

Vers un Méta-Modèle de Système d'Aide à la Décision: application au domaine médical

Ali Ayadi*, Salma Sassi*
Anis Tissaoui*

*FSJEG, Université de Jendouba, Tunisie
aliayadi1@gmail.com,sassisalma@yahoo.fr,tissaouianis@yahoo.fr

Résumé. Les systèmes d'aide à la décision connaissent un essor important en raison de leur capacité à supporter efficacement les analyses sur les données disponibles dans les organisations. Cependant, les systèmes actuels ne suivent pas les besoins d'analyses et ne s'intéressent qu'à l'analyse partielle des données (données numériques ou textuelles). Nous proposons dans cet article un Méta-Modèle de système intelligent d'aide à la décision qui peut être implémenté dans n'importe quel domaine d'activité, permettant une analyse qualitative et quantitative de tous les documents complexes rendant ainsi possible l'évolution des contextes d'analyses. Cette approche place l'agrégation des données et leurs représentations au centre du dispositif et remplace l'analyse unidimensionnelle de la donnée par une approche multidimensionnelle en se basant conjointement sur le processus d'entreposage (OLAP), l'ontologie et la théorie de l'aide à la décision.

Mots clefs: Système décisionnel, OLAP, entrepôt de métadonnées, agrégation des données, ontologie

1 Introduction

Un Système d'Aide à la Décision (SAD) est l'ensemble des outils informatiques (matériels et logiciels) qui permettent l'analyse des données opérationnelles issues du système d'information des organisations. Ces SADs sont présents dans de nombreux domaines et ont pour objectif d'aider le décideur dans sa tâche en lui fournissant tous les éléments pertinents pour la prise de décision. Cette dernière est devenue récemment un thème actif et important dans l'extraction et la gestion des connaissances.

Cependant, les modèles de décision traditionnels sont confrontés à d'énormes problèmes tels que l'hétérogénéité des données, leur analyse sémantique, leur représentation multidimensionnelle, le problème de personnalisation et celui de la visualisation. Pour résoudre ces problèmes, il serait intéressant de proposer une approche traitant tous les types de données, permettant leurs analyses sémantique et leurs représentation multidimensionnelle. Ces données répondant aux besoins du décideur, sont transformées en une vision orientée décideur puis présentées au moyen d'une interface graphique significative basée sur une technique métaphorique 3D.

Dans le domaine médical, les Systèmes d'Aide à la Décision Médicale (SADM) sont "des applications informatiques dont le but est de fournir aux cliniciens en temps et lieux utiles les informations décrivant la situation clinique d'un patient ainsi que les connaissances appropriées à cette situation, correctement filtrées et présentées afin d'améliorer la qualité des soins et la santé des patients"(Lobach et al., 2013).

Les documents médicaux sont le cœur des systèmes décisionnels dans la médecine. En effet, ces ressources ne cessent d'évoluer. Cette évolution croissante engendre, ainsi, des difficultés prévenant de trois problèmes : (i) les volumes des données médicales sont difficiles à manipuler, mais également elles sont de nature hétérogène. En effet, ces ressources sont difficilement stockées et réutilisées car elles n'ont pas été conçues à cet effet. (ii) Les outils et les méthodes actuels utilisés dans la représentation multidimensionnelle et l'analyse sémantique des données telle que la technologie OLAP (OnLine Analytical Processing), sont adaptés à la gestion des données numériques qui ne représentent que 20% des données d'un SI. Les données restantes sont hors de la portée du système d'analyse multidimensionnelle. Ceci mène sûrement à l'omission d'informations pertinentes, ce qui nuit par la suite à la prise de décision médicale vu l'absence de l'analyse sémantique. (iii) Les modèles actuels ne sont orientés qu'aux besoins d'analyse et ne concernent pas le contexte. En effet, ces modèles classiques ayant une flexibilité limitée, ne permettent pas l'évolution des contextes d'analyse et ne représentent pas les connaissances du domaine nécessaires aux décideurs.

C'est dans ce contexte que nos travaux de recherche visent à développer des infrastructures sémantiques de communications adaptées aux systèmes d'informations initialement hétérogènes mais aussi aux besoins des utilisateurs. Ces infrastructures forment un méta-modèle sémantique générique et instanciable qui permet de résoudre le problème d'hétérogénéité dans la présentation et dans la sémantique des données ainsi que la représentation multidimensionnelle des données. Ce Méta-Modèle d'Aide à la Décision (M²SAD) est basé sur l'approche des entrepôts de données, de la technologie OLAP ainsi que l'ontologie. On a développé ces thèmes en les illustrant sur un terrain d'applications qui présente de nombreux facteurs de complexité : les systèmes d'information dans le domaine médical.

Dans cet article, nous présenterons un court état de l'art sur quelques travaux liés aux problèmes déjà cités ainsi que les solutions existantes. Nous détaillerons ensuite notre méthodologie M²SAD proposée. Puis nous montrerons quelques aspects d'expérimentations menées. Un résumé des apports de notre contribution ainsi que les travaux futurs concluront cet article.

2 Les systèmes d'aide à la décision et le domaine médical : un bref état de l'art

Depuis plusieurs années, les travaux de recherche sur le thème des systèmes d'aide à la décision (SAD) se sont multipliés. En effet, ils visent à étudier et améliorer l'ensemble des techniques permettant, pour une personne donnée, d'opter pour la meilleure prise de décision possible. Ces SADs représentent un énorme potentiel en médecine. Par conséquent, nous présentons dans cette section les principaux travaux effectués dans ce domaine.

CAMD (Conforti, 1999) permet le traitement d'images basées sur la morphométrie cellulaire. Ce SADM est composé d'un classificateur automatique basé sur un outil de programmation mathématique. La phase de classification constitue le cœur du processus de son diagnostic

médical. Ce système aide le médecin à distinguer entre les cellules bénignes et malignes pour la détection prématurée du cancer du sein. Ce SADM n'a l'avantage que de classer les images d'une mammographie et d'y gérer la maladie infectée statiquement. Il est spécifique au domaine de cancer du sein, ne gère que des données structurées. Ces données sont analysées mathématiquement et affichées sous forme de texte.

SAPHIRE (Laleci et al., 2008) est un système basé sur la notion des systèmes multi agents. Son principal but est de soutenir la définition, le déploiement et l'exécution des directives cliniques à un patient. Il est, en fait, composé d'un ensemble de composants (agents) collaborateurs s'exécutant dans un environnement distribué hétérogène. SAPHIRE consiste à extraire les données cliniques des dossiers de santé électroniques du patient et effectuer des analyses statistiques afin d'aboutir à une décision. Ce système traite uniquement les données structurées en leur analysant statistiquement. Il n'est pas spécifique à un domaine bien précis. SAPHIRE représente les données d'une manière unidimensionnelle et les expose au médecin sous forme textuelle.

IHMDS (Chang et Lu, 2009) qui est basé sur la notion des services web en collaboration avec les arbres de décision. Ce système, basé sur le langage XML, permet la manipulation des données structurées. L'avantage de ce système réside dans le fait qu'il incorpore l'interopérabilité et l'extensibilité de XML. Ce système permet au médecin de faire des diagnostics sur les symptômes et de générer uniquement les maladies correspondant à la thyroïde. En effet ce système effectue des analyses sémantiques d'une manière unidimensionnelle, les affichant sous forme de texte et de formulaire.

CNS (Couturier et al., 2010) est également basé sur la notion de systèmes multi agents, favorisant un meilleur diagnostic médical du patient. En effet, CNS vérifie plusieurs conditions qui pourraient être ignorées par les humains grâce à son analyse sémantique et statistique leur permettant d'améliorer les décisions pouvant être prise par des médecins. Ce système de diagnostic médical coopératif ne s'intéresse qu'aux domaines de l'endocrinologie. Il ne permet pas de traiter les données complexes non structurées, ni la représentation multidimensionnelle. Il utilise l'affichage sous forme textuelle et de formulaire.

D'autres travaux s'avèrent aussi intéressants dans ce domaine à savoir : le système FCM (Peter, 2012) utilise l'approche de cartes cognitives floues comme méthode de modélisation informatique des données médicales. En effet, ce système combine le processus d'aide à la décision médicale et la représentation de ces connaissances sur une carte cognitive. Il ne traite que les données médicales structurées en langage XML. FCM est générique dans le domaine médical, il affiche ses connaissances sous forme textuelle.

Après ce survol de l'existant, nous élaborons une étude comparative des différents travaux déjà décrits. Pour ce faire, nous dressons un tableau dans lequel des critères de comparaison sont pris en compte. Ces critères sont : (i) *la spécialité du SAD* est le secteur couvert par son activité. On distingue deux catégories : le SAD générique (gén) qui concerne tous les secteurs du domaine et le SAD spécifique (spé) qui concerne un secteur bien précis du domaine tel que le secteur cardiologique du domaine médical. (ii) *la nature des données* (données), deux types de données sont pris en compte. Les données structurées (st) disposées de façon à être traitées automatiquement par le système, mais non nécessairement par un humain. Les données non structurées (nn) qui s'adressent à l'humain et dont il n'est pas possible de prédéfinir leur structure. (iii) *le type d'analyse*, soit une analyse statistique (sta) sur des données simples, soit une analyse sémantique (sém) sur des données textuelles complexes en prenant en compte ainsi

leurs contexte.(iv) *la représentation des données*, unidimensionnelle (uni) où le sujet à étudier est analysé en une seule dimension, multidimensionnelle (mult) qui consiste à analyser un sujet selon plusieurs dimensions (axes d'analyses).(v) *la technique de visualisation* qui montre la manière dont les informations sont affichées à l'utilisateur. On distingue la visualisation textuelle (tex), la visualisation formulaire (form) et enfin la visualisation graphique (graph).

	Spécificité		Données		Analyse		Resprésentation		Visualisation		
	Spé.	Gén.	St.	Nn	Sta.	Sém.	Uni.	Mult.	Tex.	Form.	Graph.
CAMD	*		*		*		*		*		
SAPHIRE		*	*		*		*		*		
IHMDS	*		*			*	*		*	*	
CNS		*	*		*	*	*		*	*	
FCM	*		*			*	*		*		

TAB. 1 – *Tableau Comparatif des SADMs.*

Il est également évident que ces systèmes se distinguent par la différence entre eux en certains points et se ressemblent dans d'autres.

De la perspective de critique, nous arrivons à relever les constatations suivantes :

- La plupart des SADMs sont spécifiques à des sous-domaines bien précis.
- Ces SADMs étudiés manipulent essentiellement des données très structurées.
- Les analyses des données sont généralement statistiques.
- L'analyse d'un sujet n'est faite que par rapport à un seul et unique axe d'analyse.
- Les traitements sont visualisés soit sous forme textuelle soit sous forme de formulaire.
- Aucun système ne permet de :
 - Traiter toutes les spécialités d'un domaine spécifique.
 - Traiter des données structurées et non structurées.
 - Analyser statistiquement et Sémantiquement les données médicales.
 - Représenter les sujets d'analyse d'une façon multidimensionnelle.
 - Visualiser les résultats d'analyse sous plusieurs formes de visualisation selon le besoin de l'utilisateur : forme textuelle, formulaire et graphique.

En prenant en compte ces points, nous allons proposer un nouveau SADM sûr et fiable répondant à tous ces critères.

3 Méthodologie M²SAD : vers la prise en compte de la spécialisation des domaines

Dans le contexte de la prise de décision, il semble pertinent de prendre en compte la notion de la spécialité dans le domaine. Dans ce contexte, nous présentons le système M²SAD qui apparaît donc comme un méta-modèle à partir duquel seront définies des primitives permettant l'analyse sémantique des données, leurs représentations multidimensionnelles, leurs personnalisations et leurs visualisations sous différentes formes. Ces primitives sont adaptées à un domaine et à un type d'activité particulière.

Notre méta-modèle est conçu dans le but de traiter des données hétérogènes provenant de différentes sources et d'être utile pour des profils d'utilisateurs aussi bien différents. Pour ce

faire, une architecture à trois niveaux est définie. La figure 1, propose une architecture à trois niveaux de construction d'une structure partagée et collective pour guider l'utilisateur dans sa prise de décision prenant en compte la notion de spécialisation dans le domaine étudié.

- Le niveau 1 constitué par le méta-modèle qui définit les primitives du système à savoir : l'entrepôt des données, l'entrepôt des règles, l'ontologie et la technologie OLAP.
- Le niveau 2 constitué par l'instanciation des primitives déjà citées permettant ainsi de structurer un secteur d'activité bien déterminé et représenter les différents objets professionnels du domaine sous une forme unifiée et prenant en compte les différentes spécialités de ce domaine. Par exemple, dans le domaine médical ce niveau servira à définir un entrepôt de métadonnées, un entrepôt de règles, une ontologie de structure médicale et des primitives d'OLAP adaptées au domaine médical. Selon sa spécialité, le professionnel de santé crée les objets médicaux qui correspondent à son statut.
- Le niveau 3 est l'instanciation des éléments du niveau 2 qui permet de construire ; effectivement ; le système d'aide à la décision de l'entreprise.

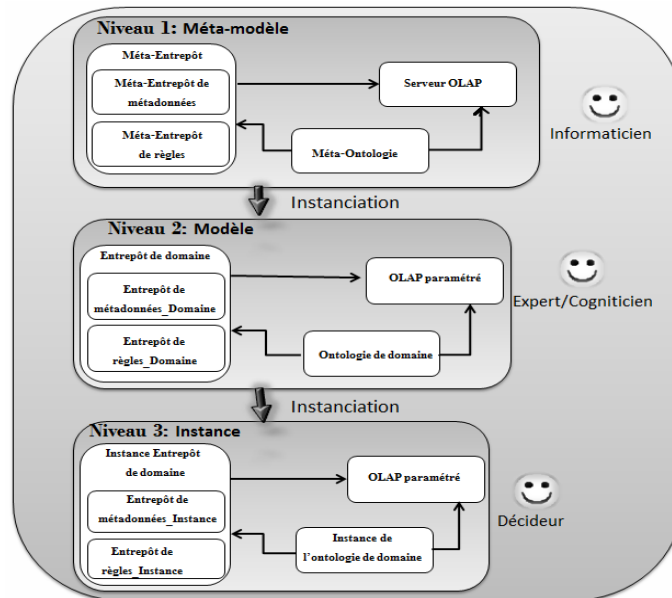


FIG. 1 – Méta-modélisation de M²SAD.

Le niveau méta-modèle est constitué par le méta-entrepôt qui contient le méta-entrepôt de métadonnées et le méta-entrepôt de règles, la méta-ontologie et le serveur OLAP. Le méta-modèle définit la description de la structure et de la sémantique de documents hétérogènes. A ce niveau, c'est l'informaticien qui intervient pour créer les concepts du méta-modèle de SAD. Les composants sont génériques pour les SADs de n'importe quel secteur d'activité. Ils serviront de patrons de modélisation de ces modèles de SAD.

Le niveau modèle définit le modèle de SAD, issu de l'instanciation de Méta-modèle, pour un domaine d'activité particulier. Le modèle de données définit l'entrepôt de métadonnées,

Vers un Méta-Modèle de système d'Aide à la Décision

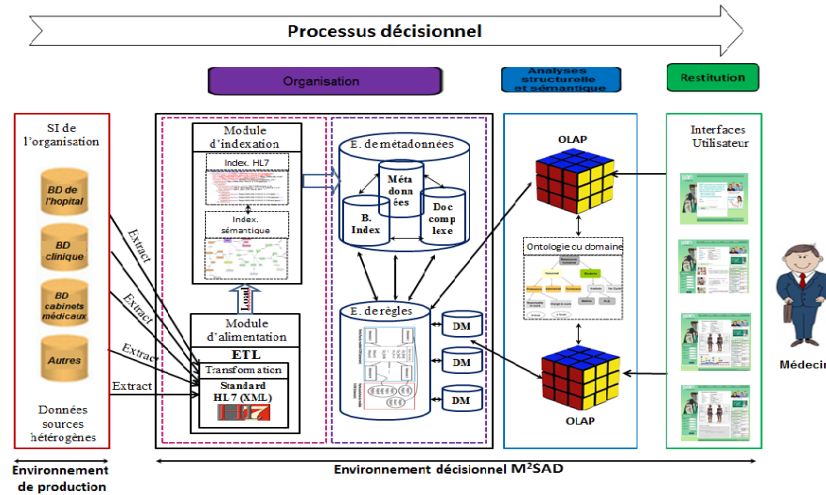


FIG. 2 – Architecture de M²SAD.

l'entrepôt de règles, l'ontologie de domaine du secteur d'activité et le serveur OLAP paramétré. Ici, c'est l'expert du domaine qui intervient et qui instancie le méta-modèle pour créer un modèle de SAD du domaine. L'instanciation correspond à limiter le méta-modèle de données aux concepts du méta-modèle de SAD autorisés par la loi du domaine d'activité.

Le niveau instance correspond à l'instanciation, par l'utilisateur final, des éléments du modèle de SAD décrits dans le niveau modèle de SAD. Cette instanciation lui permet de manipuler graphiquement, sur une interface dédiée, les différentes fonctionnalités du SAD selon ses droits et son profil. En effet, un médecin généraliste n'aura pas les mêmes fonctionnalités qu'un radiologue.

3.1 Architecture de M²SAD

L'architecture M²SAD est basée sur une organisation à trois phases, chacune d'elles étant basée sur un module de données particulier. Par exemple, le module de l'entrepôt gère l'évolution et l'indexation des données complexes alors que le module OLAP organise les données de manière multidimensionnelle pour optimiser l'analyse sémantique. La figure 2, montre de façon schématique l'architecture du Méta-Modèle de SAD proposé. Nous allons présenter, dans cette section, ces différentes phases une par une.

3.1.1 La phase d'organisation : une fusion d'entrepôt de métadonnées et d'entrepôt de règles

Cette phase a pour objectif d'assurer l'interopérabilité sémantique des données. En fait, les données provenant des sources de données hétérogènes sont restructurées, annotées, décrites, unifiées et contextualisées selon les standards de communication d'un domaine bien défini (dans le domaine médical, HL7 est utilisé pour standardiser les documents médicaux) afin de

résoudre le problème d'hétérogénéité. Cette phase permet également à travers l'entrepôt de métadonnées, de stocker de grandes masses de données et de les analyser.

Cette phase est composée de trois modules :

- Le module d'alimentation qui permet de récupérer les informations hétérogènes et pertinentes de diverses sources de données. Ce mécanisme prend en compte deux sortes d'hétérogénéité : structurelle et sémantique. L'hétérogénéité structurelle des sources se matérialise par des documents plus ou moins structurés et reposant ou non sur des standards. L'hétérogénéité sémantique est liée au fait que les documents peuvent concerner des contextes très divers (sujets, domaines, etc.). Ce module fournit des documents hétérogènes fortement structurés (standardisés par XML) en passant par deux étapes : une première étape d'extraction de structure et de contenu et une deuxième de standardisation.
- Le module d'indexation qui permet d'associer à un document un ensemble d'informations le caractérisant. Ces informations peuvent être des mots clés ou des Meta-informations. L'objectif du processus d'indexation consiste à décomposer les documents en entités informationnelles afin de les interroger et de les manipuler. Ce module se base sur trois étapes : la première étant l'étape de décomposition sert à segmenter le document en divers entités documentaires (objets ou granules). La deuxième étape est l'indexation, nous proposons à ce niveau, une indexation structurelle et sémantique. Finalement L'insertion du contenu qui consiste à insérer le contenu du document dans la base multidimensionnelle en rattachant chaque granule à l'élément de structure correspondant ainsi qu'aux métadonnées lui correspondant.
- Le module entreposage : le M²SAD repose sur deux méta-entrepôts, le premier un méta-entrepôt de métadonnées qui rend facilement accessibles, exploitables et réutilisables les données. Le deuxième méta-entrepôt étant le méta-entrepôt de règles qui rend l'analyse des données plus flexible et offre ainsi au décideur la possibilité de définir ses propres règles pour déterminer de nouvelles hiérarchies de dimensions :

(a) Méta-entrepôt de métadonnées : Un entrepôt de métadonnées est une collection de données : intégrée, orientée sujet, non volatile, historisée, résumées et disponibles pour l'interrogation et l'analyse (Inmon, 2002). L'entrepôt intègre les données pertinentes, nécessaires à la prise de décision des systèmes d'informations de l'organisation. L'objectif de ce méta-entrepôt est de stocker toutes les informations issues des documents sources (contenus, structures et métadonnées) pour effectuer des analyses décisionnelles de tous types : recherche d'information, interrogation, analyse multidimensionnelle sémantique. L'entrepôt de métadonnées repose sur une base de données contenant des métadonnées, une autre contenant des hyperliens entre les objets des documents et des liens inter-documents, et enfin une base multidimensionnelle contenant les documents complexes ainsi que leurs différentes entités. Cet entrepôt propose une richesse sémantique grâce à la méta-ontologie.

(b) Méta-entrepôt de règles : Un critère essentiel qui distingue les SADs est la flexibilité. Bien évidemment les modules classiques apportent une certaine flexibilité temporelle au modèle par la mise à jour du schéma apportant une réponse à l'évolution des données. Cependant, ils n'apportent pas de solution à l'émergence de nouveaux besoins d'analyse qui sont orientés, non pas par l'évolution des données, mais par l'expression de nouvelles connaissances. L'en-

trepôt de règles pourra ainsi contribuer à la résolution de la problématique et rendre l'analyse plus flexible. En effet, il permet de créer des hiérarchies de dimension d'une façon dynamique en renforçant l'interaction entre l'utilisateur et le SAD. Cette génération de hiérarchie à la demande permet de faire évoluer les contextes d'analyses.

Le méta-entrepôt de règles est composé de deux parties : une fixe définie en extension, comprenant une table de fait et des dimensions du premier niveau. Une partie évolutive, définie en extension par des règles de type "Si Alors", qui déterminent les niveaux de granularité dans les hiérarchies de dimensions basées sur les dimensions existantes et de nouvelles connaissances de l'utilisateur. Les règles permettent d'établir des liens sémantiques entre les données (Favre et al., 2006). Ce méta-entrepôt garantit l'évolution, la flexibilité, la performance et l'optimisation des analyses.

3.1.2 La phase d'analyse : modélisation multidimensionnelle, sémantique de données complexes

Le modèle de données multidimensionnel OLAP est le responsable de l'organisation des données dans le système. Un modèle multidimensionnel contient les informations pertinentes et les présente dans un format plus approprié pour faire leurs analyses. OLAP a pour but d'organiser les données à analyser par domaine ou thème et d'en ressortir des résultats pertinents pour le décideur. Pour garantir ce processus d'interrogation, nous avons besoin d'outils performants et conviviaux pour l'accès et l'analyse de l'information pertinente. Les systèmes OLAP fournissent des méthodes et des outils puissants permettant l'analyse de données transactionnelles (Sullivan, 2001) et (Tseng et Chou, 2006). Cependant, seul 20% des données d'un système d'information sont des données transactionnelles et peuvent être traitées par OLAP (?). Les 80% restant des données représentent les documents électroniques (Sullivan, 2001) et (Tseng et Chou, 2006).

Par exemple, dans le domaine médical, les objets médicaux sont généralement annotés, c'est-à-dire qu'ils sont principalement constitués de texte, d'images, de sons, de vidéo, etc. Ces documents restent hors de la portée des SADs faute de manque d'outils de traitement et d'analyse.

Afin de remédier à cette problématique, nous visons à ajouter des fonctions d'agrégation textuelles. En effet, l'agrégation des données textuelles permet de résumer le volume de données à visualiser lors d'une même analyse. Ainsi, en réduisant le volume de données par les méthodes de synthèse, le décideur peut avoir une vision globale du domaine qu'il analyse d'où une analyse sémantique des données. Les fonctions textuelles utilisées sont détaillées dans (Ravat et al., 2008) et (Tseng et Chou, 2006) :

- TopKeyword : fonction qui extrait les n mots-clefs jugés les plus pertinents d'un fragment de texte.
- Topic : fonction qui extrait le sujet d'un fragment de texte.
- SUMMARY : génère un résumé automatique d'un fragment de texte.

Pour créer ces fonctions, il est nécessaire de disposer d'un moyen permettant de calculer et de trouver les mots-clefs. La méta-ontologie de domaine est employée à cette fin. Bien évidemment, les indicateurs sémantiques sont utiles pour la prise de décision, car ils résument l'information et permettent d'appréhender des systèmes complexes. Nous avons intégré l'ontologie vue que de plus en plus les décideurs souhaitent pouvoir tenir compte aussi des contextes sémantiques et des connaissances lors de leurs prises de décision (l'analyse numérique seule-

ment n'est plus suffisante). Pour cette raison, il serait indispensable d'intégrer un mécanisme performant en représentation des connaissances et permettant un raisonnement sémantique qui serait la méta-ontologie. Outre, cette ontologie de domaine permettrait l'interaction des "experts" qui seront responsables d'enrichir le processus de prise de décision tel qu'ils le perçoivent par leurs connaissances. De plus, l'ontologie permet à l'entrepôt de métadonnées de révéler tout un vocabulaire lié à l'utilisation et l'organisation de la base décisionnelle documentaire des SADs, d'où la garantie de l'aspect sémantique et la consommation en temps pour les interrogations de l'utilisateur. La connaissance associée à un domaine peut être représentée de façon plus formelle au travers d'une ontologie. Pour un utilisateur, accéder à la connaissance, grâce à une ontologie, peut lui permettre de spécifier son besoin et les lacunes de sa connaissance par rapport à l'information qui lui est disponible. D'autre part, la représentation des granules d'information à partir d'une ontologie peut définir un vocabulaire contrôlé (termes et concepts) à partir duquel l'utilisateur spécifiera son besoin. La description du besoin correspond, dans ce cas-là, aux caractéristiques des granules car elles ont été indexées à partir des mêmes ressources.

3.1.3 La phase de restitution : l'interface de visualisation

Cette interface permet de mieux appréhender le résultat de l'analyse. L'utilisateur final, n'étant pas forcément un informaticien, aura plus de facilités dans des interfaces simples, basées sur la visualisation et la perception de l'œil du décideur que d'aller requêter directement dans le serveur d'analyse. Pour atteindre cet objectif, nous proposons une interface graphique utilisant des diagrammes, des courbes statiques, une visualisation métaphorique et une technique de visualisation 3D.

4 Prototype et évaluation

Afin de valider notre approche, nous avons développé un prototype d'aide à la décision dans le domaine médical. Cet outil facilite la tâche du médecin dans sa prise de décision. Il est composé essentiellement d'une interface utilisateur et d'une base décisionnelle multidimensionnelle, sémantique.

Nous détaillerons ici quelques interfaces utilisateurs (voir figure 3) graphiques permettant au médecin de (i) consulter les données hétérogènes historisées du patient, (ii) effectuer des analyses sur un sujet bien déterminé selon diverses dimensions. (iii) répondre aux requêtes du médecin de manière analytique et sémantique. (iv) effectuer les diverses analyses pour obtenir un diagnostic lié au choix du symptôme et des résultats d'analyses. L'application se charge ensuite de répondre au médecin par la pathologie associée aux symptômes et aux analyses ainsi que la présentation des traitements nécessaires. La figure 3, présente des interfaces utilisateurs représentant respectivement les principales fonctionnalités du SAD : (a) consultation du dossier médical du patient, (b) aide à la décision, (c) aide au diagnostic et (d) aide à la prescription médicamenteuse. Elles montrent aussi les divers objectifs de notre méthodologie telles que la gestion des données structurées ou non (exploitation textuelle des comptes-rendus, exploitation de vidéo d'une échographie, etc.) ainsi que leur analyse statistique et sémantique (calcul de glycémie, exploitation de connaissances, etc.).

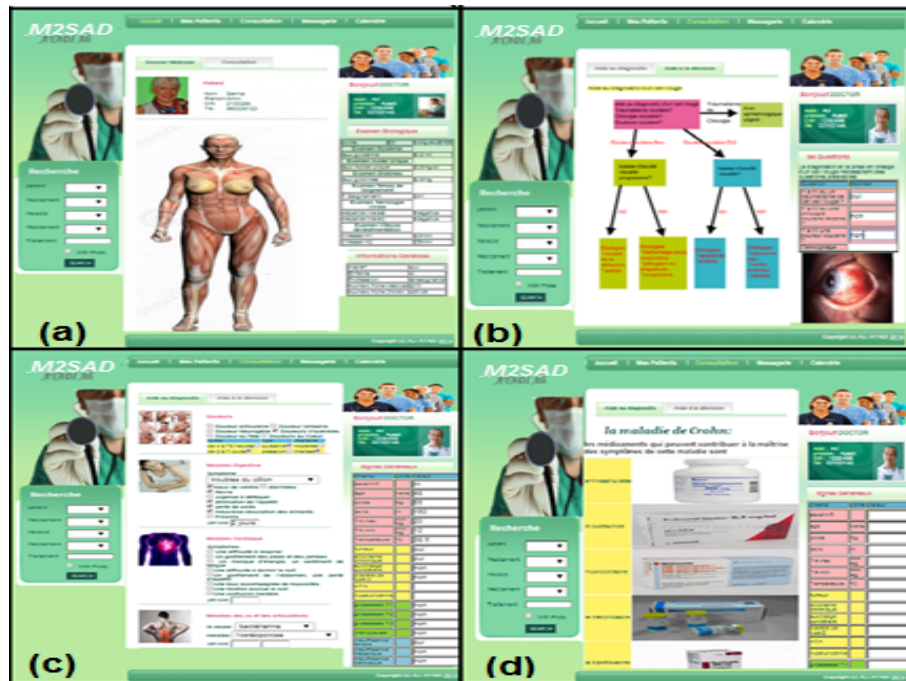


FIG. 3 – Les principales fonctionnalités du prototype.

Pour évaluer notre SADM, nous avons décidé de confronter les professionnels de santé aux fonctionnalités de ce prototype pour la pratique de prise de décision clinique et de tirer des premières recommandations sur l'utilisabilité de ses fonctionnalités. Dans cette évaluation, nous avons utilisé un corpus qui contient 23 dossiers médicaux de patient de l'hôpital régional de Jendouba (Tunisie). Nous avons mené une évaluation avec 5 cliniciens. Une étape préliminaire de l'évaluation a permis de présenter aux participants notre système. L'objectif de cette présentation est de fournir à tous les participants les connaissances nécessaires pour porter des jugements de l'utilité des services proposés par le SADM, des objectifs et de la nature concrète de notre recherche. Cette étape a permis aussi de familiariser les participants avec notre système ainsi qu'avec les notions utilisées lors de l'évaluation, afin de leur permettre de se concentrer davantage sur le test de l'utilité de notre SADM. Suite à cette démonstration, nous avons demandé à chaque participant de répondre à trois questions cliniques afin de tester sa compréhension et sa maîtrise des différentes fonctionnalités du système. Lorsque le participant pourrait répondre correctement aux questions, il a été considéré comme prêt à effectuer l'évaluation.

Cette même enquête qualitative est menée pour obtenir l'appréciation donnée par les praticiens sur l'impact perçu de notre SADM positif ou négatif par rapport à d'autres SADMs utilisés actuellement. Les métriques d'évaluation des SADMs étudiés concernent à la fois la validité des connaissances, la rapidité de la décision, la validité des solutions proposées et leurs impacts sur la qualité des soins, la satisfaction sur l'utilisateur ainsi que les technologies

déployées.

Les résultats ont démontré la faisabilité du système proposé par rapport aux autres existants. Les résultats des évaluations obtenus sont plutôt encourageants. Les cliniciens ont trouvé toutes les fonctionnalités utiles et l'interface très facile à manipuler, ils ont jugé également performant cet outil décisionnel en termes de temps de réponse, d'interopérabilité et surtout de qualité de soin. Bien que 70% des cliniciens disent ne pas avoir l'habitude de ce type d'outils, ils ont trouvé l'utilisation générale du système plutôt très simple à 90%. Seuls 10% des cliniciens ont dit avoir dû vraiment chercher des fonctionnalités (les autres les ont trouvées plus facile) et seuls 20% des cliniciens ont avoué avoir commis des erreurs de manipulation. La consultation, l'analyse, la manipulation et la visualisation des données dans son ensemble semble donc être plutôt très facile à appréhender. Ce qui était le plus encourageant, c'est que les cliniciens ont tous affirmé que ce prototype les aide énormément à élaborer des plans de soins corrects et à communiquer le plus facilement avec les autres professionnels de santé à travers les fonctionnalités du SADM.

5 Conclusion et perspectives

Dans cet article, nous avons proposé un nouveau méta-modèle générique pour l'aide à la décision. Ce méta modèle permet de trouver une syntaxe commune et partageable entre les utilisateurs du système pour décrire d'une façon identique les documents de structure différente. Ce méta-modèle assure également leur analyse structurelle et sémantique, leur représentation multidimensionnelle et leur visualisation sous différentes formes. Ce méta-modèle se base principalement sur un *méta-entrepôt de métadonnées* stockant de grosses masses de données et de différentes entités composant les documents hétérogènes grâce à la base d'index comportant les liens joignant ces entités à leurs métadonnées. Le *méta-entrepôt de règles* assure la personnalisation du SAD, procure plus d'agrégation de données et permet une analyse encore plus détaillée grâce à la multitude de dimension.

Finalement, la *base multidimensionnelle OLAP* couplée à la méta-ontologie de domaine permet d'analyser sémantiquement, contextuellement et multidimensionnellement les données décisionnelles. Enfin, nous avons validé ce méta-modèle en implémentant et évaluant un système d'aide à la Décision appliqué au domaine médical.

Bien que les résultats obtenus soient encourageants, plusieurs perspectives de travaux futurs peuvent être envisagées. Il serait intéressant de développer un formalisme qui pourra faciliter la tâche du concepteur. Nous envisageons également d'améliorer la complexité de notre approche et d'introduire d'autres mécanismes tels que le datamining qui permettra de ré-extraire et enrichir les connaissances des utilisateurs.

Ainsi, nous trouvons intéressant de développer et d'expanser les domaines d'application de notre système, par exemple l'adaptation de ce dernier dans le domaine économique, agricole, politique, etc.

Références

Chang, C. et H.-M. Lu (2009). Integration of heterogeneous medical decision support systems based on web services. In *Bioinformatics and BioEngineering, 2009. BIBE '09. Ninth IEEE*

- International Conference on*, pp. 415–422.
- Conforti, M. (1999). Decision support systems for medical diagnosis. *Information Technology Applications in Biomedicine, 1999. ITIS-ITAB '99. 1999 IEEE EMBS International Conference on*, 25–26.
- Couturier, V., M.-P. Huget, et D. Télisson (2010). Engineering agent-based information systems - a case study of automatic contract net systems. pp. 242–248. SciTePress.
- Favre, C., F. Bentayeb, et O. Boussaid (2006). A rule-based data warehouse model. *4042*, 274–277.
- Inmon, W. H. (2002). *Building the Data Warehouse, 3rd Edition*. New York, NY, USA: John Wiley & Sons, Inc.
- Laleci, G. B., A. Dogac, M. Olduz, I. Tasyurt, M. Yuksel, et A. Okcan (2008). Sapphire: A multi-agent system for remote healthcare monitoring through computerized clinical guidelines agent technology and e-health. pp. 25–44.
- Lobach, D. F., K. Kawamoto, K. J. Anstrom, G. M. Silvey, J. M. Willis, F. S. Johnson, R. Edwards, J. Simo, P. Phillips, D. R. Crosslin, et E. L. Eisenstein (2013). A randomized trial of population-based clinical decision support to manage health and resource use for medicaid beneficiaries. *J. Medical Systems* 37.
- Peter, P. (2012). The challenge of modeling decision support systems for medical problems using fuzzy cognitive maps: An overview. In *BIBE*, pp. 132–138.
- Ravat, F., O. Teste, R. Tournier, et G. Zurfluh (2008). Top_keyword: an aggregation function for textual document olap. In *Data Warehousing and Knowledge Discovery*, pp. 55–64. Springer Berlin Heidelberg.
- Sullivan, D. (2001). *Document Warehousing and Text Mining: Techniques for Improving Business Operations, Marketing, and Sales*. John Wiley & Sons.
- Tseng, F. S. et A. Y. Chou (2006). The concept of document warehousing for multidimensional modeling of textual-based business intelligence. *Decision Support Systems* 42(2), 727 – 744.

Summary

The decision support systems know an important upheaval due to their capacity to support effectively the data analysis available in the organization. However, the actual systems don't follow the organization's analysis needs evolution and are only interested in the partial data analysis (digital or textual data). We propose in this paper, a new metamodel of decision support system generic which can be implemented in any field of activity, allowing quantitative and qualitative analysis for all complex documents, makes this possible the contexts of analysis evolution. This approach puts the aggregating data and their visualization in the center of the system. It replaces the one-dimensional approach by a new multidimensional one, which is based on datawarehouse process, ontology and the decision support theory.

Keywords: *decision-making system, OLAP, data warehouse, data aggregation, ontology*