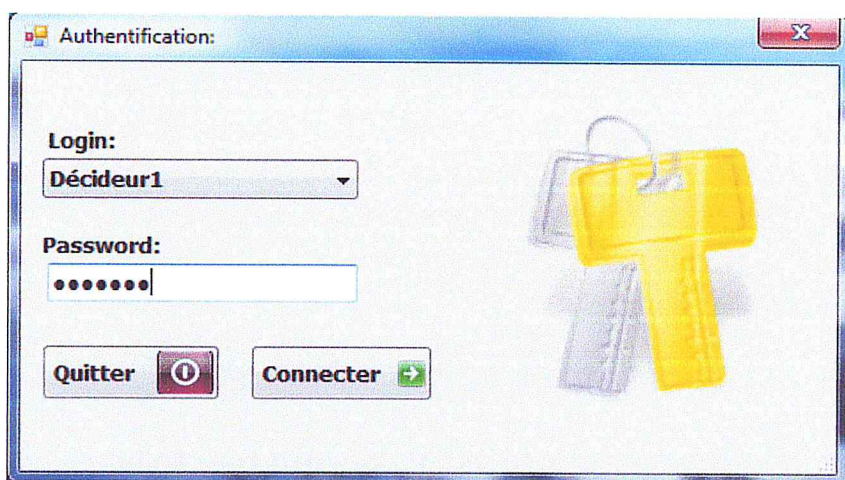


Figure 6.1 : interface Authentification.

#### 4.2. Interface Traitement profil du décideur:

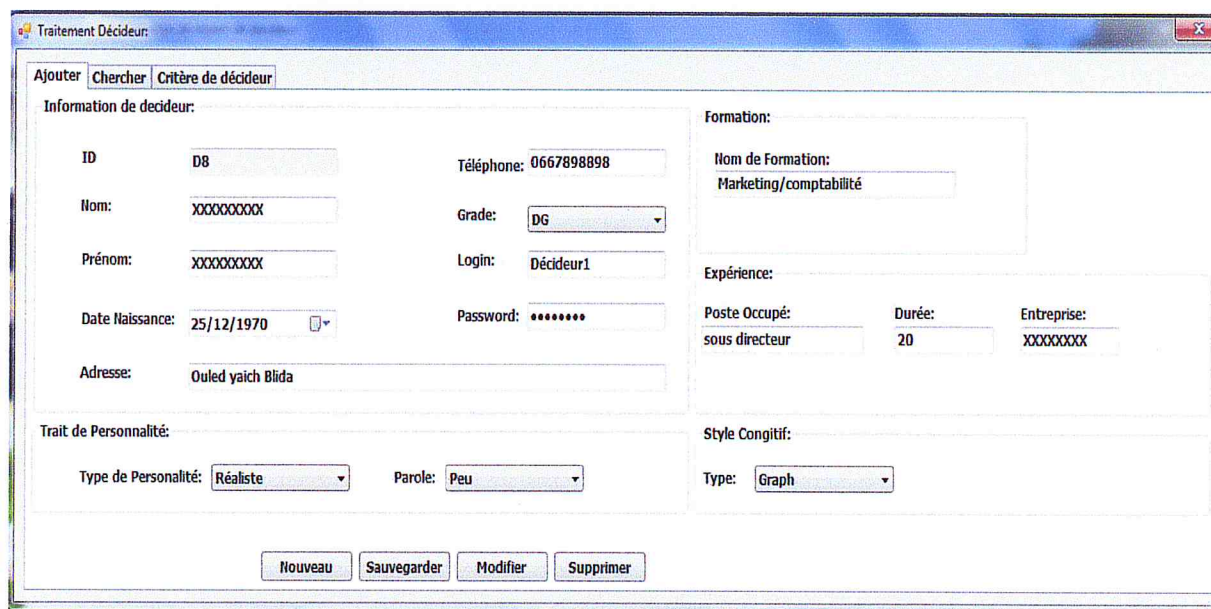


Figure 6.2 : Interface Traitement décideur selon le modèle organisation /application.

## 4.3. Interface Traitement de requête décideur :

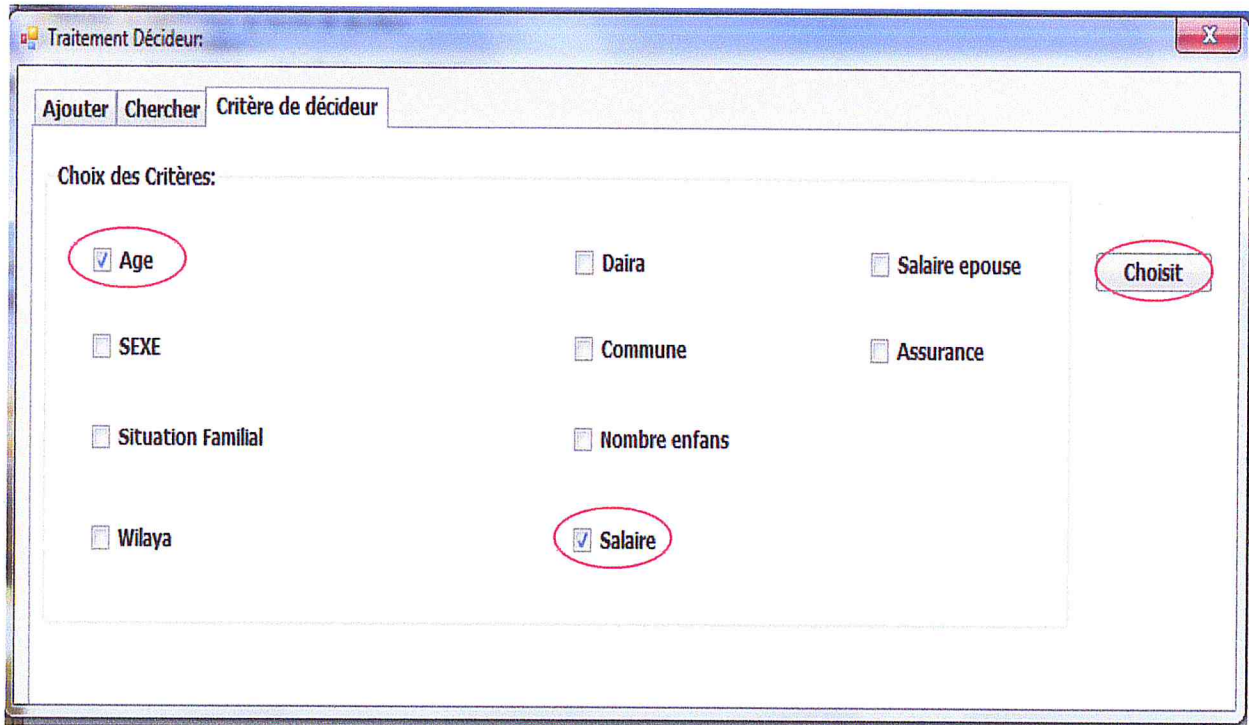


Figure 6.3 : Interface Traitement de requête décideur.

*Remarque : les interfaces ci-dessous présente les différentes phases du processus KDD*

## 4.4. Interface sélection Critère du décideur

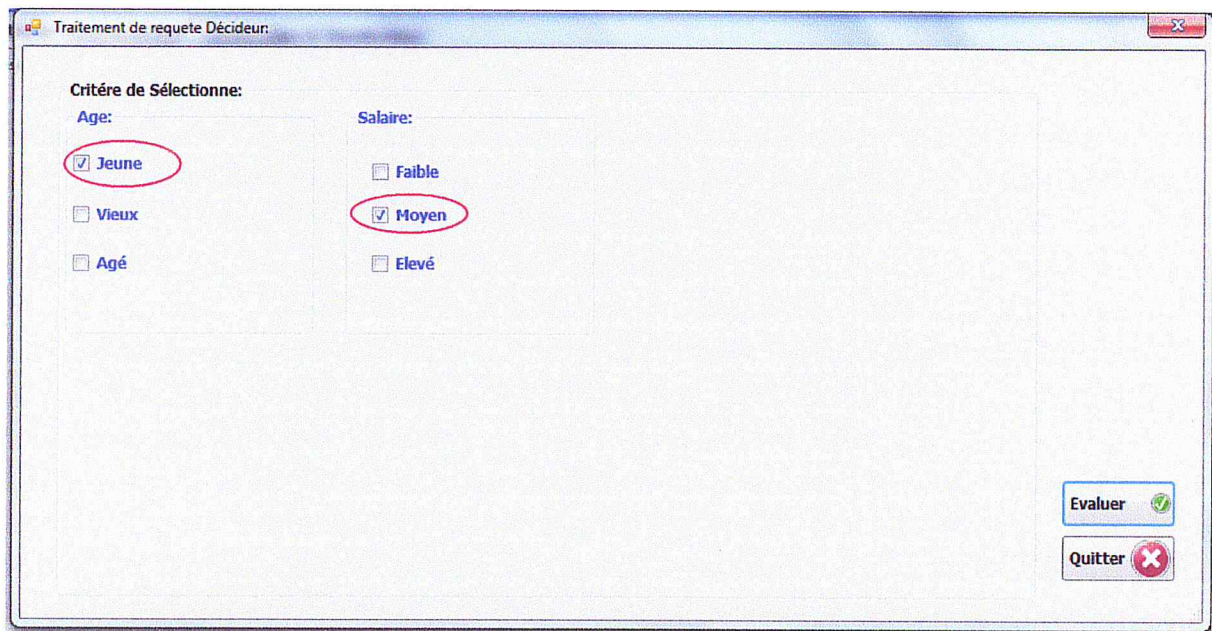


Figure 6.4 : Interface sélection Critère du décideur.



# CHAPITRE 06 : IMPLEMENTATION

## 4.5. Interface du résultat de la sélection critère.

ID Candidat	Nom Candidat	Age	Sexe	situation Familial	wilaya	Daira	Commune	Nombre Enfants	Salaire	Salaire Epouse	Assura
C1	BENZABA AZIZ	39	H	C	Bldza	MEFTAH	MEFTAH	0	40000	0	CNAS
C100	AMROUCH SA...	40	H	M	Bldza	BOUJNANE	BOUJNANE	2	30000	20000	CNAS
C1000	HADJI RACHID	25	H	M	Bldza	MEFTAH	MEFTAH	5	35000	0	CASNOS
C1003	BENATALLAH...	30	H	M	Bldza	BOUGARA	BOUGARA	1	40000	10000	CASNOS
C1005	MAKHOUDI ALI	21	H	M	Bldza	ouled yiche	ouled yiche	0	40000	0	CASNOS
C1006	AIT ZIANE AM...	38	H	M	Bldza	ARBAA	ARBAA	4	40000	25000	CNAS
C134	MESBAH FAH...	29	F	M	ALGER	AIN BENIANE	AIN BENIANE	6	30000	35000	CASNOS
C1340	HADIMI WASS...	27	F	M	ALGER	BEO	BEO	9	40000	50000	CASNOS
C1384	BOUCETTA A...	28	F	M	ALGER	BEO	BEO	3	30000	35000	CNAS
C1391	ALILAT RADIA	40	F	M	ALGER	AIN NAAJA	AIN NAAJA	10	30000	50000	CNAS
C1513	KENZOUZ KAH...	39	F	C	ALGER	BEO	BEO	0	40000	0	CNAS

Figure 6.5 : Interface du résultat de la sélection critère.

## 4.6. Interface du classification et visualisation

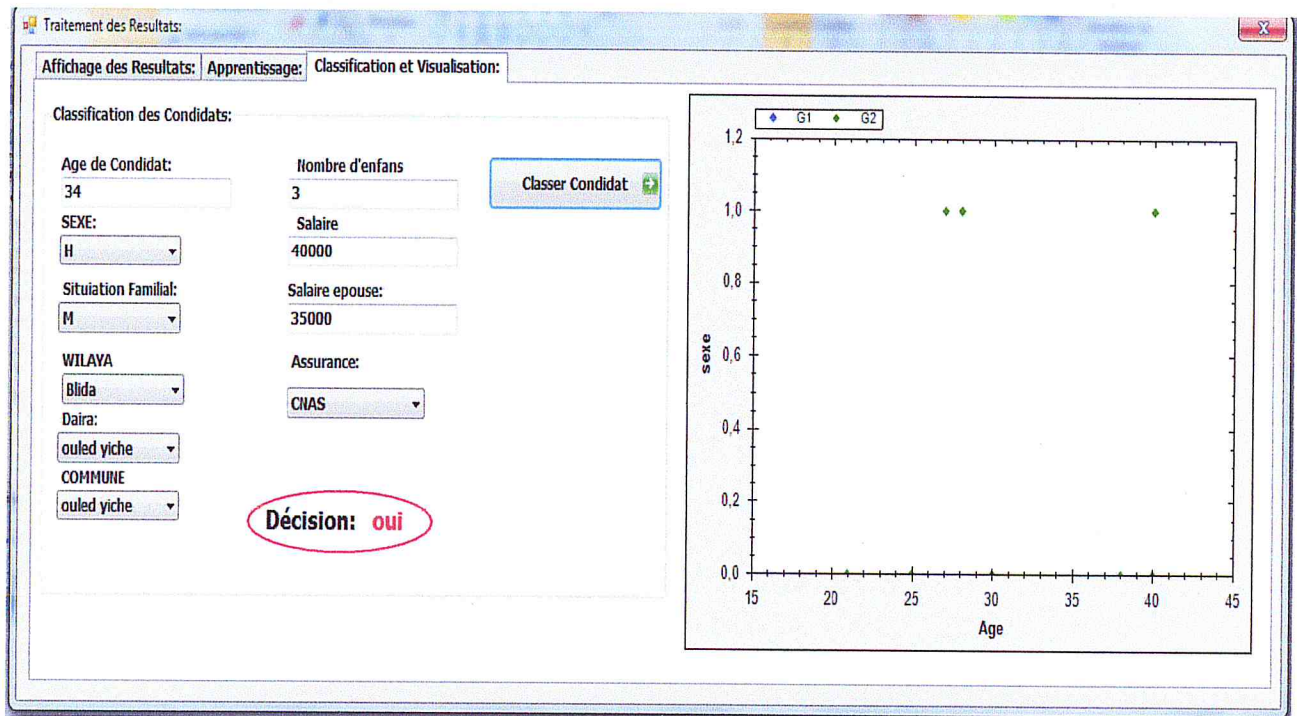


Figure 6.6 : Interface de classification et visualisation

## 5. Conclusion

Dans cette étape de notre projet, nous avons présenté les différents outils utilisés pour implémenter notre modèle ainsi les différents interfaces implémentées et le résultat du processus KDD.

# **Conclusion Générale**

## Conclusion Générale

---

### Conclusion Générale

La prise de décision est définie comme l'ensemble des actions que le décideur effectue pour prendre sa décision. Les décideurs ont sans doute besoin de support et d'aide pour augmenter leur efficacité et minimiser le risque qui peut émaner d'une mauvaise décision et pour augmenter le profit résultant d'une meilleure décision.

Notre travail de recherche intéresse au processus de découverte des connaissances à partir des bases de données KDD (KnowledgeDiscovery in Databases).

Le travail présent dans ce mémoire avait pour objectif d'incorporer la méthode KADS avec le processus KDD pour mettre à la disposition du décideur uniquement les connaissances dont il a besoin au bon moment.

Pour ce faire nous avons utilisé la méthode KADS (Knowledge Acquisition and Design Support) [Breuker 95] qui est issue de l'ingénierie de connaissance car elle propose un cadre de modélisation au niveau des connaissances. De plus elle est devenue un standard dans le développement de SBC, et la modélisation se fait grâce à la construction d'une succession de modèles. Ces modèles correspondent à plusieurs perspectives de la résolution du problème dans son environnement organisationnel et son contexte d'application.

Ainsi, il nous sera possible de situer notre objectif dans le contexte adéquat pour l'application, pour une modélisation optimale.

Pour cela nous avons proposé d'intégrer un modèle de spécification de besoins du décideur dans le processus KDD.

Le résultat de notre modèle sera comme Pré-sélection des données fiables et pertinentes dans le processus KDD.

#### **Perspectives :**

- Combiner la méthode KADS avec une autre méthode plus formelle, afin de rendre le développement des SBC plus sophistiqué et plus rigoureux.
- Ceci permettrait d'élargir le champs d'utilisation de la méthode KADS

# Références bibliographiques

## Bibliographies

[Afolabi& Thiery 05]	Afolabi B., Thiery O., Système d'intelligence économique et paramètres sur l'utilisateur : application à un entrepôt de publications, Journée sur les systèmes d'information élaborée, VSST.'05, Ile Rousse, 2005.
[Alquier 01]	Alquier A.M, Tignol M.H « Project management technique to estimate and manage risk of innovative projects », », IPMA'01 (International Project Management Association), Suède, 2001.
[Alquier 93]	Alquier A.M., « MUSIC – Management et Utilisation des Systèmes d'Information Coopératifs », Habilitation à diriger des recherches, Université Toulouse I, 1993.
[Andrassyová 02]	Eva Andrassyová and Ján Paralic. "Knowledge discovery in databases : A comparison of different views". 2002
[Angele 96]	Angele J., Fensel D. et Studer R., « Domain and task modelling in "MIKE", chapman & Hall, 1996.
[Anjewierden 87]	"The KADS system"; in proceedings of the first european workshop on knowledge acquisition for knowledge based systems; EKAW'87; Reading university; 1987
[Anton 96]	Goal based Requirements Analysis. Proceedings of the 2nd International Conference on Requirements Engineering ICRE'96, pp 136-144, 1996
[Anton 97]	Goal identification and Refinement in the Specification of Software-Based Information Systems. Phd Thesis presented to academic faculty, Georgia Institute of Technology, June 1997.
[Aussenac 05]	Aussenac-gilles Nathale, « Méthodes ascendantes pour l'ingénierie des connaissances », HDR Université Paul Sabatier Toulouse III, 1 <sup>er</sup> Décembre 2005
[Aussenac 89]	« Conception d'une méthode et d'un outil d'acquisition de connaissances expertes », Thèse de Doctorat en Informatique, Université Paul Sabatier, Toulouse, 1989.
[Aussenac 92]	N. Aussenac-Gilles, J.-P. Krivine et J. Sallantin. L'acquisition des connaissances pour les systèmes à base de connaissances. Editorial de la Revue Intelligence Artificielle, vol. 6(1-2), p. 7-18, 1992.
[Barabel 96]	" Un style de décision à la Française ? ", Revue Française de Gestion, novembre-décembre, pp. 159-170,1996.
[Bell et Thayer 76]	T.E. Bell, T.A. Thayer, Software Requirements: Are They Really a Problem ?. Proc.ICSE-2:2nd International Conference on Software Engineering, San Francisco, 1976, 61-68
[Berard 09]	Berard C., Le processus de décision dans les systèmes complexes : une analyse d'une intervention systémique. Thèse de Doctotat. Université du Quebec. 2009.
[Berry 97]	M. Berry and G. Linoff. Data Mining : techniques appliquées au marketing, à la vente et aux services clients. InterEditions, 1997.
[Bouaka 03]	Bouaka, N. : Développement d'un modèle pour l'explicitation d'un problème décisionnel: un outil d'aide à la décision dans le contexte d'intelligence économique. Thèse de doctorat. Labo Loria .2003.

[Bouaka et Al 02]	Bouaka, N., David A., Thiéry O., (2002) "Contribution to the understanding of explanatory factors for a decision-maker problem within the framework of economic intelligence, SCI'2002, Orlando, Florida, USA, 14-18 July 2002
[Boukrara & Chalal 13]	BOUKRARA A. et CHALAL R.: « specification of useful information for strategic decision support : risks based approach (RiBA) ». Int. J. Of Decision Sciences, Risk and Management, 2013.
[Bourne 97]	Bourne C., « Catégorisation et formalisation des connaissances industrielles ». In Connaissances et Savoir-faire en entreprise, ed.Hermes, p 179-197, 1997.
[Brachman 96]	J. Brachman and T. Anand. "The process of knowledge discovery in databases. Advances in knowledge discovery and data mining", pages 37-57, 1996.
[Breuker 95]	Breuker J., Wieling B., «KADS structured knowledge acquisition for expert system », 5th international workshop on expert systems and their application, Avignon, 1995.
[Breuker 94]	J. Breuker, W. Van de Velde, CommonKADS Library for Expertise Modelling – Reusable Problem Solving Components, IOS Press, Amsterdam, 1994
[Brunet 91]	E. Brunet, C. Toussaint et M. Toyé. Ascopa et Sadse: exemples d'application industrielle de la méthode KADS. In Avignon-91, Vol. 1, pp 393-406.
[Bubenko 94]	Bubenko J., Rolland C., Loucopoulos P., De Antonillis V., Facilitating 'fuzzy to formal' requirements modeling IEEE 1st Requirements Engineering, ICRE'94 pp. 154-158, 1994
[Carlsson 02]	Carlsson C. , Turban, E., « DSS: directions for the next decade », Decision Support Systems, Vol. 33 (2), pp.105-110., 2002
[Caroll 95]	The scenario Perspective on System Development In Scenario-based Design : Envisioning Work and Technology in System Development, Ed J.M Caroll, 1995
[Chalal 07]	Rachid CHALAL, « Une approche pour la capitalisation coopérative des connaissances sur les risques produit en phase initiale d'un projet industriel », thèse doctorat, ESI, Décembre 2007.
[Charlet 02]	Charlet J. «L'ingénierie des connaissances : Développements, Résultats et Perspectives pour la gestion des connaissances médicales », HDR de l'université Pierre et Marie Curie, Spécialité : Informatique, 2002.
[Cockburn 95]	Structuring use cases with goals. Technical report. Human and Technology, 7691 Dell Rd, Salt Lake City, UT 84121, HaT. TR.95.1, 1995.
[Corbel 97]	Corbel, J.C., « Méthodologie de retour d'expérience: démarche MEREX de Renault ». In « Connaissances et Savoir-faire en entreprise », ed. Hermes, p 93-110, 1997.
[Courbon 99]	Courbon J-C, TAJAN S, "Groupware et intranet Vers le partage des connaissances ", 2ème Edition, Editions Dunod, Paris 1999

[Crie 03]	Crie 01 CRIE D. «De l'extraction des connaissances au Knowledge Management » Lavoisier,Revue française de gestion /5 - n° 146 ISSN 0338-4551   pages 59 à 79. 2003.
[Dano 97]	Dano B., Briand H., Barbier F., A use case driven requirements engineering process. 3rd IEEE International symposium on Requirements Engineering RE'97, Antapolis, Maryland, IEEE. Computer Society Press, 1997
[Dardene 93]	amsweerde V.A., Dardenne A., Fickas S., Goal Directed Requirements Acquisition, and Science of Computer Programming. Vol. 20, 1993, 3-50
[Darimont 98]	Delor E., Masonnet P., Lamsweerde A.V., " GRAII/KAOS ": An Environment for Goal-Driven Requirements Engineering", Proc'98, - 20th Intl. Conf. on software Engineering. Kyoto April 98, Vol. 2, 58-62.
[David 01]	David A., Thiery O. « prise en compte du profil de l'utilisateur dans un système d'information stratégique », In veille stratégique scientifique et technique - VSST'2001.
[Davis 93]	Software requirements: objects, functions, and states. Prentice Hall, 1993
[Delotte 06]	Olivier DELOTTE : « CoCSys : une approche basée sur la construction d'un modèle comportemental pour la conception de systèmes collaboratifs mobiles », thèse de doctorat, Ecole Centrale de Lyon, 2006.
[Dieng 90]	Dieng R., «Méthodes pour la modélisation des connaissances » Rapport de Recherche N° 1624, INRIA, 1990.
[Dubois 98]	Dubois E., Yu E., Petit M., " From Early to Late Formal Requirements: A ProcessControl Case Study"Proc. IWSSD' 98, 9th IEEE CS Press, April 1998, 34-42
[Eichenbaum 94]	Eichenbaum-voline C., Malvache P., Prieur P. «La maitrise du retour d'experience avec la methode REX», Performances Humaines et techniques; Mars-avril, 1994.
[Ermine 96]	Ermine J. L. « les systemes de connaissances », Ed Organisation, France, 1996.
[Essame 02]	La méthode B et l'ingénierie système. Réponse à un appel d'offre. Technical report, IUT-Nantes, Université de Nantes, 2002
[European Software Institute 96]	European Software Institute, European User survey analysis.Report USV_EUR 2.1, ESPITI Project, 1996
[Fayyad 96]	U. M. Fayyad, G. Piatetsky-Shapiro & P. Smyth – « From data mining to knowledge discovery: an overview », in Advances in knowledge discovery and data mining (U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth & R. Uthurusamy, éd.), AAAI/MIT Press,1996, p. 1-34.
[Fredj 07]	FREDJ mounia : « composantes et modèles pour l'ingénierie des SI », thèse de doctorat, univ mohamed V, Maroc, 2007.
[Gam 08]	Inès GAM el golli : « Ingénierie des Exigences pour les Systèmes d'Information Décisionnels : Concepts, Modèles et Processus La méthode CADWE ». thèse de doctorat, univ paris 1, 2008.

[Ghomari 08]	Abdessamed Réda GHOMARI, « Approche Méthodologique d'Acquisition de Connaissances Agrégées à base d'Agents cognitifs coopérants pour les systèmes d'aide à la décision stratégique », THESE DOCTORAT, ESI, 2008.
[Ginting 00]	Rasmi GINTING : « intégration du système d'aide a la décision multicritères et du système d'intelligence économique dans l'ère concurrentielles : application dans le choix de partenaires en Indonésie » thèse de doctorat, univ Aix-Marseille, 2000.
[Gnaho & Semmak 10]	Christophe GNAHO & Farida SEMMAK : « Une extension SysML pour l'ingénierie des exigences dirigée par les buts », Actes du XXVIII° congrès INFORSID, Marseille, mai 2010.
[Greef 92]	P. De Greef, J. Breuker. Analysing system-user cooperation in KADS. In Knowledge Acquisition (1992) 4, pp 89-108
[Harel 87]	Statecharts: a Visual Formalism for Complex Systems. Sci. Computer Program, 8 pp. 231-274, 1987.
[Heisel 99]	A Method for Requirements Elicitation and Formal Specifcation. In LNCS Springer Verlag, editor, Proceedings of the 18th International Conference on Conceptua lModeling, number 1728, pages 309.324, November 1999.
[Huber 84]	Huber, G. " Issues in the design of group decision support systems ", MIS Quarterly, Vol. 8, No 3, pp. 195-204. 1984.
[IEEE 90]	. IEEE Std 610.12-1990 Standard Glossary of Software Engineering Terminology. IEEE, 1990.
[Inmon 00]	Inmon B., «Re-Engineering and the Data Warehouse», <a href="http://www.billinmon.com">www.billinmon.com</a> , consulté en Juin 2000.
[Jackson 95]	Software Requirements & Spécifications - a lexicon of practice, principles and prejudice s.ACM Press. Addison-Wesley. 1995
[Kahnemann 82]	Kahnemann, D. « Judgement under uncertainty: Heuristics and Biases », Cambridge University Press. 1982.
[Karbach 91]	W. Karbach, A. Voss, R. Schukey, U. Drouwen. Model-K: Prototyping at the knowledge level. In proceedings Expert Systems 1991, Avignon, France, pp 501-512.
[Keen 78]	Keen, P.G.W, Scott-Morton, M.S., « Decision support systems, an organizational perspective », Addison-Wesley, 1978.
[Kimball 00]	Kimball R. "Les entrepôts de données", Eyrolles, 2000.
[King 90]	King J.L,Star S.L □ Conceptual foundations for the development of Organizational Support Systems □, Proceedings of the Twentythird Annual Hawaii, International Conference on System Science, Los Alamitos, CA, IEEE Computer Society Press, 1990.
[Kodratoff 93]	Kodratoff, Y. Apprentissage de la Connaissance Experte par amélioration desexplications fournies par le système. APPRENTISSAGE SYMBOLIQUE : une approche de l'intelligence artificielle. Edité par Y. Kodratoff. Toulouse : CEPADUES EDITIONS, 1993. Tome II, p. 355-387.
[Kodratoff 97]	Y. Kodratoff – « L'extraction de connaissances à partir de données : un nouveau sujet pour la recherche scientifique », Revue Électronique sur l'Apprentissage par les Données 1 (1997), no. 1.



[Konaté 09]	Jacqueline KONATE : « Approche système pour la conception d'une méthodologie pour l'elicitation collaborative des exigences », thèse de doctorat, univ Toulouse, 2009.
[Konaté 09]	Konaté J., Approche système pour la conception d'une méthodologie pour l'elicitation collaborative des exigences. These de doctorat. Université Toulouse III - Paul Sabatier. Octobre 2009
[Lamsweerde 00]	Requirements Engineering in the year 00 : A research perspective In proceeding 22nd Int. Conf. on Software Engineering, Invited paper, ACM Press, June 2000
[Lamsweerde 01]	Lamsweerde A.V., Goal-Oriented Requirements Engineering. A Guided Tour. Proc. 5th IEEE International Symposium on Requirements Engineering, Toronto, August, 2001.
[Lamsweerde 96]	Lamsweerde V.A, Divergent Views in Goal-Driven Requirements Engineering. Proc. Workshop on Viewpoints in software Development, San Francisco, USA, October 1996.
[Lapouchnian 05]	Goal-Oriented Requirements Engineering: An Overview of the Current Research. University of Toronto. 2005.
[Lévine 89].	Lévine P., Pomerol M.J., « Systèmes interactifs d'aide à la décision et systèmes experts », Hermès, 1989.
[Mannila 94]	H. Mannila, H. Toivonen, and A. I. Verkamo. Efficient algorithms for discovering association rules. KDD'94, 181-192, Seattle, WA, July 1994.
[Mannila 97]	Heikki Mannila. "Methods and problems in data mining". In ICDT, pages 41–55, 1997.
[Maret 95]	Maret, P., « Modélisation et réutilisation des savoirs-faire. Application à l'activité de conseil aux organisations », Thèse de l'INSA, Lyon, 1995.
[Matoussi 11]	Abderrahman MATOUSSI : « Construction de spécifications formelles abstraites dirigée par les buts », thèse de doctorat, univ Paris-Est, 2011
[Mylopoulos 92]	Mylopoulos. J., Chung L., Nixon, B., Representing and Using Nonfunctional Requirements: A Process-Oriented Approach, IEEE Trans. On software. Engineering, Vol. 18 N° 6. June 1992.
[Mylopoulos 99]	Mylopoulos J., Chung L., Yu E., From Object-oriented to goal goal-oriented requirements analysis. Communication of the ACM. Vol. 42 N°1, 31-37, 1999.
[Nardi 92]	The use of scenarios in Design. SIGCHI Bulletin, 24
[Newell 82]	Newell A., « the knowledge level », Artificial intelligence, V18, PP 87-127, 1982.
[Pohl 98]	Haumer P., Pohl K., Weidenhaupt K., Requirements Elicitation and Validation with Real Wold Scenes. To appear in IEEE Transactions on Software Engineering, Vol. 24 N° 12, Special Issue on Scenario Management, 1998
[Potts 94]	Potts C., Takahashi K., Anton A.I., Inquiry-based requirements analysis in IEEE Software 11(2) pp. 21-32, 1994
[Rolland 03]	"Ingénierie des Besoins : L'Approche L'Ecritoire", Journal Techniques de l'Ingénieur

[Rolland 97]	Rolland C., Ben Achour C., Guiding the construction of textual use case specifications. Accepted to Data and Knowledge Engineering Journal, 1997.
[Rolland 99]	Rolland C., Prakash N., Benjamin A., A Multi -model View of Process Modelling. TheRequirements Engineering Journal, p. 169-187, 1999.
[Roy 85]	Roy B. " Méthodologies Multicritères d'aide à la décision ", Economica, 1985.
[Rubin 92]	Rubin K.S., Golberg A., Object Behavior Analysis, Communications of the ACM, 35(9) September 1992, pp 48-62.
[Saadoun 96].	Saadoun M., « Le Projet Groupware : des techniques de management au choix du logiciel groupware », Eyrolles, 1996.
[Schneider 96]	Schneider Daniel, 1996, Modélisation de la démarche du décideur politique dans la perspective de l'intelligence économique, Faculté des sciences économiques et social département de sciences politiques, Université de Genève, Genève
[Schreiber et al 94]	G. Schreiber, H. Akkermans, R. De Hoog, N. Shadbolt, W. Van de Velde et B. Wielinga. Knowledge Engineering and Management: The CommonKADS methodology, MIT Press, Cambridge, Massachusetts, London, 1994a.
[Sfez 92]	L. Sfez, Critique de la décision. Presses de la fondation nationale des sciences politiques, Paris, 1992.
[Simon 59]	Simon H.A. « Theories of Decision-Making in Economics and Behavioral Science » 1959.
[Simon 60]	Simon, H.A. "The New Science of Management Decision", New York: Harper and Row.1960.
[Simon 82]	Simon H. A. "The Science of Artificial ", Cambridge, The M.I.T. Press, 1982.
[Simon 97]	Simon G., «Modèles et méthodes pour conception des mémoires d'entreprise - le système DOLMEN, une application au domaine de la métallurgie», Thèse de l'université Henri Poincaré, Nancy I, 1997.
[Simoudis 96]	Evangelos Simoudis. "Reality check for data mining". IEEE Expert: Intelligent Systems and Their Applications, 11(5):26-33, 1996.
[Sprague 82]	SPRAGUE, R.H., CARLSON. E.D., « Building Effective Decision Support Systems », Prentice-Hall, 1982.
[Standish Group 99]	The Standish Group, Chaos. Standish Group Internal Report, <a href="http://www.standishgroup.com/chaos.html">http://www.standishgroup.com/chaos.html</a> , 1995.
[Studer 98]	Studer R., Benjamins V., Fensel D., Knowledge engineering: Principles and methods, In IEEE Transactions on Data and Knowledge Engineering, 25, 161-197, 1998.
[Tong 88]	X. Tong, Z. He, R. Yu. A survey of the expert system tool ZDEST-2. In Proceedings ECAI'88, Munich, pp 113-118, London: Pitman.
[Turban 01]	Turban E., « Decision support and intelligent systems », Prentice-Hall, 6 <sup>th</sup> Edition, 2001.
[Vogel 88]	Vogel C, « Génie Cognitif », Masson, 1988