



HAL
open science

Modélisation et structuration des connaissances dans les processus de télémédecine dédiés aéronautique

Alsane Sene

► To cite this version:

Alsane Sene. Modélisation et structuration des connaissances dans les processus de télémédecine dédiés aéronautique. Modélisation et simulation. Université Paul Sabatier - Toulouse III, 2018. Français. NNT : 2018TOU30251 . tel-02333755

HAL Id: tel-02333755

<https://theses.hal.science/tel-02333755>

Submitted on 25 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

En vue de l'obtention du **DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Alsane SENE

Le 17 décembre 2018

**Modélisation et structuration des connaissances dans les
processus de télémédecine dédiés aéronautique.**

Ecole doctorale : **AA - Aéronautique, Astronautique**

Spécialité : **Informatique**

Unité de recherche :
LGP - Laboratoire Génie de Production

Thèse dirigée par
Bernard KAMSU-FOGUEM et Pierre RUMEAU

Jury

Mme Sandra BRINGAY, Rapporteur
Mme Céline ROBARDET, Rapporteur
Mme Gaele CALVARY, Examineur
Mme Sylvie RANWEZ, Examineur
M. Bernard KAMSU-FOGUEM, Directeur de thèse
M. Pierre RUMEAU, Co-directeur de thèse

À mes Parents, Frères et Soeurs,

à Barth et Rébecca,

à ma deuxième famille à Thiès, Pa Alou, Tante Fama, Soua, Julo, Codou, Pa Cissé, Maman Ndèye Fatim,

à Jean Marc et Armelle,

aux proches que le destin a arrachés à la vie pendant que j'étais loin du pays :

Pape Aliou,

Thiathi Makane,

Yamak Dior,

Yateb Thioro,

je dédie ce travail.

Remerciements

J'aimerais d'abord remercier mon directeur de thèse, M. Bernard KAMSU-FOGUEM. Je lui suis reconnaissant pour le temps conséquent qu'il m'a consacré, ses qualités pédagogiques et scientifiques. J'ai appris à ses côtés à être plus autonome et je lui adresse ma gratitude la plus chaleureuse.

Je remercie aussi chaleureusement mon co-encadrant de thèse, M. Pierre RUMEAU, Docteurs ès Sciences (modélisation), Médecin des Hôpitaux, Responsable du Laboratoire de Gérontechnologie La Grave CHU de Toulouse, pour m'avoir ouvert les portes de son laboratoire, pour ses conseils avisés dans le domaine médical.

Je voudrais remercier les rapporteurs de cette thèse Madame Céline ROBARDET, Professeur des Universités, coordinatrice de l'équipe Data Mining et Machine Learning du laboratoire d'Informatique en Image et Systèmes d'information de L'INSA de Lyon, et Madame Sandra BRINGAY, Professeur des Universités, membre du Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier. La version actuelle de la thèse doit beaucoup à leurs remarques.

J'adresse de sincères remerciements aux examinateurs de cette thèse Madame Sylvie RANWEZ, Professeur de l'Institut Mines-Télécom, responsable du département d'enseignement « Informatique et Intelligence Artificielle » de l'Ecole des mines d'Alès, et Madame Gaëlle CALVARY, Professeur et Vice-Présidente Valorisation de Grenoble INP (Institut National Polytechnique) et directrice adjointe du Laboratoire d'Informatique de Grenoble, pour l'intérêt qu'elles ont porté à mon travail.

J'adresse aussi mes remerciements à la direction de l'école doctorale aéronautique astronautique de Toulouse, pour avoir par leur politique volontariste sur les travaux de recherche bi-disciplinaire permis la réalisation de cette thèse. Je remercie les membres du laboratoire Génie de Production de l'Ecole Nationale d'Ingénieurs de Tarbes de m'avoir accueillie au sein de ce laboratoire.

Un grand merci au docteur Pierre TIAKO (Tiako University, Oklahoma City, États-Unis) pour sa relecture méticuleuse.

J'associe à ces remerciements le professeur Blaise NSOM de l'Université de Bretagne Occidentale pour le temps consacré à la relecture de cette thèse.

Merci à Madame Armelle BONENFANT, responsable de la L1 informatique de l'université Paul Sabatier de Toulouse pour avoir facilité mon monitorat.

Résumé

Tout professionnel de la santé est sujet devant un patient à une incertitude inhérente à la pratique médicale. Dans le cas d'incident médical lors d'un trajet aérien, cette incertitude comporte trois sources additionnelles : (1) variabilité des conditions aéronautiques, (2) variabilité individuelle des conditions du patient, (3) variabilité individuelle des compétences de l'intervenant. Aujourd'hui les incidents médicaux dans l'avion sont estimés à 350 par jour dans le monde et lorsqu'ils surviennent, ils sont pris en charge dans 95% des cas par des professionnels de la santé passagers qui se portent volontaires. C'est souvent pour eux une première expérience.

La principale raison de la réticence des professionnels de la santé à répondre à l'appel du commandant est la nécessité d'improviser ; ayant à établir un diagnostic et évaluer la gravité de l'état du patient dans des conditions difficiles. A part l'assistance à distance par télé-médecine l'intervenant, souvent seul face à ses doutes et son incertitude, ne dispose d'aucune autre aide à bord.

Par ailleurs l'aviation civile dispose de systèmes de retour d'expérience (RETEX) pour gérer la complexité de tels processus. Des politiques de recueil et d'analyse des événements sont mises en place à l'échelle internationale, par exemple ECCAIRS (European Co-ordination Centre for Accident and Incident Reporting Systems) et ASRS (Aviation Safety Reporting System).

Dans ce travail de thèse, nous proposons tout d'abord une formalisation sémantique basée sur les ontologies pour préciser conceptuellement le vocabulaire des incidents médicaux se produisant durant les vols commerciaux. Ensuite, nous mettons en œuvre un processus d'extraction des connaissances à partir des données (bases existantes) pour identifier les structures caractéristiques (patterns) des différents groupes d'incidents majeurs. Enfin, nous proposons une architecture de Système d'Aide à la Décision Médicale (SADM) qui intègre la gestion des incertitudes présentes tant sur les données récoltées que les niveaux de compétences des professionnels médicaux intervenants.

Abstract

There is an inherent risk in the practice of medicine that can affect the conditions of medical activities (diagnostic or therapeutic purposes). The management of uncertainty is also an integral part of decision-making processes in the medical field. In the case of a medical incident during an air travel, this uncertainty includes three additional sources: (1) variability of the aeronautical conditions, (2) individual variability of the patient's conditions, (3) individual variability of the intervener's skills. Presently, medical incidents in the plane are estimated worldwide at 350 per day and when they occur, they are handled in 95 % of cases by health professionals who are passengers. It is often for them a first experience.

The main reason for the reluctance of health professionals to respond to the aircraft captain's call is the need to improvise; having to make a diagnosis and assess the severity of the patient's condition under difficult conditions. Apart from telemedicine with remote assistance, the intervener, often alone in the face of his doubts and uncertainty, has no other decision aid tool on board.

Civil aviation also has feedback systems to manage the complexity of such processes. Event collection and analysis policies are put in place internationally, for example ECCAIRS (European Co-ordination Center for Accident and Incident Reporting Systems) and ASRS (Aviation Safety Reporting System).

In this work, we first propose a semantic formalization based on ontologies to clarify conceptually the vocabulary of medical incidents occurring during commercial flights. Then, we implement a knowledge extraction process from the data available on existing databases to identify the patterns of the different groups of incidents. Finally, we propose a Clinical Decision Support System (CDSS) architecture that integrates the management of the uncertainties present on both the collected data and the skill levels of the medical professionals involved.

Table des matières

Remerciements	v
Table des figures	xii
Liste des tableaux	xiv
1 Introduction générale	1
1.1 Contexte et motivations	1
1.1.1 Incidents médicaux et prise en charge médicale en vol	1
1.1.2 Un besoin d'aide à la décision médicale en vol	1
1.1.3 Des bases règlementaires et procédurales pour le recueil d'information	2
1.1.4 Objectifs de la thèse	3
1.2 Approche et positionnement scientifique	3
1.3 Structure du mémoire	6
1.4 Publications	6
1.4.1 Revues internationales avec comité de lecture	6
1.4.2 Conférences nationales	7
1.4.3 Workshops nationaux	7
2 Etat de l'art sur les systèmes d'aide à la décision médicale (SADM)	9
2.1 Introduction	9
2.2 Les incidents médicaux dans l'avion	9
2.2.1 Les parties prenantes de la prise en charge médicale dans l'avion	9
2.2.1.1 Les organismes de gestion aéronautique	11
2.2.1.2 Personnel Navigant Commercial (PNC)	12
2.2.1.3 Service médical au sol (SMS)	12
2.2.1.4 Volontaire professionnel de la santé	12
2.3 Démarche de décision et systèmes d'aide à la décision (SAD)	14
2.3.1 Processus de décision	14
2.3.2 Systèmes d'aide à la décision (SAD)	14
2.4 Systèmes d'aide à la décision médicale (SADM)	14
2.4.1 Approche numérique ou décider à partir des données	17
2.4.2 Approche symbolique (IA) ou décider à partir des connaissances	17

2.5	Ingénierie des connaissances pour les SADM	19
2.5.1	Formalismes de représentation des connaissances	20
2.5.2	Acquisition de connaissances et raisonnement	21
2.6	Conclusion	24
3	La prise en charge médicale dans l'avion : de la télémédecine au SADM	27
3.1	Introduction	27
3.2	Les facteurs déterminants de la prise en charge médicales dans l'avion	28
3.3	Télémédecine en aéronautique	30
3.4	Les systèmes de report d'évènement dans l'aéronautique	31
3.5	Extraction de connaissances à partir de données (ECD)	32
3.5.1	Processus d'ECD	33
3.5.2	Extraction de connaissances à partir de textes (ECT)	34
3.5.2.1	Représentation et codage de textes	36
3.5.2.2	L'analyse formelle de concepts	37
3.5.2.3	Sémantique vectorielle	38
3.6	Modélisation thématique pour la définition de l'incident médical dans l'avion	38
3.6.1	Matériels	42
3.6.2	Méthode	43
3.6.2.1	Création du corpus	43
3.6.2.2	Construction de la matrice d'occurrences	43
3.6.2.3	Construction de la matrice pondérée	44
3.6.2.4	Analyse des composantes principales	45
3.7	Exploration de SNOMED CT (les ontologies des incidents)	47
3.7.1	Matériels	48
3.7.2	Méthode	49
3.8	Conclusion	50
4	Une approche de modélisation de l'incertitude pour la conception de la décision médicale dans l'avion	51
4.1	Introduction	51
4.2	L'incertitude médicale	51
4.2.1	Incertitude médicale dans l'avion	53
4.3	L'approche probabiliste de formalisation de la décision médicale sous incertitude	54
4.4	L'approche crédibiliste de formalisation de la décision sous incertitude	55
4.4.1	La théorie de Dempster-Shafer	55
4.4.2	Les concepts formels de la théorie de Dempster-Shafer	56
4.4.2.1	Les règles de combinaisons de fonctions de masse	57
4.4.3	Base de données évidentielles	58
4.4.4	La BDE comme les mondes possibles	60
4.4.5	Corps d'évidence d'une transaction	61
4.4.6	Corps d'évidence d'une BDE	62

4.5	Un modèle de représentation de l'incertitude comme cadre formel de décision médicale dans l'avion	63
4.5.1	Construction de la BDE de la prise en charge médicale dans l'avion	65
4.5.1.1	Contexte formel de la prise en charge médicale dans l'avion	66
4.5.1.2	Le cadre de discernement de la décision médicale dans l'avion	68
4.5.1.3	La BDE de la prise en charge médicale dans l'avion	68
4.5.2	Intégration de l'incertitude de l'intervenant	68
4.5.3	Création du cadre formel de décision	68
4.6	Le cadre formel de décision et l'aide à la décision	70
4.6.1	Le Cadre formel de décision comme l'ensemble des états de la nature	72
4.6.2	La fiabilité des déroutements	73
4.7	Le cadre formel de décision et le raisonnement clinique	74
4.7.1	Formalisation du Chemin clinique	74
4.7.1.1	Le chemin clinique comme ontologie	75
4.7.1.2	Le chemin clinique comme Base de Préférences Conditionnelles	77
4.8	Une architecture d'un SADM pour les incidents médicaux dans l'avion	78
4.8.1	Description du modèle de données et de connaissances et des technologies pour la prise en charge des incidents médicaux dans l'avion	78
4.8.1.1	Ontologie d'incident : cas de Vomissement	79
4.8.1.2	Utilisation du concept Procédure pour l'intégration de chemin clinique	81
4.8.1.3	Intégration du cadre formel de décision dans le modèle de connaissances	82
4.8.2	Outils de raisonnement pour la prise en charge des incidents médicaux dans l'avion	82
4.9	Conclusion	83
5	Conclusion générale	85
	Bibliographie	89

Table des figures

1.1	Variabilité, selon Hollis, du nombre d'incidents médicaux dans l'avion mesuré par le nombre de fiches d'incidents remplies, l'ouverture de la trousse de secours ou l'appel au service de télémédecine [Hol16]	2
1.2	Organisation du mémoire	7
2.1	Les parties prenantes de la gestion des incidents médicaux dans l'avion	11
2.2	Diagramme de classe des axes de la taxonomie des SADM	16
2.3	Richesse sémantique et degré d'objectivation des données et connaissances [Hor00]	16
2.4	Les principales méthodes de développement de SADM [Fra+15]	18
2.5	Les composantes d'un SADM [IHT17]	19
2.6	SNOMED CT dans l'application	25
2.7	Les relations entre concepts dans SNOMED CT	25
3.1	Formalisation de la prise en charge médicale dans l'avion	28
3.2	Les actes de la télémédecine [CNP12]	31
3.3	Téléassistance médicale	31
3.4	Taxonomie de la télémédecine et du SADM	32
3.5	Processus d'ECD [Fay+96]	33
3.6	Les composants de SNOMED CT [CT14]	39
3.7	Exemple de description de concept selon SNOMED CT : syncope de la toux (perte temporaire de conscience après des accès de toux) avec différents termes synonymes en anglais	40
3.8	Définition de l'incident médical dans l'avion	43
3.9	Processus d'exploration de SNOMED CT	44
3.10	Matrice d'occurrences (Types de concepts/Incidents)	44
3.11	Création de la matrice d'étude à partir de la matrice d'occurrences	45
3.12	Cercle de corrélations variables composantes	46
3.13	Plan factoriel des deux premières des composantes	46
3.14	Pourcentage de variance cumulée des composantes	47
3.15	Exploration de SNOMED CT	48
3.16	Relation de l'ontologie Syncope	49
3.17	Concepts de l'ontologie Syncope	50

4.1	Positionnement du médecin par rapport à l'incertitude - (MG = Médecine Générale, FMC = Formation Médicale Continue) [Blo08]	52
4.2	Facteurs déterminants de l'incertitude	53
4.3	Tolérance à l'incertitude médicale	54
4.4	Facteurs déterminants de la prise en charge médicale sous incertitude	64
4.5	Processus de création du cadre formel de décision	65
4.6	Cadre formel et chemin clinique	74
4.7	Modèle ontologique du chemin clinique – formalisation orientée objectif	75
4.8	Modèle ontologique du chemin clinique - formalisation orientée action	77
4.9	Architecture du système d'aide à la décision pour incidents médicaux dans l'avion	79
4.10	Structure de l'ontologie Dempster-Shafer	79
4.11	Les métriques de l'ontologie de l'incident vomissement	80
4.12	Structure des ontologies d'incidents	80
4.13	Incident Vomissements	81
4.14	Liens sémantiques entre Incident, Chemin clinique et Procédure	81
4.15	Transformation des facteurs déterminants de l'incertitude en concepts Dempster-Shafer	82

Liste des tableaux

1.1	Une approche de gestion d'urgences médicales dans l'avion [Pet+13]	4
2.1	Catégorisation et mappage sémantique des incidents médicaux dans l'avion	10
2.2	Les aspects médicaux de la formation du PNC	13
2.3	Classes de SADM	15
2.4	Les concepts de plus haut niveau dans SNOMED CT	23
2.5	Concepts SNOMED CT habituellement utilisés dans les dossiers cliniques sans autre élaboration [gui14]	24
3.1	Les caractéristiques des études	29
3.2	Affectation de thèmes pour chaque document	47
4.1	Types et sources d'incertitudes	55
4.2	Base de données évidentielles, SP = Syncope-Pre-syncope ; CS = Cardiac symptom ; S = Seizures ; RS = Respiratory symptom ; ND = Non diversion ; D = Diversion	59
4.3	Un monde possible imprécis	60
4.4	les mondes possibles imprécis	60
4.5	Les états précis d'un monde possible imprécis	61
4.6	Corps évidentiels de transaction	62
4.7	Corps évidentiels d'une BDE	63
4.8	Itemsets du cadre formel de décision	64
4.9	Les types d'incidents médicaux dans l'avion	67
4.10	Domaines de discernement des variables de décision	69
4.11	BDE de la prise en charge médicale dans l'avion	70
4.12	Incertitude de l'intervenant	71
4.13	Incertitude de l'intervenant	71
4.14	Cadre formel de décision	71
4.15	Matrice de décision	72
4.16	Un exemple de matrice de décision	73
4.17	Un exemple de matrice de décision basée sur le support	73
4.18	Un exemple de matrice de décision basée sur le support et la pertinence des actions	73
4.19	Structure du chemin clinique	75

4.20 Incidents cibles dans le cas d'une gastro-entérite aiguë 76

Introduction générale

1.1 Contexte et motivations

1.1.1 Incidents médicaux et prise en charge médicale en vol

Alors que les compagnies de transport aérien convoient environ 3 milliards de personnes par an [Nao+16], les avions des vols commerciaux sont des endroits où la couverture médicale est rarement optimale. Une forte augmentation de passagers a majoré le nombre d'incidents médicaux. La prise en charge de ces incidents qui découle d'une succession de processus de prise de décisions est devenue un sujet numériquement significatif. Dans l'avion, elle est assurée dans 95% des cas par des passagers volontaires qui sont des professionnels de santé, mais qui ne sont pas toujours formés à la médecine d'urgence, à fortiori aéronautique. Et ils le font souvent pour la première fois [Cos15]. Cet aspect inédit de la prise en charge médicale dans l'avion impacte les deux premières activités du processus de décision : le renseignement et la conception. Le professionnel de santé n'a pas souvent la maîtrise de l'environnement aéronautique nécessaire à la gestion de la situation. Il lui manque aussi des données et des connaissances pour optimiser les alternatives de mise en œuvre des soins appropriés. Dans ce contexte de doute sur les interventions médicales à mettre en œuvre, le principe du « primum non nocere » renforcé par l'insécurité juridique pousse les professionnels de santé volontaire à une « expectative armée » en attente de l'arrivée au sol ; le déroutement devient de fait l'élément majeur de la prise en charge [Gou13][Hol16].

1.1.2 Un besoin d'aide à la décision médicale en vol

La principale raison de la réticence des professionnels de la santé à répondre à l'appel du commandant de bord est due à la complexité de l'analyse pour établir un diagnostic et évaluer la gravité de l'état du patient dans des conditions difficiles [Cos15]. Il peuvent parfois bénéficier d'une assistance par télé-médecine, par exemple depuis 1973 les avions long-courriers d'Air France sont en liaison téléphonique avec le SAMU 75 [RF11], mais des études récentes [Hol16] montrent un écart conséquent entre le nombre de cas gérés à distance et la demande médicale à bord (figure 1.1). Dans ce diagramme représente l'évolution annuelle des nombres : en bleu d' "Appels à MedLink " (événements pour lesquels une notification est faite à un service médical de permanence à distance), en orange des occurrences "Kit médical ouvert"(évènement ayant nécessité l'ouverture du kit médical de bord) , en gris des événements médicaux rapportés par les personnels navigants dans le journal électronique "Cognos Feed ". Nous voyons que le nombre de cas où

une intervention ayant compris une ouverture du kit est maximum du tiers des cas rapportés dans le journal de bord, que le recours à la télémédecine survient dans la moitié des cas d'ouverture du kit. Ces données sont à interpréter en fonction du contenu du kit qui ne contient pas uniquement des médicaments de l'urgence vitale mais également des conditions qui vont influencer l'appel à une aide extérieure (confiance du pilote dans le professionnel volontaire, confiance du professionnel dans ses capacités, confiance du professionnel dans le service de télémédecine...). Un dispositif d'aide à la décision médicale accessible en vol pourrait agir sur les freins organisationnels et techniques limitant la prise en charge interventionnelle des accidents médicaux en vol. Il pourrait aider à la structuration des échanges de télémédecine et au retour d'expérience.

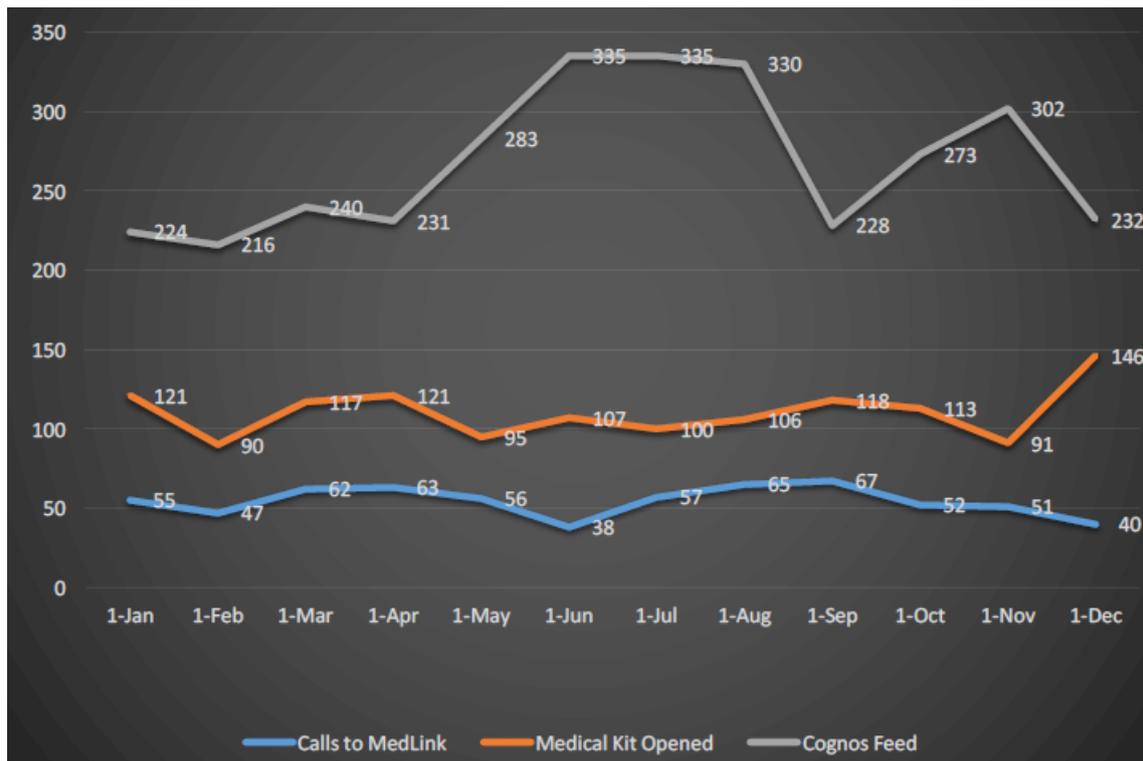


FIGURE 1.1 : Variabilité, selon Hollis, du nombre d'incidents médicaux dans l'avion mesuré par le nombre de fiches d'incidents remplies, l'ouverture de la trousse de secours ou l'appel au service de télémédecine [Hol16]

1.1.3 Des bases règlementaires et procédurales pour le recueil d'information

Dans le monde très codé de l'aviation, des systèmes de notification d'incidents utilisés pour le retour d'expérience (RETEX) et donc pour une sorte d'aide à la décision existent. Par les leviers d'organismes nationaux, régionaux, et continentaux, les gouvernements pour qui la sécurité du transport aérien est un enjeu majeur, se sont organisés à un niveau international pour mettre en place des systèmes de notification d'incidents et d'accidents visant à qualifier et prévenir les incidents et accidents en vol. Le registre Européen ECCAIRS prend en compte : « tout événement relatif à la sécurité qui met en danger ou qui, s'il n'est pas corrigé ou traité, pourrait mettre en danger un aéronef, ses occupants ou toute autre personne, et qui comprend en particulier les accidents et les incidents graves » [14a]; les événements médicaux rentrent dans ce cadre par leur effet sur les occupants, voire l'aéronef lui-même selon le recensement de

cet incident (affections somatiques ou troubles psychosomatiques touchant le personnel navigant ou les passagers). Des travaux appuyés sur des études de prévalence des types d'incidents avec taux de mortalité, de déroutement, selon les typologies de transport aérien [Cos15] ont permis la descriptions de protocoles de prise en charge d'urgences vitales (cf. tableau 1.1 [Pet+13]) et l'amélioration des valises médicales à bord.

1.1.4 Objectifs de la thèse

Nous avons montré la pertinence d'une aide à la décision médicale embarquée pour les incidents médicaux en vol. La réglementation permet la mise à disposition de bases de données pouvant servir de corpus d'analyse. Les protocoles ont été proposés sur des données épidémiologiques et en fonction des connaissances médicales. Nous souhaitons contribuer à la réalisation d'un système d'aide à la décision embarqué permettant en temps réel d'adapter l'attitude de l'intervenant à bord et en différé, d'optimiser les prises en charges futures.

Ce travail de thèse propose une méthode de modélisation et de structuration des connaissances pour les processus de décision médicale en vol grâce notamment à l'extraction de connaissances à partir de données (ECD) et à la formalisation ontologique.

Nous allons d'abord nous efforcer de comprendre les besoins de l'intervenant dans le processus décisionnel. Ensuite, nous gérerons l'incertitude inhérente à la pratique médicale appliquée au corpus des décisions possibles de l'intervenant. Pour finir nous allons définir les fondements nécessaires à la construction du cadre global de modélisation et de raisonnement pour l'intégrer dans un système décisionnel dédié.

1.2 Approche et positionnement scientifique

Pour atteindre nos objectifs, nous nous sommes intéressés d'une part à l'aide à la décision non structurée par télé-assistance (guidance par un médecin expert connecté par télé-médecine), d'autre part à l'aide à la décision structurée grâce aux systèmes d'aide à la décision médicale (SADM) dont le principal objectif est de faciliter le processus de décision. Les SADM sont des modules, fonctions, fonctionnalités ou systèmes informatiques conçus pour 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 [Gre14a].

Dans le cadre de la télé-médecine, Doumbouya [Dou16] montre l'importance d'intégrer la modélisation sémantique et le raisonnement argumentatif dans les activités de télé-expertise. La démarche adoptée pour soutenir l'architecture proposée s'inscrit dans une perspective d'intégration dans un SADM, facilite la collaboration et assure une traçabilité des actes médicaux souvent programmés à l'avance. Dans ce travail nous voulons couvrir le cas où l'intervenant, situé auprès du patient, se trouve isolé et avec des moyens limités. Mais il dispose a) de moyens de télécommunication lui permettant un accès non programmé à un expert, b) d'un environnement avionique pouvant héberger des moyens informatiques de capture, d'analyse et de traitement de données médicales.

Les SADM peuvent être envisagés sous l'angle de plusieurs domaines informatiques : traitement statistique des données, modélisation des processus, intelligence artificielle, ontologie, extraction de connaissances à partir de données, raisonnement sous incertitude, web, données multimédias, . . . Ces

Approche générale de gestion des urgences médicales en vol

- Identifier vous et préciser votre niveau de formation médicale à l'équipage de conduite.
- Évaluation des patients :
 1. Identifier le principal problème du patient et sa durée.
 2. Identifier les symptômes associés et à haut risque (p. Ex. Douleur thoracique, essoufflement, nausées ou vomissements ou faiblesse ou engourdissement unilatéral).
 3. Obtenir des signes vitaux (pouls et tension artérielle). Si vous êtes incapable d'évaluer la tension artérielle au moyen de l'auscultation, évaluer la en palpant le pouls radial.
 4. Évaluer l'état mental du patient et déterminer s'il existe des déficits neurologiques focaux.
- Si le patient est en arrêt cardiaque, procurer vous et appliquer un défibrillateur automatisé externe (DAE). Pour les patients avec un pouls mais un problème cardiaque suspecté, envisager d'utiliser un DAE s'il a des capacités de surveillance. (La compagnie aérienne peut exiger un contact avec un consultant au sol avant de l'utiliser.)
- Demander à un membre de l'équipage d'obtenir le kit médical d'urgence (KMU) et administrer de l'oxygène si nécessaire.
- Initier une consultation avec le consultant au sol si ce n'est déjà fait par l'équipage de conduite.
- Les recommandations d'interventions, telles que l'administration de médicaments ou de liquides intraveineux, devraient être discutées avec le consultant basé au sol.
- Le déroutement des aéronefs, l'assistance médicale au sol ou les deux devraient être coordonnés avec des consultations au sol.
- Documenter la présentation clinique et les soins rendus. Cette information doit être fournie au personnel médical à l'arrivée, à destination avec le transfert des soins.

Prise en charge de la douleur thoracique ou des palpitations

- Vérifier les signes vitaux.
- Fournir de l'oxygène.
- Si la douleur thoracique peut être d'origine cardiaque, envisager l'administration d'aspirine.
- Si la tension artérielle systolique est supérieure à 100 mm Hg, envisager l'administration de nitroglycérine par voie sublinguale toutes les 5 minutes. Vérifier la pression artérielle après chaque dose.
- Si le DAE a des capacités de surveillance, son utilisation permet d'évaluer le rythme cardiaque et la preuve des modifications du segment S-T dans les dérivations des membres.
- Si les symptômes se résolvent avec les mesures ci-dessus, le déroutement d'avion n'est généralement pas nécessaire. La consultation au sol peut aider à prendre des décisions de réacheminement.

approches possèdent toutes des caractéristiques recherchées pour un SADM dans le contexte avionique, l'acquisition de connaissances et la facilitation de la conversion des données médicales en évidences pour l'extraction de connaissances à partir de données (ECD), la formalisation des connaissances et le raisonnement pour les ontologies. Elles ont aussi des limitations. Pour Rosier [Ros15], quels que soient les bénéfices de l'approche basée ontologies en termes de raisonnement, il persiste une limitation dans la méthode d'acquisition de la connaissance à modéliser et il est établi que les médecins, utilisateurs des SADM, n'acceptent que très peu les systèmes pour lesquels les conclusions sont issues d'un processus caché. Ce qui pousse à intégrer ces approches de « technicien » dans une méthode globale de développement et d'analyse qui est importante pour la fiabilité, l'évolutivité, la réutilisabilité, la portabilité des systèmes proposés [Lti11] en facilitant leur adéquation aux besoins et leur « affordance » ou caractère intuitif.

Les différentes approches de SADM dans la littérature ne s'opposent pas forcément, mais elles offrent des prismes d'analyse relativement parcellaires qui ne permettent d'obtenir qu'une vision fractionnée de la gestion des processus décisionnels. Le concepteur se doit de faire la revue de la littérature sur les SADM, en s'efforçant d'organiser une vision fractionnée, plus que contradictoire, de la gestion des processus décisionnels. Ltifi [Lti11] s'intéresse à des Systèmes Interactifs d'Aide à la Décision (SIAD) basés sur l'extraction de connaissances à partir de données (l'ECD) pour l'analyse des données temporelles répertoriées concernant les actes que subissent les patients. Berner [Ber07] propose une catégorisation des SADM où il oppose les systèmes à base de connaissances (déterministes) aux systèmes qui utilisent l'apprentissage (expérience) et l'extraction de patterns (motifs).

Dans cette thèse nous proposons une approche mixte ou nous considérons des retours d'expérience comme le point de départ d'un continuum donnée-information-connaissance. La modélisation des données de l'expérience permet un traitement des données (raisonnement) pour optimiser les étapes du processus de décision. Les informations expérimentales proviennent des systèmes de RETEX, des travaux sur les incidents, des outils de raisonnement clinique, des terminologies médicales et sont obtenues par élicitation ou par ECD.

Pour Hinkelbein et al. [Hin+17], cette approche offre plus de souplesse sur les choix des outils de formalisation, de raisonnement, et d'acquisition de connaissances, ce qui est nécessaire dans le contexte des incidents médicaux dans l'avion pour améliorer la gestion des ressources techniques et humaines. Elle permet aussi de mettre l'accent sur l'intervenant et le processus de décision.

Les contributions de notre travail sont :

- i La caractérisation de la prise en charge médicale en vol et la définition des modèles de connaissances nécessaires. Nous avons d'abord étudié les facteurs déterminants de la prise en charge médicale dans l'avion. Ceci nous a permis de préciser la définition des attributs de l'incident médical dans l'avion à travers un champ sémantique défini grâce aux concepts de la terminologie médicale SNOMED CT. Nous avons enfin défini une structure de connaissances correspondante. Un travail préalable ([SKR15]) sur la prise en charge médicale dans un environnement proche de celui des incidents médicaux en vol, l'EHPAD (établissements d'hébergement pour personnes âgées dépendantes), a été nécessaire. La formalisation du "cas" dans l'approche basée sur le Raisonnement à Partir de Cas (RaPC) a inspiré notre structure de connaissances.
- ii La conception d'un cadre formel de la décision médicale dans l'avion grâce à la modélisation de l'incertitude inhérente à la pratique médicale : nous avons transformé les facteurs déterminants de la

décision médicale en domaine de discernement et conçu une base de données évidentielles (BDE) utilisant la théorie de Dempster-Shafer [Sha76]. A partir de cette BDE, nous avons construit un cadre formel de décision contenant toutes les décisions possibles de l'intervenant. Ce cadre formel nous a permis de créer une matrice de décision pour fournir une aide décisionnelle.

- iii Nous avons ensuite suggéré une architecture de SADM intégrant les propositions formulées sur la modélisation sémantique et le cadre formel de décision ce qui facilite l'utilisation de ces nouvelles données pour l'amélioration de la pertinence de l'aide décisionnelle.

1.3 Structure du mémoire

Le premier chapitre discute du contexte et de la pertinence de cette thèse. Nous y présentons aussi les contributions scientifiques qui se sont notamment traduites par des publications dans les journaux internationaux et des présentations orales dans les colloques nationaux.

Dans le chapitre 2, nous faisons l'état de l'art sur les incidents médicaux dans l'avion, sur l'aide à la décision qui est l'objet de ce travail et sur l'ingénierie des connaissances pour introduire des aspects importants des SADM comme les formalismes de représentation des connaissances, l'acquisition de connaissances, et le raisonnement.

Dans le chapitre 3, nous avons caractérisé la notion d'incident médical que nous avons liée aux concepts de SNOMED CT. Ceci a permis de construire des ontologies dédiées à la prise en charge médicale dans l'avion.

La première partie du chapitre 4 décrit une approche de modélisation de l'incertitude basée sur la théorie de Dempster-Shafer. Nous y proposons un cadre formel de décision comprenant toutes les décisions que peut prendre un intervenant. Ce cadre comporte notamment le raisonnement clinique, et les conditions de déroutement. Il permet de construire une matrice de décision, qui utilise le raisonnement fondé sur les fonctions de croyance, pour choisir et proposer les actions de prise en charge de l'incident médical. Dans la deuxième partie, nous avons proposé une architecture de SADM et construit le modèle de données et de connaissances de la prise en charge médicale dans l'avion.

La figure 1.2 montre sous forme d'ovales les apports de base que nous sommes allés chercher dans l'existant et sous forme de rectangles les principales contributions de cette thèse sous forme de processus réalisés. Les flèches représentent les liens entre les sorties d'activités et les entrées d'autres activités de la démarche de processus suivie.

1.4 Publications

Notre activité de publication scientifique au cours de la thèse de doctorat se résume comme suit :

1.4.1 Revues internationales avec comité de lecture

[A1] A. Sene, B. Kamsu-Foguem, P. Rumeau. Data mining for decision support with uncertainty on the airplane. *Data & Knowledge Engineering*, <https://doi.org/10.1016/j.datak.2018.06.002>, Available online 26 July 2018. Journal - Elsevier - Impact Factor : 1.467

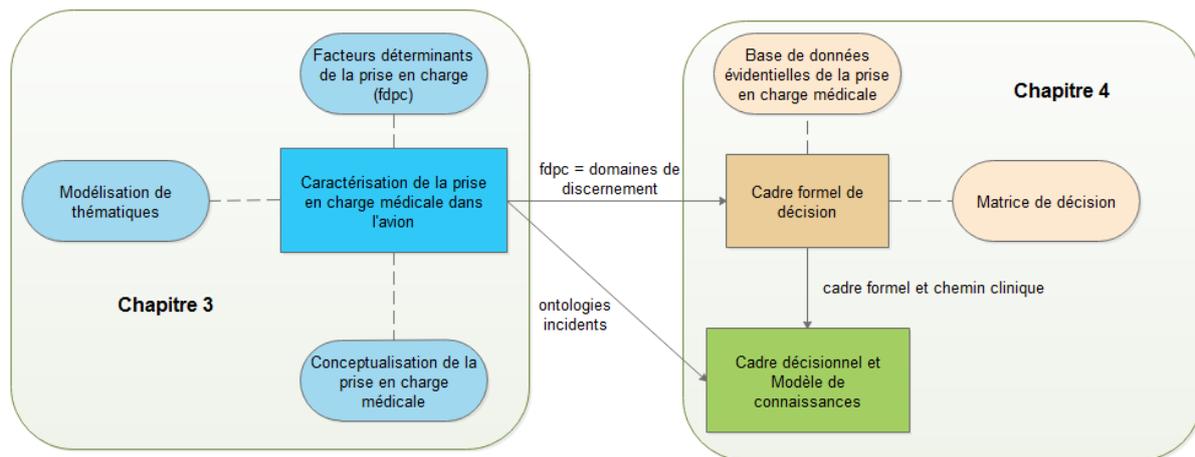


FIGURE 1.2 : Organisation du mémoire

[A2] Alsane Sene, Bernard Kamsu-Foguem, Pierre Rumeau. Discovering frequent patterns for in-flight incidents. Cognitive Systems Research, Volume 49, June 2018, Pages 97-113, Journal - Elsevier - Impact Factor : 1.425.

[A3] Alsane Sene, Bernard Kamsu-Foguem, Pierre Rumeau. Decision support system for in-flight emergency events. Decision support system for in-flight emergency events. Cognition, Technology and Work (2018), May 2018, Volume 20, Issue 2, pp 245–266, Journal - Springer -Impact Factor 1.260.

[A4] A. Sene, B. Kamsu-Foguem, P. Rumeau. Telemedicine Framework using Case-based Reasoning with Evidences. Computer Methods and Programs in Biomedicine, Volume 121, Issue 1, August 2015, Pages 21-35, Journal - Elsevier - Impact Factor : 2.674.

1.4.2 Conférences nationales

[B1] A. Sene, B. Kamsu-Foguem, P. Rumeau. Knowledge modeling and structuring in telemedicine process, 21e Journées sciences et techniques de la production du GDR MACS, 26-27 Novembre 2015, Nantes, France.

[B2] Sene A., Rumeau P., Savoldelli M., Leclercq C., Kamsu-Foguem B. Modélisation des EHPAD : travail préliminaire en vue de l'analyse de l'activité de télémédecine. 4ème édition des journées annuelles de la Société Française des Technologies pour l'Autonomie et de Gérontechnologie (SFTAG), 24 - 25 novembre 2014, Paris, France. Second Prix de la meilleure communication affichée.

1.4.3 Workshops nationaux

[C1] A. Sene, B. Kamsu-Foguem, P. Rumeau. La prise en charge des malaises lors des vols réguliers, Colloque Transfert des Technologies Aéronautiques et Spatiales au Service des Hommes et des Territoires, Jeudi 2 Février 2017, Tarbes, France.

[C2] A. Sene, B. Kamsu-Foguem, P. Rumeau. Modélisation et structuration des connaissances dans le processus de télémédecine, Journées IODE (Ingénierie des Organisations Distribuées) des 28 et 29 Mai 2015, Tarbes, France.

Etat de l'art sur les systèmes d'aide à la décision médicale (SADM)

2.1 Introduction

Nous avons évoqué dans l'introduction la nécessité de décrire et de modéliser les connaissances nécessaires à la prise en charge des incidents médicaux dans l'avion pour proposer une aide à la décision à bord des vols commerciaux. Dans la première partie de ce chapitre, nous allons décrire les incidents médicaux dans l'avion et les parties prenantes de la prise en charge médicale. Dans la deuxième partie nous nous intéressons aux systèmes d'aide à la décision (SAD) et en particulier aux systèmes d'aide à la décision médicale (SADM). Les différentes approches de SADM seront présentées. Nous nous intéresserons aussi à l'ingénierie des connaissances pour étudier les aspects d'acquisition, de formalisation très importants pour un SADM.

2.2 Les incidents médicaux dans l'avion

Marine Costa [Cos15] rapporte qu'il n'y a pas de définition claire d'un incident médical à bord. Le tableau 2.1 montre une catégorisation des incidents médicaux dans l'avion selon 3 études publiées en langue anglaise ou française ([Pet+13], [Mah+11], [Cos15]). Nous avons dans le tableau rapproché des incidents sémantiquement identiques et représentés sous des dénominations différentes. Les regroupements réalisés ne respectent pas la taxonomie de l'Organisation Mondiale de la Santé (Classification Internationale des Maladies dixième version), aggravant l'hétérogénéité des dénominations. Il s'agit principalement des études statistiques, recherchant des relations entre les variables déterminantes dans la gestion des incidents (type d'incident, formation de la personne qui prend en charge, déroutement de l'avion, ...). Diane Naouri et al. [Nao+16] ont estimé à 350/jour soit 1 pour 14000 à 40000 passagers le nombre des incidents médicaux dans l'avion.

2.2.1 Les parties prenantes de la prise en charge médicale dans l'avion

En vol nominal, l'avion est un système fermé ne pouvant pas échanger « physiquement » avec l'extérieur à l'exclusion des moyens de télécommunication. La figure 2.1 montre ceux qui peuvent constituer les parties prenantes directes de la prise en charge des incidents médicaux dans l'avion. Il y a le personnel navigant

Peterson			Mahony			Costa		
Incident	Id	%	Incident	Id	%	Incident	Id	%
Syncope or pre-syncope	2	37,4	Unconscious : rapidly recovered	2	41,1	Syncope	2	46,4
Respiratory symptoms	3	12,1	Unconscious : slow or delayed recovery	2	1,5	respiratory symptoms	3	8,7
Nausea or vomiting	4	9,5	Unconscious : no recovery	2	1,2	Nausea or vomiting	4	15,9
Cardiac symptoms	5	7,7	Breathing difficulty	3	15,9	Cardiac symptoms	5	2,9
Seizures	6	5,8	Nausea / vomiting / diarrhea	4	19,5	Epileptic seizure	6	2,9
Abdominal pain	7	4,1	Seizure	6	1,1	pain (chest)	7	5,8
Infectious disease	8	2,8	Chest pain	7	2,3	Infectious disease	8	4,3
Agitation or psychiatric symptoms	9	2,4	Behavioral	9	3,2	Agitation or psychiatric symptoms	9	4,3
Allergic reaction	10	2,2	Allergic / skin	10	2,3	Allergic reaction	10	1,4
Possible stroke	11	2	Injuries	12	2,5	Injury, wound	15	2,9
Trauma, not otherwise specified	12	1,8	Collapse-injury sustained	15	0,6	Obstetrical or gynecologic symptoms	16	1,4
Diabetic complication	13	1,6	Limb swelling	15	0,5	thrombosis	20	2,9
Headache	14	1	Pregnant/in labor	16	0,1			
Arm or leg pain or injury	15	1	Pain-no other symptoms	17	5			
Obstetrical or gynecologic symptoms	16	0,5	Congested/ fever	20	1,1			
Ear pain	17	0,4	Bleeding no injury	20	0,7			
Cardiac arrest	18	0,3	Limb/ facial weakness	20	0,2			
Laceration	19	0,3	Urinary	20	0,1			
Other	20	6,9	Advice/ undetermined	21	1			
Unknown	21	0,1						

TABLE 2.1 : Catégorisation et mappage sémantique des incidents médicaux dans l'avion

commercial (PNC), un éventuel service médical au sol (SMS) et un éventuel volontaire professionnel de la santé dans l'avion. Le processus de cette prise en charge dépend aussi de la législation applicable au vol et des organismes de gestion aéronautique qui définissent notamment le type de matériel médical à bord.

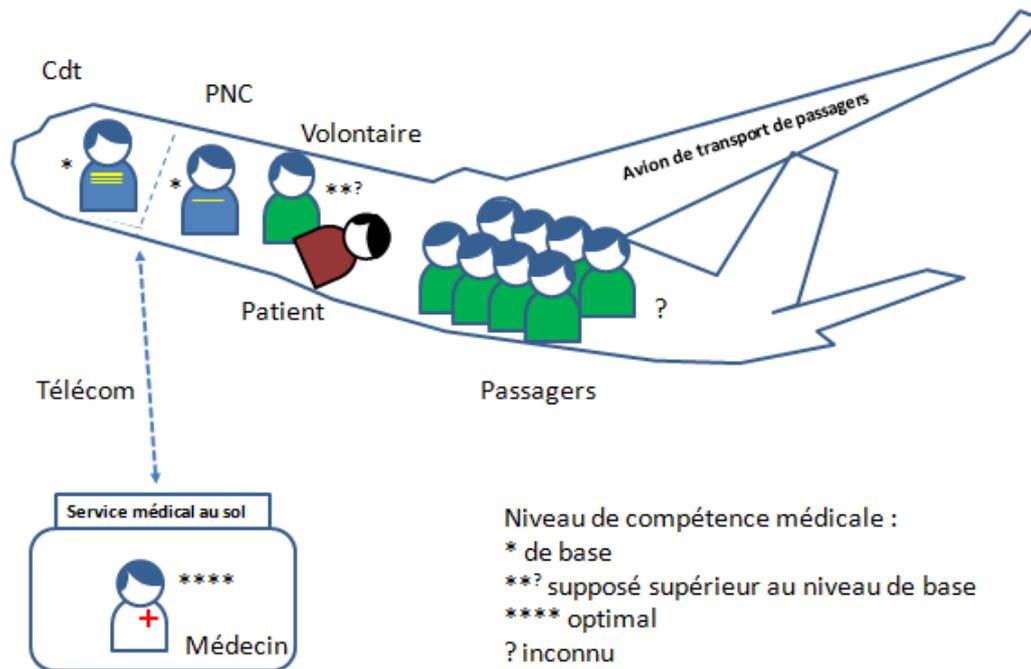


FIGURE 2.1 : Les parties prenantes de la gestion des incidents médicaux dans l'avion

2.2.1.1 Les organismes de gestion aéronautique

Les organismes de gestion aéronautique sont soit nationaux (Direction Générale de l'Aviation Civile), soit internationaux (Organisation de l'aviation civile internationale, Association Internationale du Transport Aérien). En plus de ces organismes gouvernementaux ou liés à des traités internationaux, l'Association Médicale Aérospatiale (AsMA) regroupant des professionnels dans le domaine de la santé, l'aéronautique et l'espace est focalisée sur le côté médical.

→ Direction Générale de l'Aviation Civile (DGAC)

En France la DGAC garantit la sécurité et la sûreté du transport aérien. Elle regroupe les services de l'État chargés de réglementer et de superviser les activités de l'aviation civile, de réguler le marché du transport aérien, de garantir la qualité de la formation des personnels [DGA17].

→ Organisation de l'aviation civile internationale (OACI)

L'organisation de l'aviation civile internationale qui compte 191 pays et qui dépend des Nations Unies a été créée en 1944 [OACI]. Elle compte sept bureaux régionaux (Bangkok, Dakar, Le Caire, Lima, Mexico, Nairobi et Paris). Les objectifs stratégiques de l'OACI sont la sécurité, la sûreté, le développement économique et la protection de l'environnement. Elle propose les normes régissant les vols internationaux.

→ Association Internationale du Transport Aérien (AITA)

L'AITA est une organisation commerciale internationale fondée en 1945 à Cuba et qui regroupe des sociétés de transport aérien. Cette association qui compte en 2017 deux-cent-soixante-quinze (275) compagnies aériennes (83 % du trafic aérien mondial) est utilisée comme chambre de compensation.

L'AITA édite chaque année un manuel médical et est très impliquée dans la santé du passager [IAT]. Les sections de ce manuel sont entre autres :

- Rôle et responsabilité des compagnies aériennes sur la santé du passager.
- Normes médicales dans l'avion
- Santé du passager
- Suivi des incidents médicaux dans l'avion

→ **Association Médicale Aérospatiale (AsMA)**

L'AsMA est une association à de fins caritatives, éducatives et scientifiques. Elle regroupe des personnes impliquées dans les transports aériens et spatiaux. Ses membres sont des spécialistes de la médecine aérospatiale, des infirmières en vol, des physiologistes, des psychologues, des spécialistes des facteurs humains, des assistants médicaux et des chercheurs dans ce domaine. La plupart de ses activités concernent l'industrie, les organismes de réglementation de l'aviation civile, les services de défense et les services militaires, les compagnies aériennes, les programmes spatiaux et les universités.

L'AsMA a comme mission la sensibilisation sur la santé des personnes impliquées sur les activités aérospatiales [HHH15]. Elle fournit son expertise aux organismes cités plus haut, notamment sur la composition du matériel médical à bord.

Ce matériel dont la composition et la présence dépendent des pays est organisé en trois trousse. Il y a la trousse de premiers secours accessible au PNC, la trousse de prévention universelle pour la désinfection et le nettoyage et la trousse médicale que seul un médecin avec l'accord du commandant peut ouvrir.

2.2.1.2 Personnel Navigant Commercial (PNC)

Initialement les premières hôtesses de l'air étaient toutes infirmières de formation. Aujourd'hui pour disposer de la licence de vol, le PNC a besoin d'un Certificat de Formation à la Sécurité (CFS) délivré en France par le DGAC et comprenant un important volet secourisme [Cos15]. Dans la version consolidée au 15 octobre 2017 de l'Arrêté du 25 septembre 2007 relatif au certificat de formation à la sécurité [Arr07], les aspects médicaux et matériels médicaux de cette formation sont décrits de manière explicite. Cette formation (cf. contenu dans le tableau 2.2) est essentiellement orientée vers le secourisme et la prise en charge des urgences vitales avec prise en compte des particularités barométriques.

2.2.1.3 Service médical au sol (SMS)

À travers la télémédecine des services spécialisés en urgence médicale sont disponibles pour l'assistance de l'équipage ou de professionnels de la santé à bord. Ce sont des services fournis par des organismes gouvernementaux comme le SAMU (Service d'Aide Médicale Urgente) ou des sociétés privées comme MedAire qui a été le pionnier du concept de solutions globales d'atténuation des risques médicaux liés aux voyages [Jez98] [RP97].

2.2.1.4 Volontaire professionnel de la santé

Les cas d'absence totale de professionnel de la santé dans un vol commercial sont très rares et de ce fait les professionnels de santé, par priorité médicaux, se retrouvent souvent en première ligne lors des

Aspects médicaux
<p>Notions d'anatomie et de physiologie : système cardio-vasculaire</p> <ul style="list-style-type: none"> ● l'appareil digestif ● l'appareil respiratoire
<p>Matériel disponible à bord et utilisation :</p> <ul style="list-style-type: none"> ● trousse de premier secours (nombre et contenu) ● trousse de prévention universelle ● trousse médicale d'urgence ● oxygène ● défibrillateur semi-automatique ● tensiomètre.
<p>Protocole d'intervention :</p> <ul style="list-style-type: none"> ● analyse d'une situation d'urgence : observations, antécédents ● évaluation, bilan : respiration, conscience, circulation, lésion ● surveillance, contrôle ● chaîne de secours ● médecin (présence ou non) ● sol (service médical d'urgence régional : SMUR et SAMU)
<p>Gestes de premiers secours : ensemble des techniques médicales d'aide d'urgence prodiguées aux blessés et aux malades par une personne qui n'est pas médecin. Ces techniques ont pour but d'assurer la survie de la personne en attendant l'arrivée d'une personne compétente.</p> <ul style="list-style-type: none"> ● arrêts cardio-respiratoires (ACR), état de choc, coma, détresses respiratoires avec pratique de la réanimation cardio-pulmonaire (utilisation d'un mannequin spécialement conçu à cet effet), épilepsie, hémorragies graves, obstructions complètes, et troubles de conscience ● urgences pédiatriques (obstruction, ACR, convulsions)
<p>Réactions physiologiques dues aux vols en altitude :</p> <ul style="list-style-type: none"> ● caractères physiques et chimiques de l'atmosphère ● effets sur les cavités semi-closes (ORL, dents, tube digestif) ● effets liés aux variations rapides de pression ou décompressions

TABLE 2.2 : Les aspects médicaux de la formation du PNC

incidents médicaux aériens. Ces volontaires qui doivent, en théorie, fournir une carte professionnelle ont des capacités, des formations et une connaissance de la médecine aéronautique variables [Hol16]. La moitié des volontaires sont des médecins et plus de 20% sont des infirmiers. Marine Costa [Cos15] rapporte qu'une partie des médecins qui proposent leur aide ne se sont pas pour autant sentis utiles. Pour améliorer cela, elle préconise une amélioration du matériel médical pour le rendre plus adapté aux situations d'urgences et aux conditions en vol dans un avion commercial. Elle préconise également une bonne coordination et un soutien mutuel avec le PNC, et plus de reconnaissance de la part des compagnies aériennes.

2.3 Démarche de décision et systèmes d'aide à la décision (SAD)

La décision dans la prise en charge des incidents médicaux dans l'avion est multicritère, avec des contraintes temporelles et des contraintes de ressources. En général, la démarche de décision consiste à choisir une solution en comparant plusieurs possibilités, ce n'est pas une phase univoque et clairement identifiable [Akh04], mais c'est un cheminement dont le déroulement peut comprendre un ou plusieurs processus. Comme définition de la décision, Kadri [Kad14] propose : « la décision constitue la réponse d'un décideur, dans un contexte donné, qui effectue un choix entre plusieurs solutions susceptibles de résoudre le problème et la situation auxquels il est confronté ». La décision, est fondée sur un processus de décision et, le cas échéant, l'aide à la décision.

2.3.1 Processus de décision

Selon Herbert Simon [Sim77], une décision suit un processus dont les étapes sont la perception de la nécessité d'une décision (Renseignement ou Information), l'inventaire et l'analyse des choix possibles (Conception ou Modélisation), la sélection (Choix) puis la mise en œuvre et l'évaluation d'une des possibilités sélectionnées (Evaluation). Ce processus met à disposition un ensemble de solutions pour la résolution d'un problème donné en tenant compte de contraintes diverses.

2.3.2 Systèmes d'aide à la décision (SAD)

L'aide à la décision est l'activité qui, en prenant appui sur des modèles, 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 à recommander un comportement de nature à accroître la cohérence entre l'évolution du processus et les objectifs de l'acteur [RB93]. Les SAD sont des systèmes informatiques qui facilitent le processus de décision. Ils sont aussi divers et variés que leurs caractéristiques et leurs outils de mises en œuvre (tableau 2.3). Dans ce travail nous nous intéressons aux systèmes d'aide à la décision médicale.

2.4 Systèmes d'aide à la décision médicale (SADM)

Les systèmes d'aide à la décision médicale (SADM) sont des outils informatiques capables de traiter l'ensemble des caractéristiques d'un patient donné afin de générer les diagnostics probables de son état clinique (aide au diagnostic) et/ou les traitements qui lui seraient adaptés (aide à la thérapeutique) [SB14] [Syb16]. Ils représentent un moyen potentiel pour améliorer la qualité, la sécurité et l'efficacité des soins lorsqu'ils sont accessibles aux médecins pendant leurs activités de soins au moyen de leurs outils métier

Critère de classification	Exemples
Domaine	Médical [Fra+15] [Qat+17], Energie (Schillings et al., 2012), Développement durable [RO17], Transport [BM18]
Approche	Numérique, Intelligence Artificielle (IA), Recherche opérationnelle, Extraction de Connaissances à partir de Données (ECD)
Exploitation de la connaissance	Documentaire, Automatique, Alerte, centrée patient

TABLE 2.3 : Classes de SADM

(dossier médical et prescription informatisée) et s'ils sont correctement intégrés au processus de travail clinique [DGM14].

L'ampleur et la complexité des connaissances médicales, telles que les données physiopathologiques et les moyens diagnostiques et thérapeutiques, obligent le médecin à gérer toujours plus d'informations pour soigner un patient. La taxonomie des SADM proposée par Berlin et al. [BSS06] permet de voir leur diversité. Elle est composée de 26 axes qui chacun représente un aspect de diversification. Les 26 axes sont divisés en 5 catégories.

↪ **Contexte** : Cette catégorie d'axe décrit les objectifs du système ainsi que les facteurs liés au contexte d'usage, elle définit le cadre médical du SADM.

↪ **Source de données et de connaissances** : Cette catégorie va permettre de définir la source des connaissances de base sur les aspects pratiques et théoriques, la source des données en entrée, les choix d'encodage, de degré de customisation et les mécanismes de mises à jour des connaissances.

↪ **Support de décision** : Les axes dans cette catégorie caractérisent les méthodes de raisonnement utilisées, le niveau d'urgence médicale qu'on aura à traiter, les types de recommandations et réponses. On traitera ici aussi l'aspect logistique de l'offre de soins.

↪ **Restitution des recommandations** : Cette catégorie décrit comment l'aide sera fournie au destinataire.

↪ **Modèle de communication** : Cette catégorie décrit l'interaction entre le système et l'utilisateur final, et les intermédiaires s'il en existe.

Nous résumons les différents axes présentés sous forme d'un diagramme de classes avec les liens entre les différentes classes au sein du développement des processus d'exécution dans la figure 2.2.

Les SADM ont généralement quatre caractéristiques fonctionnelles [Fra+15]. Ils peuvent (1) générer une alerte (prescription ou dosage), (2) permettre l'entrée de données (résultats de laboratoire), (3) proposer des interventions (actions à partir du chemin clinique), (4) offrir le choix (choix de thérapies). Ces fonctionnalités sont assurées grâce à plusieurs méthodes de traitement de l'information (Apprentissage, Représentation de connaissances, Visualisation d'informations, Fouille de données textuelles, ...).

Ce sont les mêmes méthodes qui sont utilisées pour développer les deux principales approches de SADM : (1) les approches numériques et (2) les approches symboliques. Dans l'approche numérique, on décide à partir de la transformation des données alors que dans l'approche symbolique, on décide à partir de la manipulation des connaissances. Ces approches qui définissent une catégorisation basée

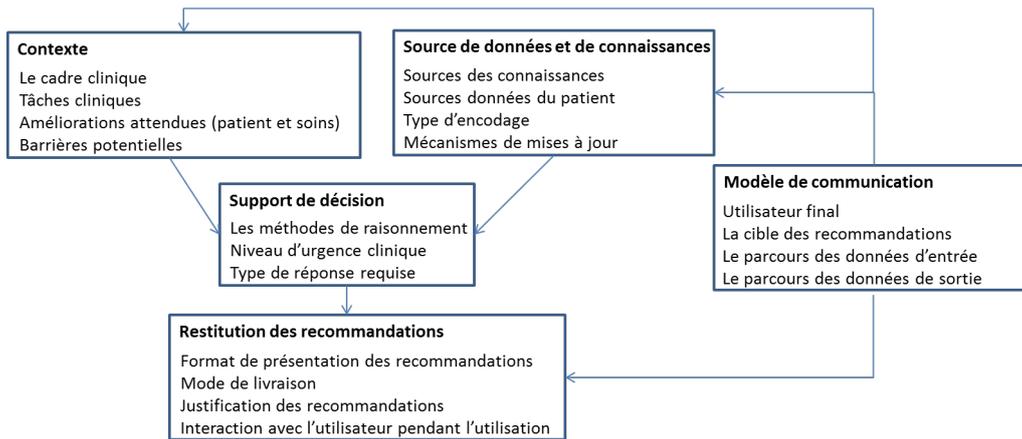


FIGURE 2.2 : Diagramme de classe des axes de la taxonomie des SADM

sur la nature de l'information traitée par les SAD ont marqué l'évolution des SADM. Pour étudier les particularités du traitement automatique de l'information, et donc des SADM, il est nécessaire de cerner la notion d'information, et notamment de pouvoir la distinguer de notions voisines comme celles de données et de connaissances. Les données et les connaissances peuvent être comparées selon deux dimensions (figure 2.3), (1) le degré d'objectivation ou de subjectivation et (2) la richesse sémantique [Hor00]. Dans ce travail, on fera la différence entre données et connaissances surtout au niveau sémantique.

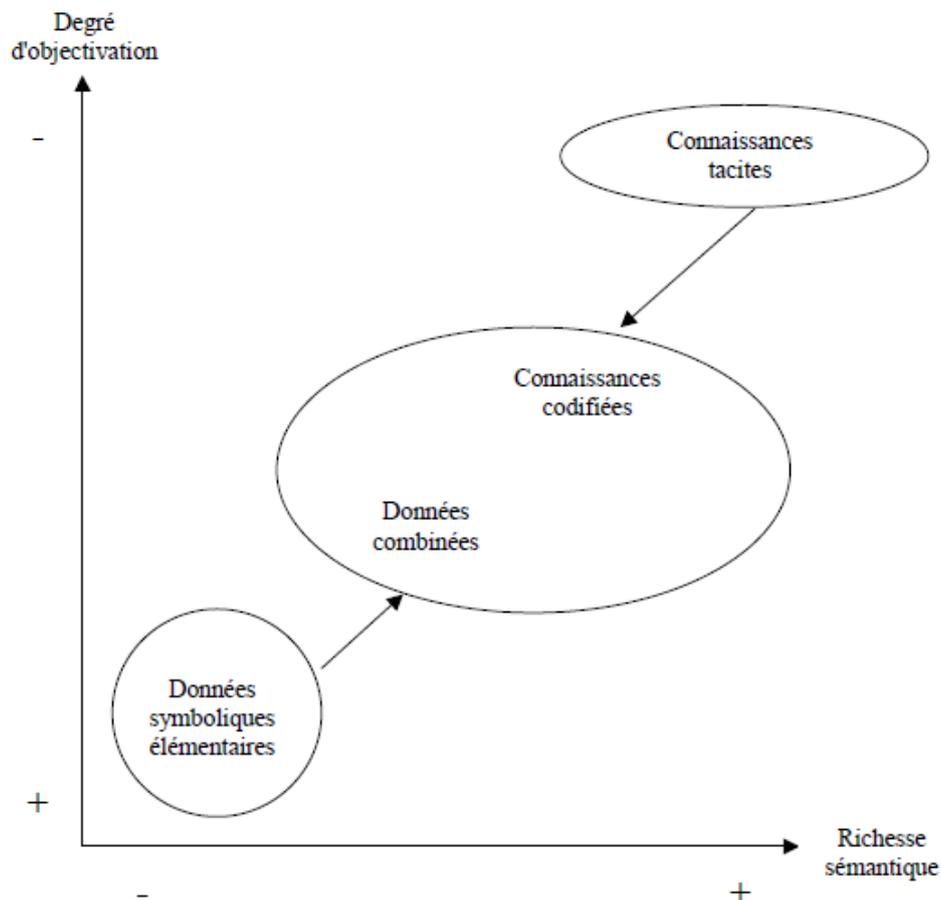


FIGURE 2.3 : Richesse sémantique et degré d'objectivation des données et connaissances [Hor00]

La donnée

Les données symboliques élémentaires sont diverses : lettres, chiffres, et pictogrammes. Elles constituent une réalité objective à un double niveau : d'une part, elles sont inscrites sur un support matériel et peuvent circuler indépendamment de leur créateur ; d'autre part leur signification, dépourvue d'ambiguïté, est dépendante d'un code purement conventionnel et dont l'espace de validité peut être extrêmement vaste (universel dans certains cas) [Hor00]. Leur contenu sémantique est par contre limité aux symboles ou éléments qui les différencient. L'objectivité de la donnée justifie son utilisation comme base de recherches, et d'études.

Les données brutes appelées aussi données primaires sont des données non interprétées émanant d'une source primaire, ayant des caractéristiques liées à la nature de la source et qui n'ont été soumises à aucun traitement ou toute autre manipulation. Elles sont souvent quantitatives.

L'information

L'information naît de la compréhension des relations qui peuvent exister entre plusieurs données [Bal02] [Paq06]. Elle est donc porteuse d'un contenu sémantique plus riche que les données qu'elle intègre et marquée par la subjectivité du producteur de l'information et/ou les caractéristiques propres du processus automatisé de production. Elle est une donnée qui acquiert du sens, un "signifié" porté par la donnée [Hor00].

La connaissance

La connaissance est une information à laquelle une compétence a été attachée et qui est donc capable de générer une action ou une autre connaissance [Dir09] [Sch00]. Elle peut, par conséquent, être présentée sous la forme suivante : "Si [conditions], alors [actions]". La connaissance naît de la compréhension et l'assimilation des règles qui régissent les modules ou les schémas mentaux sous-jacents aux relations entre les données, permettant ainsi de comprendre comment la situation évoluera si les données se modifient [Bal02].

2.4.1 Approche numérique ou décider à partir des données

L'approche numérique propose une modélisation commune des données d'un problème à résoudre [SB14]. Elle enrichit ainsi le contenu sémantique de ces données grâce à l'application principalement d'analyses statistiques. L'approche numérique est utilisée aussi bien en aide à la décision de diagnostic qu'à la décision thérapeutique. Comme nous allons le voir par la suite, les méthodes informatiques utilisées dans l'approche numérique peuvent aussi être utilisées dans l'approche symbolique. La différence entre ces approches réside sur la façon dont le SADM résultant agit sur le raisonnement. L'approche symbolique peut être considérée comme une évolution de l'approche numérique dans sa recherche de la nature de la connaissance capturée par les données.

2.4.2 Approche symbolique (IA) ou décider à partir des connaissances

L'approche symbolique vise à développer un système informatique ayant un comportement rationnel. Cette approche est basée sur l'intelligence artificielle qui est une discipline de l'informatique qui vise à modéliser les processus de la pensée humaine [Héo14]. L'IA est née dans les années 50 avec les premiers systèmes informatisés et les travaux de Turing. Les systèmes experts puis les systèmes à base de connaissances ont marqué son évolution. Mireille Cléret et al. en parlent de la manière suivante [CBD01] : « On peut définir le système expert comme étant un système qui imite la démarche de la

personne compétente dans un domaine donné, quel que soit la méthode de raisonnement qu'elle utilise. De plus, il doit être interactif, capable de dialoguer avec ses utilisateurs et d'expliquer ses raisonnements. Les systèmes experts ont laissé la place aux systèmes à base de connaissances (SBC), dont les termes sont moins restrictifs. Dans ces systèmes, les connaissances sont séparées du raisonnement mais il apparaît que la représentation des connaissances est étroitement liée à la méthode de raisonnement ».

Les premiers SADM ont été donc des systèmes experts. Par la suite, conformément au principe de la médecine fondée sur les données probantes, l'expertise subjective des systèmes experts a été remplacée par les connaissances objectives des recommandations de bonne pratique dans les SADM modernes [SB14]. Les SBC constituent aujourd'hui une grande partie de la littérature des SADM et la création de connaissances (Apprentissage, représentation de connaissances, et fouille de données textuelle) est l'une des principales activités comme le montre Fraccaro et al. [Fra+15] (figure 2.4).

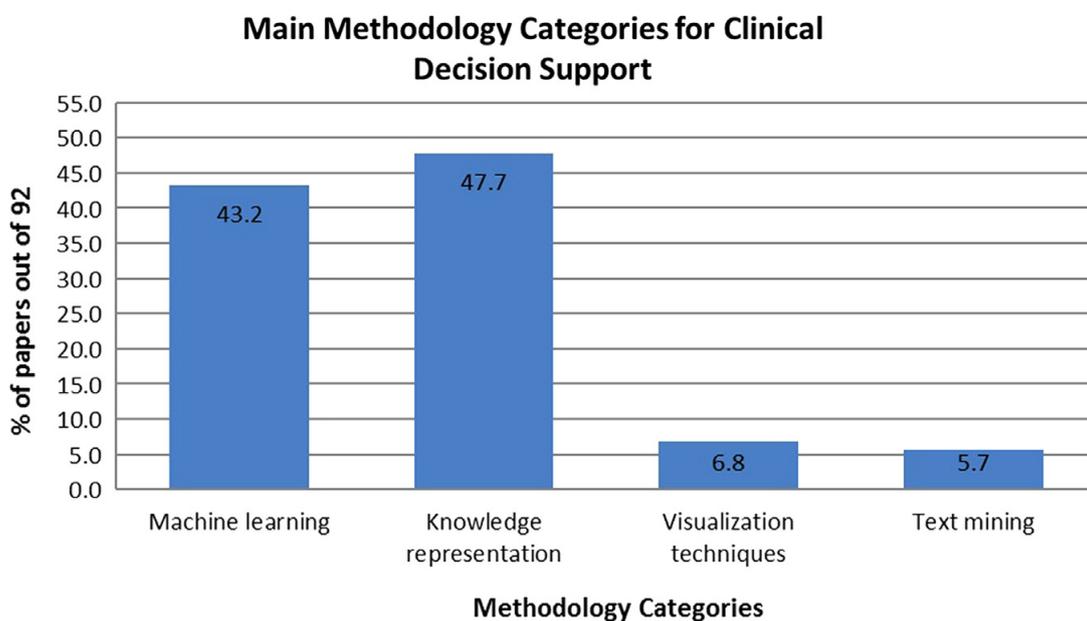


FIGURE 2.4 : Les principales méthodes de développement de SADM [Fra+15]

Un SBC met à disposition de l'utilisateur un ensemble de connaissances et un moyen automatique de raisonner sur ces connaissances [BMB10]. Dans de tels systèmes, la connaissance est centrale mais elle n'est rien tant qu'elle n'est pas convertie en décision, puis en action [Lti11], une fonction assurée par le raisonneur. Un SBC est donc constitué d'une base de connaissance et d'un raisonneur. Pour en faire un SADM, il faut en plus de choisir le bon raisonneur, et intégrer une Interface Homme Machine (IHM) (figure 2.5). La base de connaissances contient toutes sortes de règles permettant de prendre une décision. Le moteur d'inférence (raisonneur) accède à des retours d'expériences organisés sous forme d'enregistrements électroniques (comme le Dossier Médical Partagé : DMP) et exécute la logique de décision grâce à la base de connaissances. Le système de communication relie l'interface utilisateur au reste du système.

La représentation des connaissances est une partie prenante dans la réalisation des SADM. Elle ouvre également la voie aux techniques d'acquisition des données et est étroitement liée à la méthode de raisonnement [CBD01], d'où l'intérêt de l'ingénierie des connaissances pour le développement des SADM.

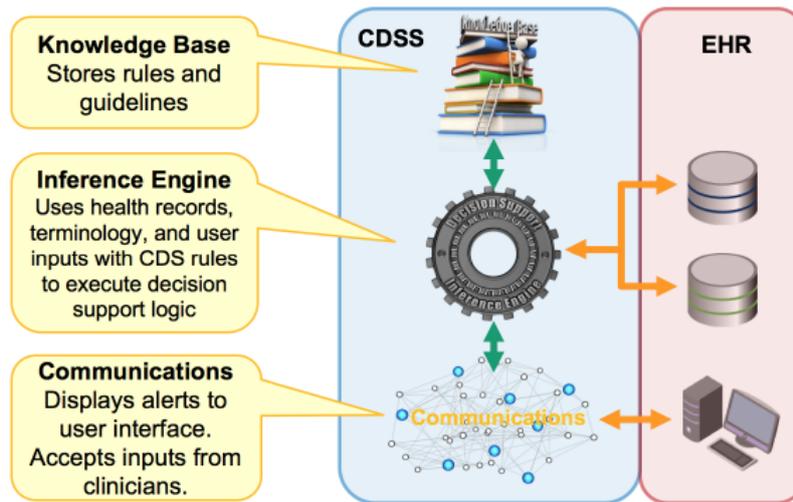


FIGURE 2.5 : Les composantes d'un SADM [IHT17]

2.5 Ingénierie des connaissances pour les SADM

L'ingénierie des connaissances se définit comme la partie de l'IA qui étudie et propose des concepts, méthodes, techniques et outils permettant d'acquérir, de modéliser, et de formaliser des connaissances dans le but d'opérationnalisation, de structuration ou de gestion au sens large. Elle trouve ainsi ses champs d'application dans les domaines où l'objectif est de modéliser les connaissances et les mettre à disposition comme support à une activité ou à un raisonnement. Les applications concernées sont celles liées à la gestion des connaissances, à la recherche d'information (notamment en exploitant les ressources sémantiques), et à l'aide à la décision [Mar+14].

L'opérationnalisation des connaissances dans les SADM, c'est-à-dire le passage des connaissances d'un état abstrait (humain), à un état compréhensible par la machine nécessite plusieurs étapes [PP07]. Dans ce sens, l'acquisition et l'Ingénierie des connaissances ont repris à leur compte les travaux de Newell [New82] qui, le premier, a différencié les connaissances à représenter dans un système et l'implémentation de ce système. Newell a fait apparaître la nécessité d'un niveau de description des systèmes qui ne soit pas celui des symboles et langages informatiques, le niveau des connaissances [Cha02]. Le support pour rendre compte de ce niveau a été appelé modèle conceptuel [Mar+14]. Le modèle conceptuel est vu comme une construction abstraite finalisée qui permet de réduire la complexité de l'expertise en se focalisant sur certains aspects des connaissances [NA18] [KLV90]. Il contient les connaissances du domaine et les connaissances de raisonnement. Cette modularisation répond au besoin de représenter par un modèle les structures de données utilisés par les programmes informatiques lors du développement d'applications informatiques [Héo14]. Le modèle conceptuel est le support à l'acquisition de connaissances, à la compréhension des méthodes de résolution à utiliser, à la communication, à la prédiction, à la médiation et à la conception.

L'idée de mieux construire et plus rapidement les modèles en réutilisant le plus possible des composants génériques, que ce soit au niveau du raisonnement ou des connaissances du domaine a conduit à l'utilisation d'ontologies en Ingénierie des connaissances [Cha02]. Définie à l'origine en 1993 par Thomas Gruber [Gru93] comme une spécification formelle explicite d'une conceptualisation, l'ontologie est conçue pour le partage des connaissances. L'ontologie contribue à une clarification sémantique et

contient une description hiérarchique des concepts et des relations qui peuvent exister pour une entité ou une communauté d'entités. Le développement d'ontologies est essentiel pour formaliser un vocabulaire commun pour la collaboration distribuée dans un domaine particulier.

Les ontologies informatiques ont habituellement trois rôles majeurs qui sont : 1) renforcer la communication à l'échelle humaine et organisationnelle 2) fournir l'interopérabilité sémantique au niveau machine et système et 3) aider dans le raisonnement et la résolution de problèmes. Le choix des ontologies au niveau conceptuel ouvre beaucoup de possibilités sur les choix des formalismes de représentation, et des méthodes de raisonnement.

2.5.1 Formalismes de représentation des connaissances

Trois modes principaux de représentation des connaissances biomédicales ont émergé à travers l'histoire de l'intelligence artificielle des 40 dernières années [Ros15] :

- les systèmes à base de règles, également appelées règles de production (business rules)
- les langages formels de représentation des guides de bonne pratique clinique
- les « ontologies » au sens des sciences de l'information, parfois appelées ontologies appliquées, afin de les différencier de l'Ontologie philosophique

Les règles de production étaient utilisées dans l'un des premiers systèmes experts (MYCIN) pour représenter les connaissances issues d'experts humains. Elles étaient sous la forme « SI <condition> ALORS <conséquence> ».

De nos jours, conformément au principe de la médecine fondée sur les données probantes, les connaissances objectives des recommandations de bonne pratique sont complétées par des Chemins cliniques décrivant pour une pathologie donnée, tous les éléments du processus de prise en charge du parcours du patient [SB14]. La formalisation nécessaire à l'intégration de ces connaissances dans des systèmes informatiques a toujours engendré de nombreux travaux [Hua+16] [Geo06] [Zha+16].

De nombreux langages de modélisation de guides de bonnes pratiques existent [Pel13] [Pel+03] : GLIF [Box+04], ASBRU [ESF14]. Par rapport à ces formalismes Galopin [Gal15] a fait ce constat : « Ces formats ne sont pas open-source et il n'est donc pas possible d'en avoir une idée objective et précise, ce qui a poussé les chercheurs à développer leurs propres formalismes en fonction de leurs besoins et attentes. On a ainsi observé pendant longtemps des initiatives multiples, mais qui finalement s'avèrent assez similaires ».

Aujourd'hui les ontologies constituent les principaux outils de conceptualisation des connaissances médicales [EE15] [Lee+13]. Deux principales familles de formalisme logique sont souvent utilisées pour représenter et raisonner avec des ontologies : les logiques de description [N+03] et les graphes conceptuels [Dou+15]. Dans le domaine médical, il est particulièrement frappant de constater une large utilisation du langage OWL (Web Ontology Language) qui est basé sur les logiques de description et est devenu une recommandation du World Wide Web Consortium (W3C) pour le web sémantique.

Les Logiques de Description sont conçus pour la représentation des connaissances d'un domaine d'application, ce qui en fait des candidats idoines pour la description des connaissances ontologiques. La dénomination Logique de Description vient du fait que les notions importantes du domaine sont des descriptions de concepts et ils diffèrent de leurs prédécesseurs, comme les réseaux sémantiques et les

cadres (frames), parce qu'ils utilisent une sémantique reposant sur une logique qui peut être transcrite en logique des prédicats du premier ordre [Bar09]. Toutefois, les limites de la logique classique concernant la représentation de certains concepts a entraîné la création d'autres logiques [Bra90]. En effet, le besoin de saisir la nécessité, le temps, la possibilité, les croyances des agents, etc. a donné naissance aux logiques modale, épistémique, conditionnel, déontique, dynamique, des prédicats etc. De nombreux travaux se sont donc intéressés à définir des extensions des logiques pour permettre la manipulation de notions plus complexes. Si ces logiques gagnaient en expressivité, elles n'apporteraient pas une réponse aux problématiques posées par Minsky [Min74]. Ce dernier soutenait que les formalismes de représentation des connaissances devraient être flexibles et informels et que les notions logiques d'exactitude et d'exhaustivité pouvaient poser problème dans un contexte de représentation des connaissances.

Les travaux autour des logiques de description et des ontologies se sont eux intéressés à définir des fragments plus simples de ces logiques. Ces simplifications visent autant l'aspect « acquisition des connaissances » (c'est-à-dire la facilité à formaliser les connaissances, à les retrouver, à les partager, à les maintenir. . .) que l'aspect « calculatoire » (c'est-à-dire la complexité des mécanismes de raisonnement associés) [Leh92] [Mar+14].

2.5.2 Acquisition de connaissances et raisonnement

La plupart des SADM disponibles aujourd'hui impliquent des systèmes basés sur les connaissances avec des règles, des guides de bonne pratique ou d'autres connaissances compilées dérivées de la littérature médicale [BT07]. Traditionnellement, les connaissances sous-jacentes à l'aide à la décision clinique proviennent de trois sources possibles [Gre14b] :

- « Human intensive » : le débriefing, l'organisation et la formalisation des connaissances des experts humains
- « Literature derived » : le passage en revue et la synthèse de la littérature scientifique. Les guides de bonne pratique et les chemins cliniques en sont des exemples [Ros15].
- « Data intensive » : l'extraction de connaissances à partir de données.

Ces trois manières d'acquérir la connaissance marque l'évolution des SADM et ont permis de développer trois formes prédominantes de méthodes de raisonnement pour l'aide à la décision médicale : les règles de décision [Jen14], les guides de bonnes pratiques et chemins cliniques [PG14] et les ontologies et terminologies médicales telles que SNOMED CT, et ICD [Huf+14].

Les guides de bonnes pratiques

L'objectif principal est d'offrir au médecin l'ensemble des informations qui lui permettront de prendre des décisions de manière éclairée en lui évitant de réaliser lui-même une étude poussée de la littérature, dont le caractère chronophage est incompatible avec sa pratique quotidienne de la médecine [Gal15].

Le chemin clinique

Le chemin clinique fournit une description efficace de la prise en charge médicale du patient pour une pathologie donnée. C'est un outil qui non seulement formalise la prévision des actes à réaliser mais précise aussi leur sens : risque infectieux ou risque d'anxiété [PG13], ce qui est indispensable dans notre contexte. En effet la formalisation de la prévision des actes est nécessaire dans les cas d'urgences tels que ceux relatifs à un incident médical dans un avion et surtout en l'absence d'un médecin. Cette adéquation

des chemins cliniques avec notre contexte ressort à travers sa différence avec les directives et autres algorithmes comme le soulignent les auteurs de [Bell1] : « les chemins cliniques sont différents des guides de bonne pratique, des protocoles ou algorithmes en cela qu'ils : (1) sont utilisés par une équipe pluridisciplinaire; (2) sont orientés vers la qualité et la coordination des soins. » Le chemin clinique détaille le processus de soins contenant des interventions cliniques concrètes pour des résultats escomptés et s'inscrit dans le cadre de l'interdisciplinarité [Ye+09], ce qui le rend intéressant dans le cadre d'une application de la télémédecine.

SNOMED CT

SNOMED Clinical Terms ou SNOMED CT est une terminologie clinique complète. Elle a été créée par la fusion de SNOMED RT (Reference Terminology) développée par le Collège des pathologistes américains et la troisième version des «Clinical Terms» développée par le British National Health Service. SNOMED CT a une base ontologique avec des concepts actifs (plus de 310000) et des relations. Elle est utilisée pour le codage de l'information clinique, l'indexation et la description pragmatique des documents médicaux. SNOMED CT présente une nomenclature multiaxiale couvrant les domaines de la médecine humaine et de la dentisterie. Elle est utilisée pour capturer et représenter les données des patients à des fins cliniques et administratives. SNOMED CT est actuellement disponible dans différentes langues, y compris l'anglais, l'espagnol, et le français. Elle peut servir de structure au dossier électronique du patient au cours des soins et peut donc être utilisée pour développer des outils de prise de décision. Les concepts du niveau le plus élevé dans SNOMED CT sont structurés en 19 éléments conceptuels différents (tableau 2.4). Chacun des éléments couvre différentes facettes de la santé

Dans ce contexte, nous fournissons une brève description de certains concepts les plus élevés de SNOMED CT. Ils sont sélectionnés en fonction de leur utilisation générale. Le tableau 2.5 montre certains concepts pertinents qui sont utilisés dans les dossiers cliniques avec des indications illustratives. Par exemple, le concept de « Clinical Finding » couvre tous les troubles cliniques et les résultats de l'examen. Les caractéristiques particulières d'un élément conceptuel peuvent être utilisées pour exprimer des relations sémantiques avec d'autres concepts. Par exemple, pour décrire certains attributs dans la caractérisation de l'état du patient, le concept relation « Finding site » indique la localisation d'une maladie sur le corps (par exemple, la pneumonie a pour localisation la structure pulmonaire).

SNOMED CT a une réelle capacité d'encodage automatique dès le stade du diagnostic. Ses composants de base (codes, descriptions, relations et ensembles de référence) permettent la caractérisation du patient et du traitement, et ils couvrent les signes et symptômes d'un patient. Depuis sa création, SNOMED-CT a cherché à couvrir le maximum de termes cliniques pour exprimer des scénarios cliniques [Sam+10]. Pour certains travaux, elle est considérée comme une base intéressante pour évaluer la richesse des terminologies médicales [YBB12]. Depuis Juillet 2003, SNOMED-CT est accessible gratuitement aux pathologistes américains et le Système d'information clinique du Centre médical Cedars-Sinai (États-Unis d'Amérique) a permis d'obtenir un ensemble de diagnostics de problèmes médicaux. L'intégration de SNOMED CT dans une application informatique est bien défini dans SNOMED CT Technical Implementation Guide [CT14] (figure 2.6). Le modèle formel sous-jacent à SNOMED CT assure le suivi du contenu sémantique et la représentation explicite de la connaissance médicale.

En raison de ses capacités de représentation des connaissances, SNOMED CT peut être utilisé dans différents contextes :

- La représentation de l'information clinique des patients [Dua+14]

Concept	Traduction	Concept	Traduction
Clinical Finding	Constatacion clinique	Physical object	Objet Physique
Procedure	Procédure	Physical force	Force Physique
Observable entity	Entité observable	Event	Événement
Body structure	structure corporelle	Environment	Environnement
Organism	Organisme	Social context	Contexte social
Substance	Substance	Situation with explicit content	Situation avec un contenu explicite
Pharmaceutical / biologic product	Produit pharmacologique / biologique	Staging and scales	Satisfaction et mesure
Specimen	Spécimen	Linkage concept	Concept de relation
Qualifier value	Valeur qualifiée	Special concept	Concept spécial
Record artifact	Enregistrement Artefact		

TABLE 2.4 : Les concepts de plus haut niveau dans SNOMED CT

- La représentation de connaissances cliniques
- L'analyse des systèmes d'informations médicales

Comme pour chaque outil de représentation de connaissances dans un domaine médical, SNOMED CT a besoin d'une couverture parfaite des concepts médicaux (figure 2.7), des concepts bien structurés et des relations qui garantissent la formalisation sémantique des modèles. Dans sa structure, SNOMED CT définit deux principaux types de relations entre concepts. La relation de subsomption (IS-A) qui définit le niveau d'expressivité (spécialisation / agrégation). La relation « attribut » qui permet la caractérisation des concepts en leur définissant des attributs. Ces relations visent principalement trois objectifs : expliciter la sémantique, automatiser la classification et permettre la post-coordination [Mer10]. La post-coordination est un procédé de création d'expressions plus riches à partir des concepts. Elle réduit considérablement le nombre de concepts qui peut permettre d'utiliser le minimum de causes primaires pour expliquer un phénomène.

Hiérarchie	Description	Exemples
Constatation clinique (Clinical finding)	<i>Quel phénomène est observé ?</i> Contient les sous-hiérarchies représentant le résultat d'une observation clinique, d'une évaluation ou d'un jugement. Important pour documenter les troubles cliniques, les symptômes et les signes.	<ul style="list-style-type: none"> • Gonflement du bras (constatation) • Pneumonie (maladie)
Procédure (Procedure)	<i>Que se passe-t-il ?</i> Activités utiles dans la prestation de soins de santé.	<ul style="list-style-type: none"> • Biopsie pulmonaire • Endoscopie diagnostique • Radiographie thoracique
Évènement (Event)	<i>Ce qui se passe ?</i> Décrit la situation autour de l'individu à un moment précis qui est pertinent pour ses soins de santé. Cela n'inclut pas les procédures ou les interventions qui se trouvent dans la hiérarchie « Procedure ».	<ul style="list-style-type: none"> • Crue brutale • Accident automobile • Exposition au virus de la rougeole
Entités observables (Observable Entity)	<i>Une observation quantitative ?</i> Les termes utilisés pour enregistrer les mesures, les lectures, les résultats numériques ou les dates et ont toujours une entrée de valeur associée.	<ul style="list-style-type: none"> • Poids • Taille de la tumeur • Date d'admission • Mesure de l'abduction de hanche
Produits pharmaco-logiques/biologiques (Pharmaceutical/biologic product)	<i>Un médicament ou autre substance utilisée pour traiter un patient ?</i> Cette hiérarchie est distincte de la hiérarchie des substances afin de distinguer clairement les produits médicamenteux (Products) des constituants chimiques (Substances) et des produits pharmaceutiques.	<ul style="list-style-type: none"> • Tamoxifen (product) • Tramadol (product) • Suppléments multivitaminés en comprimé (product) • Produit hormonal

TABLE 2.5 : Concepts SNOMED CT habituellement utilisés dans les dossiers cliniques sans autre élaboration [gui14]

2.6 Conclusion

Dans ce chapitre nous avons présenté le contexte de prise en charge des incidents médicaux dans l'avion. Nous avons aussi présenté les SADM dont nous avons décrit dans l'introduction générale comment ils peuvent améliorer la prise en charge médicale dans l'avion.

En considérant la multitude d'approches de SADM qui existent, nous avons cherché laquelle serait

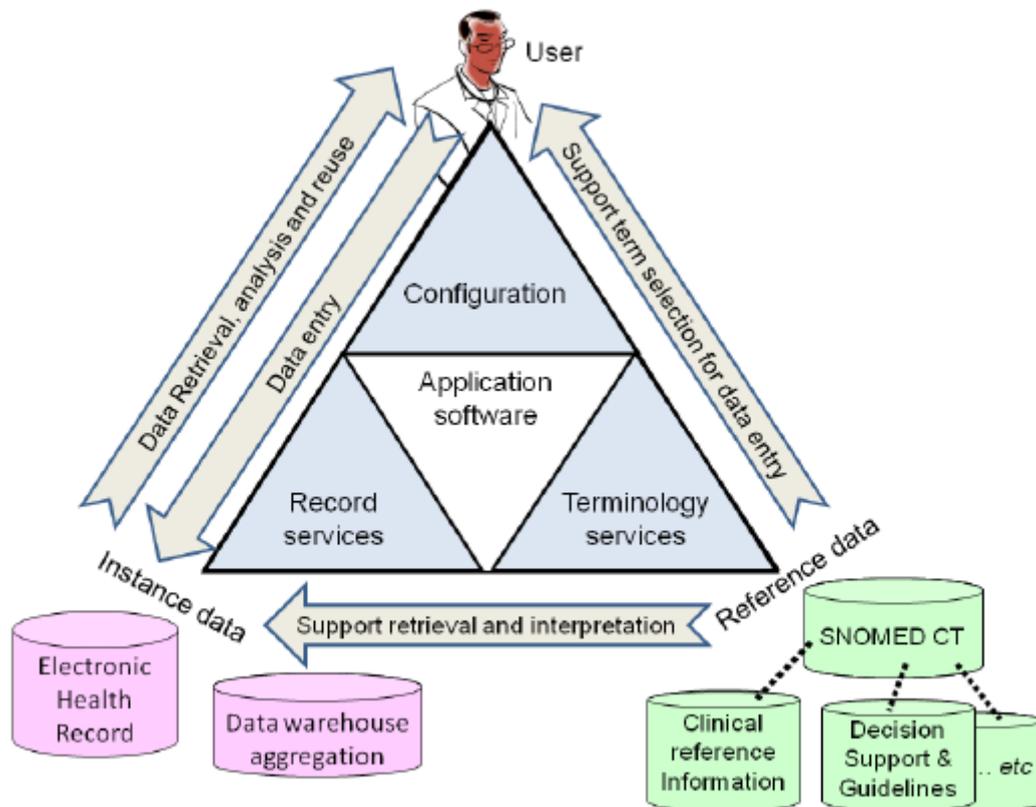


FIGURE 2.6 : SNOMED CT dans l'application

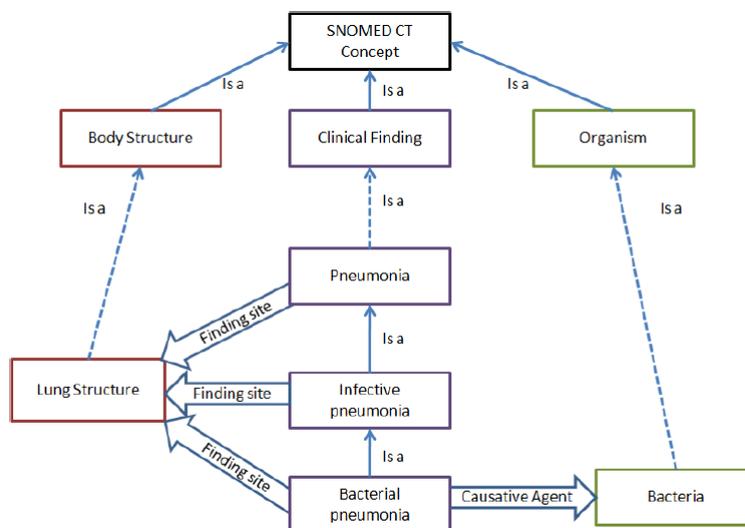


FIGURE 2.7 : Les relations entre concepts dans SNOMED CT

la mieux adaptée dans le contexte d'urgence médicale dans l'avion. Les deux approches qui ont marqué l'évolution des SADM sont les approches purement numériques et les approches symboliques avec l'IA. Des approches résultant du rapprochement entre SADM et d'autres domaines existent (ontologie, ECD, raisonnement sous incertitude, web, données multimédias, ...). C'est ainsi que l'intérêt a d'abord été porté sur les formalismes de représentation de connaissances dans lesquels le langage OWL basé sur les logiques de description semble bien positionné dans divers projets applicatifs. Ensuite, les différents

modes de raisonnement (les guides de bonnes pratiques, le chemin clinique et la terminologie SNOMED CT) ont été décrits de manière assez détaillée. Des précisions supplémentaires sur le déroulement de ses modes de raisonnement et les principes d'analyse sous-jacents seront apportées dans les chapitres ultérieurs.

La prise en charge médicale dans l'avion : de la télémédecine au SADM

3.1 Introduction

Dans ce chapitre, grâce aux travaux sur les incidents médicaux dans l'avion, nous avons analysé le questionnement et l'attente des intervenants dans la prise en charge médicale dans l'avion. Nous avons matérialisé cette analyse par une formalisation sémantique basée sur les ontologies pour préciser conceptuellement le vocabulaire des incidents médicaux se produisant durant les vols commerciaux. Nous allons décrire les étapes de la démarche utilisée qui est visualisable sur la figure 3.1.

- Nous avons dans un premier temps défini puis extrait tous les facteurs déterminants de la prise en charge médicale dans l'avion. La typologie des incidents médicaux qui constituent la base de ces facteurs a nécessité un mappage sémantique. Nous avons aussi lié les incidents médicaux aux rapports d'événements médicaux dans le système ASRS (Aviation Safety Reporting System).
- Ensuite à partir de l'analyse de la télémédecine, nous avons étudié le cadre d'interactions et d'échange de données dans lequel les facteurs déterminants de la prise en charge médicale dans l'avion sont définis.
- Dans une troisième activité nous avons rattaché l'incident médical dans l'avion à un concept médical en utilisant SNOMED CT. L'objectif de cette activité est de réduire cette ambiguïté terminologique de l'incident médical à bord pour lequel il n'y a pas de définition claire [Cos15].
- Pour finir, cette conceptualisation des incidents médicaux et sa spécification explicite avec la structure de SNOMED CT a permis de créer une ontologie de la prise en charge médicale dans l'avion.

Ces étapes ont été validées par deux articles scientifiques acceptés et publiés ([SKR18c] [SKR18b]).

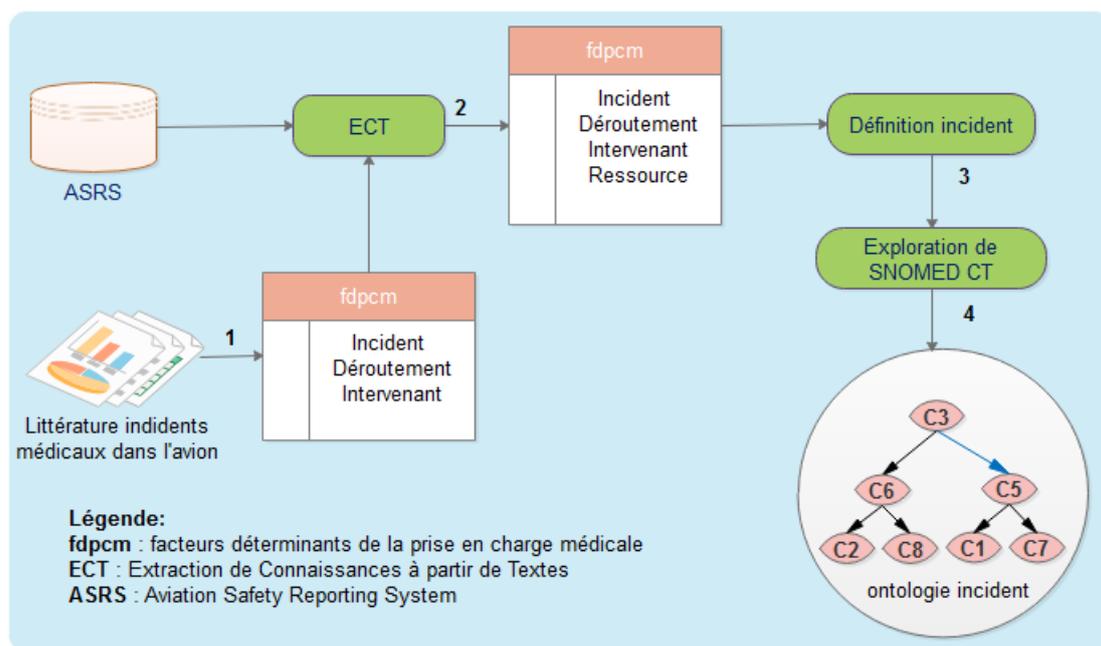


FIGURE 3.1 : Formalisation de la prise en charge médicale dans l'avion

3.2 Les facteurs déterminants de la prise en charge médicales dans l'avion

Définition 1. Nous considérons comme facteurs déterminants de la prise en charge médicale dans l'avion le type d'incident médical et tout autre facteur pouvant être relié à ce dernier.

Les facteurs déterminants de la prise en charge médicale dans l'avion définissent le questionnement de l'intervenant. Pour les trouver, nous avons mené une analyse faite d'études quantitatives et qualitatives sur les urgences médicales à bord. Dans notre cas, il s'agit de d'articles scientifiques et des manuscrits de thèses.

Les études sur les incidents médicaux dans l'avion (1) s'organisent de façons différentes, (2) donnent des noms différents à des incidents qui sont sémantiquement identiques, (3) utilisent des langages différents, (4) et traitent de problématiques différentes.

Nous avons d'abord effectué une recherche de journaux scientifiques sur PubMed (principal moteur de recherche des domaines de spécialisation de la biologie et de la médecine) du 01/10/2014 au 06/07/2017. Chandra et Conry [CC13] rapportent plusieurs recueils de données entre 1980 et 2010. Le travail décrit par [San+09] apporte des éléments de réflexion pertinents et intéressants, avec une plus grande population dans son échantillon d'étude.

Nous avons ensuite effectué une nouvelle recherche, cette fois-ci en ciblant les recherches originales et essais cliniques entre 2010 et 2017 dans PubMed et ScienceDirect (site web géré par l'éditeur Elsevier qui fournit une riche base de données de la recherche scientifique et médicale). Les études les plus pertinentes en termes de population, de dates, de variables et de citations sont fournies par Mahony [Mah+11], Peterson [Pet+13], Costa [Cos15] et leurs collègues.

D'autres études traitent des aspects particuliers comme le problème de matériels médicaux à bord [Hin+14] [Hin+17], ou de la prise en charge [BR16].

Le tableau 3.1 montre l'étendue de l'analyse des données sur les incidents médicaux dans l'avion. Sand et al. [San+09] et Mahony et al. [Mah+11] ont rétrospectivement recueilli chacun plus de 10000 cas sur des périodes assez longues (5 ans et 9 ans) et sur des compagnies aériennes situées dans des continents différents. Peterson et al. [Pet+13] recueillent directement les données auprès de cinq compagnies aériennes nationales et internationales et d'un centre de communications médicales dirigé par un médecin qui ont offert l'accès aux dossiers décrivant l'urgence médicale en vol de du 1er janvier 2008 au 31 octobre 2010. Marine Costa [Cos15] a diffusé un questionnaire aux médecins généralistes de la Réunion. Parmi les 97 réponses recueillies, 69 ont eu à s'occuper d'un passager malade lors d'un vol commercial.

Auteurs	Journal	Dates	Type d'étude	Nombre de cas
Sand [San+09]	Crit Care	01/2002-12/2007	Rétrospective de deux compagnies européennes	10,189/ 5 ans
Mahony [Mah+11]	Aviation, Space, and Environmental Medicine	01/1996-12/2004	Rétrospective de compagnies océaniques	11,326 / 9 ans
Peterson [Pet+13]	New England Journal of Medicine	01/2008-10/2010.	Rétrospective de compagnies de compagnies nationales et internationales recueillies par des médecins	11,920 /2 ans
Costa [Cos15]	Thèse	02/2015-04/2015	Réponse de 97 médecins généralistes à La Réunion	69 / 3 mois

TABLE 3.1 : Les caractéristiques des études

La plupart des études sont statistiques, et recherchent souvent des relations entre facteurs déterminants de la prise en charge médicale dans l'avion. C'est pourquoi les tableaux qu'elles contiennent constituent les meilleures sources de facteurs.

Type d'incident :

La catégorisation des incidents est l'aspect le plus maîtrisé dans la gestion des incidents médicaux dans l'avion. Malgré le manque de normalisation, les différentes études convergent sur la typologie et la prévalence des incidents (tableau 2.1).

Déroutement :

Il arrive parfois que la condition du passager victime d'un incident dans l'avion nécessite un déroutement. Ce déroutement qui doit être validé par le commandant de bord n'est pas possible dans tous les cas. En effet, pour qu'un déroutement puisse être validé, outre une raison médicale solide, plusieurs conditions doivent être réunies dont les suivantes :

- Les conditions sur les procédures de vol et d'atterrissage : le délai de mise en œuvre d'un atterrissage

est au moins égal à 20 minutes. On doit pouvoir se délester du carburant si nécessaire. Ces conditions sont totalement maîtrisées par le commandant et peuvent être intégrées sous forme de programme informatique dans le système.

- La disponibilité du centre médical adapté : selon la zone survolée dans le monde, la capacité de prise en charge d'un patient qui nécessite un déroutement n'est pas la même. La société internationale de services SOS qui propose des conseils de prévention médicale, d'accès aux soins et de sécurité édite chaque année une carte des risques médicaux et sécuritaires [Ris17]. Cette carte visualise la capacité de prise en charge médicale selon la zone géographique survolée. Certaines compagnies comme Air France classent les aéroports en fonction de la capacité à gérer les urgences médicale [Cos15].

3.3 Télémédecine en aéronautique

Les facteurs déterminants de la prise en charges médicale dans l'avion sont définis dans un cadre d'interactions, d'échanges de données entre les acteurs de la prise en charge. Ces interactions sont aussi des aspects importants de la prise en charge médicale et nous avons essayé de les comprendre à partir de la télémédecine aéronautique qui est le principal outil d'assistance dans l'avion.

La télémédecine est une pratique médicale à distance utilisant des technologies de l'information et de la communication. C'est une approche médicale inscrite dans la loi française [Tél10] et qui contribue à l'amélioration de l'accès aux soins de santé à travers cinq actes principaux [CNP12] : téléconsultation, téléexpertise, télésurveillance, téléassistance, régulation médicale.

Les systèmes de télémédecine mettent en jeu deux types d'interfaces : (1) une interface « Patient – Soignant » et une interface « Soignant – Soignant » (figure 3.2). Dans la prise en charge médicale dans l'avion avec une assistance médicale au sol, les deux types d'interfaces existent en même temps avec comme acteurs l'intervenant à bord, l'équipe médicale au sol, le patient auxquels on peut ajouter le PNC.

Le figure 3.3 décrit le cadre de la prise en charge médicale dans l'avion avec une téléassistance.

Comme dans le cadre général de la télémédecine l'interaction entre soignant se fait par visioconférence ou audioconférence. L'interaction entre le patient et le professionnel de la santé à bord est par contre directe. Les soignants reçoivent des données (mesures, signaux physiologiques, imagerie, ...) des dispositifs médicaux et peuvent dans le cadre du dossier médical partagé, accéder au dossier patient (compte-rendu d'hospitalisation ou de consultation, prescriptions médicamenteuses). Les dispositifs médicaux font partie du matériel de secours à bord. Ce matériel est organisé en plusieurs types les trousse dont le contenu et la présence dépend des autorités réglementaires. Un autre aspect intéressant de la télémédecine par rapport aux SADM est comme on le voit sur la figure 3.4, que sa taxonomie est incluse dans celle des SADM. Cela permet de déduire le contexte d'assistance médicale dans l'avion, les moyens de restitution des recommandations et l'infrastructure de communication d'un SADM à partir de l'expérience sur les systèmes de télémédecine. Dans le cadre d'un système d'aide à la décision médicale faisant usage de la taxonomie de la télémédecine, nous avons présenté un processus de raisonnement à partir des cas qui a été enrichi avec la prise en compte des preuves [SKR15]. Même si l'environnement considéré n'est pas identique à celui de l'avion, nous avons pu avoir un aperçu d'une application de la prise en charge médicale à distance sous contraintes de ressources, avec des informations instructives sur les moyens de restitution de recommandations et l'infrastructure de communication.

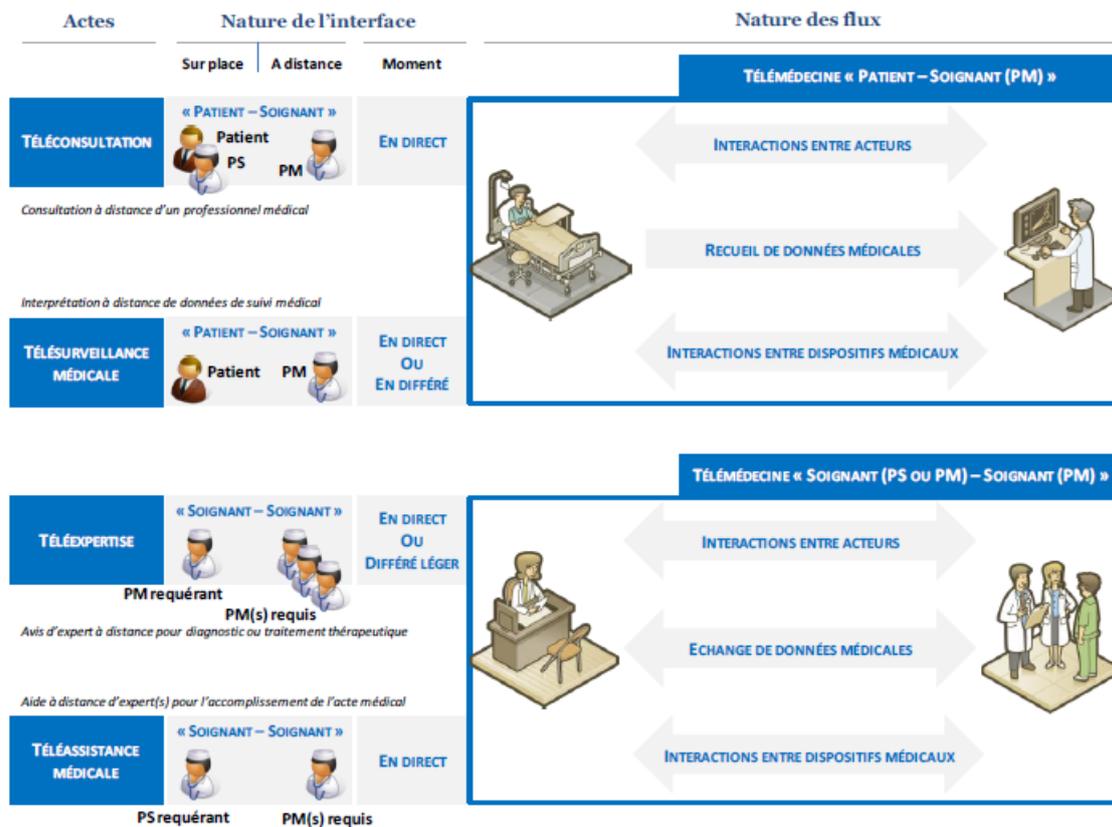


FIGURE 3.2 : Les actes de la télémédecine [CNP12]

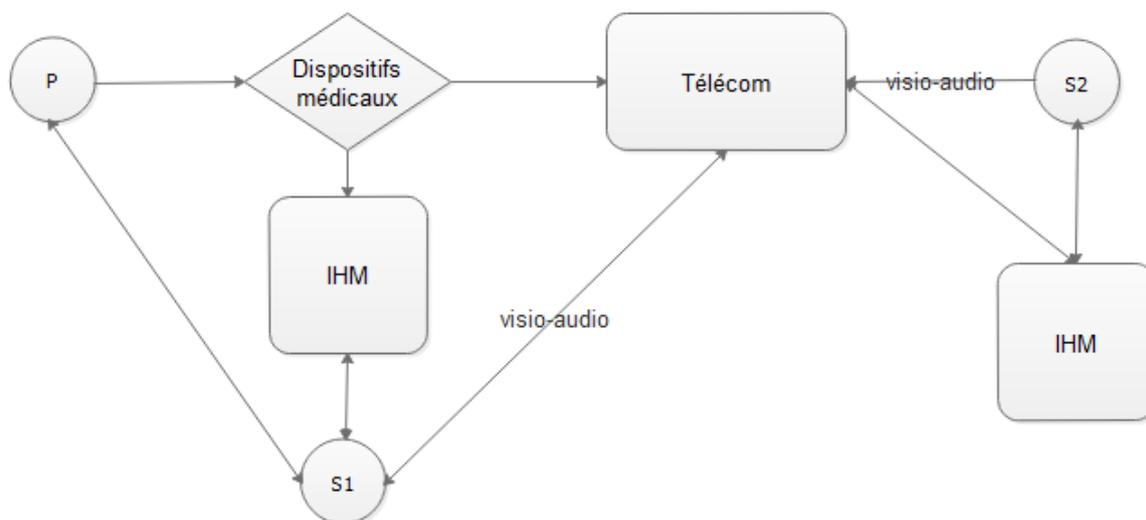


FIGURE 3.3 : Téléassistance médicale

3.4 Les systèmes de report d'évènement dans l'aéronautique

Les deux principaux systèmes de reports d'évènement dans l'aéronautique sont ASRS (Aviation Safety Report System) et ECCAIRS (European Co-ordination Centre for Accident and Incident Reporting Systems) [Tan+16].

ASRS est le plus ancien et le plus célèbre programme de déclaration d'incidents pour l'aviation.

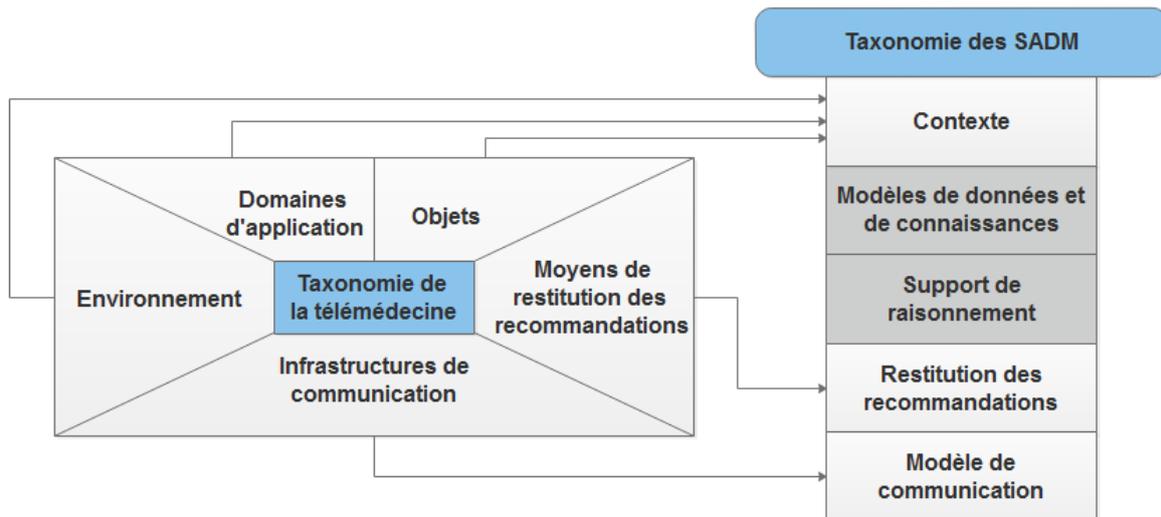


FIGURE 3.4 : Taxonomie de la télémédecine et du SADM

Il est géré par l'Administration fédérale américaine de l'aviation (FAA) et consiste en un système de déclaration volontaire et confidentiel dans l'intérêt d'améliorer la sécurité aérienne. Opérationnel depuis 1976, ASRS a traité plus d'un million de rapports d'incidents et fait en moyenne 6736 demandes par mois (322 par jour), avec un taux croissant au cours des années. Ce système cible plusieurs types d'événements provenant de différents types de rapports : les rapports généraux des pilotes, les rapports des contrôleurs du trafic aérien, les rapports de maintenance des techniciens mécaniques et les rapports du personnel de cabine.

Le système ECCAIRS est régi par une réglementation de l'Union Européenne numéro 376/2014 [14b]. C'est un effort continu visant à normaliser la collecte et l'échange de données sur les accidents et incidents au sein de l'Union européenne. Développé par le Centre commun de recherche de la Commission européenne, ECCAIRS a pour mission «d'assister les entités de transport nationales et européennes dans la collecte, le partage et l'analyse de leurs informations de sécurité afin d'améliorer la sécurité des transports publics». Il prend la forme d'une plate-forme logicielle couvrant la majeure partie de la collecte, de l'indexation et de l'interrogation des rapports d'incidents. Il est actuellement utilisé par plusieurs agences nationales, parmi lesquelles la DGAC.

Ces politiques de recueil et d'analyse des événements liés à l'aviation sont mis en place partout dans le monde et leur exploitation pourrait être bénéfique à la gestion des incidents médicaux dans l'avion. En enrichissant les incidents médicaux extraits des travaux de la littérature effectués par Mahony [Mah+11], Peterson [Pet+13] et Costa [Cos15] via l'usage de leurs synonymes dans SNOMED CT, nous avons montré que les 1493 rapports d'événements médicaux extraits de la base de données en ligne de ASRS sont chacun au moins lié à un de ces incidents.

3.5 Extraction de connaissances à partir de données (ECD)

L'ECD (Knowledge Discovery in Databases (KDD) en anglais) extrait des informations implicites valides, nouvelles et potentiellement utiles à partir des données. Elle vise à découvrir des modèles cachés et des connexions dans les données et fait référence au processus global de découverte de connaissances utiles à partir de données [All+17]. Même si certaines définitions utilisent interchangeablement les termes «

ECD » et « fouille de données », cette dernière est en réalité une étape spécifique du processus d'ECD. La fouille de données (l'exploration de données) est l'application d'algorithmes particuliers pour l'extraction de modèles (patterns) à partir de données.

En terme de positionnement scientifique, la fouille de données se situe à l'intersection de l'informatique et des statistiques mais contrairement à cette dernière l'objectif de la fouille de données n'est pas d'évaluer la vraisemblance d'hypothèses mais d'élargir le champs de connaissances à travers tous les outils et modèles de représentation disponibles et considérer des questions plus ouvertes que celles posées antérieurement par l'analyse statistique de données [Pen09].

Les données peuvent être structurées comme des éléments issus des bases de données, mais aussi non structurées comme des données décrites dans un simple fichier texte. Dans notre travail, on s'est intéressé aux rapports d'événements médicaux dans l'aviation civile sous forme de texte (articles, thèses, partie narrative de retour d'expérience). C'est pourquoi nous allons particulièrement nous intéresser dans ce qui suit à l'extraction de connaissances à partir de données textuelles.

3.5.1 Processus d'ECD

L'ECD est un processus qui nécessite d'abord la compréhension des données et l'identification des objectifs. Il implique plusieurs étapes à appliquer à l'ensemble de données afin d'en extraire des modèles utiles [All+17]. Ces étapes sont itératives et interactives et peuvent nécessiter des décisions de la part de l'utilisateur. L'étape de fouille de données constitue l'étape centrale mais souvent la plus transparente du point de vue de l'utilisateur qui est souvent aussi l'expert d'un domaine d'application considéré [Pen09]. Le processus commence par la sélection, le pré-traitement et la transformation des données en vue de leurs appliquer des méthodes de fouilles de données. Comme fouiller les données ne sert à rien si les résultats ne sont pas interprétables, le processus global d'ECD (figure 3.5) comprend l'évaluation et l'interprétation des patterns (fouille des données) afin de déterminer lesquels peuvent être considérés comme de nouvelles «connaissances» [Fay+96].

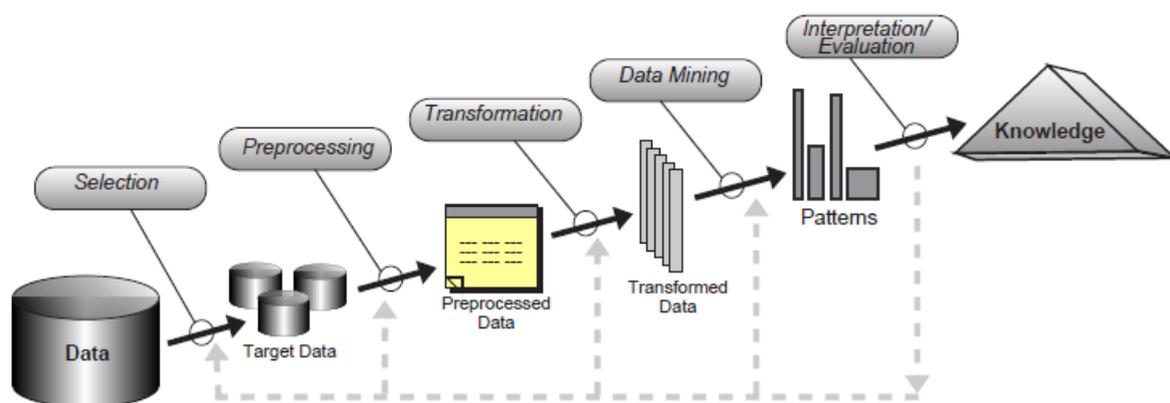


FIGURE 3.5 : Processus d'ECD [Fay+96]

Même si la décomposition en tâches du processus d'ECD peut légèrement varier d'une étude à une autres [BA94], [FPS96], [Lti11], [All+17], l'objectif de découvrir des nouvelles connaissances reste le même. En combinant ces travaux, nous distinguons huit tâches :

1. poser le problème : avant d'engager une action, il faut d'abord bien définir le problème ainsi que son contexte.
2. recherche des données : elle consiste à identifier les informations et les sources, vérifier leur qualité ainsi que leur facilité d'accès (documents papier, supports électroniques, fichiers internes ou externes, Data Warehouse ou Data Mart, etc.).
3. sélection des données : les données ayant un rapport avec l'analyse demandée sont retrouvées dans la base.
4. nettoyage des données : le nettoyage des données a pour objectif de corriger des inexactitudes ou erreurs de données.
5. transformation des données : les données sont regroupées, normalisées, et transformées dans un format qui les prépare à la fouille.
6. fouille de données : elle est définie comme l'art d'extraire des connaissances à partir des données. Il s'agit d'appliquer des méthodes intelligentes dans le but d'extraire des motifs (patterns).
7. évaluation du résultat : elle consiste à mesurer l'intérêt des patterns générés. L'évaluation est généralement qualitative (facilité de compréhension du résultat) et quantitative (tel que l'intervalle de confiance, etc.). Cette étape peut également impliquer la visualisation des patterns extraits, ou la visualisation des données selon des modèles extraits.
8. Consolider les connaissances découvertes : incorporer ces connaissances dans un système décisionnel, ou simplement les documenter et les signaler aux parties intéressées. Cela inclut également la vérification et la résolution des conflits potentiels avec les connaissances déjà connues (ou extraites).

3.5.2 Extraction de connaissances à partir de textes (ECT)

L'ECD consiste en une collection de techniques et d'outils qui détectent automatiquement des tendances, des modèles, des corrélations cachées dans les données et, en utilisant des techniques statistiques et d'Intelligence Artificielle, avec aussi la capacité d'identifier ce qui vaut la peine d'être examiné plus avant et ce qui peut être laissé de côté [Lti11]. Pour souligner la spécificité des données textuelles, Feldman et Dagan [FD95] introduisent l'extraction de connaissances à partir de textes. La revue de la littérature relative au domaine de l'ECT montre que les contributions proposées traitent de deux types d'approches [ZLW12], (1) les approches d'extraction et de représentation de connaissances à partir de textes et (2) les approches d'implémentation de solutions algorithmiques performantes, applicables à de grandes masses de données textuelles. Notre étude concerne plutôt le premier type d'approches.

L'ECT couvre un large éventail de sujets et d'algorithmes d'analyse de textes, couvrant diverses communautés, telles que la recherche d'information, le traitement du langage naturel, l'exploration de données, l'apprentissage automatique et de nombreux domaines d'applications tels que les sciences biomédicales [All+17].

Traitement automatique du langage naturel

On regroupe sous le vocable de traitement automatique du langage naturel (TAL), l'ensemble des recherches et développements visant à modéliser et reproduire, à l'aide de programmes informatiques, la

capacité humaine à produire et à comprendre des énoncés linguistiques dans des buts de communication [Yvo10]. Le traitement du langage naturel est un sous-domaine de l'informatique, de l'intelligence artificielle et de la linguistique. Un survol rapide des domaines qui se croisent avec l'ECT, montre que le Traitement du Langage Naturel est le domaine le plus privilégié [Lat13]. C'est un outil intéressant pour l'amélioration des procédures d'extraction d'informations. La plupart des algorithmes d'exploration de textes font largement appel aux techniques de traitement du langage naturel, telles que l'analyse syntaxique, l'étiquetage automatique des parties de la parole, l'extraction de termes pour prétraiter les données textuelles [Tan+16] [RB98]. Parmi les applications de l'ECT avec des techniques de TAL, on peut citer l'extraction d'informations à partir du texte.

Extraction d'informations à partir du texte

L'extraction d'informations sert généralement de point de départ pour d'autres algorithmes d'exploration de textes. On peut extraire des entités nommées (Reconnaissance d'entités nommées), des termsets (séquence de termes) par analogie à la terminologie itemset [Lat13], des termsets fermés, des termsets fermés fréquents, des règles d'association, des séquences fréquentes.

Une entité nommée est une séquence de termes qui identifie une entité du monde réel et la tâche de la reconnaissance d'entités nommées (Name Entity Recognition : NER) est de localiser et de classer les entités nommées d'un texte dans des catégories prédéfinies telles que la personne, la gestion, l'emplacement, etc [All+17].

Les termsets, les règles d'association, les séquences fréquentes sont des motifs (patterns) que peut contenir un texte. La notion de termset fermé provient de l'utilisation de l'Analyse Formelle de Concepts (AFC) dans l'extraction de motifs fréquents à partir de textes [Wil89]. L'AFC est une application de la théorie de treillis et particulièrement de treillis de Galois. Un termset est fréquent quand sa fréquence dans un ensemble de documents est supérieure à un seuil fixé.

Méthodes d'apprentissage automatique

Les méthodes D'ECT font appel aussi aux méthodes d'apprentissage automatique — dont l'objet est la conception d'algorithmes capables d'apprendre à résoudre un problème à partir d'exemples de solutions — pour induire des modèles qui soient représentatifs des données tout en étant robustes aux erreurs, et à la fois fournir les moyens algorithmiques efficaces pour pouvoir traiter de grandes quantités de données [Pen09].

Les méthodes d'apprentissage supervisé permettent d'inférer une fonction ou d'apprendre un classificateur à partir des données d'apprentissage afin d'effectuer des prédictions sur des données non vues. Il existe un large éventail de méthodes supervisées telles que l'apprentissage statistique, l'apprentissage par réseaux neuronaux, l'apprentissage par similarité (le raisonnement à partir de cas).

Les méthodes d'apprentissage non supervisées sont des techniques qui tentent de trouver une structure cachée à partir de données non labélisées. Elles n'ont pas besoin de phase d'apprentissage, et peuvent donc être appliquées à n'importe quelle donnée de texte sans effort manuel. Le clustering et la modélisation thématique sont les deux algorithmes d'apprentissage non supervisés couramment utilisés dans le contexte des données textuelles.

Fouille de données textuelles biomédicales

Il existe une quantité énorme de données textuelles sous formes de dossiers des patients, et de données sur l'assurance santé dans le domaine médical. La fouille de données textuelles au-delà de permettre aux acteurs du domaine médical d'accéder efficacement et d'extraire les connaissances dans de

grandes volumes de données, facilite et stimule la découverte biomédicale en augmentant l'extraction d'autres données biomédicales telles que les séquences de génomes et les structures de protéines [Gut+15]. L'extraction de connaissances à partir de données facilite la conversion de données médicales en preuves, ce qui très important dans le cadre de la médecine basée sur les faits.

Dans le processus d'ECT, le nettoyage et le prétraitement des données sont l'une des étapes les plus fastidieuses, car ils nécessitent des méthodes spéciales pour convertir les données textuelles en un format approprié pour les algorithmes d'exploration de données à utiliser [All+17]. Ces étapes réalisent un travail comparable à celui effectué grâce aux méthodes SIFT/HOG (Scale-invariant feature transform/ Histogram of oriented gradients) dans le traitement de l'image [Li+18] ainsi qu'au calcul des coefficients cepstraux (MFCC : Mel-frequency cepstral coefficients) dans le traitement de la parole [YPP15].

3.5.2.1 Représentation et codage de textes

Les données textuelles peuvent être traitées à différents niveaux de profondeurs [Lau16] :

- Identification de mots appartenant à des listes de mots (bag of words)
- Identification de chaînes de mots (phrases ou expressions)
- Identification d'éléments sémantiques (reconnaissance des mots par leur sens)

Ces niveaux de profondeur définissent la nature des connaissances extraites à partir de textes. Ces connaissances sont des patterns (motifs) extraits d'un corpus textuel et peuvent être déclinées sous différentes formes [Lat13] :

- ❖ un ensemble de termes (termset) fréquents
- ❖ une séquence fréquente de termes
- ❖ une règle d'association entre termes validée par des mesures statistiques, tels que le support et la confiance [AS94].

Le contexte d'extraction de connaissances à partir de données, appelé aussi contexte formel est un triplet $(\mathcal{D}, \mathcal{I}, \mathcal{R})$, où \mathcal{D} et \mathcal{I} sont deux ensembles finis et \mathcal{R} une relation binaire entre \mathcal{D} et \mathcal{I} , $\mathcal{R} \subseteq \mathcal{D} \times \mathcal{I}$. Dans la fouille de données traditionnelle, \mathcal{D} est un ensemble de transactions et \mathcal{I} , un ensemble d'attributs appelés aussi items. Dans le cadre de l'extraction de connaissances à partir de données textuelles, nous avons des documents à la place des transactions et les items (attributs) sont des termes. Par commodité nous utiliserons la notation $(\mathcal{D}, \mathcal{T}, \mathcal{R})$ pour le contexte.

- $\mathcal{D} = \{d_1, d_2, \dots, d_n\}$ est un ensemble de documents textuels
- $\mathcal{T} = \{t_1, t_2, \dots, t_m\}$ est l'ensemble des termes dans les documents, sans duplication
- $\mathcal{R} \subseteq \mathcal{D} \times \mathcal{T}$ est une relation binaire d'incidence. Chaque couple $(d, t) \in \mathcal{R}$ indique que le document $d \in \mathcal{D}$ contient le terme $t \in \mathcal{T}$.

Pour construire l'ensemble \mathcal{T} , le corpus de textes subit de nombreuses transformations successives.

La tokenisation

La tokenisation est une analyse lexicale qui convertit les textes en une liste de mots ou expressions. Certains caractères comme les signes de ponctuations sont souvent ignorés.

La racinisation (stemming) et la lemmatisation

Les mots obtenus de la tokenisation sont transformés en leur racine (stems) ou en forme canonique (lemmes).

Le Filtrage

Le filtrage permet de supprimer certains mots jugés inutiles. C'est souvent le cas des « mots vides » (stop words). Les « mots vides » sont les mots qui apparaissent fréquemment dans le texte sans avoir beaucoup d'informations sur le contenu (les prépositions, les conjonctions, etc.).

Toutes ces transformations du corpus permettent de construire l'ensemble \mathcal{T} qui est ainsi composé de stems ou de lemmes. Nous allons présenter ci-dessous deux cadres théoriques appliqués au contexte d'ECT que nous utiliserons, l'analyse de concepts formels (AFC) [BTN10] et le modèle vectorielle (sémantique vectorielle) avec l'analyse sémantique latente (ASL). L'AFC est un excellent cadre pour l'extraction de connaissances (termsets et règles d'association) [Lat13] et la sémantique vectorielle est un cadre pour l'extraction de caractéristiques (feature selection), la réduction de dimensions, l'apprentissage, etc.

3.5.2.2 L'analyse formelle de concepts

Le contexte d'extraction peut être caractérisé par la notion de concept. Un concept est défini par un ensemble d'objets (extension), auquel s'applique un ensemble d'attributs (intention). Dans le cadre de l'ECT, l'extension est un ensemble de documents et l'intention un ensemble de termes ou de séquences de termes. Les concepts seraient de type $c = (D, T)$, avec $D \in \mathcal{D}$ un ensemble de documents et $T \in \mathcal{T}$ un termset. Le concept $c = (D, T)$ est formel s'il vérifie la condition suivante : si $D \times T \subseteq D' \times T' \subseteq \mathcal{R}$ alors $D = D'$ et $T = T'$ [Sam14]. Cette condition qui fait correspondre un ensemble de documents à un ensemble de termes est déterminée par les opérateurs de la connexion de Galois [MZB06]. Une connexion de Galois est un couple de fonctions (ψ, ϕ) définie de la manière suivante :

$$\psi(T) = \{d | d \in D \wedge \forall t \in T, (d, t) \in \mathcal{R}\}$$

$$\phi(D) = \{t | t \in T \wedge \forall d \in D, (d, t) \in \mathcal{R}\}$$

La fonction $\psi(T)$ représente l'ensemble des documents contenant tous les termes de T et $\phi(D)$, l'ensemble des termes apparaissant dans tous les documents de D . La cardinalité de $\psi(T)$, appelé aussi support de T est noté $supp(T)$.

Et ainsi, $c = (D, T)$ est un concept formel s'il vérifie la condition suivante : $\phi(D) = T$ et $\psi(T) = D$. Ces fonctions permettent de vérifier les propriétés des documents [LHH12]. La fonction $\Omega = \phi \circ \psi$, appelée opérateur de fermeture de Galois associe à un termset T l'ensemble de tous les termes qui apparaissent dans les documents ou les termes de T co-occurrent. Elle définit ainsi un ensemble maximal de termes commun à un ensemble de documents. Si l'image de T par l'opérateur de fermeture est égale à lui-même $\Omega(T) = T$, alors T est un termset fermé.

Sans entrer dans les détails, on voit bien que les notions de concept formel, de termset fermé, et d'autres notions qui leurs sont associées (générateur, et classe d'équivalence) ainsi que les propriétés

qui en découlent définissent un cadre d'extraction de connaissances comme les termsets fréquents, et les règles d'association. Il suffit de fixer un seuil permettant de déterminer lesquels des termsets sont fréquents ou pas. Une règle d'association quant à elle est une union de termsets fermés. En effet une règle associative est une relation causale de la forme $r = T_1 \implies T_2$, où T_1 (prémisse) et T_2 (conclusion) sont des termsets distincts. La règle r est générée à partir du termset fermé fréquent $T = T_1 \cup T_2$.

$$supp(r) = supp(T) = supp(T_1 \cup T_2) = |\psi(T_1 \cup T_2)|$$

$$conf(r) = \frac{supp(T)}{supp(T_1)} = \frac{supp(T_1 \cup T_2)}{supp(T_1)} = \frac{|\psi(T_1 \cup T_2)|}{|\psi(T_1)|}$$

3.5.2.3 Sémantique vectorielle

Une autre manière différente de l'analyse formelle de concepts de caractérisation du contexte d'extraction textuelle est de représenter les documents comme des vecteurs et d'attribuer une importance numérique (poids) à chaque terme [All+17]. Dans la sémantique vectorielle, les documents sont convertis en vecteurs numériques, qui prennent en compte le nombre d'occurrences de chaque terme mais ignore l'ordre. Deux méthodes de pondération sont particulièrement utilisées :

- La méthode booléenne : pour tout document $d \in D$, un poids $p_{dt} > 0$ est attribué au terme t si $(d, t) \in \mathcal{R}$. Et si $(d, t) \notin \mathcal{R}$ alors $p_{dt} = 0$.
- La méthode Term Frequency-Inverse Document Frequency (TF-IDF) : La méthode TF-IDF est basée sur la loi de Zipf qui est une observation empirique de la fréquence des mots dans les documents [AA17] et dont l'interprétation théorique fait appel à la notion d'entropie. Le poids d'un terme dans un document est d'autant plus grand que le nombre d'occurrences du terme est élevé dans le document et que le terme est rare en général [Lau16]. Pour calculer le poids d'un terme dans un document, la méthode TF-IDF combine la fréquence de ce terme à la fréquence inverse du document. Cette dernière consiste à calculer le logarithme (en base 10) de l'inverse de la proportion de documents du corpus qui contiennent le terme. Soit n_{dt} le nombre de termes t dans d ; n_{dT} le nombre total de termes dans d ; N_{Dt} le nombre de documents contenant t ; $|D|$ le nombre total de documents.

$$p_{dt} = \frac{n_{dt}}{n_{dT}} * \frac{|D|}{N_{Dt}}$$

3.6 Modélisation thématique pour la définition de l'incident médical dans l'avion

Il n'y a pas de définition claire d'un incident médical à bord, et cela conduit à des données peu exploitables. Chaque incident décrit est probablement le produit d'une combinaison de facteurs, notamment les comorbidités du patient, l'environnement en vol, l'alcool et les drogues [Cos15]. L'objectif de cette section est de réduire cette ambiguïté de l'incident médical à bord en l'associant à un concept médical. Et pour cela nous allons utiliser SNOMED CT (figure 3.6).

SNOMED CT est la terminologie internationale de santé la plus détaillée pour décrire un dossier médical électronique [Mer10]. Elle facilite la caractérisation du patient dans les SADM. Son implémentation automatise l'accès au dossier patient et permet la mise en place de la médecine fondée sur des faits.

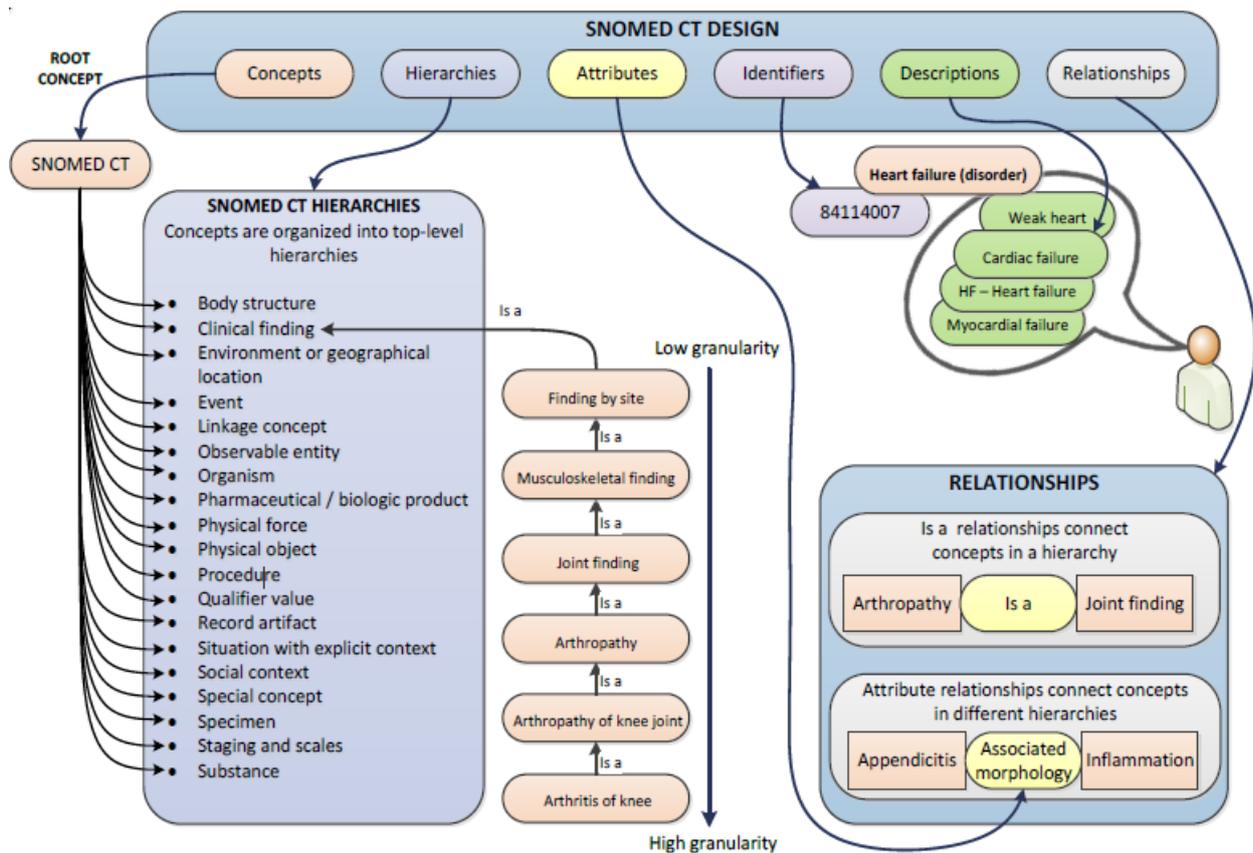


FIGURE 3.6 : Les composants de SNOMED CT [CT14]

La représentation du contenu de SNOMED CT est réalisée grâce à trois types de composants [CT14] :

- Concepts : les concepts sont des expressions utilisées dans le domaine médical. Ils sont organisés dans une hiérarchie et possède chacun un identifiant numérique unique.
- Descriptions : Pour faciliter leur compréhension, chaque concept possède plusieurs descriptions qui sont des synonymes représentant les mêmes concepts (figure 3.7). Chaque description possède un identifiant numérique unique.
- Relations : SNOMED CT fournit une sémantique à ses concepts en utilisant deux types de relations. La relation de subsomption « IS-A » et les relations attribuées.

Nous nous sommes posés la question (Q1) suivante :

« Selon la conceptualisation du domaine médical dans SNOMED CT, qu'est ce qui ferait de l'incident médical dans l'avion un élément conceptuel sémantiquement mieux identifiable ? »

La première solution pour répondre à cette question est la lecture exhaustive de la documentation de SNOMED CT associée à la compréhension du contexte des incidents médicaux dans l'avion. Nous proposons l'utilisation de la sémantique vectorielle et de la modélisation thématique.

La représentation vectorielle peut être analysée avec des algorithmes d'apprentissage, de classification, de clustering, de calcul de similarité mais surtout de réduction de dimensions. La réduction de dimensions est souvent utilisée en amont des autres algorithmes. Dans son étude de cinq techniques de réduction de

```

1 SELECT conceptId, term FROM `Description`
2 WHERE conceptId=90129003

```

	conceptId	term
1	90129003	Tussive syncope
2	90129003	Laryngeal vertigo
3	90129003	Cough syncope
4	90129003	Laryngeal syncope
5	90129003	Tussive syncope (disorder)
6	90129003	Charcot vertigo
7	90129003	Tussive syncope (finding)

FIGURE 3.7 : Exemple de description de concept selon SNOMED CT : syncope de la toux (perte temporaire de conscience après des accès de toux) avec différents termes synonymes en anglais

dimensions, Underhil [Und07] a montré que, pour n'importe quel nombre fixe de dimensions utilisées par un classificateur, l'exécution de tout type de réduction de dimensions améliorerait presque toujours la précision de la classification par rapport à l'utilisation des données non réduites. La réduction de dimensions a entre autres pour objectifs [Cru17] :

- La réduction du volume de données à traiter, tout en conservant au mieux « l'information utile ».
- L'amélioration du rapport signal / bruit et la suppression des variables non pertinentes.
- L'amélioration de la « lisibilité » des données en mettant en évidence des relations entre variables ou groupes de variables ou en facilitant la visualisation.
- L'atténuation de la « malédiction de la dimension » (curse of dimensionality). La « malédiction de la dimension » est utilisée pour nommer les problèmes liés à la grande dimension de la représentation des données. Dans le contexte d'extraction $(\mathcal{D}, \mathcal{T}, \mathcal{R})$, le nombre de dimensions pour représenter $d \in \mathcal{D}$ est égale à la cardinalité de \mathcal{T} .

Ceux sont les mêmes objectifs qui sont recherchés dans la modélisation de thématiques. La réduction de dimensions est un processus de transformation d'une grande quantité de données en une représentation (un espace) intermédiaire beaucoup plus petite et moins bruyante qui préserve les relations importantes des données d'origine [Und07]. En modélisation de thématiques, cet espace intermédiaire correspondant à un ensemble de thèmes (topics en anglais). Les thèmes ou concepts sont en réalité des composantes principales combinaisons linéaires des termes. Ils peuvent être calculés via la diagonalisation de la matrice de variance-covariance (ou la matrice de corrélation), par le biais de la décomposition en valeurs singulières (DVS) en Analyse Sémantique Latente (ASL) [Dub04] ou via le modèle probabiliste Allocation de Dirichlet Latente (ADL) [BNJ03].

La modélisation de thématiques

Elle est utilisée en apprentissage automatique ou en traitement automatique du langage naturel et s'appuie sur une représentation multidimensionnelle de la signification des mots dans la langue considérée. Son objectif est de faire ressortir les relations sémantiques entre termes ou entre documents [Zam05]. Pour atteindre cet objectif, l'idée est d'essayer de créer un espace intermédiaire, composé de thèmes (ou concepts) [Lau16] et dans lequel les termes exprimeraient davantage les distinctions sémantiques utiles. Dans ce nouvel espace porteur de sens, chaque terme et chaque document sera représenté par un vecteur de dimension égale au nombre de thèmes. Ce nombre de thèmes devient la dimension du problème et rend plus efficace les algorithmes de clustering, et d'apprentissage.

L'ASL est une technique de traitement du langage naturel [Kin02] développée par les laboratoires Bellcore en 1989. Elle applique une décomposition en valeurs singulières (SVD) à la matrice terme-document et permet d'en obtenir une approximation de rang inférieur. En réalité, il s'agit simplement d'une technique factorielle équivalente à l'Analyse en Composantes Principales (ACP) où les variables ne sont ni réduites, ni centrées, avec les mêmes objectifs et les mêmes outils pour évaluer la qualité de représentation [Rak13].

Comme l'ASL, l'objectif de l'ADL est construire des classes ou des hiérarchies de documents thématiquement homogènes à partir d'un corpus. Elle a été proposée dans [BNJ03] et consiste à envisager des associations non-déterministes entre classes et documents, permettant de dériver des représentations numériques synthétiques des documents, qui sont alors vus comme des points dans un espace sémantique latent [RCY07].

Analyse par composantes principales (ACP)

L'idée générale dans l'analyse par composantes principales est de créer un nouvel espace. Pour que cet espace soit de dimension plus petite, on cherche une rotation de l'espace initiale qui fasse apparaître les nuages de points de telle façon que certains axes soient presque inutiles pour décrire ce nuage de points [Lau16]. Ce nouvel espace sera formé de nouveaux axes appelés composantes principales qui passent par le maximum de variance du nuage de points. Les composantes principales sont construites séquentiellement et de façon perpendiculaire. On détermine d'abord l'axe (composante) sur lequel le nuage de points se déforme le moins en projection. Ensuite on recherche un autre axe orthogonal au premier sur lequel le nuage de points se déforme le moins en projection et ainsi de suite jusqu'à atteindre le nombre de composantes voulues.

Si on considère un contexte d'extraction $(\mathcal{D}, \mathcal{T}, \mathcal{R})$ et la matrice $M = \mathcal{D} \times \mathcal{T}$ deux analyses sont donc possibles :

- analyse du nuage de points des documents
- analyse du nuage de points des termes

Le rôle de l'ACP est de trouver le sous-espace sur lequel projeter le nuage tel que la déformation subie soit la plus faible possible et permette ainsi de récupérer les liens les plus significatifs entre les termes et les documents [Reb17]. Ce sous-espace correspond à celui qui diagonalise la matrice de variance-covariance ou la matrice de corrélation [Lau16].

Pour l'analyse du nuage de points des documents, le sous-espace cherché est généré par les vecteurs propres u_α associés aux plus grandes valeurs propres λ_α de la matrice $M^T M$ qui correspond à la matrice des covariances empiriques ou à la matrice des corrélations empiriques selon que les données sont centrées

ou normées [CAB04]. Ces valeurs et vecteurs propres sont reliés aux valeurs singulières obtenues par décomposition en valeurs singulières [GZ05] de M . La matrice M peut être factorisée de la manière suivante :

$$M = D \times T = U\Sigma V^T$$

U et V sont des matrices unitaires et Σ contient les valeurs singulières de la matrice M dans ses coefficients diagonaux. Si on effectue la décomposition en valeurs singulières de la matrice de covariance, on obtient le résultat suivant :

$$\rightarrow M^T M = (U\Sigma V^T)^T (U\Sigma V^T)$$

$$\rightarrow M^T M = V\Sigma^T U^T U\Sigma V^T$$

$$\rightarrow M^T M = V\Sigma^2 V^T$$

Le résultat est une diagonalisation (décomposition en valeurs propres) de la matrice $M^T M$. Par conséquent les valeurs propres recherchées sont les carrés des valeurs singulières de M et V représente la matrice des vecteurs propres.

Pour l'analyse des nuages de termes la matrice à diagonaliser est MM^T .

$$\rightarrow MM^T = (U\Sigma V^T)(U\Sigma V^T)^T$$

$$\rightarrow MM^T = U\Sigma^T V^T V\Sigma U^T$$

$$\rightarrow MM^T = U\Sigma^2 U^T$$

Comme dans l'analyse des nuages de points des termes, les valeurs propres recherchées sont les carrés des valeurs singulières de M . Par contre la matrice U contient les vecteurs propres en ses colonnes.

Choix des composantes

Les composantes principales sont les directions importantes du nuage multidimensionnel de points formé par les données. En diagonalisant les matrices citées ci-dessus, on obtient les directions du nuage de points portés par les vecteurs propres ainsi que la quantité de variance sur chaque direction indiquée par les valeurs propres. Le vecteur propre correspondant à la plus grande valeur propre donne la direction de plus grande dispersion du nuage de points. Et comme on sait que la somme des valeurs propre d'une matrice est égale à la variance contenue dans l'ensemble des données, on peut calculer le pourcentage de la variance cumulée par un nombre donné de valeurs propres. Si ce pourcentage est jugé suffisant (principe de Pareto), les vecteurs propres correspondants à ces valeurs propres deviennent les composantes principales.

3.6.1 Matériels

PyMedTermino (Medical Terminology for python) [LVD15]

C'est un module Python pour faciliter l'accès et la transformation des terminologies médicales généralement disponibles sous forme de fichiers texte. Les principales caractéristiques de PyMedTermino sont les suivantes : (i) Une interface de programmation (API) pour accéder à toutes les terminologies, (ii) une recherche optimisée de textes associés à des terminologies, (iii) un accès aux termes, synonymes

et traductions, (iv) la gestion les concepts et les relations entre concepts, et (v) des mappages entre les terminologies.

SQLite [SQL]

C'est un système de gestion de base de données relationnelle qui implémente la plupart des normes SQL (Structured Query Language). Il stocke la base de données complète (définitions, tableaux, indices et les données elles-mêmes) en tant que fichier multi-plateforme unique. SQLite prend en charge les principales fonctions SQL, les fonctions de date et heure, les fonctions agrégées et est connecté aux programmes de l'environnement de développement intégré Eclipse. Il facilite l'utilisation des fonctions de manipulation de données dans les fichiers SNOMED CT.

3.6.2 Méthode

La figure 3.8 décrit la méthode utilisée pour associer aux incidents médicaux dans l'avion un concept médical de SNOMED CT.

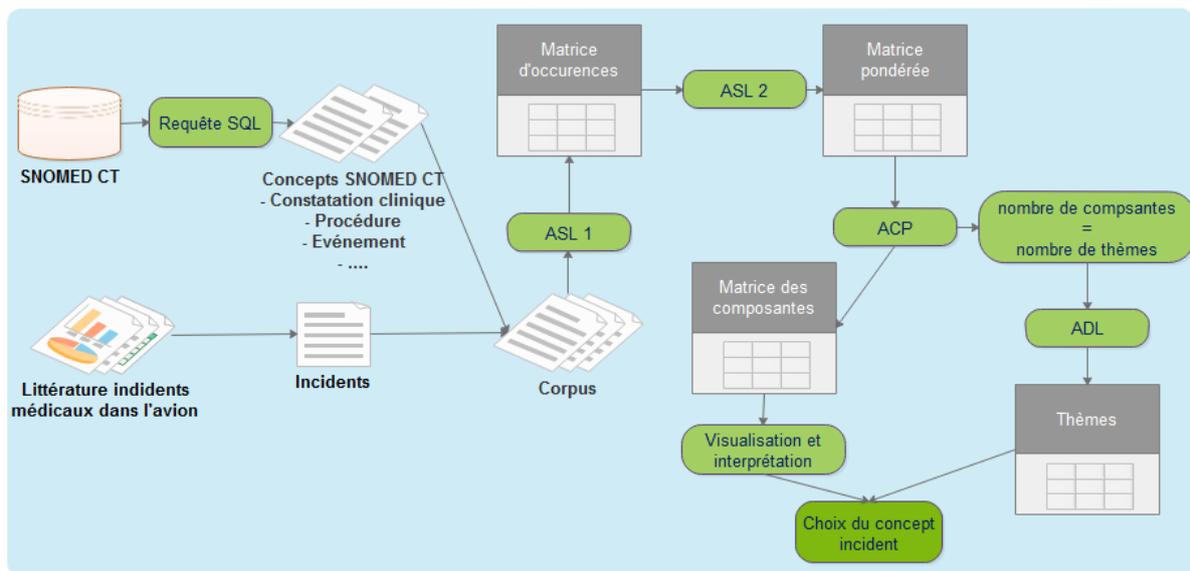


FIGURE 3.8 : Définition de l'incident médical dans l'avion

3.6.2.1 Création du corpus

Le corpus est composé d'un document texte contenant les incidents médicaux dans l'avion avec en plus pour chaque concept de SNOMED CT un document contenant sa description. La description des concepts est obtenue par interrogation d'une base de données SQL contenant tous les concepts et les relations de SNOMED CT. Cette base de données d'environ un demi-gigaoctet de mémoire a été construite grâce à l'outil pyMedTermino (figure 3.9) et à partir des données proposées par le National Library of Medicine (NLM pour Bibliothèque nationale de médecine) des États-Unis d'Amérique.

3.6.2.2 Construction de la matrice d'occurrences

Nous avons utilisé le corpus créé précédemment pour construire le contexte d'extraction de connaissance $c_e = (D_c, t/t \in T_i, \mathcal{R})$ utilisé pour établir les composantes principales. Le contexte est composé des

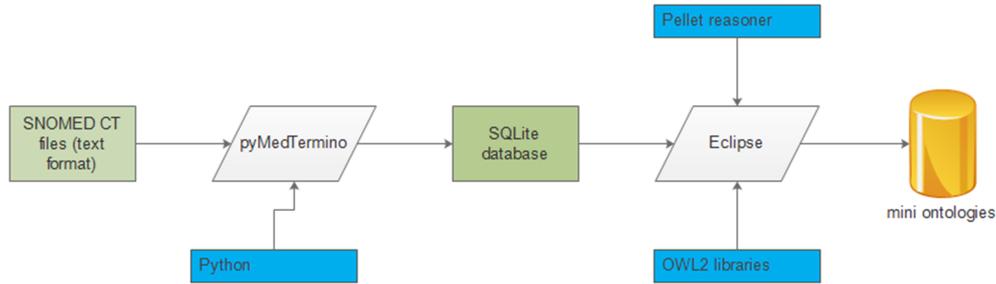


FIGURE 3.9 : Processus d'exploration de SNOMED CT

documents décrivant les concepts SNOMED CT (D_c) et des termes décrivant les incidents médicaux en vol (T_i). La relation binaire d'incidence \mathcal{R} donne la matrice d'occurrences (figure 3.10). Cette relation définit les couples (d, t) tel que le document $d \in D_c$. L'idée étant de construire à partir de la description de concepts dans SNOMED CT et de la description des incidents dans la littérature, la matrice croisant les degrés d'appartenance de chaque incident en vol et les types de concepts de SNOMED CT.

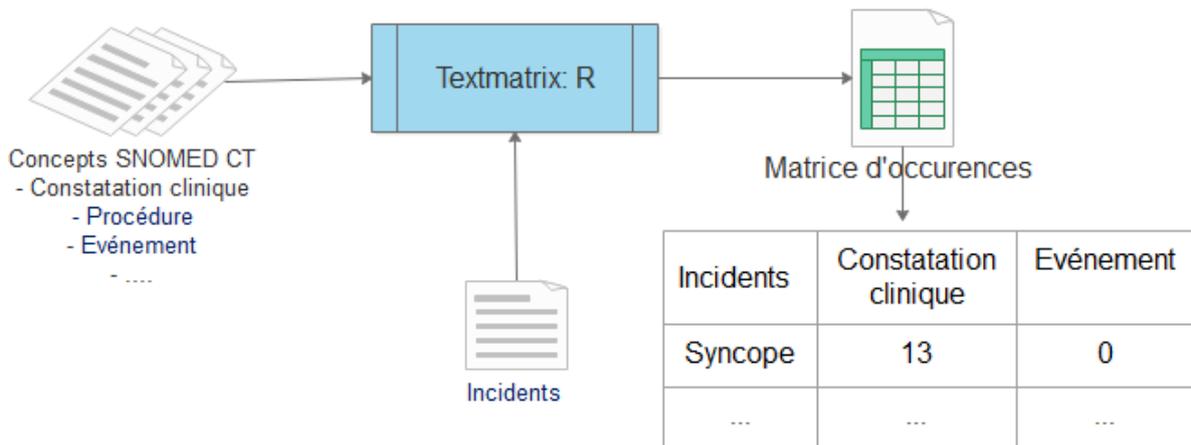


FIGURE 3.10 : Matrice d'occurrences (Types de concepts/Incidents)

3.6.2.3 Construction de la matrice pondérée

Pour trouver les concepts SNOMED CT correspondant aux incidents médicaux dans l'avion, nous nous posons la question suivante dans l'introduction ? « *Selon la conceptualisation du domaine médical dans SNOMED CT, quel est l'élément conceptuel le plus sémantiquement proche ou représentatif de l'incident médical observé dans l'avion ?* ». Nous recherchons donc les relations entre les incidents et concepts SNOMED CT. Ces relations représentées dans la matrice pondérée sont construites à appliquer à la matrice d'occurrences une fonction qui répond à la question précédemment posée. Une réponse littérale à cette question est : « *Les incidents doivent être présents dans la hiérarchie de ce concept et moins présents dans celles des autres* » et la fonction correspondante est TF-IDF (Term Frequency-Inverse Document Frequency). Mais dans la conceptualisation du domaine médical dans SNOMED CT plutôt que la fréquence, la présence d'un incident dans la hiérarchie du concept suffit pour définir son appartenance à ce type de concept. C'est pourquoi nous binarisons la matrice avant d'appliquer la fonction TF-IDF. La

fonction que nous appliquons est $(TF - IDF) \circ BinTF$. L'application de la fonction de pondération sur la matrice des occurrences donne la matrice d'étude illustrée ci-dessous (figure 3.11).

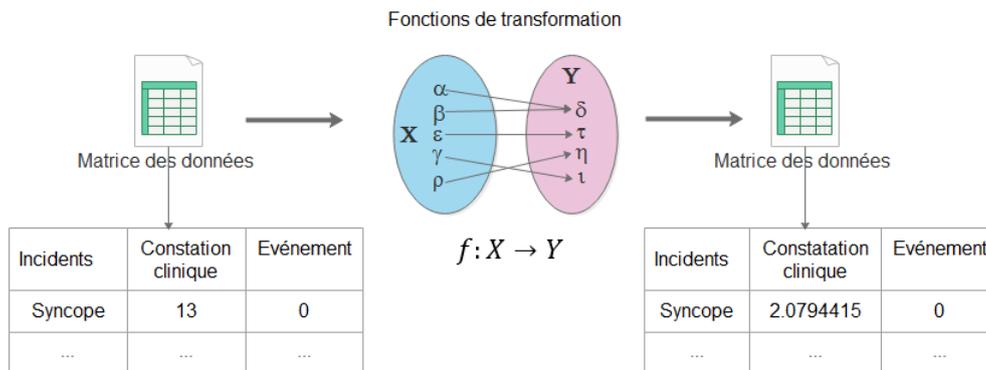


FIGURE 3.11 : Création de la matrice d'étude à partir de la matrice d'occurrences

3.6.2.4 Analyse des composantes principales

Dans la matrice pondérée, les termes décrivant les incidents médicaux dans l'avion sont décrits dans l'espace des concepts de SNOMED CT. Chaque terme est donc décrit par tous les concepts qui forment les axes. Nous cherchons une rotation des axes, un nouvel espace, qui rendrait presque inutiles certains concepts pour décrire les termes. Ce nouvel espace dans lequel les axes passent par le maximum de variance est celui des composantes.

Les outils d'interprétation de l'analyse réalisée ici sont entre autres la qualité de représentation des termes et des documents sur les facteurs, les contributions des termes et des documents aux facteurs. L'Analyse en Composantes Principales (ACP) est une approche statistique recherchant des axes indépendants expliquant au mieux la variabilité – la variance – des données et qui offre aussi une illustration géométrique de ces axes ou composantes selon des directions d'inertie maximale [Bau18]. Sur plan factoriel de l'ACP, on peut afficher le cercle de corrélations fournissant une observation visuelle intéressante. Une illustration est fournie avec le cercle de corrélations de la figure 3.12 et F1 et F2 constituent ici les composantes et les points en couleurs sont les projections des variables. Plus la projection d'une variable initiale est proche du cercle, plus cette variable est proche du plan de projection, donc bien représentée par ce plan.

- Point rouge : une variable devait être parfaitement corrélée avec l'axe F1. Cette situation ne se vérifie presque jamais.
- Point bleue : l'inertie d'une variable est complètement absorbée par les deux premières composantes principales.
- Point vert : la corrélation est parfaite avec la première composante, mais elle est négative.
- Point rose : l'inertie de la variable n'est presque pas absorbée par les deux premiers axes mais par un troisième ou par les suivants, non représentés. Ceci signifie que la corrélation est faible et que cette variable ne doit pas être prise en compte à ce niveau-ci de l'analyse.

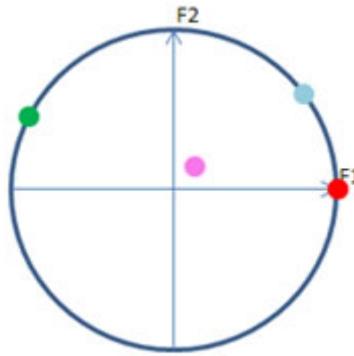


FIGURE 3.12 : Cercle de corrélations variables composantes

Le plan factoriel sur la figure 3.13 est formé par les deux premières composantes (Dim_1, Dim_2) qui représentent à elles deux 60% de l'information. On y voit les 7 variables représentant différents concepts de SNOMED CT (Constatations cliniques (Finding), Evènements (Events), Organisme (Organism), Substances (Substances), Spécimen (Specimen), Structure du corps (Body structure), Objet physique (Physical object)). Les variables « Constatation clinique » et « Evènement » y sont bien représentées, « Organisme » moins et pour les autres variables, nous observons une faible corrélation négative avec la composante Dim_2 .

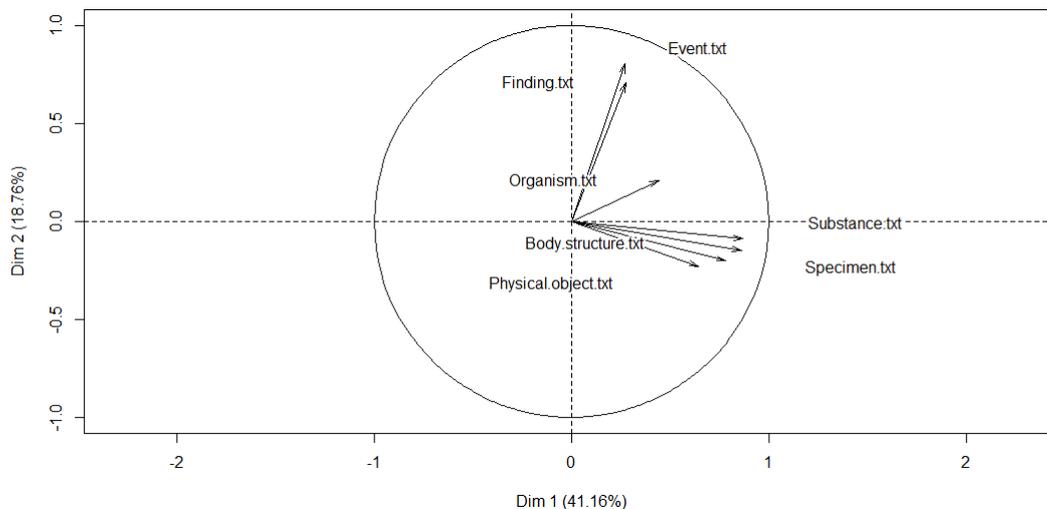


FIGURE 3.13 : Plan factoriel des deux premières des composantes

L'interprétabilité des matrices formées par ASL n'est pas toujours facile, surtout au niveau des apports des termes et des documents sur les thèmes (topics) où l'on peut avoir des poids négatifs. C'est ainsi que nous lui associons l'ADL. Le nombre de thèmes (topic) est défini à partir de L'ASL (ou l'ACP) et représente le nombre de composantes nécessaires pour représenter 80% de l'information (principe de Pareto). Dans notre cas il est égal à 4, puisque les quatre premières composantes représentent 86% (41,15 + 18,75+16,45+10,09) de l'information totale comme le montre la figure 3.14.

L'affectation des thèmes trouvés par l'ADL pour chaque document est visible dans le tableau 3.2 et elle est cohérente par rapport à la position des documents sur les cercles des corrélations.

Dans la compréhension qu'on peut avoir des concepts « Constatations cliniques » et « Evènements »

	eigenvalue	percentage of variance	cumulative percentage of variance
comp 1	2.8809956	41.157080	41.15708
comp 2	1.3130993	18.758562	59.91564
comp 3	1.1516776	16.452537	76.36818
comp 4	0.7066167	10.094525	86.46270
comp 5	0.5228881	7.469829	93.93253
comp 6	0.2846554	4.066506	97.99904
comp 7	0.1400673	2.000961	100.00000

FIGURE 3.14 : Pourcentage de variance cumulée des composantes

de SNOMED CT et sans Traitement automatique du langage naturel, ils constituent des candidats pour représenter les incidents médicaux dans l'avion. Nous avons confirmé cela grâce à la modélisation de thématiques. Par ailleurs la présence des incidents médicaux dans l'avion dans le concept « Constatations cliniques » est la plus homogène et c'est le seul concept où on trouve tous les incidents, c'est pourquoi nous l'avons choisi comme concept contenant les incidents médicaux dans l'avion.

Documents	Thème
Objet physique	1
Organisme	2
Substances	2
Constatations cliniques	3
Évènements	3
Spécimen	4
Structure du corps	4

TABLE 3.2 : Affectation de thèmes pour chaque document

3.7 Exploration de SNOMED CT (les ontologies des incidents)

Ce travail se situe à l'aval d'une bonne catégorisation des incidents médicaux dans l'avion [Mah+11] et permet d'assurer l'efficacité des processus de soins. Pour chaque type d'incident médical dans l'avion, nous construisons grâce à la bibliothèque OWL2 de java une ontologie contenant tous les concepts liés à ce type d'incident. On appelle cette ontologie, « ontologie d'incident ».

Une description schématique courante du processus de décision médical différencie la phase de

diagnostic, préalable à la décision, de la phase de décision proprement dite, c'est-à-dire de prescription [Abe99]. Dans notre travail, elles sont réunies. La première phase intègre deux ensembles d'information, celui issu des études épidémiologiques (maladies, symptômes, facteurs explicatifs, et prophylaxie), l'autre provenant des interactions avec le patient (interrogation, auscultation, examens complémentaires, dossier médical, carnet ou carte santé).

Ce que recherche le praticien lors de la phase de diagnostic, ce n'est donc pas de l'information en général, qui est surabondante, mais la meilleure information possible compte tenu du contexte d'incomplétude dans lequel il exerce. C'est pour cela que nous avons construit les ontologies d'incidents afin de décrire de façon formelle le domaine de connaissances, en identifiant les types de concepts de ce domaine, leurs propriétés et leurs relations (figure 3.15).

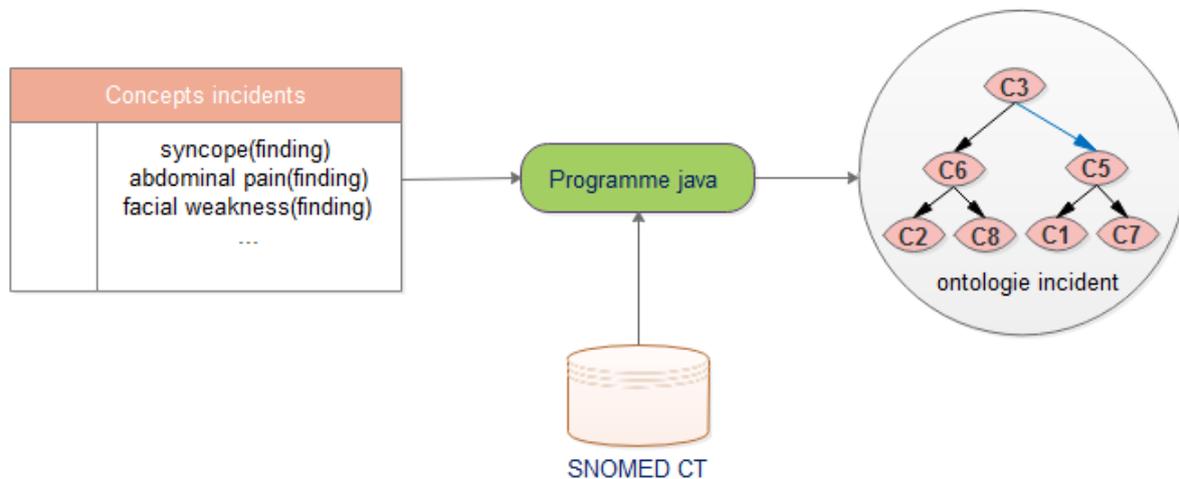


FIGURE 3.15 : Exploration de SNOMED CT

3.7.1 Matériels

Eclipse

Eclipse est un environnement de développement intégré (IDE) qui offre un espace de travail de base et un système d'extension pour l'adaptation de l'environnement. Il permet d'ajouter facilement de nouvelles fonctionnalités, par exemple, l'ajout de Pellet qui est un raisonneur ontologique capable de formaliser une description d'ontologie avec un langage spécifique (p. ex. Web Ontology Language (OWL)). Les ontologies OWL 2 fournissent des classes, des propriétés, des individus et des valeurs de données et sont mises en mémoire sous forme de documents du Web sémantique. Pellet est un raisonneur rigoureux et fondé sur le langage OWL-DL (sigle pour Ontology Web Language Description Logics). OWL-DL est une version décidable du langage informatique OWL qui permet d'effectuer des tâches de raisonnement telles que la vérification de cohérence des modèles cibles ou la vérification d'une relation de subsomption entre les concepts. Afin de construire les mini-ontologies contextuelles, les relations de subsomption sont calculées entre les causes et les conséquences des résultats de raisonnements médicaux pour récupérer les correspondances sémantiques. Cela peut justifier l'application de règles spécifiques à ces résultats.

3.7.2 Méthode

Ci-dessous on a la description simplifiée de l'algorithme de création d'ontologie d'incidents.

Soit i l'incident dont on veut construire l'ontologie

1. On récupère les concepts symptômes (cs) liés à i
2. Pour chaque cs , trouver les concepts liés
3. Récursivité à partir de l'étape 2

Ontologie Syncope : Cet incident constitue 37,4% des incidents médicaux dans l'avion selon Peterson et al. [Pet+13].

On utilise les résultats de la définition de l'incident médical dans l'avion. On recherche les incidents médicaux de l'avion dans le contenu du concept « *Constatation clinique* ».

Les relations et les concepts de l'ontologie Syncope sont visibles sur la figure 3.16 et la figure 3.17. Nous avons obtenu une bonne caractérisation de l'incident Syncope :

- Les symptômes
- D'autres signes qui peuvent être associés
- Les parties du corps concernées
- L'interprétation de l'incident

The screenshot displays the 'Interprets (attribute)' property in an ontology editor. The left pane shows the 'Object property hierarchy' with the following entries:

- owl:topObjectProperty
- "Associated_with_(attribute)" ≡ 47429007
- "Clinical_course_(attribute)" ≡ 263502005
- "Due_to_(attribute)" ≡ 42752001
- "Finding_site_(attribute)" ≡ 363698007
- "Has_interpretation_(attribute)" ≡ 36371300
- "Interprets_(attribute)" ≡ 363714003
- "Part_of_(attribute)" ≡ 123005000
- 123005000 ≡ "Part_of_(attribute)"
- 263502005 ≡ "Clinical_course_(attribute)"
- 363698007 ≡ "Finding_site_(attribute)"
- 363713009 ≡ "Has_interpretation_(attribute)"
- 363714003 ≡ "Interprets_(attribute)"
- 42752001 ≡ "Due_to_(attribute)"
- 47429007 ≡ "Associated_with_(attribute)"

The right pane shows the 'Annotations' and 'Object Property Usage' for the property. The 'Annotations' section includes two `rdfs:comment` entries with the value "Interprets (attribute)". The 'Object Property Usage' section shows the property's characteristics (Functional, Inverse functional, Transitive, Symmetric, Asymmetric, Reflexive, Irreflexive) and its domain and ranges. The domain is "Vasovagal_syncope_(finding)" and the ranges are 248217000, 18373002, and 6942003.

FIGURE 3.16 : Relation de l'ontologie Syncope

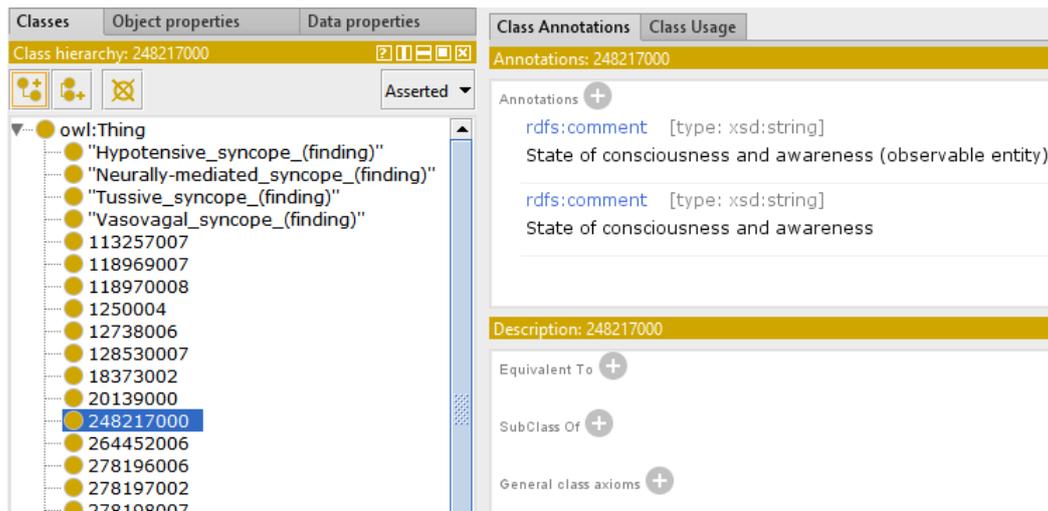


FIGURE 3.17 : Concepts de l'ontologie Syncope

3.8 Conclusion

Dans ce chapitre, nous nous sommes intéressés à la caractérisation et la définition des structures de données pour la prise en charge médicale des passagers dans l'avion. Pour cela, une description des facteurs déterminants de la prise en charge médicales dans l'avion a d'abord été effectuée. Par la suite, nous avons relié les incidents médicaux aux rapports d'événements médicaux dans le système ASRS et dont l'exploitation pourrait améliorer la prise en charge médicale dans l'avion. Puis une modélisation thématique est effectuée pour la définition de l'incident médical dans l'avion en utilisant notamment les techniques d'analyse sémantique. Enfin, les connaissances ontologiques permettant de spécifier explicitement les éléments conceptuels contenus dans les textes et documents de références comme ceux de SNOMED CT et de l'ASRS qui servent de sources d'informations pour enrichir et améliorer la réalisation d'ontologies en l'occurrence des incidents médicaux.

Ces ontologies pourront être utilisées pour faire des inférences complexes et pour dériver des règles nécessaires à l'interprétation sémantique et récupérer des informations textuelles connexes. Dans notre cas, cela facilitera la recherche des types d'informations suivantes : (1) une relation de synonymie entre termes (par exemple, deux termes décrivant sémantiquement le même incident), (2) une relation hiérarchique entre divers termes (par exemple, une catégorie d'incident et ses sous-catégories), (3) une relation dans le chemin clinique entre un problème et ses moyens de résolution (par exemple, un incident et ses actions de prise en charge).

Ce chapitre nous a permis de définir les facteurs de décision dans le cadre de la prise en charge médicale dans l'avion. On y réalise la première étape du cycle de vie de toutes les décisions qu'un professionnel de la santé pourrait être amené à prendre.

Une approche de modélisation de l'incertitude pour la conception de la décision médicale dans l'avion

4.1 Introduction

Dans le chapitre précédent, nous avons concrétisé l'analyse de la première étape de la décision médicale dans l'avion par une conceptualisation des facteurs déterminants de la décision médicale à bord. Dans ce chapitre nous proposons une formalisation de la décision grâce à la modélisation de l'incertitude de l'intervenant traitée dans l'un nos articles scientifiques ([SKR18a]). L'objectif est de concevoir un cadre formel contenant toutes les décisions possibles, sur lequel nous pourrions nous appuyer pour proposer une aide dans la gestion des incidents médicaux dans l'avion. Pour cela nous associons la notion de tolérance à l'incertitude proposée par Lorenzo [Lor13] au cadre de discernement défini dans la théorie de Dempster-Shafer utilisant les fonctions de croyance [Sha76]. Nous avons transformé les facteurs déterminants de la décision médicale (type d'incident, choix de déroutement et type d'intervenant) en domaines de discernement contenant l'ensemble des choix possibles. Ce qui facilite la représentation des choix de l'intervenant par des hypothèses dont on peut évaluer la crédibilité et la plausibilité. Cette évaluation nous permet sans quantifier le niveau d'incertitude de l'intervenant, d'agir sur les facteurs influençant son seuil de tolérance à l'incertitude.

4.2 L'incertitude médicale

L'incertitude est selon le dictionnaire Larousse « l'état de quelqu'un qui ne sait quel parti prendre, ou l'état plus ou moins préoccupant de quelqu'un qui est dans l'attente d'une chose incertaine ». Pour Fox [Fox88] à qui l'on doit le développement de cette notion, l'incertitude médicale est une caractéristique centrale de la pratique médicale. Elle est une conséquence inévitable de la rapide progression du savoir scientifique et médical, redoublée, dans le contexte de la situation clinique, par les demandes émotionnelles des patients [Fer08]. Pour représenter les différents positionnements possibles des médecins par rapport à l'incertitude dans un espace commun ordonné, Géraldine Bloy [Blo08] utilise deux axes (figure 4.1) :

- le premier (axe horizontal) correspond à l'incertitude introduite par la forme « brute » de la plainte

du patient. La considération accordée à cette plainte, c'est-à-dire le fait qu'elle puisse être entendue dans sa formulation profane et que le professionnel accepte d'y voir un objet à part entière du travail médical, est d'autant plus importante qu'on se déplace vers la droite du graphique. Sur la gauche se déploient des pratiques de réduction ou fermeture par rapport aux éléments d'informations relatifs au patient et non immédiatement pertinents pour le médecin.

- le deuxième (axe vertical) correspond à l'incertitude liée aux questions de maîtrise du savoir scientifique biomédical. Plus on se déplace vers le haut, plus la vigilance des praticiens par rapport à l'évolution des connaissances biomédicales et au régime contemporain d'administration de la preuve en médecine est grande. C'est cette incertitude qui est le plus balisée. Renée Fox [Fox88] distingue l'incertitude attribuable au caractère inachevé de la maîtrise du savoir médical par le praticien, de l'incertitude liée aux limites de la science elle-même.

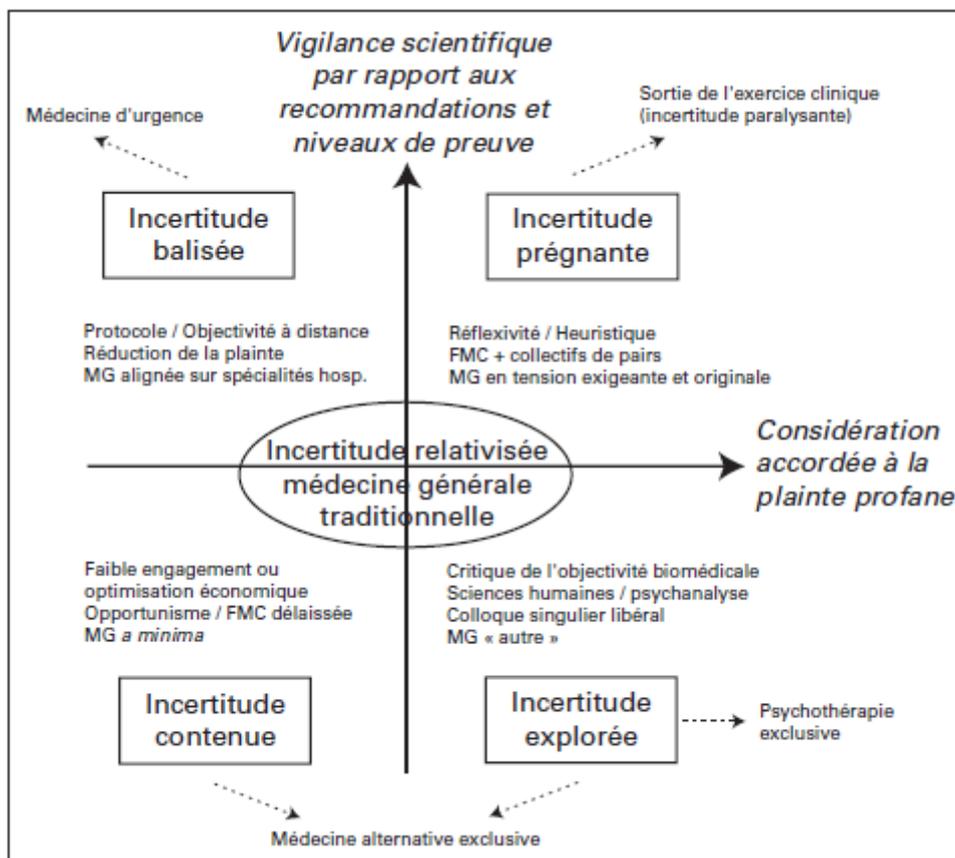


FIGURE 4.1 : Positionnement du médecin par rapport à l'incertitude - (MG = Médecine Générale, FMC = Formation Médicale Continue) [Blo08]

Il s'agit globalement d'une typologie des postures des médecins par rapport à l'incertitude. Pour cela Lorenzo [Lor13] utilise la notion de tolérance à l'incertitude que nous utilisons dans notre travail. Il remplace les axes par les facteurs déterminants de l'incertitude. Sur la figure 4.2 nous avons une catégorie de ces facteurs, le contexte qui regroupe entre autres, le lieu, le mode d'exercice, etc. La tolérance à l'incertitude dépendra de tous ces facteurs.

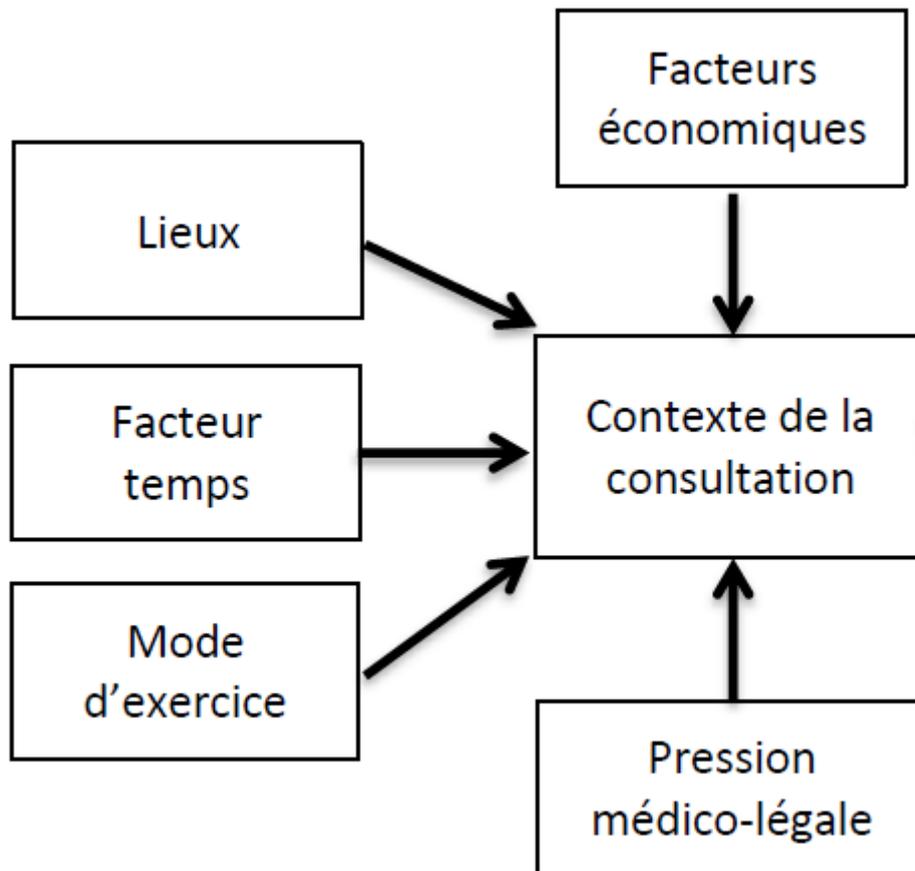


FIGURE 4.2 : Facteurs déterminants de l'incertitude

4.2.1 Incertitude médicale dans l'avion

La définition de l'incertitude médicale renvoie à la notion de décision. Ce qui fait des facteurs déterminants de la prise en charge médicale dans l'avion (définition 1), des facteurs déterminants de l'incertitude. En effet décider dans le domaine médical c'est aussi gérer cette incertitude inhérente à la pratique médicale.

La perception de l'incertitude médicale est comportementale, cognitive et émotionnelle [Hil+17]. Lorenzo [Lor13], Karim [San16], Hillen et ses collègues [Hil+17] s'accordent sur le problème de quantification complète et d'utilisation de l'incertitude par les systèmes informatiques. C'est pourquoi nous utilisons la notion de tolérance à l'incertitude (figure 4.3) défini par Lorenzo [Lor13] comme la capacité du médecin à travailler et mener à bien sa consultation sans connaître avec exactitude les mécanismes physiques, biologiques et psychologiques à l'œuvre chez son patient, et dont il a proposé une théorisation ancrée [Pai94] en se basant sur les déterminants de l'incertitude. Les facteurs déterminants de l'incertitude sont divisés en 5 groupes, (1) le contexte de la prise en charge, (2) les facteurs liés au médecin, (3) les facteurs liés au patient, (4) la relation entre le médecin et le patient, (5) le raisonnement clinique et l'estimation du risque.

Chaque médecin utilise pour travailler un raisonnement clinique. Celui-ci est défini comme l'ensemble des processus cognitifs basés sur l'expérience permettant d'évaluer le problème de santé du patient puis de le prendre en charge [Lor13]. Les guides de bonnes pratiques et le Chemin clinique sont des outils de raisonnement clinique.

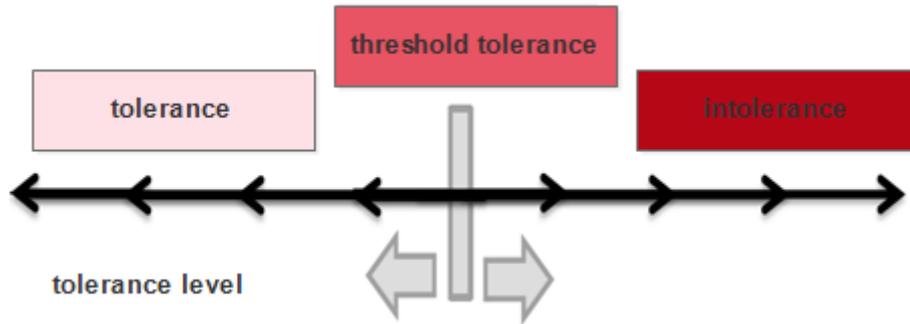


FIGURE 4.3 : Tolérance à l'incertitude médicale

En situation d'incertitude, le raisonnement clinique intègre une estimation de risque [Eva05]. L'estimation du risque de la situation conduit à différencier un risque acceptable d'un risque inacceptable et donc à caractériser la tolérance à l'incertitude. Elle dépend de tous les autres facteurs déterminants de la prise en charge médicale (facteurs déterminants de l'incertitude).

4.3 L'approche probabiliste de formalisation de la décision médicale sous incertitude

Auparavant, comme dans tous les domaines, la théorie des probabilités était utilisée pour modéliser presque tous les aspects de l'incertitude médicale. Un problème de prise de décision est représenté par un quadruplet $\{E, C, A, f\}$ [FS05].

- E est un ensemble d'états de la nature. Un élément $e \in E$ représente une décision que peut prendre la Nature et susceptible d'influencer les conséquences de l'exécution de l'une, au moins, des actions de A .
- C est un ensemble de conséquences.
- A est ensemble d'actions. Un élément $a \in A$ est une fonction $g : E \rightarrow C$ et représente une action qu'il est possible de mettre à exécution par le décideur.
- f est une fonction définissant une relation de préférence sur A , généralement avec les propriétés de réflexivité, de transitivité et de totalité (un pré-ordre total).

Dans l'approche probabiliste basée sur le modèle « Utilité Espérée » [Neu44]; [Sav54], on utilise le quadruplet suivant $\{E, C, p, u\}$. L'incertitude est représentée par une distribution de probabilités p sur E . La relation de préférence est une fonction d'utilité $u : C \rightarrow [0, 1]$. Le critère de décision est tout simplement la maximisation de l'espérance de la satisfaction procurée par les décisions (fonction d'utilité). Les actions sont classées selon leurs « Utilité Espérée » UE_{pu} .

$$\forall a \in A, UE_{pu}(a) = \sum_{e \in E} p(e) \cdot u(a(e))$$

Ce modèle a été beaucoup discuté dans la littérature de l'économie comme par exemple dans Mathematical Economics [Cha16] et dans Health Economics [Att+16]. Ceci a d'ailleurs engendré un nouveau

champ d'étude : l'économie de la santé. Mais beaucoup s'interroge sur l'applicabilité pratique de ces modèles sur les soins de santé [FGS13]. L'aspect global des déterminants majeurs de la prise de décision médicale reste toujours mal caractérisée. D'ailleurs les modèles qui fonctionnent le mieux sont les modèles qui en plus de la variable coût de la santé du patient ajoutent d'autres variables comme la qualité des soins et le délai de prise en charge.

4.4 L'approche crédibiliste de formalisation de la décision sous incertitude

Au cours des trois dernières décennies, la communauté mathématique s'est rendu compte que l'incertitude peut englober une grande variété d'aspects et que les probabilités ne peuvent pas toujours y faire face [Bel13]. De nouvelles théories ont donc été développées. Ces aspects concernent les types et les sources d'incertitudes que Dubois [Dub11] organise comme le montre le tableau 4.1.

Type d'incertitude	Incertainité aléatoire	Incertainité épistémique	Incohérence
Origines	Variabilité, aléatoire 	Ignorance, Incomplétude 	Conflit entre les sources 

TABLE 4.1 : Types et sources d'incertitudes

L'incertitude aléatoire est induite par les résultats d'expériences aléatoires (répétables). On l'appelle souvent l'aspect aléatoire d'un phénomène observé, puisque cette incertitude est inhérente au phénomène [Bel13]. Un exemple typique de l'aléatoire est le résultat du lancer d'une pièce ou d'un dé.

L'incertitude épistémique constitue le deuxième type d'incertitude. Elle est due au manque de connaissances et est caractérisée par l'ignorance ou l'incomplétude. L'ignorance peut conduire à l'imprécision, qui peut donc être qualifiée d'incertitude épistémique.

L'autre type d'incertitude est l'incohérence qui peut résulter de la multiplicité des sources. L'incertitude médicale dans l'avion présente un aspect épistémique, c'est pourquoi dans ce qui suit nous vous présenterons une approche adaptée à ce type d'incertitude, la théorie de Dempster-Shafer que nous allons utiliser dans notre travail. D'autres raisons motivant ce choix seront expliquées dans les prochaines sections.

4.4.1 La théorie de Dempster-Shafer

La théorie de l'évidence, également connue sous le nom de théorie des fonctions de croyance, a été développée par Arthur P. Dempster et Glenn Shafer. Le modèle de croyance transférable [Sme90]; [SK94] est une contribution majeure dans le contexte de cette théorie. La théorie de Dempster-Shafer [Sha76] peut

être considérée comme une généralisation des approches probabilistes ou possibilistes. L'un des principaux avantages de la théorie de Dempster-Shafer sur la théorie des probabilités est de permettre de spécifier un degré d'ignorance dans une situation au lieu d'être obligé de fournir des probabilités aléatoirement avec les faits antérieurs [Bel13]. C'est un outil qui permet de modéliser aussi bien l'incertitude que l'imprécision. Il prend en compte des ambiguïtés, des conflits entre les sources, le principe du monde ouvert (il peut exister certaines possibilités non recensées initialement). Sous l'hypothèse du monde ouvert, plus généralement, une masse positive (valeur appelée conflit) sur l'ensemble vide quantifie la discordance entre les sources de croyance. Les origines de ce conflit peuvent être multiplés : (i) le manque d'exhaustivité du cadre de discernement, (ii) la nature de la règle de la combinaison conjonctive employée ou le non-respect des hypothèses sous-jacentes à son application. Le passage du monde ouvert au monde fermé consiste à redistribuer (par exemple avec un processus de normalisation) la masse de l'ensemble vide sur les autres sous-ensembles du cadre de discernement.

4.4.2 Les concepts formels de la théorie de Dempster-Shafer

Pour chaque problème, la théorie des fonctions de croyance définit un cadre de discernement $\Theta = \{H_1 \dots H_n\}$ [Dut15]. Θ un ensemble fini non-vide incluant n hypothèses exhaustives et mutuellement exclusives. L'ensemble des parties 2^Θ qui contient les sous-ensembles possibles de Θ constitue le référentiel de définition utilisé pour évaluer la véracité d'une proposition. Une connaissance imparfaite sur Θ est représentée par une masse de croyance (en anglais basic belief assignment ou bba) [Mas05], définie par la fonction $m : 2^\Theta \rightarrow [0, 1]$ qui vérifie les propriétés suivantes.

- La masse de l'élément vide est égale à zéro.

$$m(\emptyset) = 0$$

- La somme des masses de tous les éléments est égale à un.

$$\sum_{A \subset \Theta} m(A) = 1$$

$m(A)$, avec $A \subset \Theta$ désigne la quantité de preuve soutenant l'évènement A [JLH02]. Elle représente la partie du degré de croyance placée sur A et qui n'a pas été distribuée aux sous-ensembles de A [TY11]. Les éléments A de 2^Θ pour qui $m(A) \neq 0$ sont appelés « éléments focaux ». Soit F l'ensemble des éléments focaux :

Définition 2. (*corps d'évidence*) : Le triplet (Θ, F, m) est appelé corps d'évidence.

Deux autres mesures de croyance, à savoir la crédibilité (*bel*) et la plausibilité (*pl*), peuvent de façon équivalente représenter la bba m .

Comme on la vu ci-dessus la masse $m(A)$ représente la partie du degré de croyance placée exactement sur A . La fonction de croyance (*bel*) calcule toute la croyance placée sur A . Elle est définie à partir de la fonction de masse m et donc prend en compte tous les sous-ensembles de A .

$$bel(A) = \sum_{B \subset A, B \neq \emptyset} m(B)$$

La fonction de plausibilité (pl) est la fonction duale de la fonction de croyance, elle mesure l'intensité avec laquelle une proposition A peut être considérée comme vraie. La plausibilité de A est donc la somme des masses de tous les éléments focaux qui sont compatibles avec A .

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

Pour assister l'intervenant dans une prise en charge médicale dans l'avion nous combinons des connaissances provenant de différentes sources. Il y a d'abord les connaissances provenant des retours d'expériences de la gestion d'incidents et les connaissances ou ignorances de l'intervenant. Pour le stockage et la fusion de ces connaissances nous utiliserons une base de données évidentielles.

4.4.2.1 Les règles de combinaisons de fonctions de masse

L'un des principaux aspects de la théorie des fonctions de croyance est la possibilité de combiner des sources de croyances distinctes. Plusieurs modes de combinaison existent. Ils peuvent prendre en compte, la fiabilité des sources, le conflit sur les sources, l'hypothèse du monde ouvert, etc. Smet [Sme93] a introduit les règles conjonctive et disjonctive pour combiner des fonctions de masses de sources fiables et non fiables sur un même domaine de discernement. Soit les *bba* m_1 et m_2 fournis par des sources différentes sur un même cadre de discernement Θ . Les combinaisons conjonctive (on suppose que les sources sont toutes deux fiables) et disjonctive (on suppose que l'une au moins des deux sources est fiable) sont définies comme suit :

$$m_1 \odot m_2(Z) = \sum_{X, Y \subseteq \Theta; X \cap Y = Z} m(X) \times m(Y)$$

$$m_1 \oplus m_2(Z) = \sum_{X, Y \subseteq \Theta; X \cup Y = Z} m(X) \times m(Y)$$

La règle conjonctive de combinaison fusionne les *bba* m_1 et m_2 fournis par différentes sources ;

Les éléments focaux du *bba* combinés par la règle conjonctive sont l'intersection de ceux de m_1 et m_2 . Pour la règle disjonctive, ils sont l'union des éléments focaux de m_1 et m_2 .

Les règles décrites ci-dessus sont construites selon l'hypothèse que les *bba* sont définies sur le même cadre de discernement. Cette hypothèse limite les applications pratiques. Dans certaines situations on est amené à combiner les croyances définies sur des cadres de discernement différents. Smet [Sme90], dans son Modèle de Croyances Transférables (MCT) a introduit le principe d'engagement minimal ou principe du minimum d'information (Principle of Minimal Commitment en anglais). Ce principe au cœur d'un grand nombre de mécanismes développés dans le MCT impose de choisir parmi un ensemble d'affectation de masses *bba* possibles, celle qui est la moins engagée ou, autrement dit, la moins informative [Ram07]. Cette notion de fonction de croyance la moins informative (au sens d'un certain ordre) peut avoir une interprétation contextuelle variable [Mas05] et est à la base d'opérations comme l'extension vide et la marginalisation qui permettent le changement d'espaces. Ces opérations consistent en 3 étapes :

- D'abord il faut construire le cadre de discernement commun.
- Ensuite, le principe d'engagement minimal est utilisé pour exprimer les croyances sur le cadre de discernement commun.

- Et pour finir, les croyances obtenues dans l'étape précédente et donc partageant le même domaine de discernement peuvent être combinées en utilisant les règles de combinaison classiques.

La marginalisation qui a une opération analogue dans les lois de probabilités permet de réduire un cadre de discernement. L'extension vide par contre n'existe pas en probabilité. Elle permet de combiner des *bba* définies sur des cadres de discernement disjoints.

Soit maintenant deux *bba* m_1 et m_2 sur deux cadres différents Θ_1 et Θ_2 . L'extension vide de m_1 sur $\Theta = \Theta_1 \times \Theta_2$ est définie pour tout $C \subseteq \Theta$ comme suit :

$$m^{\Theta_1 \uparrow \Theta}(C) = \begin{cases} m_1(B) & \text{si } C = B \times \Theta_2 \text{ pour un } B \subseteq \Theta_1 \\ 0 & \text{sinon} \end{cases}$$

La masse initialement allouée à la proposition $B \subseteq \Theta_1$ est réallouée à $B \times \Theta_2$ après extension vide minimisant ainsi tout a priori sur Θ_2 (aucune hypothèse de Θ_2 n'est favorisée car la source de croyance fournissant m_1 ne connaît rien de Θ_2) [Ram07]. La transformation $B \subseteq \Theta_1 \mapsto B \times \Theta_2$ est appelée extension cylindrique de B sur Θ . L'extension vide de $C \subseteq \Theta$ consiste donc à étendre chaque élément focal par son extension cylindrique sur Θ notée $B^{\uparrow \Theta}$. La *bba* étendue est ensuite appliquée sur les extensions cylindriques des éléments focaux. La combinaison des *bba* est donnée par l'expression ci-dessous :

$$m_1 \odot m_2(D) = m^{\Theta_1 \uparrow \Theta} \odot m^{\Theta_2 \uparrow \Theta}(D) = \begin{cases} m_1(B) \cdot m_2(C) & \text{si } D = B \times C, B \subseteq \Theta_1, C \subseteq \Theta_2 \\ 0 & \text{sinon} \end{cases}$$

4.4.3 Base de données évidentielles

L'apport de la théorie des fonctions de croyance sur la représentation des données imparfaites dans les bases de données a été étudié [Sam14] [Mas05]. Les méthodes algorithmiques traditionnelles en fouille de données supposent que les données manipulées sont certaines et précises pour constituer des bases de données booléennes. Néanmoins, dans plusieurs situations pratiques, la collecte de tels types de données est difficilement réalisable puisque de nombreuses données acquises pourraient souffrir d'imperfection (imprécision ou incertitude). C'est ainsi que Suk Kyoon Lee a proposé le principe de fouille de données évidentielles [Lee92] qui s'appuie sur la théorie de Dempster-Shafer utilisant les fonctions de croyance pour traiter de l'imperfection pouvant se manifester dans les données.

Les bases de données évidentielle (BDE), également appelées bases de données D-S (Dempster-Shafer), sont utilisées pour stocker à la fois des données précises, imprécises, ou incertaines contrairement aux bases de données classiques utilisées pour stocker uniquement des données précises [Bou+15]. Les valeurs incertaines dans les bases de données évidentielles sont exprimées par la théorie de Dempster-Shafer.

Une BDE contient D attributs et N transactions (objets). Chaque attribut i ($1 \leq i \leq D$) a un domaine Θ_i de valeurs discrètes. La valeur d'un attribut i pour un objet j est dite valeur évidentielle et notée V_{ij} . Elle correspond à une *bba* qui est définie comme suit [TY11] :

Soit V_{ij} une valeur évidentielle qui correspond à la colonne i et à la ligne j . V_{ij} correspond à un corps d'évidence (définition 2) composé du cadre de discernement Θ_i (le domaine de la colonne i), de l'ensemble F_{ij} des éléments focaux et de la fonction de masse m_{ij} qui est définie comme suit :

$$m_{ij} : 2^{\Theta_i} \rightarrow [0, 1] \text{ avec } m_{ij}(\emptyset) = 0 \text{ et } \sum_{x \subseteq \Theta_i} m_{ij}(x) = 1$$

Le tableau 4.2 est une base de données évidentielles. Elle montre aussi le nombre de types de données qui peuvent être modélisé par la théorie de Dempster-Shafer.

- L'élément focal pour le domaine de discernement « Déroutement » dans la transaction T1 est unique. Sa *bba* est certaine. La BDE contient des données parfaites.
- La valeur de l'attribut « Déroutement » dans la transaction T2 est probabiliste. En effet tous les éléments focaux sont des singletons.
- Lorsque les éléments sont imbriqués comme dans le domaine de discernement « Type d'incident » pour la transaction T1, on peut parler de valeur possibiliste. La mesure de la possibilité de SP est ainsi égale à 1 alors que celle de CS est égale à 0,7.
- Lorsque que le seul élément focal dans le domaine de discernement correspond au domaine de discernement lui-même et que sa masse soit l'unité, on est en totale ignorance.
- Lorsqu'aucun des cas précédents n'est présent, nous sommes dans le cas d'une information évidentielle générique telle que le type d'incident dans la transaction T2.
- Pour la conception d'un système de diagnostic basé sur la logique floue Jones et ses collègues [JLH02] montrent que l'utilisation du cadre de modélisation évidentielle permet tout en gardant une structure de modèle relationnelle floue, de contourner certains problèmes comme le choix des symptômes pour utiliser des opérateurs flous et d'utiliser toutes les informations disponibles.
- Lorsqu'on se passe de la condition $m(\emptyset) = 0$ qui a été imposée par Smet [Sme90], on accepte l'hypothèse du monde ouvert qui suppose que le cadre de discernement peut être incomplet.

Transaction	Type d'incident	Déroutement
T1	SP (0,3) {CS,SP} (0,7)	D (1)
T2	{SP,S} (0,3) RS (0,7)	D(0,5) ND (0,5)

TABLE 4.2 : Base de données évidentielles, SP = Syncope-Pre-syncope ; CS = Cardiac symptom ; S = Seizures ; RS = Respiratory symptom ; ND = Non diversion ; D = Diversion

4.4.4 La BDE comme les mondes possibles

La BDE est dans sa définition classique une forme compacte d'un ensemble de N transactions et D attributs comme indiqué en 4.4.3. La forme non compacte proposée dans [Bou+15] propose une autre vue intéressante. Une BDE y est définie comme un ensemble fini de mondes possibles imprécis tels que $BDE = \{MI_1, MI_2, \dots, MI_k\}$. Chaque monde possible imprécis comprend N transactions, où chaque transaction contient un élément focal par attribut. Les mondes possibles sont qualifiés d'imprécis à cause des valeurs imprécises comme $\{CS, SP\}$ qu'ils peuvent contenir. Le tableau 4.3 montre un exemple de monde possible imprécis.

Transaction	Type d'incident	Déroutement
T1	{CS,SP}	D
T2	RS	ND

TABLE 4.3 : Un monde possible imprécis

Le nombre de mondes possibles imprécis dépend des ensembles d'éléments focaux F_{ij} . Formellement un monde possible imprécis MI est défini ainsi :

$$MI \in \prod_{j=1}^N \prod_{i=1}^D F_{ij}$$

Le tableau 4.4 montre les mondes possibles imprécis de la BDE décrit dans le tableau 4.2.

MI_1	MI_2	MI_3	MI_4
(T_1, SP, D) $(T_2, \{SP, S\}, D)$	(T_1, SP, D) $(T_2, \{SP, S\}, ND)$	(T_1, SP, D) (T_2, RS, D)	(T_1, SP, D) (T_2, RS, ND)
MI_5	MI_6	MI_7	MI_8
$(T_1, \{CS, SP\}, D)$ $(T_2, \{SP, S\}, D)$	$(T_1, \{CS, SP\}, D)$ $(T_2, \{SP, S\}, ND)$	$(T_1, \{CS, SP\}, D)$ $T_2, RS, D)$	$(T_1, \{CS, SP\}, D)$ (T_2, RS, ND)

TABLE 4.4 : les mondes possibles imprécis

Chaque monde possible imprécis peut être divisé en un ensemble d'états précis appelés mondes possibles. Le tableau 4.5 montre les mondes possibles de MI_1 .

MI_1	
(T_1, SP, D) (T_2, SP, D)	(T_1, SP, D) (T_2, S, D)

TABLE 4.5 : Les états précis d'un monde possible imprécis

Par analogie à une base de données classique, une BDE est une union de ses transactions évidentielles. Les corps d'évidence des transactions et de la BDE sont construits à partir des notions d'items et d'itemsets évidentiels.

Définition 3. (*Item évidentiel*)

Un item évidentiel noté iv est un élément focal d'un domaine de discernement. Les items évidentiels sont donc les éléments des ensembles F_{ij} des corps évidentiels correspondant aux valeurs évidentielles V_{ij} ($iv \in 2^{\Theta_i}$). SP et $\{SP, S\}$ dans le tableau 4.2 sont des items

Définition 4. (*Itemset évidentiel*)

Un itemset évidentiel est un ensemble d'items évidentiels inclus dans l'union des produits cartésiens des ensembles F_{ij} dans chaque transaction. Si X est un itemset évidentiel alors :

$$X \in \bigcup_{j=1}^N \prod_{i=1}^D F_{ij}$$

Les items évidentiels de chaque itemset évidentiel correspondent à des attributs différents.

4.4.5 Corps d'évidence d'une transaction

Le corps d'évidence d'une transaction évidentielle est déduit des valeurs évidentielles qui la composent. Comme ces dernières correspondent à des corps d'évidence composés de domaines de discernement différents, nous avons utilisé l'extension vide (4.4.2.1).

Soit une transaction évidentielle j constitué des valeurs évidentielle $V_{ij} = (\Theta_i, F_{ij}, m_{ij})$, $i = 1, \dots, D$. Son cadre de discernement noté Θ est le produit cartésien des domaines des attributs de la base $\Theta = \prod_{i=1}^D \Theta_i$. Les éléments focaux sont les sous-ensembles de Θ . F_j contient donc des vecteurs de la forme $X = \{x_1, x_2, \dots, x_k\}$ où $x_i \subseteq \Theta_i$. Sa fonction de masse m_j est la combinaison conjonctive de toutes les valeurs probantes étendues au cadre commun.

$$m_j(X) = \bigodot_{x_i \in X} m^{\Theta_i \uparrow \Theta}(x_i^{\uparrow \Theta}) = \bigodot_{x_i \in X} m_i(x_i)$$

Le tableau 4.6 montre les corps évidentiels des transactions du tableau 4.2.

La représentation non compacte en mondes possibles imprécis d'une BDE (tableau 4.4) est un produit cartésien des ensembles d'éléments focaux de ses transactions (tableau 4.6).

Transactions	Θ_j	F_j	m_j
T1	Θ	(SP, D) $(\{CS, SP\}, D)$	$(SP, D) = 0,3 \times 1 = 0,3$ $(\{CS, SP\}, D) = 0,7 \times 1 = 0,7$
T1	Θ	$(\{SP, S\}, D)$ $(\{SP, S\}, ND)$ (RS, D) (RS, ND)	$(\{SP, S\}, D) = 0,3 \times 0,5 = 0,15$ $(\{SP, S\}, ND) = 0,3 \times 0,5 = 0,15$ $(RS, D) = 0,7 \times 0,5 = 0,35$ $(RS, ND) = 0,7 \times 0,5 = 0,35$

TABLE 4.6 : Corps évidentiels de transaction

4.4.6 Corps d'évidence d'une BDE

Comme nous l'avons rappelé précédemment, une BDE est une union de ses transactions évidentiels. Par conséquent comme les transactions qui le composent le cadre de discernement d'une BDE est aussi $\Theta = \prod_{i=1}^D \Theta_i$. L'ensemble F_{BDE} de ses éléments focaux est l'union des éléments focaux de toutes les transactions c'est-à-dire tous les itemsets évidentiels existants dans la base. Soit X un itemset évidentiel dans la BDE de N transactions. La *bba* de la BDE est définie comme suit :

$$m_{BDE}(X) = \frac{1}{N} \sum_{j=1}^N m_j(X)$$

Comme dans le cas d'une *bba* sur un seul cadre de discernement, la crédibilité et la plausibilité sont aussi définies.

$$Bel_{BDE}(X) = \sum_{Y \subseteq X} m_j(Y)$$

Le calcul de la croyance ci-dessus utilise l'inclusion d'itemset.

Définition 5. (inclusion d'itemset)

L'inclusion d'itemset est définie comme suit : $Y \subseteq X$ si et seulement si : $\forall y_i \in Y, y_i \subseteq x_i$ avec x_i et y_i respectivement les items de X et Y relatif au cadre Θ_i . Exemple : $(SP, D) \subseteq (\{CS, SP\}, D)$

$$pl_{BDE}(X) = \sum_{Y \cap X \neq \emptyset} m_j(Y)$$

La masse d'un itemset évidentiel X dans une base de données évidentielles EDB est la croyance partielle attribuée exactement à X . Par contre, la somme des masses des sous-ensembles de X reflète la croyance totale de X dans BDE qui n'est autre que la fonction de croyance $Bel_{EDB}(X)$ [TY11].

Le support de X dans BDE, soit la fréquence de X dans la base, coïncide donc avec sa croyance dans le corps d'évidence de BDE [TYM08].

Définition 6. (support d'un itemset évidentiel)

Le support d'un itemset évidentiel est sa croyance dans la BDE. $support(X) = Bel(X)$

Le tableau 4.7 montre le corps d'évidence d'un BDE.

Θ_{BDE}	F_{BDE}	m_{BDE}	$Bel_{BDE} = support_{BDE}$
Θ	(SP, D)	$0,3 \times 1/2 = 0,15$	0,15
	$(\{CS, SP\}, D)$	$0,7 \times 1/2 = 0,35$	$0,35 + 0,15 = 0,45$
	$(\{SP, S\}, D)$	$0,3 \times 0,5/2 = 0,075$	0,075
	$(\{SP, S\}, ND)$	$0,3 \times 0,5/2 = 0,075$	0,075
	(RS, D)	$0,7 \times 0,5/2 = 0,175$	0,175
	(RS, ND)	$0,7 \times 0,5/2 = 0,175$	0,175

TABLE 4.7 : Corps évidentiels d'une BDE

4.5 Un modèle de représentation de l'incertitude comme cadre formel de décision médicale dans l'avion

Dans ce travail nous proposons une implémentation de la théorisation ancrée [Pai94] de la tolérance à l'incertitude proposée par Lorenzo (section 4.2) [Lor13]. Nous avons remplacé l'estimation de risque par un cadre formel de décision médicale dans l'avion basé sur la théorie de Dempster-Shefer (figure 4.4) et qui va décrire l'incertitude de l'intervenant. Pour cela nous avons transformé en domaines de discernement tous les autres facteurs déterminants de l'incertitude. La lecture de l'incertitude faite par Lorenzo est proche de celle de Schmidt [Sch10] qui considère le risque et l'incertitude plutôt comme des notions complémentaires que mutuellement exclusives. Il désigne par "risque", l'évaluation d'un aléa (non nécessairement par des probabilités) et par "incertitude" le jugement de vérité porté par le décideur sur cet aléa [Abe99].

Définition 7. (cadre formel de décision)

Un cadre formel de décision est un ensemble $c = (X_k, C_k)$ d'itemsets et leurs masses associées sont issues du corps d'évidence (Θ, F, m) d'une BDE de deux transactions dont l'une des transactions contient l'incertitude de l'intervenant et l'autre est issue des données de retour d'expériences de la gestion des incidents médicaux dans l'avion. Le cadre de discernement de ce BDE est le produit cartésien de tous les facteurs déterminants de la décision médicale dans l'avion. Les itemsets dans le cadre formel de décision sont inclus dans le produit cartésien des singletons dans la transaction décrivant l'incertitude de

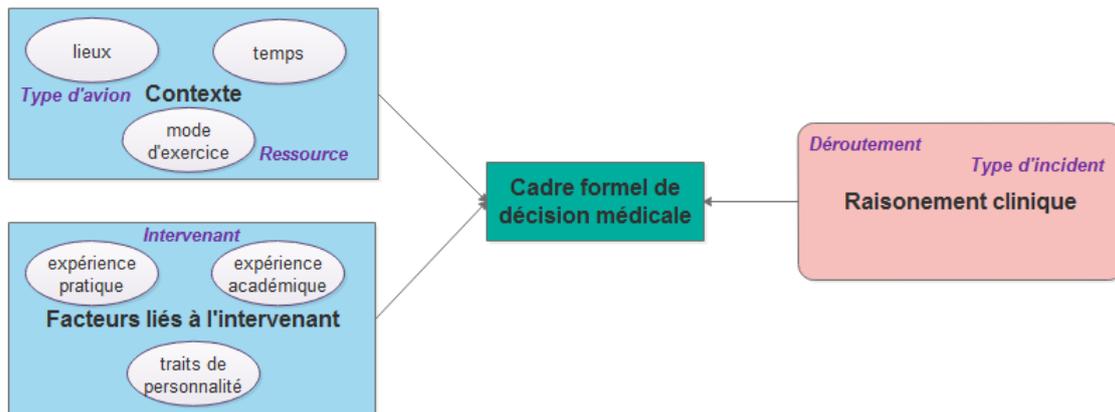


FIGURE 4.4 : Facteurs déterminants de la prise en charge médicale sous incertitude

l'intervenant. Chaque itemset dénote ainsi une décision possible selon l'intervenant. Le cadre formel de décision contient donc toutes les décisions possibles.

Tout itemset dans le cadre est inclus (définition 5) dans un itemset évidentiel, élément focal de la BDE. La relation entre les itemsets du cadre formel et les itemsets évidentiels est la même entre les mondes possibles imprécis dans la forme non compacte de la BDE et des états précis qui en découlent (section 4.4.4).

Soit T2 la transaction qui contient l'incertitude de l'intervenant dans la BDE décrit dans le tableau 4.2. Alors voici dans le tableau 4.8 l'ensemble des itemsets du cadre formel :

Itemsets du cadre	Itemsets évidentiels
(SP, D)	$(\{SP, S\}, D)$
(S, D)	$(\{SP, S\}, D)$
(SP, ND)	$(\{SP, S\}, ND)$
(S, ND)	$(\{SP, S\}, ND)$
(RS, D)	(RS, D)
(RS, ND)	(RS, ND)

TABLE 4.8 : Itemsets du cadre formel de décision

Nous considérons les itemsets du cadre formel comme « itemset d'utilité élevée » (en anglais high-utility itemset (HUI)) [Lan+14] [Lin+16]. L'itemset d'utilité élevée est une extension du problème

d'itemset fréquent et prend en compte le fait que ce ne sont pas toujours les itemsets les plus fréquents qui sont les plus importants [Lan+14] [Lin+16]. Pour cela, on a introduit la notion de poids pour les items. L'utilité et le support ne dépendent plus seulement de la fréquence des itemsets mais aussi des poids des items. Le seuil défini par rapport au support des itemsets permet de trouver les itemsets d'utilité élevée.

L'utilité de nos itemsets n'est pas quantitative comme dans les travaux de Zhang et ses collègues [Zha+17]. Toutefois, cette notion d'utilité joue le même rôle qui est de distinguer les itemsets selon leur degré d'importance.

Le processus de création du cadre formel de décision médicale dans l'avion est décrit dans la figure 4.5 :

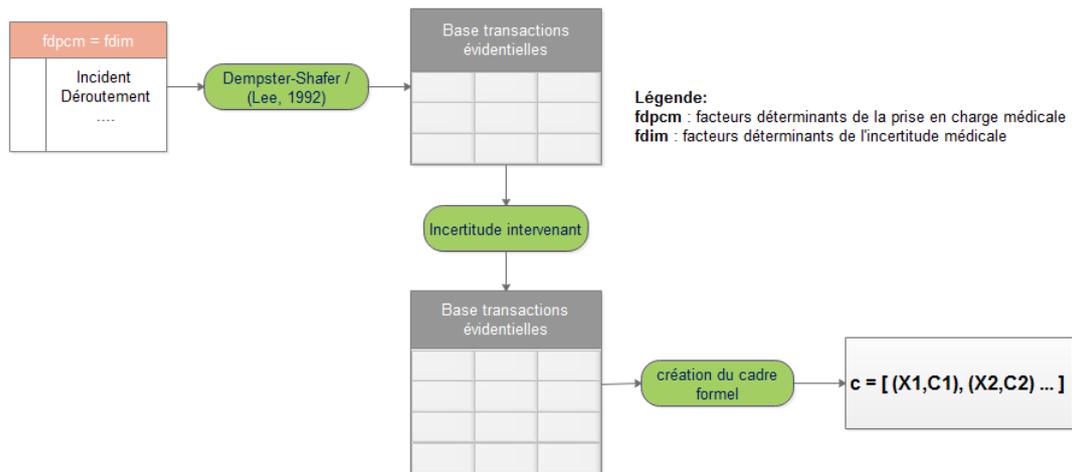


FIGURE 4.5 : Processus de création du cadre formel de décision

4.5.1 Construction de la BDE de la prise en charge médicale dans l'avion

La construction de la BDE s'appuie sur la théorie de Dempster-Shafer qui porte aussi la dénomination de théorie de l'évidence ou théorie des croyances. Plusieurs raisons ont motivé notre choix d'une approche basée sur la théorie de Dempster-Shafer.

- L'inadéquation des méthodes probabilistes basées sur le modèle la théorie de l'Utilité Espérée pour la prise en charge médicale dans l'avion. Ces modèles plutôt destinés à l'économie ne sont pas toujours adaptés aux problèmes de décisions médicales où les variables ne sont pas toujours quantitatives. Ils ont inspiré des travaux sur la décision médicale sous incertitude [BBE13]; [ELS85]; [PK75]; [Dju+14]. Ces méthodes utilisent en général 4 types de décisions : « Malade » et « pas malade » s'il s'agit de diagnostic, « Traiter » et « Pas traiter » lors de décisions thérapeutiques. Elles ne proposent donc pas un accompagnement au décideur dans l'acceptation ou le refus de chaque action médicale. Aussi, elles quantifient l'état de santé en nombre d'années d'espérance de vie, ce qui est inadapté dans notre contexte.
- La capacité de l'implémentation du contexte de fouille de données des incidents médicaux dans l'avion grâce notamment au mécanisme de l'extension vide qui à la différence de la marginalisation n'existe pas en probabilités. Le problème d'implémentation du contexte formel de prise en charge médicale dans l'avion est expliqué à la section (4.5.1.1).

- Cette approche permet aussi de saisir l’aspect épistémique de l’information, ce qui empêche de réduire l’imperfection de cette dernière à son caractère aléatoire.

4.5.1.1 Contexte formel de la prise en charge médicale dans l’avion

La prise en charge médicale consiste à choisir chaque facteur déterminant (définition 1). Nous avons donc défini comme contexte formel de prise en charge médicale dans l’avion le triplet $cp = (P, F, R)$. P est l’ensemble des incidents médicaux pris en charge dans l’avion. F est un ensemble d’éléments de type facteurs déterminants, il contient donc tous les types d’incidents, le déroutement, le non déroutement, etc. R est une relation binaire entre P et F .

L’implémentation de ce contexte correspond au travail de prétraitement des données nécessaire avant l’application d’algorithmes de fouille de données. Il donne une base de transactions. Chaque transaction est une prise en charge et les items décrivant les attributs des transactions sont des éléments de type facteurs déterminants.

Dans ce cas précis, le prétraitement consiste d’abord à un mappage sémantique des incidents qui est nécessaire à cause du manque de normalisation, suivi d’une caractérisation de chaque transaction (prise en charge). La caractérisation des transactions consiste à définir la relation R du contexte, autrement dit, définir pour chaque transaction les éléments de type facteur déterminant liés. Ce travail est loin d’être facile à cause de la documentation inadaptée des incidents rapportés. Les études font souvent une description globale des transactions, mais pas une description transaction par transaction qui permettrait de mieux exploiter ces données. Les bases de transactions qu’on peut construire contiennent les mêmes types de facteurs et la plupart du temps c’est le « type d’incident ».

Ci-dessous nous avons les étapes d’un prétraitement des données contenues dans trois études phares de la littérature sur le sujet ([Mah+11], [Pet+13], [Cos15]). Nous construisons l’entrée d’une bibliothèque Open Source (SPMF) spécialisée dans l’exploration de données [Fou17] et qui propose des implémentations de plus d’une centaine d’algorithmes.

→ Le mappage sémantique effectué manuellement donne le tableau 2.1.

→ L’extrapolation conduit à la multiplication de chaque pourcentage par 10. On obtient environ 1000 transactions pour chaque étude et donc une base de 3000 transactions.

Le tableau 4.9 montre le support et la confiance de chaque type d’incident obtenus grâce à l’application de l’algorithme « Apriori Association Rules » qui réalise un travail un peu plus approfondi que ceux réalisés dans les études des incidents médicaux dans l’avion. Le chiffre 1 indique l’occurrence d’un incident et les autres chiffres représentent le type d’incident. Ce tableau compare les trois études considérées ([Mah+11], [Pet+13], [Cos15]) en fonction des indices de support et de confiance des patterns générés pour chaque type d’incident.

L’objectif de section est de parler des difficultés de l’implémentation d’un contexte formel de prise en charge complète de l’incident médical dans l’avion. L’algorithme « Apriori Association Rules » utilisé réalise le même travail de catégorisation que la quasi-totalité des études sur les incidents médicaux dans l’avion. Ces études ont permis aujourd’hui que des éléments comme typologie et la prévalence de ces incidents soient maîtrisées. Elles ont aussi permis l’amélioration des différents kits médicaux à bord. Mais cela ne suffit pas pour proposer une aide à un intervenant dans une prise en charge médicale dans l’avion.

Pattern	#SUP :	#CONF :	Incident	Peterson	Mahony	Costa
1 \implies 2	1276.0	42,5901202	Syncope or presyncope	37,4	43,8	46,4
1 \implies 3	367.0	12,2496662	Respiratory symptoms	12,1	15,9	8,7
1 \implies 4	449.0	14,9866489	Nausea or vomiting	9,5	19,5	15,9
1 \implies 5	106.0	3,53805073	Cardiac symptoms	7,7		2,9
1 \implies 6	98.0	3,27102804	Seizures	5,8	1,1	2,9
1 \implies 7	122.0	4,07209613	Abdominal pain	4,1	2,3	5,8
1 \implies 8	71.0	2,36982644	Infectious disease	2,8		4,3
1 \implies 9	99.0	3,30440587	Agitation or psychiatric symptoms	2,4	3,2	4,3
1 \implies 10	59.0	1,96929239	Allergic reaction	2,2	2,3	1,4
1 \implies 11	20.0	0,66755674	Possible stroke	2		
1 \implies 12	43.0	1,435247	Trauma, not otherwise specified	1,8	2,5	
1 \implies 13	16.0	0,53404539	Diabetic complication	1,6		
1 \implies 14	10.0	0,33377837	Headache	1		
1 \implies 15	50.0	1,66889186	Arm or leg pain or injury	1	1,1	2,9
1 \implies 16	20.0	0,66755674	Obstetrical or gynecological symptoms	0,5	0,1	1,4
1 \implies 17	54.0	1,8024032	Ear pain	0,4	5	
1 \implies 18	3.0	0,10013351	Cardiac arrest	0,3		
1 \implies 19	3.0	0,10013351	Laceration	0,3		
1 \implies 20	119.0	3,97196262	Other	6,9	2,1	2,9
1 \implies 21	11.0	0,36715621	Unknown	0,1	1	

TABLE 4.9 : Les types d'incidents médicaux dans l'avion

Pour cela il faut relier les facteurs déterminants de la prise de décision et c'est l'objectif de la BDE de prise en charge médicale dans l'avion que nous allons construire.

4.5.1.2 Le cadre de discernement de la décision médicale dans l'avion

De nombreux travaux de catégorisation des facteurs déterminants de la prise en charge médicale dans l'avion ont été effectués et ces travaux convergent. On peut donc considérer que quelle que soit la décision d'un intervenant dans la prise en charge médicale dans l'avion, elle est incluse dans une des catégories proposée dans ces études. C'est pourquoi dans ce travail nous considérerons l'incertitude de l'intervenant comme une imprécision disjonctive que nous allons représenter grâce à la théorie de Dempster-Shafer. Une imprécision disjonctive est incarnée par de multiples hypothèses singletons où l'une d'entre elles devrait représenter le phénomène réel, mais nous ne savons pas lequel [Bel13].

Nous avons donc transformé les facteurs déterminants de la prise en charge médicale en domaines de discernement. Le tableau 4.10 montre la transformation de trois variables de décision extraites de travaux de Peterson [Pet+13]. Les catégories dans chaque facteur, le type de chaque d'incident, le déroutement de l'avion, le non déroutement de l'avion, le type d'intervenant, et autres deviennent des hypothèses singletons, ou des items.

4.5.1.3 La BDE de la prise en charge médicale dans l'avion

Dans ce travail nous utilisons la théorie fondée sur les cas passés pour proposer une aide à l'intervenant. Cette théorie ne prévoyant comme seule forme de sélection de l'information, la similitude du cas passé avec le problème présent [Abe99], les résultats des travaux de catégorisation peuvent être considérés comme une affectation d'une masse de croyance générale aux différents hypothèses (tableau 4.11).

4.5.2 Intégration de l'incertitude de l'intervenant

Comme nous l'avons souligné au début de la section 4.5, l'incertitude renvoie à un jugement de vérité porté par le décideur sur un aléa. Le décideur est incertain lorsqu'il est dans l'incapacité de réaliser ce jugement. Dans le cas d'une ignorance totale, le décideur ne connaît pas tous les états de la nature, on dit que l'incertitude est « pure », ou « radicale ».

Lors d'une prise en charge médicale, l'intervenant doit choisir un item (une hypothèse) dans chaque domaine de discernement. Il est incertain lorsque dans au moins un domaine de discernement, il ne sait pas choisir entre deux ou plusieurs items. Deux cas peuvent être envisagés. Dans le premier cas, son incertitude est « radicale ». Il lui manque des informations sur un domaine de discernement donné. Dans le deuxième cas il distribue la masse de croyance entre un certain nombre d'hypothèses. Le tableau 4.12 montre deux transactions caractérisant les deux cas.

4.5.3 Création du cadre formel de décision

Les itemsets du cadre sont déduits de la transaction contenant l'incertitude de l'intervenant mais le calcul des masses qui leurs sont affectées nécessite la BDE complète. Nous avons vu dans la définition du cadre formel (définition 7) que chaque itemset du cadre formel est inclus dans un itemset évidentiel.

Définition 8. *(La masse d'un itemset du cadre formel)*

La masse affectée à un itemset du cadre formel de décision correspond au support (définition 6) de l'itemset évidentiel qui l'inclut. Elle reflète donc la croyance totale affectée à l'itemset dans le cadre général dont il est une hypothèse. Soit la transaction T contenant l'incertitude de l'intervenant (tableau

Type of incident	Action	Provider
<ul style="list-style-type: none"> • Syncope or pre-syncope : SP • Respiratory symptoms : RS • Nausea or vomiting : NV • Cardiac symptoms : CS • Seizures : S • Abdominal pain : AP • Infectious disease : ID • Agitation or psychiatric symptoms : APS • Allergic reaction : AR • Possible stroke : PS • Trauma, not otherwise specified : T • Diabetic complication : DC • Headache : H • Arm or leg pain or injury : ALP • Obstetrical or gynecologic symptoms : OG • Ear pain : EP • Cardiac arrest : CA • Laceration : L • Other : O • Unknown : U 	<ul style="list-style-type: none"> • Diversion : DD • Non Diversion : ND 	<ul style="list-style-type: none"> • Physician : PP • Nurses : PN • EMS : PEMS • Other Health care professionals : PO • In-flight Crew : PIC

TABLE 4.10 : Domaines de discernement des variables de décision

Type of incident	Action	Provider
<ul style="list-style-type: none"> • SP (0,374) • RS (0,121) • NV (0,095) • CS (0,077) • S (0,058) • AP (0,041) • ID (0,028) • APS (0,024) • AR (0,022) • PS (0,02) • T (0,018) • DC (0,016) • H (0,01) • ALP (0,01) • OG (0,005) • EP (0,004) • CA (0,003) • L (0,003) • O (0,069) • U (0,01) 	<ul style="list-style-type: none"> • DD (0,073) • ND (0,927) 	<ul style="list-style-type: none"> • PP (0,481) • PN (0,201) • PEMS (0,044) • PO (0,037) • PIC (0,237)

TABLE 4.11 : BDE de la prise en charge médicale dans l'avion

4.13). Les itemsets du cadre formel sont : (SP, D, PP) , (SP, ND, PP) , (S, D, PP) , (S, ND, PP) . Le cadre formel résultant de son association avec la BDE de la prise en charge médicale (tableau 4.11) est décrit dans le tableau 4.14

4.6 Le cadre formel de décision et l'aide à la décision

Les fonctions de croyance offrent plusieurs outils pour décider [Ram07] :

- ✓ On peut décider à partir d'une distribution de probabilités pignistiques obtenue à partir d'une bba.

Transactions	Type d'incident	Déroutement	Intervenant
Incertitude	$\{SP, S\}(1)$	$D(0, 5)$ $ND(0, 5)$	$PP(1)$
Ignorance totale	$RS(1)$	$\Theta_{Diversi\text{on}}(1) = \{D, ND\}(1)$	$PP(1)$

TABLE 4.12 : Incertitude de l'intervenant

Transactions	Type d'incident	Déroutement	Intervenant
T	$SP(0, 7)$ $S(0, 3)$	$D(0, 5)$ $ND(0, 5)$	$PP(1)$

TABLE 4.13 : Incertitude de l'intervenant

Cadre formel de décision	
Itemsets	$Bel_{Itemset} = support_{Itemset}$
(SP, D, PP)	$((0, 374 \times 0, 073 \times 0, 481) + (0, 7 \times 0, 5 \times 1)) \div 2 = 0, 1815$
(SP, ND, PP)	$((0, 374 \times 0, 927 \times 0, 481) + (0, 7 \times 0, 5 \times 1)) \div 2 = 0, 2585$
(S, D, PP)	$((0, 058 \times 0, 073 \times 0, 481) + (0, 3 \times 0, 5 \times 1)) \div 2 = 0, 076$
(S, ND, PP)	$((0, 058 \times 0, 927 \times 0, 481) + (0, 3 \times 0, 5 \times 1)) \div 2 = 0, 088$

TABLE 4.14 : Cadre formel de décision

- ✓ C'est aussi possible de décider à partir de la crédibilité ou de plausibilité. On est respectivement dans ces cas pessimiste ou optimiste.

Dans ce travail nous avons construit une matrice de décision associée aux supports (définition 6) pour proposer une aide à la décision.

4.6.1 Le Cadre formel de décision comme l'ensemble des états de la nature

Comme nous l'avons rappelé dans la section 4.3, un problème de prise de décision peut être représenté par un quadruplet $t = E, C, A, f$, où E est un ensemble d'états de la nature, C un ensemble de conséquences, A un ensemble des actions possibles et f , une relation de préférence sur A .

Toute action est sujet au préalable d'une décision. On peut donc dire que les actions sont les conséquences des décisions du décideur et les états de la nature sont les conséquences des décisions de la nature. Le tableau 4.15 est une matrice de décision de trois états de nature (e_1, e_2, e_3) et deux actions (a_1, a_2).

f	e_1	e_2	e_3
a_1	$f(e_1, a_1)$	$f(e_1, a_2)$	$f(e_1, a_3)$
a_2	$f(e_1, a_2)$	$f(e_2, a_2)$	$f(e_3, a_2)$

TABLE 4.15 : Matrice de décision

Dans certains domaines de discernement qui constitue le cadre général de la prise en charge médicale dans l'avion, les décisions sont effectuées par l'intervenant. C'est le cas « Déroutement ». Dans d'autres domaines comme le « type de l'incident », le « type d'intervenant », c'est la nature qui décide. La nature décide dans le sens « ce n'est pas sous le contrôle de l'intervenant ». Nous pouvons donc construire les états de la nature et les actions de la prise de décision médicale dans l'avion en fonction des domaines de discernement.

Définition 9. (les actions de la prise en charge médicale dans l'avion)

Nous considérons comme action toute hypothèse dans un domaine de discernement ou c'est l'intervenant qui décide. Les principales actions sont issues du cadre « Déroutement ». L'utilisation des ressources peut aussi être considérés comme une action mais elle n'est pas assez documentée dans les études pour qu'on plus en faire un domaine de discernement. Formellement une action a est définie comme suit : $a \subseteq \Theta_{Déroutement}$.

Définition 10. (les états de la nature de la prise en charge médicale dans l'avion)

Un état de nature est le produit cartésien de tous les domaines de discernement qui ne sont pas sous le contrôle de l'intervenant. Il s'agit dans ce travail de tous les domaines excepté le « Déroutement ». Formellement une état de la nature e est définie comme suit : $e \subseteq \Theta - \Theta_{Déroutement}$.

Le cadre formel de décision contient toutes les décisions qu'un intervenant dans la prise en charge médicale dans l'avion peut prendre. Il suffit donc pour construire la matrice de décision. Le tableau 4.16 montre la matrice de décision correspondante au cadre de discernement construit à partir de la transaction décrite dans le tableau 4.13. Les états de la nature sont (SP, PP) et (SP, PP) . D et ND constituent les actions.

D'après les définition 9 et 10 des actions et des états de la nature, le couple $A \times E$ correspond au cadre général Θ . Ce qui fait des fonctions définies sur la BDE de la prise en charge médicale dans l'avion (bba, support, plausibilité), des fonctions candidates pour construire la matrice de décision. Le tableau 4.17

f	SP, PP	S, PP
D	$f(SP, PP, D)$	$f(S, PP, D)$
ND	$f(SP, PP, ND)$	$f(S, PP, ND)$

TABLE 4.16 : Un exemple de matrice de décision

montre que de la même manière que le support affecté à chaque itemset dans le cadre formel (définition 8), la fonction f est définie sur le cadre formel. La fonction support définit un ordre basé sur les croyances. Pour qu'elle soit une relation de préférence, nous lui associons la fiabilité des actions (déroutement).

f	SP, PP	S, PP
D	$f(SP, PP, D) = 0,1815$	$f(S, PP, D) = 0,076$
ND	$f(SP, PP, ND) = 0,2585$	$f(S, PP, ND) = 0,088$

TABLE 4.17 : Un exemple de matrice de décision basée sur le support

4.6.2 La fiabilité des déroutements

Selon Air France, sur les 213 déroutements entre 2004 et 2010 [Rod+12] [RF11], 40% n'étaient pas pertinents. La pertinence varie aussi selon le profil de l'intervenant. Elle est à 85% si c'est le SAMU qui le propose, descend à 50% si c'est un médecin et n'est plus qu'à 14% si c'est l'équipage. Le tableau 4.18 montre la matrice de décision avec la prise en compte de la fiabilité des déroutements.

f	SP, PP	S, PP
D	$(0,1815 \setminus 0,5)$	$(0,076 \setminus 0,5)$
ND	$(0,2585 \setminus 0,5)$	$(0,088 \setminus 0,5)$

TABLE 4.18 : Un exemple de matrice de décision basée sur le support et la pertinence des actions

4.7 Le cadre formel de décision et le raisonnement clinique

L'une des principales failles des systèmes de décisions médicales est l'absence de prise en compte des interactions entre les actions complémentaires ou successives de la thérapie [Abe99], autrement dit l'absence de prise en compte du raisonnement clinique dans son ensemble. Ces actions successives d'une thérapie sont bien décrites et ordonnées dans des outils de raisonnement clinique comme le chemin clinique.

En situation d'incertitude, le raisonnement clinique intègre d'autres actions qui peuvent être des examens supplémentaires, des demandes d'avis à des experts. Dans le contexte avionique la principale action qui intègre le raisonnement clinique est le déroutement. Le cadre formel intègre le déroutement et toutes les autres actions du raisonnement clinique qui sont regroupées dans l'action « non déroutement ».

Notre processus de décision s'appuie sur un cadre formel de décision associé au chemin clinique pour générer les actions autres que le déroutement, avec les actions médicales (figure 4.6). D'où la formalisation du chemin proposée ci-dessous.

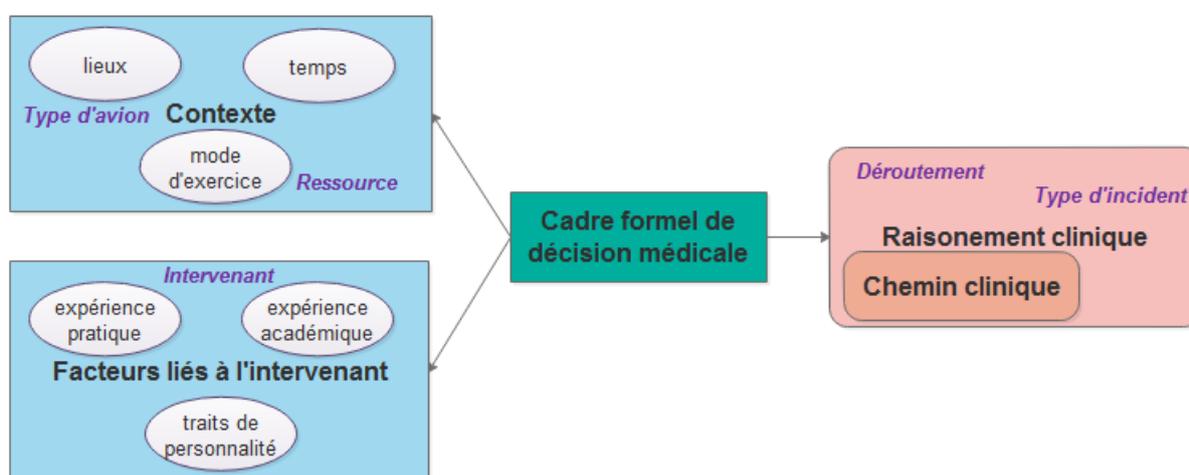


FIGURE 4.6 : Cadre formel et chemin clinique

4.7.1 Formalisation du Chemin clinique

Nous utilisons le type de Chemin Clinique trifocal proposé par Psiuk et Gouby [PG13]. Il est construit à l'aide d'un Plan de Soins Type et définit les cibles à traiter à partir des domaines cliniques. Ce choix s'explique par la richesse et la clarté de leur démarche avec notamment la séparation des domaines cliniques en trois catégories (au lieu de deux) : (i) les réactions humaines, (ii) les complications liées à la pathologie et (iii) les effets secondaires du traitement.

Le tableau 4.19 montre la structure du chemin clinique. Il est constitué d'incidents cibles correspondant aux problèmes qu'il faut prendre en charge. Ces problèmes sont résolus grâce à ces actions qui elles, nécessitent des ressources. Les actions ont des dépendances et produisent des résultats. Les incidents cibles constituent les composants principaux du chemin clinique et représentent les problèmes à traiter lorsqu'un incident survient.

Pour toute personne victime d'un incident médical ou atteinte d'une pathologie donnée, on définit trois domaines cliniques qui contiennent les incidents cibles :

Chemin clinique		
Cible 1	Action 1	Début, durée, période
	Action 2	Début, durée, période
	Action 3	Début, durée, période
Cible 2	Action 2	Début, durée, période
	Action 4	Début, durée, période
...

TABLE 4.19 : Structure du chemin clinique

- Signes et symptômes
- Complications liées à la pathologie et aux effets secondaires de traitement
- Réactions humaines physiques et physiologiques

Les deux derniers domaines sont définis à partir du premier. Le Chemin Clinique organise les interventions définies pour la prise en charge des observations dans chaque domaine clinique. Les observations dans le cas d'un enfant atteint de gastro-entérite aiguë peuvent correspondre aux éléments descriptifs fournis dans le tableau 4.20.

4.7.1.1 Le chemin clinique comme ontologie

Pour faciliter l'informatisation du chemin clinique, Ye et al. [Ye+09] ont proposé une modélisation ontologique. Leur formalisation du chemin clinique est orientée « objectif ». C'est-à-dire que pour chaque intervention, on part des objectifs pour arriver aux actions (figure 4.7). L'ordre des relations est Chemin Clinique → Objectif → Intervention (Action).

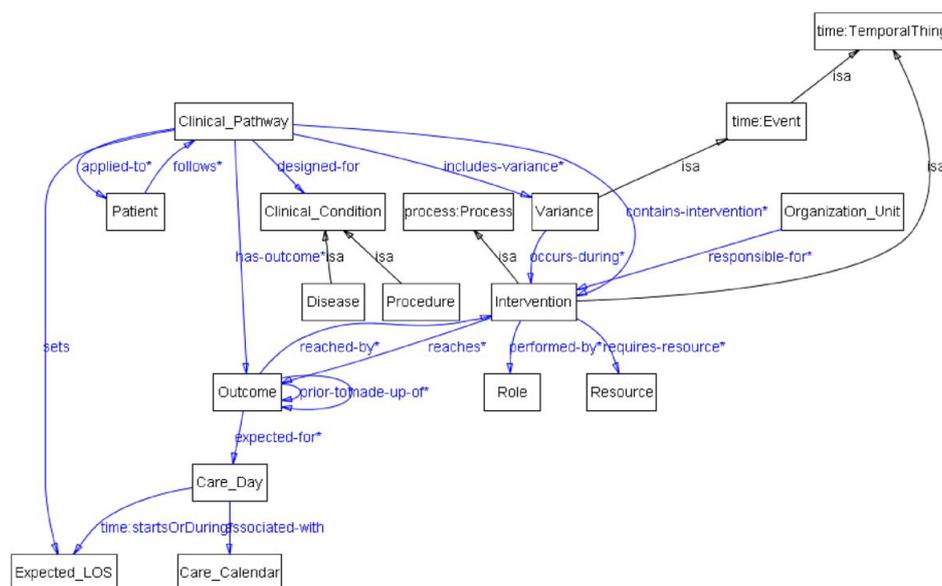


FIGURE 4.7 : Modèle ontologique du chemin clinique – formalisation orientée objectif

Signes et symptômes	Complications liées à la pathologie et aux effets secondaires de traitement	Réactions humaines physiques et physiologiques
<ul style="list-style-type: none"> ○ Diarrhées ○ Vomissements ○ Perte de poids ○ Douleurs abdominales ○ Fièvre ○ Déshydratation 	<ul style="list-style-type: none"> ○ Risque de transmission d'infection nosocomiale ○ Risque de contracter d'infection nosocomiale ○ Risque de contaminer des parents, fratrie ○ Risque choc (hypovolémique, toxi-infectieux) ○ Choc (hypovolémique, toxi-infectieux) ○ Risque d'extravasation ○ Extravasation ○ Troubles neurologiques sur troubles ioniques ○ Risque de SHU (syndrome hémolytique et urémique) ○ Risque de difficulté d'abord veineux ○ Difficulté d'abord veineuse ○ Irritation du siège ○ Risque de douleur 	<ul style="list-style-type: none"> ○ Anxiété des parents ○ Risque de perturbation de la dynamique familiale et perturbation de la dynamique familiale ○ Pleurs de l'enfant ○ Risque d'agitation de l'enfant ○ Agitation de l'enfant ○ Malaise des parents ○ Capacité des parents à gérer le retour à domicile

TABLE 4.20 : Incidents cibles dans le cas d'une gastro-entérite aiguë

Pour mieux modéliser le chemin clinique trifocal, nous avons utilisé une formalisation orientée « action ». L'ordre des relations sera donc Chemin Clinique → Incident cible → Action → Objectif (Résultat) (figure 4.8). Ce qui correspond mieux à la structure du chemin clinique trifocal.

La formalisation orientée action du chemin clinique structure le problème médical en un ensemble d'incidents cibles pris en charge grâce à des actions médicales. Elle résout ainsi le problème de la granularité de la décision médicale et du type de réponse requise dans le cadre de la prise en charge des incidents médicaux dans l'avion.

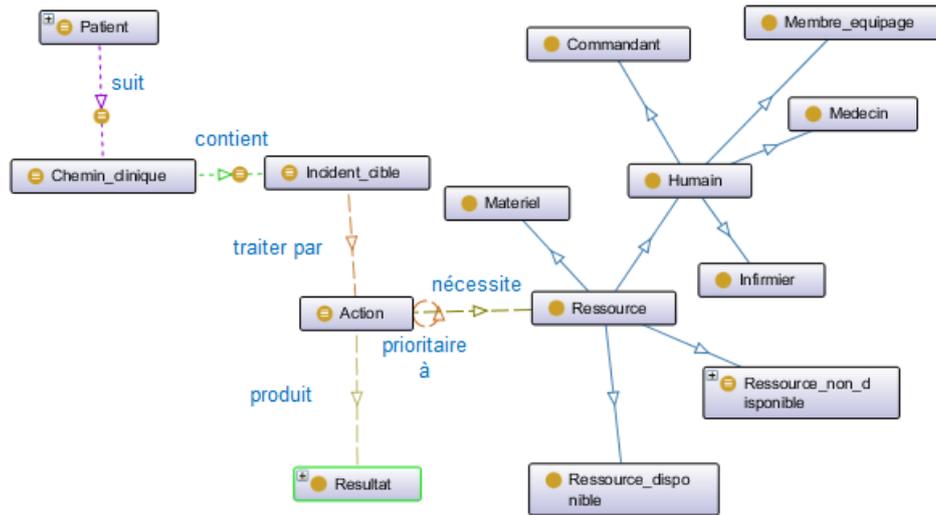


FIGURE 4.8 : Modèle ontologique du chemin clinique - formalisation orientée action

4.7.1.2 Le chemin clinique comme Base de Préférences Conditionnelles

Le problème de la décision médicale dans l'avion avec une formalisation ontologique du chemin clinique consiste en un ordonnancement des objets de la classe Action. Une manière efficace pour décrire le chemin clinique est de le considérer comme une base de préférences conditionnelles.

Une base de préférence conditionnelle est un triplet $PB = (T, A, P)$ où T est une base de connaissances en logique de description, A un ensemble fini de concepts, et P un ensemble fini de préférences conditionnelles [LS07].

Une préférence conditionnelle est de la forme $(\alpha|\varphi)[s]$, où α et φ sont des concepts (appelés tête et corps, respectivement), et s (appelé sa force) un nombre entier compris entre 1 et 100. Les bases de préférences conditionnelles consistent en une base de connaissances en logiques de description et un ensemble fini de préférences conditionnelles, qui sont des énoncés de la forme "généralement, dans le contexte φ , la propriété α est préférée à la propriété $\neg\alpha$ avec la force s ". Dans notre cas, les chemins cliniques définissent déjà un classement par préférence des actions médicales et donc on peut se passer de cette puissance, car ceci n'apporte pas d'information additionnelle à la formalisation. Dans ce travail on se passe de la puissance s et la préférence conditionnelle sera donc de la forme $(\alpha|\varphi)$. $(\alpha|\varphi)$ exprime que généralement, parmi les objets satisfaisant φ , ceux satisfaisant α sont préférés à ceux qui satisfont $\neg\alpha$.

Intuitivement, les bases de préférences conditionnelles $PB = (T, A, P)$ encodent l'ordonnancement d'un ensemble d'individus o . Dans notre cas T est l'ontologie du chemin clinique en logique de description, et A contient le concept Action.

$T =$

- {Matériel \subset Ressource ;
- Humain \subset Outcome ;
- Médecin \subset Humain ;
- Membre équipage \subset Humain ;

suit(Patient, Chemin clinique);
 prioritaire à (Action, Action);
 contient(Chemin clinique, Incident cible);
 traiter par(Incident cible, Action);
 produit(Action, Résultat);
 nécessite(Action, Ressource);
 ...}

$A = Action$

$P = \cup_{Action\ prioritaire}(Action_p, Action) | Action$

Cette formule signifie que parmi les instances qui sont des actions, on préfère celles qui sont prioritaires. On définit ainsi un ordonnancement des actions, car cette relation est aussi transitive.

4.8 Une architecture d'un SADM pour les incidents médicaux dans l'avion

De même qu'une décision politique se réfère à un modèle politique, une décision informatique se réfère à un modèle de données et/ou de connaissances. En l'absence de modèle les décisions deviennent aléatoires [DF91]. Comme nous l'avons indiqué dans l'introduction générale, ce travail de thèse s'inscrit dans le cadre de la proposition d'un modèle de données et de connaissances pour les processus de décision médicale dans l'avion.

La réalisation de modèle de données ou de connaissances associe aux données enregistrées un modèle de représentation puis de signification des données. Si dans le chapitre précédent, nous avons cherché à définir et à conceptualiser le problème à résoudre dans la prise en charge médicale dans l'avion, dans la première partie de celui-ci, nous avons construit un cadre formel contenant toutes les décisions et les actions qu'un intervenant peut entreprendre ou exécuter. L'architecture du SADM proposée définit le modèle de données et de connaissances, il consiste à l'intégration de la conceptualisation du problème de la prise en charge médicale et du cadre formel de décision associé à la formalisation du chemin clinique. La figure 4.9 montre cette architecture.

4.8.1 Description du modèle de données et de connaissances et des technologies pour la prise en charge des incidents médicaux dans l'avion

Le modèle de données et de connaissances est basé sur les ontologies et alimente un moteur d'inférence. Il intègre en plus des ontologies d'incidents (section 3.7), du chemin clinique (section 4.7.1.1), une ontologie Dempster-Shafer (DS-Ontology) proposée par Bellenger [Bel13] (figure 4.10).

L'ontologie Dempster-Shafer basée sur la théorie des fonctions de croyance et est facilement importable dans une ontologie de domaine. La caractérisation de l'incertitude est flexible et facile. Comme l'incertitude est représentée à l'aide des fonctions de croyance, tout concept avec un aspect incertain est un

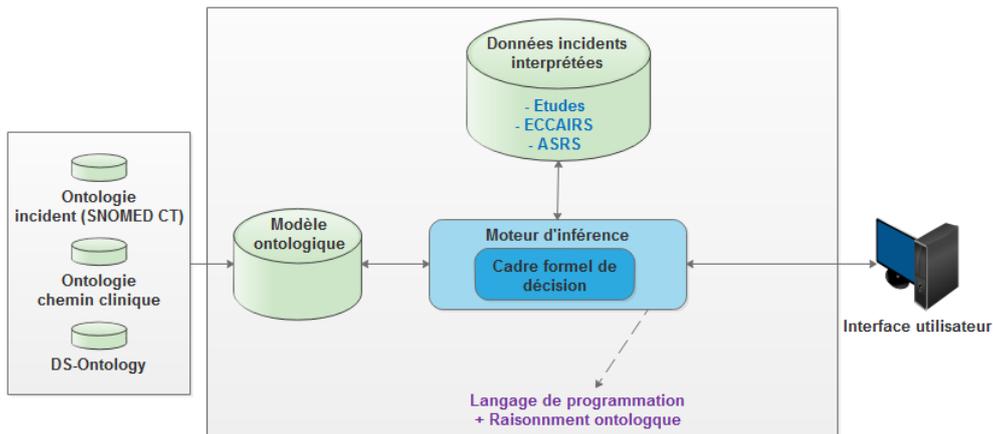


FIGURE 4.9 : Architecture du système d'aide à la décision pour incidents médicaux dans l'avion

DS concept (relation hasDS_concept). Les hypothèses sont reliées au domaine de discernement grâce à la relation hasDS_hypothesisElement. Chaque concept incertain (DS Concept) a une masse, une crédibilité et une plausibilité.

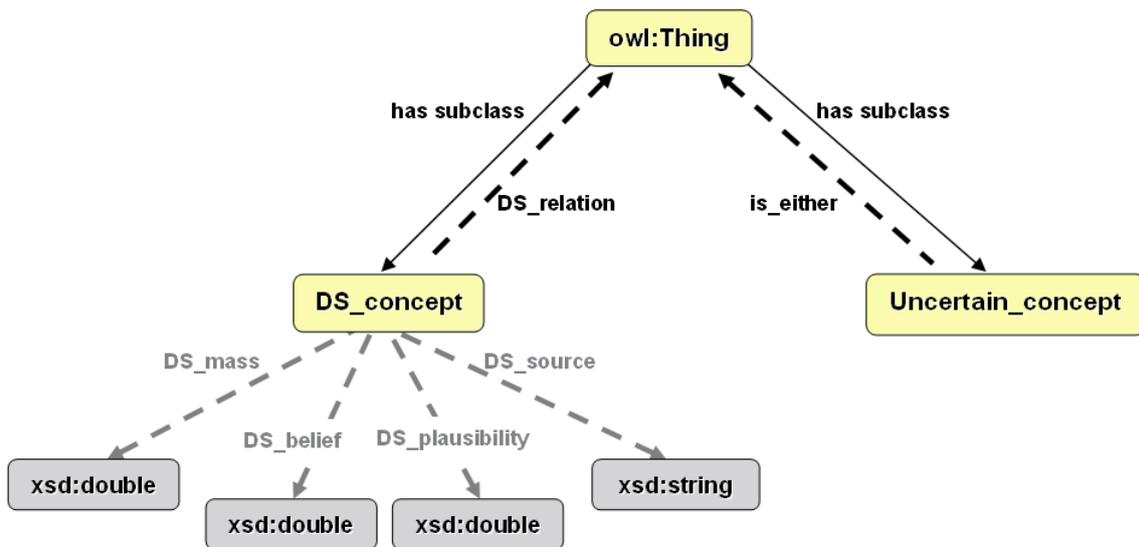


FIGURE 4.10 : Structure de l'ontologie Dempster-Shafer

L'intégration de ces parties d'ontologies est facilitée par les concepts qui leurs sont communs en particulier, le type d'incidents. Tout concept qui est facteur déterminant de prise en charge médicale dans l'avion est transformé en « concept Dempster-Shafer » (incident, action (déroutement/non déroutement), etc.

4.8.1.1 Ontologie d'incident : cas de Vomissement

La caractérisation de l'incident Vomissement dans SNOMED CT fournit 61 concepts (champs « Class count » dans la figure 4.11) et 12 relations (champs « Object property count » dans la figure 4.11).

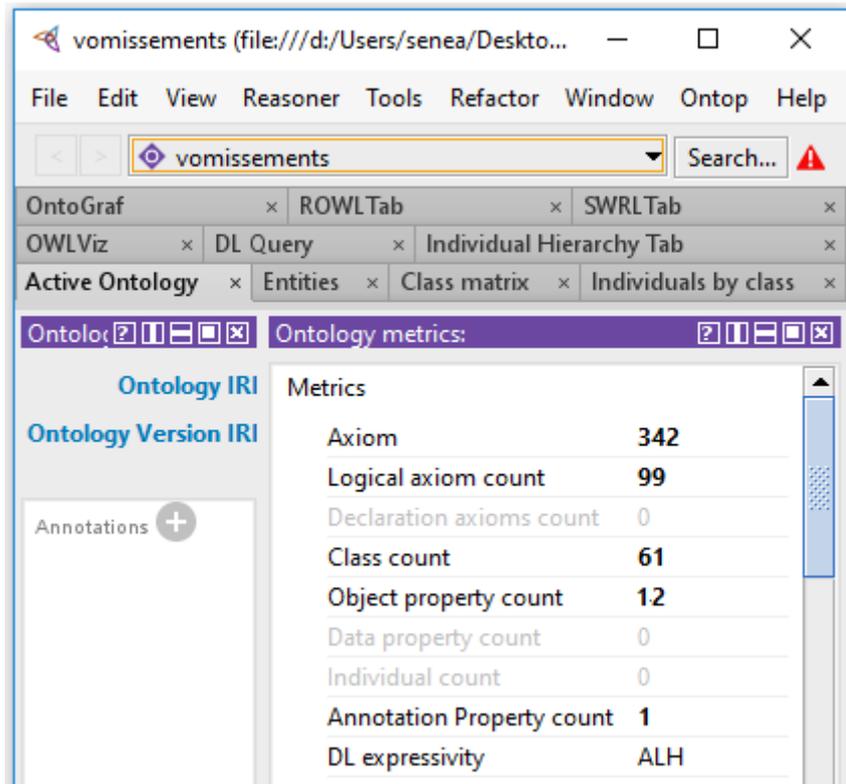


FIGURE 4.11 : Les métriques de l'ontologie de l'incident vomissement

La figure 4.12 montre la structure des concepts qui caractérisent les incidents médicaux dans l'avion. Pour chaque concept de type « Constatation clinique » relié à un incident, on récupère son arborescence.

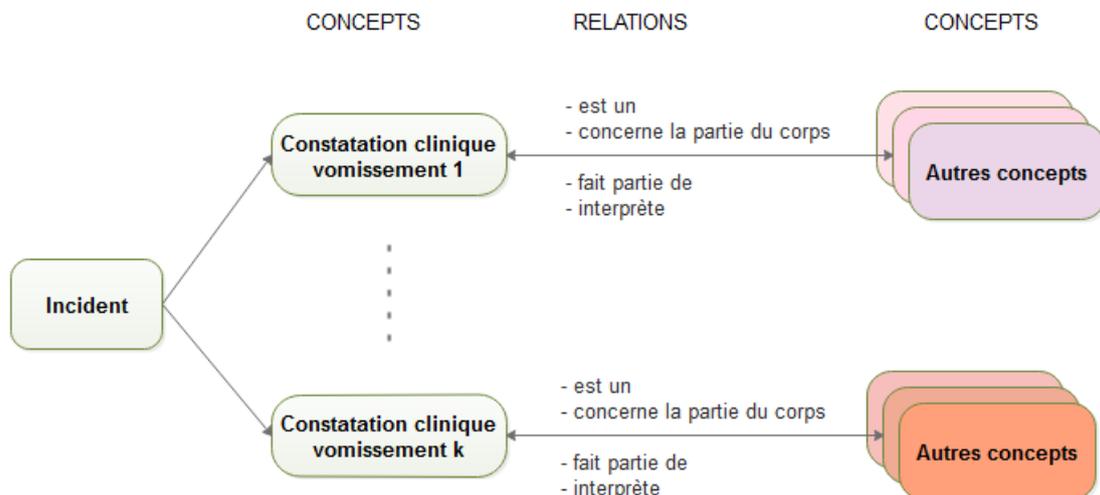


FIGURE 4.12 : Structure des ontologies d'incidents

La figure 4.13 montre une partie des concepts de type constatation clinique concernant l'incident « vomissement ». Elle montre également des interprétations possibles du Concept « Diarrhée et vomissement » et les parties du corps qui peuvent être concernées.

Les ontologies d'incidents préparent l'intervention et facilitent la compréhension des incidents. Elles

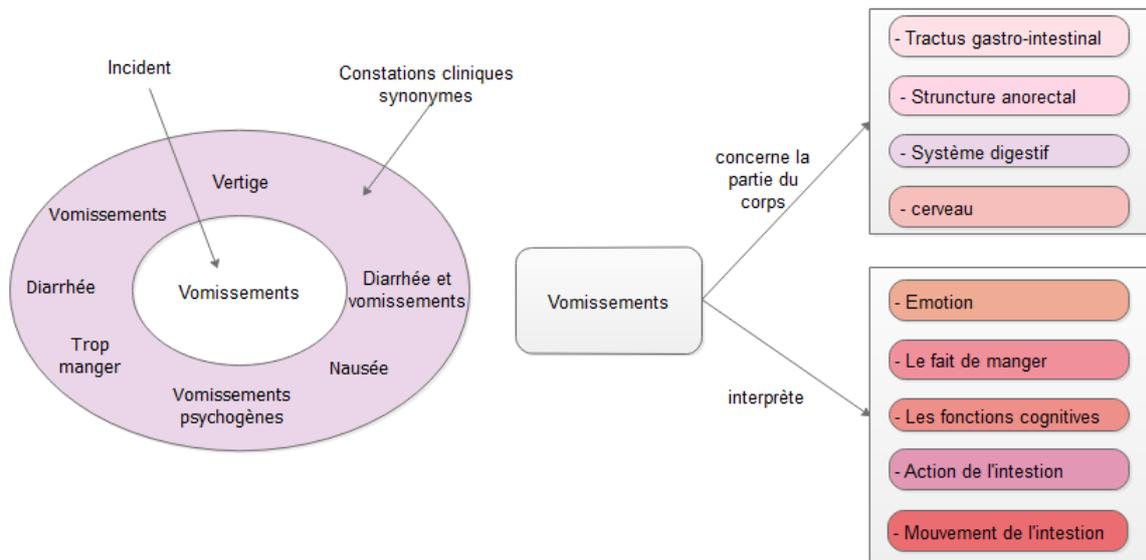


FIGURE 4.13 : Incident Vomissements

permettent de faire la correspondance entre les dénominations différentes d'un même incident et facilitent une formalisation sémantique pour des opportunités de raisonnements.

4.8.1.2 Utilisation du concept Procédure pour l'intégration de chemin clinique

SNOMED CT intègre un concept « Procédure » qui décrit les activités utiles dans la prestation de soins de santé. Le chemin clinique fournit une description efficace d'une procédure qui n'est autre que la prise en charge médicale du patient pour une pathologie donnée. Le concept « Procédure » subsume donc le concept « Chemin clinique » (figure 4.14). Nous utilisons la relation « Procédure associée » (« Associated procedure » dans SNOMED CT) pour relier un incident au chemin clinique correspondant.

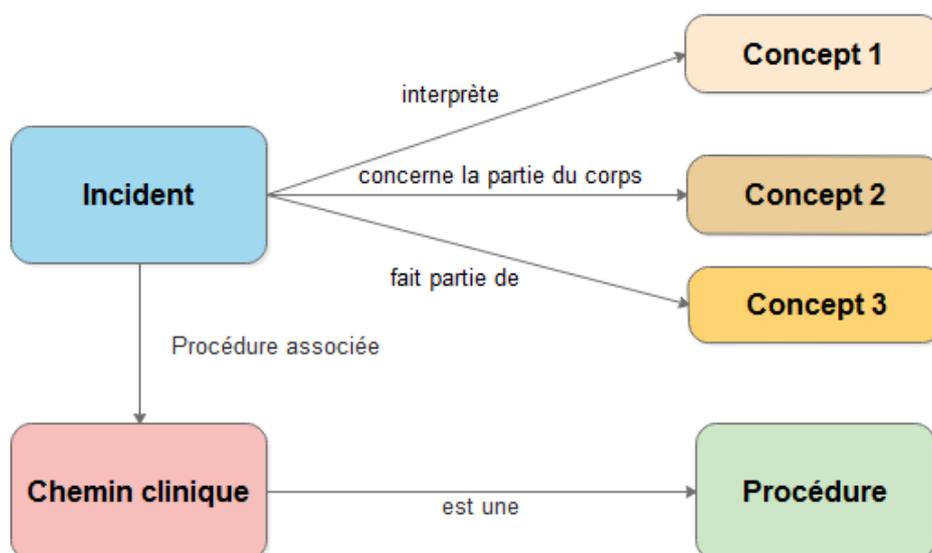


FIGURE 4.14 : Liens sémantiques entre Incident, Chemin clinique et Procédure

4.8.1.3 Intégration du cadre formel de décision dans le modèle de connaissances

Dans le cadre d'une procédure normale sans incertitude, l'incident médical relié à un chemin clinique qui est un outil de raisonnement clinique pourrait assurer l'aide à la décision pour la prise en charge. En situation d'incertitude le raisonnement clinique intègre d'autres actions comme les demandes d'analyses complémentaires. Dans le contexte de la prise en charge médicale dans l'avion, il s'agit surtout de l'action de déroutement qui peut être assez coûteuse financièrement.

Dans la définition de l'incertitude médicale, nous avons vu que le type d'incident et le raisonnement (chemin clinique) sont des facteurs déterminants de l'incertitude et de la prise en charge. Pour intégrer l'incertitude de l'intervenant dans le modèle de connaissances, nous avons relié à l'incident tous les autres facteurs déterminant de l'incertitude (type d'intervenant, type d'avion, ...), puis transformé tous ces facteurs en « DS-Concept » (figure 4.15).

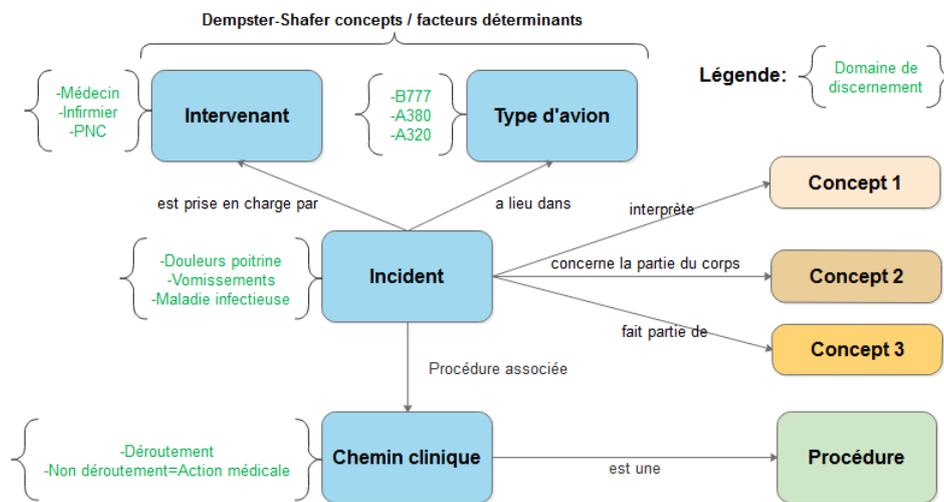


FIGURE 4.15 : Transformation des facteurs déterminants de l'incertitude en concepts Dempster-Shafer

Le modèle de données et de connaissances contient à ce stade tous les concepts nécessaires pour construire le cadre formel de décision.

4.8.2 Outils de raisonnement pour la prise en charge des incidents médicaux dans l'avion

Dans les modèles de données et de connaissances basées sur les ontologies, le raisonnement est géré par le moteur d'inférence. Ce dernier est bien aidé dans cette tâche par les ontologies qui non seulement contiennent des connaissances du domaine mais aussi des connaissances de raisonnement. Les ontologies fournissent comme mécanismes de raisonnement :

- ❖ La subsomption (classification)
- ❖ La vérification de consistance
- ❖ La satisfaisabilité de concept (vérification d'instance)
- ❖ La réalisation : établir la classe la plus spécifique

Ces outils de raisonnement assurent la consistance et la cohérence des données et des connaissances. Ils augmentent ainsi la fiabilité des décisions. La subsumption permet de prendre en compte les appellations multiples du même incident.

En plus des mécanismes de raisonnement fournis par les ontologies, la prise en charge médicale dans l'avion nécessite la réalisation du cadre formel permettant le déploiement des calculs qui ont été effectués manuellement dans la section 4.5.3. Pour automatiser ces calculs, il est nécessaire d'interfacer les ontologies avec les langages de programmation. Ce que nous avons déjà effectué pour la construction des ontologies d'incidents et avec le langage java. D'autres travaux proposent l'utilisation d'autres API comme Apache Jena [BM12] [BG11] ou d'autres langages comme python [Lam17]. En intégrant le cadre formel de décision, le moteur d'inférence convertit les données et connaissances en éléments d'aide à la décision, puis en support aux actions à mettre en œuvre.

4.9 Conclusion

Dans ce chapitre nous avons proposé une approche de modélisation de l'incertitude basée sur la théorie de Dempster-Shafer et une architecture de SADM qui définit un modèle de données et de connaissances ontologiques nécessaires pour la prise en charge médicale dans l'avion.

Nous avons indiqué que le chapitre précédent réalise la première étape du processus de décision (Renseignement). Dans ce chapitre nous avons réalisé les deux étapes suivantes. Le cadre formel de décision contient toutes les décisions qu'un intervenant peut prendre, et il réalise donc l'étape de conception de la décision. L'étape qui suit la conception est le choix des options pertinentes (par exemple, tâches à réaliser ou à déléguer) et la matrice de décision contribue à sa réalisation.

Le cadre formel de décision permet de considérer toute l'incertitude du processus de décision et d'estimer les conditions spécifiques d'une situation à risque. Chaque itemset dans ce cadre dénote une décision contenant l'état de la nature et l'action qui lui est associée, ce qui nous a permis de proposer pour chaque cadre formel de décision une matrice de décision. A la différence des modèles de décision basés sur la théorie de l'Utilité Espérée, nos actions ne se limitent pas à « traiter » et « non traiter » ; elles couvrent tous les actes médicaux pertinents définis dans le processus de soins considéré.

Dans ce travail, nous proposons une approche basée sur les connaissances pour soutenir les processus décisionnels critiques et contribuer à fournir des soins médicaux appropriés pour les incidents médicaux en vol. Ces processus de prise de décision sont complexes en raison des incertitudes externes et internes liées aux procédures relatives aux passagers malades, aux compagnies aériennes, à la gestion du trafic aérien et aux communications à distance. Il existe donc plusieurs variables (type d'incident, action médicale, ou type d'intervenant) qui doivent être prises en compte à la fois pour des raisons scientifiques et pragmatiques, en ce qui concerne la santé des passagers des transports aériens.

Néanmoins, le fondement juridique de l'aide d'urgence dans les incidents médicaux en vol est souvent faible ou variable selon les pays, frustrant les médecins volontaires du manque d'informations et de connaissances claires. Même si le manque de certitude juridique et scientifique ne doit pas empêcher une proposition de solutions possibles pour répondre aux besoins médicaux critiques des passagers dans le transport aérien.

Conclusion générale

Contributions

Le travail présenté dans ce document concerne l'aide à la décision en santé avec une approche informatique qui prend en compte à la fois les caractéristiques du domaine médical et celles de l'environnement aéronautique. La problématique traitée examine particulièrement les moyens d'amélioration de la gestion des incidents médicaux en vol en utilisant les sciences informatiques pour améliorer la transmission des acquis de l'expérience, notamment par la mise en œuvre d'un ensemble de théories et de techniques du domaine de l'intelligence artificielle.

Même s'il existe des systèmes de retour d'expérience dans l'aéronautique, le problème des incidents en plein vol est souvent confiné à une analyse épidémiologique au sein des sciences médicales et ceci s'avère insuffisant pour saisir ses différentes facettes avec la complexité qui est sous-jacente. Pour bien saisir cette complexité, il a été au préalable nécessaire de mener des investigations pour acquérir une meilleure compréhension des situations concernant les conditions médicales et la pratique des fournisseurs des soins traitant des conditions aiguës et chroniques de leur patient durant le vol sur les lignes aériennes commerciales.

Notre travail de thèse se situe dans une volonté d'apporter un éclairage nouveau à ce problème de gestion des incidents médicaux à bord des avions via la combinaison des modèles et méthodes de représentation de connaissances avec les techniques de fouilles de données.

Les principaux apports de notre contribution sont les suivants :

□ Premier résultat

La définition et la conceptualisation du problème de prise en charge médicale dans l'avion ont requis l'établissement d'une démarche de structuration progressive. Tout d'abord, il a été nécessaire d'effectuer l'identification des facteurs déterminants de la prise en charge médicale dans l'avion à partir de la littérature des incidents médicaux dans l'avion.

L'étude du principal outil d'assistance médicale dans l'avion, la télémédecine, dans le cadre d'un SADM a permis d'étudier le contexte d'assistance médicale dans l'avion, les moyens de restitution des recommandations et l'infrastructure de communication qui sont des dimensions communes des taxonomies de télémédecine et des SADM.

L'étude des systèmes de report d'événements dans l'aéronautique a permis de faire le lien avec les incidents médicaux dans l'avion. Ces systèmes pourtant normalisés et dont certains sont même régis par la loi (ECCAIRS) ne sont toujours pas pris en compte dans la gestion des incidents médicaux dans l'avion.

La modélisation de thématiques pour la définition de l'incident médical dans l'avion a permis de conceptualiser ce dernier en le reliant aux concepts de SNOMED CT grâce à l'analyse sémantique latente et à l'allocation de Dirichlet latente.

L'exploration de SNOMED CT pour chaque incident permet d'établir des liens sémantiques entre concepts synonymes (relation IS-A) ou pas (autres relations) et facilite ainsi la première phase de clarification sémantique du vocabulaire de la prise en charge médicale.

□ Deuxième résultat

La construction d'un cadre formel de décision qui définit les éléments au service de la conception de l'aide à la décision médicale dans l'avion. Ce cadre contient l'ensemble des décisions et actions qu'un intervenant peut entreprendre. Il est basé sur la théorie de Dempster-Shafer et a nécessité un certain nombre de tâches.

La création d'une base de données évidentielles de la prise en charge médicale dans l'avion. Pour cela, il a fallu transformer les facteurs déterminants de la prise en charge médicale dans l'avion en domaines de discernement et définir des mécanismes de combinaisons de croyances provenant de domaines de discernement distincts. La prise en compte de l'incertitude de l'intervenant permet de définir les itemsets dans le cadre formel et chaque itemset dénote une décision possible.

La construction d'une matrice de décision à partir du cadre formel de décision. Lors de la prise en charge médicale dans l'avion, l'intervenant fixe pour chaque facteur déterminant de la prise en charge médicale (domaine de discernement) une hypothèse (un état). Pour certains domaines de discernement comme le « Déroutement », c'est lui-même qui décide et dans d'autres domaines comme pour le type d'incident et le type d'intervenant, c'est la nature qui décide. Ces domaines de discernement constituent donc respectivement les actions et les états de la nature de la matrice de décision.

La définition de la fonction de préférence sur les actions à partir de croyance affectée à chaque itemset dans le cadre formel. Chaque intersection entre une action et un état de la nature est un itemset, ce qui permet de déterminer le niveau de croyance associé en lien avec le support des itemsets afin de définir un ordre sur les actions envisagées. Une relation de préférence est ainsi établie en y associant la fiabilité des déroutements.

□ Troisième résultat

La description architecturale synthétise notre proposition de spécification d'une aide à la décision pour incidents médicaux dans l'avion. L'architecture du SADM proposée relie la caractérisation du patient et de la prise en charge à l'incertitude de l'intervenant et permet de définir un modèle de données et de connaissances pour la décision médicale dans l'avion. Le modèle de connaissances offre une interopérabilité syntaxique et sémantique, facilite l'acquisition de nouvelles connaissances, intègre des connaissances de domaines mais aussi de raisonnement, ce qui facilite le travail du moteur d'inférence.

Enfin, le cadre architectural SADM est extensible pour permettre l'intégration de différentes méthodes algorithmiques capables d'extraire des patterns fréquents incertains à partir de telles données incertaines d'une manière plus efficace. Il est donc possible de pouvoir l'enrichir de manière flexible pour l'utilisation d'une ou plusieurs stratégies de raisonnement permettant des améliorations continues dans l'aide à la décision en situation critique.

Perspectives opérationnelles

La principale perspective concerne la caractérisation de plus de facteurs déterminants de la prise en charge médicale dans l'avion pour en faire des domaines de discernement et ainsi augmenter la fiabilité de

la décision.

Positionner le SADM sur le cadre de la prévention nous paraît aussi avantageux. En effet la prévention peut potentiellement être la meilleure méthode pour réduire les incidents médicaux en vol et après le vol en identifiant les patients les plus à risque afin de leur fournir les recommandations permettant les actions nécessaires pour atténuer les incidents futurs ou l'accroissement des facteurs conduisant aux conditions des incidents actuels.

Cela nécessitera une éducation des passagers qui ont tendance à retenir des informations médicales afin d'éviter un rejet par les compagnies aériennes. Il existe des divisions médicales au sein de certaines compagnies aériennes internationales qui offrent un service de vérification des mesures sanitaires préventives.

Une standardisation des problèmes médicaux dans l'avion permettra de fournir des pistes pour améliorer la sécurité des passagers et déterminer les moyens des technologies de l'information et de la communication (comme ceux de la télémédecine) au service de la gestion des incidents précis en vol. D'autres questions ouvertes sont relatives aux progrès de la recherche médicale sur l'approfondissement des connaissances relatives au rôle précis de l'altitude sur la santé et le bien-être de l'homme pour une meilleure compréhension des leviers d'actions préventifs, curatifs ou prédictifs dans la gestion opportune des incidents médicaux des passagers durant un transport aérien.

Travail restant à réaliser

A ce stade d'élaboration, nous avons proposé une méthode de conception de modèle de données et de connaissances pour la prise en charge médicale dans l'avion et une solution aux verrous présentés par la représentation sémantique et la gestion de l'incertitude, la conception reste à réaliser et donnera un SADM qui est un dispositif médical selon la réglementation EU 2017/745. Il devra donc être validé pour son utilisabilité (développement d'une interface homme-machine prenant en compte les conditions aéronautiques et le stress de l'utilisateur), faire l'objet d'une démarche d'analyse et de réduction des risques (ISO 14971) prenant notamment en compte les conditions avioniques.

La Section 17 de cette réglementation EU 2017/745 stipule que tout logiciel destiné à être utilisé avec des plates-formes informatiques mobiles doit être conçu et fabriqué en tenant compte des spécificités de la plate-forme mobile (par exemple, la taille et le rapport de contraste de l'écran) et les facteurs externes liés à leur utilisation (environnement variable en ce qui concerne le niveau de lumière ou de bruit). En l'occurrence, l'idée est d'améliorer les capacités de diagnostic et de vérification en vol afin d'orienter et d'améliorer la collaboration avec les services médicaux sur le terrain. En utilisant les listes de contrôle et les formulaires proposés par le système considéré, les agents de bord peuvent établir et enregistrer les données personnelles du patient (sexe, âge, taille et poids, etc.), ses antécédents médicaux (problèmes de santé établis, médicaments, allergies connues, maladies antérieures ou chirurgies), leurs signes vitaux actuels (pouls, tension artérielle, etc.), les symptômes et signes qu'ils présentent.

L'approche proposée est définie pour décrire les caractéristiques du système sous-jacent afin d'améliorer la gestion des urgences médicales en vol, en fournissant un travail de base pour la réconciliation sémantique entre les trois types de connaissances ontologiques (ontologie de tâche, ontologie d'inférence et ontologie de domaine). Ceci est fondamental pour développer des principes opérationnels et pour fournir des conseils pratiques concernant les ajustements appropriés pour les liens sémantiques utiles entre la vue du processus informatique et la vue de l'interaction humaine. Cette approche de recherche fournissant un système utilisant l'arrangement d'ontologies multicouches peut avoir des avantages et

des inconvénients. Les avantages sont entre autres la forte cohésion, le faible couplage et la facilité de modification ; tandis que les inconvénients sont souvent liés à la navigabilité du code et à l'incohérence possible des représentations multiples.

Le produit obtenu sera enfin évalué sur son service médical rendu éventuellement dans un premier temps dans le cadre des EHPAD (établissements d'hébergement pour personnes âgées dépendantes) (situation comparable aux conditions avioniques près) en vue de contrer l'effet statistique de la rareté des accidents médicaux en vol sur la durée de l'étude, puis enfin validé en conditions de médecine aéronautique. Un suivi après mise en production sera réalisé sous forme d'une vigilance réglementaire (effets indésirables) et de gestion des risques (suivi de l'opération du service).

Bibliographie

- [14a] *règlement (UE) No 376/2014*. 2014. URL : <http://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32014R0376&from=FR> (visité le 03/12/2017).
- [14b] *règlement (UE) No 376/2014*. Avec la coll. d'Européenne UNION. 2014. URL : <http://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32014R0376&from=FR> (visité le 03/12/2017).
- [AA17] Fawaz S. AL-ANZI et Dia ABUZEINA. "Toward an enhanced Arabic text classification using cosine similarity and Latent Semantic Indexing". In : *Arabic Natural Language Processing : Models, Systems and Applications* 29.2 (1^{er} avr. 2017), p. 189–195. ISSN : 1319-1578. DOI : 10.1016/j.jksuci.2016.04.001. URL : <http://www.sciencedirect.com/science/article/pii/S1319157816300210>.
- [Abe99] Philippe ABECASSIS. "Incertitude médicale et processus de décision". In : *Communication au Ateliers Thématiques en Économie de la Santé*. Rationalité et processus de décision, FORUM-LATEC. Dijon, 1999. URL : https://www.researchgate.net/profile/Philippe_Abecassis/publication/242560362_Incertitude_medicale_et_processus_de_decision/links/0deec52a040dc48e9f000000.pdf (visité le 11/04/2017).
- [Akh04] Abdellah AKHARRAZ. "Acceptabilité de la décision et risque décisionnel : un système explicatif de fusion d'informations par l'intégrale de Choquet". Thèse de doct. Université Savoie Mont Blanc, 1^{er} jan. 2004. URL : <http://www.theses.fr/2004CHAMS027> (visité le 13/02/2018).
- [All+17] Mehdi ALLAHYARI et al. "A Brief Survey of Text Mining : Classification, Clustering and Extraction Techniques". In : *arXiv :1707.02919 [cs]* (10 juil. 2017). arXiv : 1707.02919. URL : <http://arxiv.org/abs/1707.02919> (visité le 02/02/2018).
- [Arr07] ARRÉTÉ. *Arrêté du 25 septembre 2007 relatif au certificat de formation à la sécurité*. 2007. (Visité le 17/08/2017).
- [AS94] Rakesh AGRAWAL et Ramakrishnan SRIKANT. "Fast Algorithms for Mining Association Rules in Large Databases". In : *Proceedings of the 20th International Conference on Very Large Data Bases*. VLDB '94. San Francisco, CA, USA : Morgan Kaufmann Publishers Inc.,

- 1994, p. 487–499. ISBN : 978-1-55860-153-6. URL : <http://dl.acm.org/citation.cfm?id=645920.672836> (visité le 15/05/2017).
- [Att+16] Arthur E. ATTEMA et al. “An elicitation of utility for quality of life under prospect theory”. In : *Journal of Health Economics* 48 (1^{er} juil. 2016), p. 121–134. ISSN : 0167-6296. DOI : [10.1016/j.jhealeco.2016.04.002](https://doi.org/10.1016/j.jhealeco.2016.04.002). URL : <http://www.sciencedirect.com/science/article/pii/S0167629616300200>.
- [BA94] Ronald J. BRACHMAN et Tej ANAND. *The Process of Knowledge Discovery in A First Sketch*. AAAI Tech. Rep. WS-94-03, 1994.
- [Bal02] Gilles BALMISSE. *Gestion des connaissances : outils et applications du knowledge management*. Google-Books-ID : 2RL_PAAACAAJ. Vuibert, 2002. 266 p. ISBN : 978-2-7117-8697-8.
- [Bar09] Patrick BARLATIER. “Conception et implantation d’un modèle de raisonnement sur les contextes basée sur une théorie des types et utilisant une ontologie de domaine”. Thèse de doct. Université de Savoie, 2009. URL : <https://tel.archives-ouvertes.fr/tel-00678447/> (visité le 22/07/2016).
- [Bau18] JY BAUDOT. *Cercle des corrélations*. 2018. URL : <http://www.jybaudot.fr/Analdonnees/cerclecor.html> (visité le 03/04/2018).
- [BBE13] Loïc BERGER, Han BLEICHRODT et Louis EECKHOUDT. “Treatment decisions under ambiguity”. In : *Journal of Health Economics* 32.3 (mai 2013), p. 559–569. ISSN : 01676296. DOI : [10.1016/j.jhealeco.2013.02.001](https://doi.org/10.1016/j.jhealeco.2013.02.001). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0167629613000088> (visité le 18/04/2017).
- [Bel11] Rinaldo BELLOMO. “Recent advances in critical care medicine relevant to cardiac surgery”. In : *Heart, Lung and Circulation* 20.3 (2011), p. 170–172. URL : <http://www.sciencedirect.com/science/article/pii/S144395061001382X> (visité le 17/06/2016).
- [Bel13] Amandine BELLENGER. “Aide à la décision sémantique pour la diffusion d’informations”. Thèse de doct. INSA de Rouen, 2013. URL : <https://tel.archives-ouvertes.fr/tel-00845918/> (visité le 07/06/2017).
- [Ber07] Eta S. BERNER, éd. *Clinical Decision Support Systems : Theory and Practice*. 2^e éd. Health Informatics. New York : Springer-Verlag, 2007. ISBN : 978-1-4419-2223-6. URL : <http://www.springer.com/br/book/9781441922236> (visité le 02/03/2018).
- [BG11] Amandine BELLENGER et Sylvain GATEPAILLE. “Uncertainty in ontologies : Dempster-shafer theory for data fusion applications”. In : *arXiv preprint arXiv :1106.3876* (2011). URL : <https://arxiv.org/abs/1106.3876> (visité le 07/06/2017).
- [Blo08] Géraldine BLOY. “L’incertitude en médecine générale : sources, formes et accommodements possibles”. In : *Sciences sociales et santé* 26.1 (2008), p. 67. ISSN : 0294-0337, 1777-5914. DOI : [10.3917/sss.261.0067](https://doi.org/10.3917/sss.261.0067). URL : <http://www.cairn.info/revue-sciences-sociales-et-sante-2008-1-page-67.htm> (visité le 05/06/2017).

- [BM12] Alexandros BATZIOS et Pericles A. MITKAS. “WebOWL : A Semantic Web search engine development experiment”. In : *Expert Systems with Applications* 39.5 (1^{er} avr. 2012), p. 5052–5060. ISSN : 0957-4174. DOI : [10.1016/j.eswa.2011.11.034](https://doi.org/10.1016/j.eswa.2011.11.034). URL : <http://www.sciencedirect.com/science/article/pii/S0957417411015776> (visité le 16/05/2018).
- [BM18] Luciano BARCELLOS DE PAULA et Fernando Augusto Silva MARINS. “Algorithms applied in decision-making for sustainable transport”. In : *Journal of Cleaner Production* 176 (1^{er} mar. 2018), p. 1133–1143. ISSN : 0959-6526. DOI : [10.1016/j.jclepro.2017.11.216](https://doi.org/10.1016/j.jclepro.2017.11.216). URL : <http://www.sciencedirect.com/science/article/pii/S0959652617328895> (visité le 14/02/2018).
- [BMB10] Jamal BENTAHAR, Bernard MOULIN et Micheline BÉLANGER. “A taxonomy of argumentation models used for knowledge representation”. In : *Artificial Intelligence Review* 33.3 (mar. 2010), p. 211–259. ISSN : 0269-2821, 1573-7462. DOI : [10.1007/s10462-010-9154-1](https://doi.org/10.1007/s10462-010-9154-1). URL : <http://link.springer.com/10.1007/s10462-010-9154-1> (visité le 02/03/2016).
- [BNJ03] David M. BLEI, Andrew Y. NG et Michael I. JORDAN. “Latent Dirichlet Allocation”. In : *Journal of Machine Learning Research* 3 (Jan 2003), p. 993–1022. ISSN : ISSN 1533-7928. URL : <http://jmlr.csail.mit.edu/papers/v3/blei03a.html> (visité le 29/04/2018).
- [Bou+15] Fatma Ezzahra BOUSNINA et al. “A New Formalism for Evidential Databases”. In : *Foundations of Intelligent Systems - 22nd International Symposium, ISMIS 2015, Lyon, France, October 21-23, 2015, Proceedings*. Sous la dir. de Floriana ESPOSITO et al. T. 9384. Lecture Notes in Computer Science. Springer, 2015, p. 31–40. ISBN : 978-3-319-25251-3. DOI : [10.1007/978-3-319-25252-0_4](https://doi.org/10.1007/978-3-319-25252-0_4).
- [Box+04] Aziz A BOXWALA et al. “GLIF3 : a representation format for sharable computer-interpretable clinical practice guidelines”. In : *Journal of Biomedical Informatics* 37.3 (juin 2004), p. 147–161. ISSN : 15320464. DOI : [10.1016/j.jbi.2004.04.002](https://doi.org/10.1016/j.jbi.2004.04.002). URL : <http://linkinghub.elsevier.com/retrieve/pii/S1532046404000334> (visité le 14/06/2016).
- [BR16] Josh H. BUKOWSKI et John R. RICHARDS. “Commercial Airline In-Flight Emergency : Medical Student Response and Review of Medicolegal Issues”. In : *The Journal of Emergency Medicine* 50.1 (jan. 2016), p. 74–78. ISSN : 07364679. DOI : [10.1016/j.jemermed.2015.09.026](https://doi.org/10.1016/j.jemermed.2015.09.026). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0736467915010215> (visité le 30/03/2016).
- [Bra90] Ronald J. BRACHMAN. “The Future of Knowledge Representation”. In : *Proceedings of the Eighth National Conference on Artificial Intelligence - Volume 2. AAAI’90*. Boston, Massachusetts : AAAI Press, 1990, p. 1082–1092. ISBN : 978-0-262-51057-8. URL : <http://dl.acm.org/citation.cfm?id=1865609.1865661> (visité le 24/02/2018).

- [BSS06] Amy BERLIN, Marco SORANI et Ida SIM. “A taxonomic description of computer-based clinical decision support systems”. In : *Journal of Biomedical Informatics* 39.6 (déc. 2006), p. 656–667. ISSN : 15320464. DOI : [10.1016/j.jbi.2005.12.003](https://doi.org/10.1016/j.jbi.2005.12.003). URL : <http://linkinghub.elsevier.com/retrieve/pii/S1532046405001383> (visité le 17/06/2016).
- [BT07] Eta S. BERNER et La Lande TONYA J. “Overview of Clinical Decision Support Systems”. In : *Clinical Decision Support Systems : Theory and Practice*. Health Informatics. 2007. ISBN : 978-1-4419-2223-6.
- [BTN10] Rokia BENDAOU, Yannick TOUSSAINT et Amedeo NAPOLI. “L’analyse Formelle de Concepts au service de la construction et l’enrichissement d’une ontologie”. In : *Revue des Nouvelles Technologies de l’Information* (2010), p. 133–164.
- [CAB04] Michel CRUCIANU, Jean-Pierre ASSELIN DE BEAUVILLE et Romuald BONÉ. *Méthodes factorielles pour l’analyse des données : méthodes linéaires et extensions non-linéaires*. OCLC : 254143804. Paris : Hermes science publ. [u.a.], 2004. 288 p. ISBN : 978-2-7462-0921-3.
- [CBD01] Mireille CLÉRET, Pierre Le BEUX et Franck Le DUFF. “Les systèmes d’aide à la décision médicale”. In : *Les Cahiers du numérique* 2.2 (2001), p. 125–154. ISSN : 1622-1494. URL : <https://www.cairn.info/revue-les-cahiers-du-numerique-2001-2-page-125.htm> (visité le 14/02/2018).
- [CC13] Amit CHANDRA et Shauna CONRY. “In-flight Medical Emergencies”. In : *Western Journal of Emergency Medicine* 14.5 (17 sept. 2013), p. 499–504. ISSN : 1936900X, 19369018. DOI : [10.5811/westjem.2013.4.16052](https://doi.org/10.5811/westjem.2013.4.16052). URL : <http://www.escholarship.org/uc/item/3706n3px> (visité le 17/02/2017).
- [Cha02] Jean CHARLET. “L’ingénierie des connaissances : développements, résultats et perspectives pour la gestion des connaissances médicales”. thesis. Université Pierre et Marie Curie - Paris VI, 10 déc. 2002. URL : <https://tel.archives-ouvertes.fr/tel-00006920/document> (visité le 03/01/2019).
- [Cha16] G. CHARLES-CADOGAN. “Expected utility theory and inner and outer measures of loss aversion”. In : *Journal of Mathematical Economics* 63 (1^{er} mar. 2016), p. 10–20. ISSN : 0304-4068. DOI : [10.1016/j.jmateco.2015.12.001](https://doi.org/10.1016/j.jmateco.2015.12.001). URL : <http://www.sciencedirect.com/science/article/pii/S0304406815001457>.
- [CNP12] CNPT. *Plan national de déploiement de la télémédecine : Comité National de Pilotage de la Télémédecine (CNPT)*. 2012. URL : http://solidarites-sante.gouv.fr/IMG/pdf/Recommandations_mise_en_oeuvre_projet_telemedecine.pdf.
- [Cos15] Marine COSTA. “In-flight medical emergencies : experience and challenges faced by private general practitioners of Reunion answering the call of the crew during a long-haul flight serving the island”. Thèse de doct. Juin 2015. 104 p. URL : <https://dumas.ccsd.cnrs.fr/dumas-01199919>.
- [Cru17] Michel CRUCIANU. *Réduction du volume de données*. 2017. URL : <http%20://cedric.cnam.fr/vertigo/Cours/RCP216/>.

- [CT14] SNOMED CT. *SNOMED CT - Starter Guide, Version : 2014-07-31*. 2014. URL : www.snomed.org/resource/resource/41.
- [DF91] Patrice DEGOULET et Marius FIESCHI. *Traitement de l'information médicale : méthodes et applications hospitalières*. Google-Books-ID : hbjAAAACAAJ. Masson, 1991. 295 p. ISBN : 978-2-225-82514-9.
- [DGA17] DGAC. *Direction générale de l'Aviation civile*. In : *Wikipédia*. Page Version ID : 139228053. 25 juil. 2017. URL : https://fr.wikipedia.org/w/index.php?title=Direction_g%C3%A9n%C3%A9rale_de_l%27Aviation_civile&oldid=139228053.
- [DGM14] Stefan DARMONI, Nicolas GRIFFON et Philippe MASSARI. “Les systèmes d’aide à la décision médicale”. In : *Annales des Mines - Réalités industrielles* Novembre 2014.4 (19 déc. 2014), p. 47–50. ISSN : 1148-7941. URL : <https://www.cairn.info/revue-realites-industrielles1-2014-4-page-47.htm> (visité le 07/01/2019).
- [Dir09] Driss DIRI. *Élaboration d'un système de maintien de vérité : une approche orientée objet*. 2009. URL : <http://theses.ulaval.ca/archimede/fichiers/26049/ch03.html> (visité le 01/03/2016).
- [Dju+14] Benjamin DJULBEGOVIC et al. “How do physicians decide to treat : an empirical evaluation of the threshold model”. In : *BMC medical informatics and decision making* 14.1 (2014), p. 47. URL : <https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-14-47> (visité le 18/04/2017).
- [Dou+15] Mamadou Bilo DOUMBOUYA et al. “A framework for decision making on teleexpertise with traceability of the reasoning”. In : *IRBM* 36.1 (2015), p. 40–51. URL : <http://www.sciencedirect.com/science/article/pii/S1959031814001146> (visité le 04/09/2017).
- [Dou16] Mamadou Bilo DOUMBOUYA. “Système d’argumentation pour la collaboration en télé-médecine”. Thèse. 8 déc. 2016. URL : <http://www.theses.fr/2016TOU10030> (visité le 26/03/2018).
- [Dua+14] Júlio DUARTE et al. “Improving Quality of Electronic Health Records with SNOMED”. In : *CENTERIS 2014 - Conference on ENTERprise Information Systems / ProjMAN 2014 - International Conference on Project MANagement / HCIST 2014 - International Conference on Health and Social Care Information Systems and Technologies* 16 (Supplement C 1^{er} jan. 2014), p. 1342–1350. ISSN : 2212-0173. DOI : [10.1016/j.protcy.2014.10.151](https://doi.org/10.1016/j.protcy.2014.10.151). URL : <http://www.sciencedirect.com/science/article/pii/S2212017314003788>.
- [Dub04] R. DUBOIS. “Application des nouvelles méthodes d’apprentissage à la détection précoce d’anomalies cardiaques en électrocardiographie”. Thèse de doct. Université Pierre et Marie Curie - Paris VI, 27 jan. 2004. URL : <https://pastel.archives-ouvertes.fr/pastel-00000571/document> (visité le 17/03/2018).
- [Dub11] Didier DUBOIS. “Uncertainty Theories : a Unied View”. In : (2011), p. 6.

- [Dut15] Palash DUTTA. “Uncertainty Modeling in Risk Assessment Based on Dempster–Shafer Theory of Evidence with Generalized Fuzzy Focal Elements”. In : *Fuzzy Information and Engineering* 7.1 (1^{er} mar. 2015), p. 15–30. ISSN : 1616-8658. DOI : [10.1016/j.fiae.2015.03.002](https://doi.org/10.1016/j.fiae.2015.03.002). URL : <http://www.sciencedirect.com/science/article/pii/S1616865815000229> (visité le 14/06/2017).
- [EE15] Shaker EL-SAPPAGH et Mohammed ELMOGY. “An encoding methodology for medical knowledge using SNOMED CT ontology”. In : *Journal of King Saud University - Computer and Information Sciences* (oct. 2015). ISSN : 13191578. DOI : [10.1016/j.jksuci.2015.10.002](https://doi.org/10.1016/j.jksuci.2015.10.002). URL : <http://linkinghub.elsevier.com/retrieve/pii/S1319157815000919> (visité le 01/04/2016).
- [ELS85] L. EECKHOUDT, T. LEBRUN et J. C. SAILLY. “Risk-aversion and physicians’ medical decision-making”. In : *Journal of Health Economics* 4.3 (1^{er} sept. 1985), p. 273–281. ISSN : 0167-6296. DOI : [10.1016/0167-6296\(85\)90033-5](https://doi.org/10.1016/0167-6296(85)90033-5). URL : <http://www.sciencedirect.com/science/article/pii/0167629685900335> (visité le 18/04/2017).
- [ESF14] Claudio ECCHER, Andreas SEYFANG et Antonella FERRO. “Implementation and evaluation of an Asbru-based decision support system for adjuvant treatment in breast cancer”. In : *Computer Methods and Programs in Biomedicine* 117.2 (nov. 2014), p. 308–321. ISSN : 01692607. DOI : [10.1016/j.cmpb.2014.06.021](https://doi.org/10.1016/j.cmpb.2014.06.021). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0169260714002557> (visité le 17/06/2016).
- [Eva05] Kevin W. EVA. “Ce que tout enseignant devrait savoir concernant le raisonnement clinique”. In : *Pédagogie Médicale* 6.4 (nov. 2005), p. 225–234. ISSN : 1625-6484, 1627-4784. DOI : [10.1051/pmed:2005027](https://doi.org/10.1051/pmed:2005027). URL : <http://www.pedagogie-medicale.org/10.1051/pmed:2005027> (visité le 09/04/2018).
- [Fay+96] Usama FAYYAD et al. “Knowledge Discovery and Data Mining : Towards a Unifying Framework”. In : AAAI Press, 1996, p. 82–88.
- [FD95] Ronen FELDMAN et Ido DAGAN. “Knowledge Discovery in Textual Databases (KDT).” In : *KDD*. T. 95. 1995, p. 112–117.
- [Fer08] Isabelle FERONI. “Avant-propos”. In : *Sciences sociales et santé* 26.1 (2008), p. 5–7. ISSN : 0294-0337. DOI : [10.3917/sss.261.0005](https://doi.org/10.3917/sss.261.0005). URL : <https://www.cairn.info/revue-sciences-sociales-et-sante-2008-1-page-5.htm> (visité le 16/04/2018).
- [FGS13] Sherman FOLLAND, Allen C. GOODMAN et Miron STANO. *The economics of health and health care*. 7th ed. Upper Saddle River, N.J : Pearson, 2013. 602 p. ISBN : 978-0-13-277369-0.
- [Fou17] FOURNIER-VIGER. *SPMF : A Java Open-Source Data Mining Library*. 2017. URL : <http://www.philippe-fournier-viger.com/spmf/index.php?link=algorithms.php> (visité le 25/07/2017).
- [Fox88] Renée C. FOX. *L’incertitude médicale*. Paris, France, Belgique : L’Harmattan, 1988. 202 p. ISBN : 978-2-87085-141-8.

- [FPS96] Usama FAYYAD, Gregory PIATETSKY-SHAPIO et Padhraic SMYTH. “From Data Mining to Knowledge Discovery in Databases”. In : *AI Magazine* 17.3 (15 mar. 1996), p. 37. ISSN : 0738-4602. DOI : [10.1609/aimag.v17i3.1230](https://doi.org/10.1609/aimag.v17i3.1230). URL : <https://www.aaai.org/ojs/index.php/aimagazine/article/view/1230> (visité le 02/03/2018).
- [Fra+15] Paolo FRACCARO et al. “Behind the screens : Clinical decision support methodologies – A review”. In : *Health Policy and Technology* 4.1 (mar. 2015), p. 29–38. ISSN : 22118837. DOI : [10.1016/j.hlpt.2014.10.001](https://doi.org/10.1016/j.hlpt.2014.10.001). URL : <http://linkinghub.elsevier.com/retrieve/pii/S2211883714000616> (visité le 17/02/2016).
- [FS05] Hélène FARGIER et Régis SABBADIN. “Qualitative decision under uncertainty : back to expected utility”. In : *Artificial Intelligence* 164.1 (mai 2005), p. 245–280. ISSN : 00043702. DOI : [10.1016/j.artint.2004.12.002](https://doi.org/10.1016/j.artint.2004.12.002). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0004370205000020> (visité le 20/04/2017).
- [Gal15] Alexandre GALOPIN. “Modélisation ontologique des recommandations de pratique clinique pour une aide à la décision à niveaux d’abstraction variables”. Thèse de doct. Université Pierre et Marie Curie-Paris VI, 2015. URL : <https://tel.archives-ouvertes.fr/tel-01226450/> (visité le 13/06/2016).
- [Geo06] Gersende GEORG. “Analyse Informatique de Guides de Bonnes Pratiques Cliniques”. Thèse de doct. Université Pierre et Marie Curie - Paris VI, 6 sept. 2006. URL : <https://tel.archives-ouvertes.fr/tel-00102511/document> (visité le 13/06/2016).
- [Gou13] Celine GOUNDER. “Medical Emergencies at 40,000 Feet”. In : *The Atlantic* (4 avr. 2013). ISSN : 1072-7825. URL : <https://www.theatlantic.com/health/archive/2013/04/medical-emergencies-at-40-000-feet/274623/>.
- [Gre14a] Robert A. GREENES. “Chapter 2 - A Brief History of Clinical Decision Support : Technical, Social, Cultural, Economic, and Governmental Perspectives”. In : *Clinical Decision Support (Second Edition)*. Oxford : Academic Press, 2014, p. 49–109. ISBN : 978-0-12-398476-0. URL : <http://www.sciencedirect.com/science/article/pii/B9780123984760000026> (visité le 17/02/2016).
- [Gre14b] “Section III. Sources of Knowledge for Clinical Decision Support”. In : *Clinical Decision Support (Second Edition)*. Sous la dir. de Robert A. GREENES. Oxford : Academic Press, 2014, p. 283–284. ISBN : 978-0-12-398476-0. DOI : [10.1016/B978-0-12-398476-0.00039-7](https://doi.org/10.1016/B978-0-12-398476-0.00039-7). URL : <https://www.sciencedirect.com/science/article/pii/B97801239847600000397> (visité le 28/02/2018).
- [Gru93] Thomas R. GRUBER. “A translation approach to portable ontology specifications”. In : *Knowledge Acquisition* 5.2 (1^{er} juin 1993), p. 199–220. ISSN : 1042-8143. DOI : [10.1006/knac.1993.1008](https://doi.org/10.1006/knac.1993.1008). URL : <http://www.sciencedirect.com/science/article/pii/S1042814383710083> (visité le 23/02/2018).
- [gui14] SNOMED CT User GUIDE. *SNOMED CT® Release Format 1 (RF1) Guide*. 2014. URL : http://doc.ihtsdo.org/download/doc_RF1Guide_Current-en-US_INT_20140731.pdf.

- [Gut+15] Juan B. GUTIERREZ et al. “From Within Host Dynamics to the Epidemiology of Infectious Disease : Scientific Overview and Challenges”. In : *Mathematical biosciences* 270.0 (déc. 2015), p. 143–155. ISSN : 0025-5564. DOI : [10.1016/j.mbs.2015.10.002](https://doi.org/10.1016/j.mbs.2015.10.002). URL : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4607538/> (visité le 05/03/2018).
- [GZ05] Jing GAO et Jun ZHANG. “Clustered SVD strategies in latent semantic indexing”. In : *Information Processing & Management* 41.5 (1^{er} sept. 2005), p. 1051–1063. ISSN : 0306-4573. DOI : [10.1016/j.ipm.2004.10.005](https://doi.org/10.1016/j.ipm.2004.10.005). URL : <http://www.sciencedirect.com/science/article/pii/S0306457304001384>.
- [Héo14] Michel HÉON. *Web sémantique et modélisation ontologique (avec G-OWL) guide du développeur Java sous Eclipse*. OCLC : 880347455. St Herblain : Éd. ENI, 2014. ISBN : 978-2-7460-8869-6.
- [HHH15] Michael HUERTA, F. A. A. HEADQUARTERS et Dear Administrator HUERTA. “Aerospace Medical Association”. In : (2015). URL : <https://www.asma.org/asma/media/AsMA/pdf-policy/2014/Third-Class-Medical-Letter-to-FAA-Administrator-4-9-2014.pdf> (visité le 17/08/2017).
- [Hil+17] Marij A. HILLEN et al. “Tolerance of uncertainty : Conceptual analysis, integrative model, and implications for healthcare”. In : *Social Science & Medicine* 180 (mai 2017), p. 62–75. ISSN : 02779536. DOI : [10.1016/j.socscimed.2017.03.024](https://doi.org/10.1016/j.socscimed.2017.03.024). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0277953617301703> (visité le 22/06/2017).
- [Hin+14] Jochen HINKELBEIN et al. “Emergency Medical Equipment On Board German Airlines”. In : *Journal of Travel Medicine* 21.5 (1^{er} sept. 2014), p. 318–323. ISSN : 1195-1982, 1708-8305. DOI : [10.1111/jtm.12138](https://doi.org/10.1111/jtm.12138). URL : <https://academic.oup.com/jtm/article-lookup/doi/10.1111/jtm.12138> (visité le 21/02/2017).
- [Hin+17] Jochen HINKELBEIN et al. “In-flight medical emergencies during airline operations : a survey of physicians on the incidence, nature, and available medical equipment”. In : *Open Access Emergency Medicine* Volume 9 (fév. 2017), p. 31–35. ISSN : 1179-1500. DOI : [10.2147/OAEM.S129250](https://doi.org/10.2147/OAEM.S129250). URL : <https://www.dovepress.com/in-flight-medical-emergencies-during-airline-operations-a-survey-of-ph-peer-reviewed-article-OAEM> (visité le 21/03/2017).
- [Hol16] Brett F. HOLLIS. “Medical emergencies on commercial airlines : An Analysis of Onboard Medical Incidents, Treatment versus Prevention”. Thèse de doct. Brandman University, 2016. URL : <http://search.proquest.com/openview/69598832138577246987f22ba91fa03c1?pq-origsite=gscholar&cbl=18750&diss=y> (visité le 26/07/2017).
- [Hor00] François HORN. “L’économie du logiciel”. thesis. Lille 1, 1^{er} jan. 2000. URL : <http://www.theses.fr/2000LIL12017> (visité le 03/01/2019).
- [Hua+16] Zhengxing HUANG et al. “Predictive monitoring of clinical pathways”. In : *Expert Systems with Applications* 56 (sept. 2016), p. 227–241. ISSN : 09574174. DOI : [10.1016/j.eswa.2016.02.052](https://doi.org/10.1016/j.eswa.2016.02.052). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0957417416300926> (visité le 27/05/2016).

- [Huf+14] Stanley M. HUFF et al. “Chapter 17 - Ontologies, Vocabularies and Data Models”. In : *Clinical Decision Support (Second Edition)*. Sous la dir. de Robert A. GREENES. Oxford : Academic Press, 2014, p. 465–498. ISBN : 978-0-12-398476-0. DOI : [10.1016/B978-0-12-398476-0.00017-8](https://doi.org/10.1016/B978-0-12-398476-0.00017-8). URL : <https://www.sciencedirect.com/science/article/pii/B9780123984760000178> (visité le 28/02/2018).
- [IAT] IATA. *IATA - Home*. URL : <http://www.iata.org/Pages/default.aspx> (visité le 17/08/2017).
- [IHT17] IHTSDO. *Decision Support with SNOMED CT - Decision Support with SNOMED CT - SNOMED Confluence*. 2017. URL : <https://confluence.ihtsdotools.org/display/DOCCDS> (visité le 03/07/2017).
- [Jen14] Robert A. JENDERS. “Chapter 15 - Decision Rules and Expressions”. In : *Clinical Decision Support (Second Edition)*. Sous la dir. de Robert A. GREENES. Oxford : Academic Press, 2014, p. 417–434. ISBN : 978-0-12-398476-0. DOI : [10.1016/B978-0-12-398476-0.00015-4](https://doi.org/10.1016/B978-0-12-398476-0.00015-4). URL : <https://www.sciencedirect.com/science/article/pii/B9780123984760000154> (visité le 28/02/2018).
- [Jez98] Marlene JEZERSKI. “MedAire : Peace of mind in the skies—A flight nurse’s dream come true”. In : *Journal of Emergency Nursing* 24.1 (fév. 1998), p. 71–73. ISSN : 00991767. DOI : [10.1016/S0099-1767\(98\)90178-7](https://doi.org/10.1016/S0099-1767(98)90178-7). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0099176798901787> (visité le 20/02/2017).
- [JLH02] Richard W. JONES, Andrew LOWE et Michael J. HARRISON. “A framework for intelligent medical diagnosis using the theory of evidence”. In : *Knowledge-Based Systems* 15.1 (jan. 2002), p. 77–84. ISSN : 09507051. DOI : [10.1016/S0950-7051\(01\)00123-X](https://doi.org/10.1016/S0950-7051(01)00123-X). URL : <http://linkinghub.elsevier.com/retrieve/pii/S095070510100123X> (visité le 09/06/2017).
- [Kad14] Farid KADRI. “Contribution à la conception d’un système d’aide à la décision pour la gestion de situations de tension au sein des systèmes hospitaliers. Application à un service d’urgence.” Thèse de doct. Université de Valenciennes et du Hainaut-Cambrésis, 27 nov. 2014. URL : <https://hal.archives-ouvertes.fr/tel-01119943/document> (visité le 23/03/2017).
- [Kin02] Walter KINTSCH. “The potential of latent semantic analysis for machine grading of clinical case summaries”. In : *Journal of Biomedical Informatics* 35.1 (1^{er} fév. 2002), p. 3–7. ISSN : 1532-0464. DOI : [10.1016/S1532-0464\(02\)00004-7](https://doi.org/10.1016/S1532-0464(02)00004-7). URL : <http://www.sciencedirect.com/science/article/pii/S1532046402000047>.
- [KLV90] Werner KARBACH, Marc LINSTER et Angi VOSS. “Models, methods, roles and tasks : Many labels—one idea?” In : *Knowledge Acquisition* 2.4 (1^{er} déc. 1990), p. 279–299. ISSN : 1042-8143. DOI : [10.1016/S1042-8143\(05\)80011-1](https://doi.org/10.1016/S1042-8143(05)80011-1). URL : <http://www.sciencedirect.com/science/article/pii/S1042814305800111> (visité le 22/02/2018).

- [Lam17] Jean-Baptiste LAMY. “Owlready : Ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies”. In : *Artificial Intelligence in Medicine* 80 (juil. 2017), p. 11–28. ISSN : 09333657. DOI : [10.1016/j.artmed.2017.07.002](https://doi.org/10.1016/j.artmed.2017.07.002). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0933365717300271> (visité le 16/05/2018).
- [Lan+14] Guo-Cheng LAN et al. “Applying the maximum utility measure in high utility sequential pattern mining”. In : *Expert Systems with Applications* 41.11 (sept. 2014), p. 5071–5081. ISSN : 09574174. DOI : [10.1016/j.eswa.2014.02.022](https://doi.org/10.1016/j.eswa.2014.02.022). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0957417414000852> (visité le 12/07/2017).
- [Lat13] Chiraz LATIRI. “Extraction de Connaissances a partir de Textes : M ethodes et Applications”. thesis. Université de Lorraine, 24 juin 2013. URL : <https://tel.archives-ouvertes.fr/tel-00927238/document> (visité le 29/03/2017).
- [Lau16] Henri LAUDE. *Data scientist et langage R : guide d'autoformation à l'exploitation des Big Data*. OCLC : 959661227. 2016. ISBN : 978-2-409-00043-0.
- [Lee+13] Dennis LEE et al. “A survey of SNOMED CT implementations”. In : *Journal of Biomedical Informatics* 46.1 (fév. 2013), p. 87–96. ISSN : 15320464. DOI : [10.1016/j.jbi.2012.09.006](https://doi.org/10.1016/j.jbi.2012.09.006). URL : <http://linkinghub.elsevier.com/retrieve/pii/S1532046412001530> (visité le 06/10/2017).
- [Lee92] Suk Kyoon LEE. “Imprecise and Uncertain Information in Databases : An Evidential Approach”. In : *Proceedings of the Eighth International Conference on Data Engineering*. Washington, DC, USA : IEEE Computer Society, 1992, p. 614–621. ISBN : 978-0-8186-2545-9. URL : <http://dl.acm.org/citation.cfm?id=645477.654492> (visité le 02/07/2017).
- [Leh92] Fritz LEHMANN. “Semantic networks”. In : *Computers & Mathematics with Applications* 23.2 (1^{er} jan. 1992), p. 1–50. ISSN : 0898-1221. DOI : [10.1016/0898-1221\(92\)90135-5](https://doi.org/10.1016/0898-1221(92)90135-5). URL : <http://www.sciencedirect.com/science/article/pii/S0898122192901355> (visité le 26/02/2018).
- [LHH12] Chiraz LATIRI, Hatem HADDAD et Tarek HAMROUNI. “Towards an effective automatic query expansion process using an association rule mining approach”. In : *Journal of Intelligent Information Systems* 39.1 (1^{er} août 2012), p. 209–247. ISSN : 0925-9902, 1573-7675. DOI : [10.1007/s10844-011-0189-9](https://doi.org/10.1007/s10844-011-0189-9). URL : <https://link.springer.com/article/10.1007/s10844-011-0189-9> (visité le 06/03/2018).
- [Li+18] Zhongyu LI et al. “Large-scale retrieval for medical image analytics : A comprehensive review”. In : *Medical Image Analysis* 43 (1^{er} jan. 2018), p. 66–84. ISSN : 1361-8415. DOI : [10.1016/j.media.2017.09.007](https://doi.org/10.1016/j.media.2017.09.007). URL : <http://www.sciencedirect.com/science/article/pii/S136184151730138X> (visité le 20/03/2018).
- [Lin+16] Jerry Chun-Wei LIN et al. “High utility-itemset mining and privacy-preserving utility mining”. In : *Perspectives in Science* 7 (mar. 2016), p. 74–80. ISSN : 22130209. DOI : [10.1016/j.pisc.2015.11.013](https://doi.org/10.1016/j.pisc.2015.11.013). URL : <http://linkinghub.elsevier.com/retrieve/pii/S2213020915000580> (visité le 13/07/2017).

- [Lor13] Mathieu LORENZO. “TOLERANCE A L’INCERTITUDE EN MEDECINE GENERALE Une démarche de théorisation ancrée à partir de 14 entretiens”. Thèse de doct. UNIVERSITE DE STRASBOURG, 2013. URL : [http://psychanalyse.com/pdf/TOLERANCE%20A%20L%20INCERTITUDE%20EN%20MEDECINE%20GENERALE%20-%20THESE%202013%20\(73%20pages%20-%201,1%20mo\).pdf](http://psychanalyse.com/pdf/TOLERANCE%20A%20L%20INCERTITUDE%20EN%20MEDECINE%20GENERALE%20-%20THESE%202013%20(73%20pages%20-%201,1%20mo).pdf) (visité le 02/06/2017).
- [LS07] Thomas LUKASIEWICZ et J?rg SCHELLHASE. “Variable-strength conditional preferences for ranking objects in ontologies”. In : *Web Semantics : Science, Services and Agents on the World Wide Web 5.3* (sept. 2007), p. 180–194. ISSN : 15708268. DOI : [10.1016/j.websem.2007.06.001](https://doi.org/10.1016/j.websem.2007.06.001). URL : <http://linkinghub.elsevier.com/retrieve/pii/S1570826807000261> (visité le 16/06/2017).
- [Lti11] Hela LTIFI. “Démarche centrée utilisateur pour la conception de SIAD basés sur un processus d’Extraction de Connaissances à partir de Données, Application à la lutte contre les infections nosocomiales”. Thèse de doct. Université de Valenciennes et du Hainaut-Cambresis ; Ecole Nationale d’Ingénieurs de Sfax, 2011. URL : <https://hal.archives-ouvertes.fr/tel-00630296/> (visité le 28/05/2017).
- [LVD15] Jean-Baptiste LAMY, Alain VENOT et Catherine DUCLOS. “PyMedTermino : an open-source generic API for advanced terminology services.” In : *MIE*. 2015, p. 924–928. URL : <http://person.hst.aau.dk/ska/MIE2015/Papers/SHTI210-0924.pdf> (visité le 26/09/2017).
- [Mah+11] Paul H. MAHONY et al. “Symptom-Based Categorization of In-Flight Passenger Medical Incidents”. In : *Aviation, Space, and Environmental Medicine 82.12* (1^{er} déc. 2011), p. 1131–1137. ISSN : 00956562. DOI : [10.3357/ASEM.3099.2011](https://doi.org/10.3357/ASEM.3099.2011). URL : <http://openurl.ingenta.com/content/xref?genre=article&issn=0095-6562&volume=82&issue=12&spage=1131> (visité le 17/02/2016).
- [Mar+14] Pierre MARQUIS et al. *Représentation des connaissances et formalisation des raisonnements*. OCLC : 881842667. Toulouse : Cepaduès, 2014. ISBN : 978-2-36493-041-4.
- [Mas05] Marie-Hélène MASSON. “Apports de la théorie des possibilités et des fonctions de croyance à l’analyse de données imprécises”. Thèse de doct. 2005.
- [Mer10] Tayeb MERABTI. “Méthodes pour la mise en relations des terminologies médicales : contribution à l’interopérabilité sémantique Inter et Intra terminologique”. Thèse de doct. 2010. URL : www.chu-rouen.fr/tibs/wp-content/uploads/th%C3%A8se_version_finale_TM.pdf.
- [Min74] Marvin MINSKY. *A Framework for Representing Knowledge*. Cambridge, MA, USA : Massachusetts Institute of Technology, 1974.
- [MZB06] Rabia MARGHOUBI, Karine ZEITOUNI et Azedine BOULMAKOUL. “Utilisation des treillis de Galois pour l’extraction et la visualisation des règles d’association spatiales”. In : Hammamet Yasmine -B.P. 35 , Hammamet, 8056, 2006, p. 703–718. ISBN : 2-906855-22-7.

- [N+03] Daniele NARDI, Ronald J. BRACHMAN et al. “An introduction to description logics.” In : *Description logic handbook 1* (2003), p. 40. URL : <http://vision.unipv.it/IA2/aa2004-2005/IntroductionToDescriptionLogics-01.pdf> (visité le 24/08/2017).
- [NA18] Matta NADA et Nathalie AUSSENAC-GILLES. *Le Schéma du modèle conceptuel, étape dans la modélisation des connaissances*. 2018. URL : https://www.researchgate.net/profile/Nathalie_Aussenac-Gilles/publication/266467037_Le_Schema_du_modele_conceptuel_etape_dans_la_modelisation_des_connaissances/links/543c066c0cf2d6698be362bb/Le-Schema-du-modele-conceptuel-etape-dans-la-modelisation-des-connaissances.pdf?origin=publication_detail.
- [Nao+16] Diane NAOURI et al. “Prevention of Medical Events During Air Travel : A Narrative Review”. In : *The American Journal of Medicine* 129.9 (sept. 2016), 1000.e1–1000.e6. ISSN : 00029343. DOI : 10.1016/j.amjmed.2016.05.013. URL : <http://linkinghub.elsevier.com/retrieve/pii/S0002934316305460> (visité le 19/09/2016).
- [Neu44] John Von NEUMANN AND MORGENSTERN. *Theory Of Games And Economic Behavior*. Avec la coll. d’OSMANIA UNIVERSITY et DIGITAL LIBRARY OF INDIA. Princeton University Press., 1944. 674 p. URL : <http://archive.org/details/theoryofgamesand030098mbp> (visité le 21/07/2017).
- [New82] Allen NEWELL. “The knowledge level”. In : *Artificial intelligence* 18.1 (1982), p. 87–127.
- [OAC] OACI. *Au sujet de l’OACI*. URL : https://www.icao.int/about-icao/Pages/FR/default_FR.aspx (visité le 16/08/2017).
- [Pai94] Pierre PAILLÉ. “L’analyse par théorisation ancrée”. In : *Cahiers de recherche sociologique* 23 (1994), p. 147. ISSN : 0831-1048, 1923-5771. DOI : 10.7202/1002253ar. URL : <http://id.erudit.org/iderudit/1002253ar> (visité le 24/04/2018).
- [Paq06] Philippe PAQUET. “De l’information à la connaissance”. In : (2006), p. 24. URL : <http://www.univ-orleans.fr/log/Doc-Rech/Textes-PDF/2006-1.pdf>.
- [Pel+03] Mor PELEG et al. “Comparing computer-interpretable guideline models : a case-study approach”. In : *Journal of the American Medical Informatics Association* 10.1 (2003), p. 52–68. URL : <http://jamia.oxfordjournals.org/content/10/1/52.short> (visité le 22/06/2016).
- [Pel13] Mor PELEG. “Computer-interpretable clinical guidelines : A methodological review”. In : *Journal of Biomedical Informatics* 46.4 (août 2013), p. 744–763. ISSN : 15320464. DOI : 10.1016/j.jbi.2013.06.009. URL : <http://linkinghub.elsevier.com/retrieve/pii/S1532046413000841> (visité le 13/06/2016).
- [Pen09] Frédéric PENNERATH. “Méthodes d’extraction de connaissances à partir de données modélisables par des graphes. Application à des problèmes de synthèse organique.” Thèse de doct. Université Henri Poincaré - Nancy I, 2 juil. 2009. URL : <https://tel.archives-ouvertes.fr/tel-00436568/document> (visité le 01/03/2018).

- [Pet+13] Drew C. PETERSON et al. “Outcomes of Medical Emergencies on Commercial Airline Flights”. In : *New England Journal of Medicine* 368.22 (30 mai 2013), p. 2075–2083. ISSN : 0028-4793, 1533-4406. DOI : [10.1056/NEJMoa1212052](https://doi.org/10.1056/NEJMoa1212052). URL : <http://www.nejm.org/doi/abs/10.1056/NEJMoa1212052> (visité le 31/03/2016).
- [PG13] Thérèse PSIUK et Monique GOUBY. *Plans de soins types et chemins cliniques : 20 situations cliniques prévalentes*. OCLC : 858205807. Issy-les-Moulineaux : Elsevier Masson, 2013. ISBN : 978-2-294-71656-0.
- [PG14] Mor PELEG et Arturo GONZÁLEZ-FERRER. “Chapter 16 - Guidelines and Workflow Models”. In : *Clinical Decision Support (Second Edition)*. Sous la dir. de Robert A. GREENES. Oxford : Academic Press, 2014, p. 435–464. ISBN : 978-0-12-398476-0. DOI : [10.1016/B978-0-12-398476-0.00016-6](https://doi.org/10.1016/B978-0-12-398476-0.00016-6). URL : <https://www.sciencedirect.com/science/article/pii/B9780123984760000166> (visité le 28/02/2018).
- [PK75] Stephen G. PAUKER et Jerome P. KASSIRER. “Therapeutic Decision Making : A Cost-Benefit Analysis”. In : *New England Journal of Medicine* 293.5 (31 juil. 1975), p. 229–234. ISSN : 0028-4793, 1533-4406. DOI : [10.1056/NEJM197507312930505](https://doi.org/10.1056/NEJM197507312930505). URL : <http://www.nejm.org/doi/abs/10.1056/NEJM197507312930505> (visité le 18/04/2017).
- [PP07] Heather D. PFEIFFER et Joseph J. PFEIFFER JR. “Representation Levels Within Knowledge Representation”. In : *International Conference on Conceptual Structures*. Springer, 2007, p. 484–487. URL : http://link.springer.com/chapter/10.1007/978-3-540-73681-3_40 (visité le 19/12/2016).
- [Qat+17] Zelal QATAWNEH et al. “Clinical decision support system for venous thromboembolism risk classification”. In : *Applied Computing and Informatics* (13 sept. 2017). ISSN : 2210-8327. DOI : [10.1016/j.aci.2017.09.003](https://doi.org/10.1016/j.aci.2017.09.003). URL : <http://www.sciencedirect.com/science/article/pii/S2210832717300819> (visité le 14/02/2018).
- [Rak13] Ricco RAKOTOMALALA. *Topic model et réduction de dimensions*. 2013. URL : <http://eric.univ-lyon2.fr/~ricco/cours/slides/TM.D%20-%20reduction%20de%20dimension.pdf>.
- [Ram07] Emmanuel RAMASSO. “Reconnaissance de séquences d’états par le Modèle des Croyances Transférables. Application à l’analyse de vidéos d’athlétisme.” Thèse de doct. Université Joseph-Fourier - Grenoble I, 5 déc. 2007. URL : <https://tel.archives-ouvertes.fr/tel-00260770/document> (visité le 20/04/2018).
- [RB93] Bernard ROY et Denis BOUYSSOU. *Aide multicritère à la décision : méthodes et cas*. London School of Economics et Political Science, 1993. ISBN : 978-2-7178-2473-5. URL : <https://basepub.dauphine.fr/handle/123456789/4522> (visité le 13/02/2018).
- [RB98] Martin RAJMAN et Romaric BESANÇON. “Text mining : natural language techniques and text mining applications”. In : *Data mining and reverse engineering*. Springer, 1998, p. 50–64.

- [RCY07] Loïs RIGOUSTE, Olivier CAPPÉ et François YVON. “Inference and Evaluation of the Multinomial Mixture Model for Text Clustering”. In : *Information Processing & Management* 43.5 (sept. 2007), p. 1260–1280. ISSN : 03064573. DOI : [10.1016/j.ipm.2006.11.001](https://doi.org/10.1016/j.ipm.2006.11.001). arXiv : [cs/0606069](https://arxiv.org/abs/cs/0606069). URL : <http://arxiv.org/abs/cs/0606069> (visité le 03/01/2019).
- [Reb17] Laurence REBOUL. *Introduction à l'Analyse Des Données*. 2017. URL : <http://iml.univ-mrs.fr/~reboul/SP-ADD1.pdf>.
- [RF11] Patrick RODRIGUEZ et Vincent FEUILLIE. *La télémédecine en aéronautique*. 2011. URL : https://www.soframas.asso.fr/images/soframas/presentations/Toulon_17nov11_Rodriguez.pdf.
- [Ris17] RISKMAP. *Travel Risk Map 2017*. RiskMap 2017. 2017. URL : <http://riskmap.controlrisks.com/portfolio-view/travel-risk-map-2017/> (visité le 27/07/2017).
- [RO17] Mehdi RAHMANPOUR et Morteza OSANLOO. “A decision support system for determination of a sustainable pit limit”. In : *Journal of Cleaner Production* 141 (10 jan. 2017), p. 1249–1258. ISSN : 0959-6526. DOI : [10.1016/j.jclepro.2016.09.205](https://doi.org/10.1016/j.jclepro.2016.09.205). URL : <http://www.sciencedirect.com/science/article/pii/S0959652616315530> (visité le 14/02/2018).
- [Rod+12] Alejandro RODRÍGUEZ-GONZÁLEZ et al. “Using agents to parallelize a medical reasoning system based on ontologies and description logics as an application case”. In : *Expert Systems with Applications* 39.18 (déc. 2012), p. 13085–13092. ISSN : 09574174. DOI : [10.1016/j.eswa.2012.05.093](https://doi.org/10.1016/j.eswa.2012.05.093). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0957417412008093> (visité le 19/09/2016).
- [Ros15] Arnaud ROSIER. “Raisonnement automatique basé ontologies appliqué à la hiérarchisation des alertes en télécardiologie”. Thèse de doct. Rennes 1, 2015.
- [RP97] Craig A. ROSENBERG et Firooz PAK. “Emergencies in the air : Problems, management, and prevention”. In : *The Journal of Emergency Medicine* 15.2 (1^{er} mar. 1997), p. 159–164. ISSN : 0736-4679. DOI : [10.1016/S0736-4679\(96\)00314-9](https://doi.org/10.1016/S0736-4679(96)00314-9). URL : <http://www.sciencedirect.com/science/article/pii/S0736467996003149>.
- [Sam+10] Tara SAMPALLI et al. “An evaluation of SNOMED CT® in the domain of complex chronic conditions”. In : *International Journal of Integrated Care* 10 (2010), e038. ISSN : 1568-4156. JSTOR : {PMC}2858517. URL : <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2858517/>.
- [Sam14] Ahmed SAMET. “Théorie des fonctions de croyance : application des outils de data mining pour le traitement des données imparfaites”. Thèse de doct. Artois, 2014. URL : <http://www.theses.fr/2014ARTO0206> (visité le 13/06/2017).
- [San+09] Michael SAND et al. “Surgical and medical emergencies on board European aircraft : a retrospective study of 10189 cases”. In : *Critical Care* 13.1 (2009), R3. URL : <https://ccforum.biomedcentral.com/articles/10.1186/cc7690> (visité le 05/07/2017).

- [San16] Karim SANDID. “Déterminants des réactions face à l’incertitude des internes de médecine générale en France : l’étude DERFI-IMG”. Thèse d’exercice. France : Université Paris Diderot - Paris 7. UFR de médecine, 2016. 38 p. URL : <http://www.theseimg.fr/1/sites/default/files/Th%C3%A8se%20-%20%C3%A9terminants%20des%20r%C3%A9actions%20fac%20%C3%A0%20l'incertitude%20des%20Internes%20de%20M%C3%A9decine%20G%C3%A9n%C3%A9rale%20en%20France%20-%20l'%C3%A9tude%20DERFI-IMG.pdf>.
- [Sav54] Leonard J. SAVAGE. *The foundations of statistics*. 2d rev. ed. New York : Dover Publications, 1954. 310 p. ISBN : 978-0-486-62349-8.
- [SB14] B. SÉROUSSI et J. BOUAUD. “Systèmes informatiques d’aide à la décision en médecine : panorama des approches utilisant les données et les connaissances”. In : *Pratique Neurologique - FMC* 5.4 (déc. 2014), p. 303–316. ISSN : 18787762. DOI : [10.1016/j.praneu.2014.09.006](https://doi.org/10.1016/j.praneu.2014.09.006). URL : <http://linkinghub.elsevier.com/retrieve/pii/S1878776214001125> (visité le 03/01/2018).
- [Sch00] Guus SCHREIBER, éd. *Knowledge engineering and management : the CommonKADS methodology*. Cambridge, Mass : MIT Press, 2000. 455 p. ISBN : 978-0-262-19300-9.
- [Sch10] Christian SCHMIDT. *Neuroéconomie*. Google-Books-ID : z6sHqfWwRaUC. Odile Jacob, 25 mar. 2010. 324 p. ISBN : 978-2-7381-2444-9.
- [Sha76] Glenn SHAFER. *A mathematical theory of evidence*. OCLC : 1859710. Princeton, NJ : Princeton Univ. Press, 1976. 297 p. ISBN : 978-0-691-10042-5 978-0-691-08175-5.
- [Sim77] Herbert Alexander SIMON. *The New Science of Management Decision*. Upper Saddle River, NJ, USA : Prentice Hall PTR, 1977. ISBN : 978-0-13-616136-3.
- [SK94] Philippe SMETS et Robert KENNES. “The transferable belief model”. In : *Artificial Intelligence* 66.2 (1^{er} avr. 1994), p. 191–234. ISSN : 0004-3702. DOI : [10.1016/0004-3702\(94\)90026-4](https://doi.org/10.1016/0004-3702(94)90026-4). URL : <http://www.sciencedirect.com/science/article/pii/0004370294900264>.
- [SKR15] A. SENE, B. KAMSU-FOGUEM et P. RUMEAU. “Telemedicine framework using case-based reasoning with evidences”. In : *Computer Methods and Programs in Biomedicine* 121.1 (août 2015), p. 21–35. ISSN : 01692607. DOI : [10.1016/j.cmpb.2015.04.012](https://doi.org/10.1016/j.cmpb.2015.04.012). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0169260715001108> (visité le 17/02/2016).
- [SKR18a] A. SENE, B. KAMSU-FOGUEM et P. RUMEAU. “Data mining for decision support with uncertainty on the airplane”. In : *Data & Knowledge Engineering* (juil. 2018). ISSN : 0169023X. DOI : [10.1016/j.datak.2018.06.002](https://doi.org/10.1016/j.datak.2018.06.002). URL : <https://linkinghub.elsevier.com/retrieve/pii/S0169023X1730071X> (visité le 18/08/2018).
- [SKR18b] Alsane SENE, Bernard KAMSU-FOGUEM et Pierre RUMEAU. “Decision support system for in-flight emergency events”. In : *Cognition, Technology & Work* (10 fév. 2018). ISSN : 1435-5558, 1435-5566. DOI : [10.1007/s10111-018-0466-2](https://doi.org/10.1007/s10111-018-0466-2). URL : <http://link.springer.com/10.1007/s10111-018-0466-2> (visité le 13/02/2018).

- [SKR18c] Alsane SENE, Bernard KAMUSU-FOGUEM et Pierre RUMEAU. “Discovering frequent patterns for in-flight incidents”. In : *Cognitive Systems Research* 49 (juin 2018), p. 97–113. ISSN : 13890417. DOI : [10.1016/j.cogsys.2018.01.002](https://doi.org/10.1016/j.cogsys.2018.01.002). URL : <http://linkinghub.elsevier.com/retrieve/pii/S1389041717301997> (visité le 13/02/2018).
- [Sme90] P. SMETS. “The combination of evidence in the transferable belief model”. In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12.5 (mai 1990), p. 447–458. ISSN : 0162-8828. DOI : [10.1109/34.55104](https://doi.org/10.1109/34.55104).
- [Sme93] Philippe SMETS. “Belief functions : The disjunctive rule of combination and the generalized Bayesian theorem”. In : *International Journal of Approximate Reasoning* 9.1 (1^{er} août 1993), p. 1–35. ISSN : 0888-613X. DOI : [10.1016/0888-613X\(93\)90005-X](https://doi.org/10.1016/0888-613X(93)90005-X). URL : <http://www.sciencedirect.com/science/article/pii/0888613X9390005X> (visité le 20/04/2018).
- [SQL] SQLITE. *SQLite Home Page*. URL : <https://www.sqlite.org/> (visité le 26/09/2017).
- [Syb16] Christine SYBORD. “Big Data et conception d’un système d’information d’aide à la décision clinique : vers une gestion sociocognitive de la responsabilité médicale ?” In : *Les Cahiers du numérique*. Enjeux du big data et identification des données médicales 12.1 (mai 2016), p. 73–108. URL : <https://hal.archives-ouvertes.fr/hal-01560622> (visité le 14/02/2018).
- [Tan+16] Ludovic TANGUY et al. “Natural language processing for aviation safety reports : from classification to interactive analysis”. In : *Computers in Industry* 78 (2016), p. 80–95.
- [Tél10] TÉLÉMÉDECINE. *Décret n 2010-1229 du 19 octobre 2010 relatif à la télémédecine*. 19 oct. 2010. (Visité le 17/02/2016).
- [TY11] Mohamed Anis Bach TOBJI et Boutheina Ben YAGHLANE. “Extraction des itemsets fréquents à partir de données évidentielles : application à une base de données éducationnelles”. In : *Fouille de données complexes. Complexité liée aux données multiples*. Sous la dir. de Guillaume CLEUZIQU et al. T. E-21. RNTI. Hermann-Éditions, 2011, p. 211–232. URL : <http://editions-rnti.fr/?inprocid=1001247> (visité le 17/04/2018).
- [TYM08] MA Bach TOBJI, Boutheina Ben YAGHLANE et Khaled MELLOULI. “A new algorithm for mining frequent itemsets from evidential databases”. In : *Proceedings of IPMU*. T. 8. 2008, p. 1535–1542. URL : <http://www.gimac.uma.es/ipmu08/proceedings/papers/204-Bach.pdf> (visité le 30/06/2017).
- [Und07] David G. UNDERHILL. *Exploring dimensionality reduction for text mining*. NAVAL ACADEMY ANNAPOLIS MD, 2007.
- [Wil89] R. WILLE. “Knowledge Acquisition by Methods of Formal Concept Analysis”. In : *Proceedings of the Conference on Data Analysis, Learning Symbolic and Numeric Knowledge*. Commack, NY, USA : Nova Science Publishers, Inc., 1989, p. 365–380. ISBN : 0-941743-64-0. URL : <http://dl.acm.org/citation.cfm?id=102513.102547>.

- [YBB12] Sheng YU, Damon BERRY et Jesus BISBAL. “Clinical coverage of an archetype repository over SNOMED-CT”. In : *Journal of Biomedical Informatics* 45.3 (1^{er} juin 2012), p. 408–418. ISSN : 1532-0464. DOI : [10.1016/j.jbi.2011.12.001](https://doi.org/10.1016/j.jbi.2011.12.001). URL : <http://www.sciencedirect.com/science/article/pii/S1532046411002152>.
- [Ye+09] Yan YE et al. “An ontology-based hierarchical semantic modeling approach to clinical pathway workflows”. In : *Computers in Biology and Medicine* 39.8 (août 2009), p. 722–732. ISSN : 00104825. DOI : [10.1016/j.combiomed.2009.05.005](https://doi.org/10.1016/j.combiomed.2009.05.005). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0010482509000985> (visité le 06/06/2016).
- [YPP15] Xiao-Chen YUAN, Chi-Man PUN et C. L. PHILIP CHEN. “Robust Mel-Frequency Cepstral coefficients feature detection and dual-tree complex wavelet transform for digital audio watermarking”. In : *Information Sciences* 298 (20 mar. 2015), p. 159–179. ISSN : 0020-0255. DOI : [10.1016/j.ins.2014.11.040](https://doi.org/10.1016/j.ins.2014.11.040). URL : <http://www.sciencedirect.com/science/article/pii/S002002551401130X> (visité le 20/03/2018).
- [Yvo10] François YVON. “Une petite introduction au traitement automatique des langues naturelles”. In : *Conference on Knowledge discovery and data mining*. 2010, p. 27–36.
- [Zam05] Virginie ZAMPA. “Utilisation de l’analyse sémantique latente pour tenter d’optimiser l’acquisition par exposition à une langue étrangère de spécialité”. In : *Alsic* (Vol. 8, n 2 2 déc. 2005). ISSN : 1286-4986. DOI : [10.4000/alsic.339](https://doi.org/10.4000/alsic.339). URL : <http://journals.openedition.org/alsic/339> (visité le 14/03/2018).
- [Zha+16] Yi-Fan ZHANG et al. “Integrating HL7 RIM and ontology for unified knowledge and data representation in clinical decision support systems”. In : *Computer Methods and Programs in Biomedicine* 123 (jan. 2016), p. 94–108. ISSN : 01692607. DOI : [10.1016/j.cmpb.2015.09.020](https://doi.org/10.1016/j.cmpb.2015.09.020). URL : <http://linkinghub.elsevier.com/retrieve/pii/S0169260715002473> (visité le 14/06/2016).
- [Zha+17] Binbin ZHANG et al. “Mining of high utility-probability sequential patterns from uncertain databases”. In : *PLoS One* 12.7 (2017), e0180931. URL : <http://search.proquest.com/openview/d549836dfb4302407a539df909911cd6/1?pq-origsite=gscholar&cbl=1436336> (visité le 07/08/2017).
- [ZLW12] N. ZHONG, Y. LI et S. T. WU. “Effective Pattern Discovery for Text Mining”. In : *IEEE Transactions on Knowledge and Data Engineering* 24.1 (jan. 2012), p. 30–44. ISSN : 1041-4347. DOI : [10.1109/TKDE.2010.211](https://doi.org/10.1109/TKDE.2010.211).