



Proposition d'un outil d'aide à la décision adapté aux contraintes et aux enjeux d'un soutien informatique aux consultations médicales coutumières

Antoine Richard

► To cite this version:

Antoine Richard. Proposition d'un outil d'aide à la décision adapté aux contraintes et aux enjeux d'un soutien informatique aux consultations médicales coutumières. Recherche opérationnelle [math.OC]. Université Paris sciences et lettres, 2021. Français. ⟨NNT : 2021UPSLD031⟩. ⟨tel-03662496⟩

HAL Id: tel-03662496

<https://theses.hal.science/tel-03662496v1>

Submitted on 9 May 2022

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.



HAL Authorization



THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à Université Paris-Dauphine

**Proposition d'un outil d'aide à la décision adapté aux
contraintes et aux enjeux d'un soutien informatique aux
consultations médicales coutumières**

Soutenue par

Antoine RICHARD

Le 06.04.2021

École doctorale n°543

École Doctorale SDOSE

Spécialité

Informatique

Composition du jury :

Daniela GRIGORI

Professeur
Université Paris Dauphine - PSL

Présidente

Sylvie DESPRES

Professeur
Université Sorbonne Paris 13

Rapportrice

Pascale ZARATÉ

Professeur
Université de Toulouse 1 Capitole

Rapportrice

Pavlos DELIAS

Associate Professor
International Hellenic University

Examineur

Delphine MAUCORT-BOULCH

Professeur
Université Claude Bernard Lyon 1

Examinatrice

Brice MAYAG

Maître de Conférence
Université Paris Dauphine - PSL

Encadrant de thèse

Yves MEINARD

Chargé de Recherche
Université Paris Dauphine - PSL

Directeur de thèse

Alexis TSOUKIÀS

Directeur de Recherche
Université Paris Dauphine - PSL

Directeur de thèse

Avant-propos

Tout travail scientifique est réalisé par des êtres humains, avec leur subjectivité, leurs valeurs et leurs positionnements politiques. La présente thèse ne fait pas exception. Malgré tous les efforts que j'ai pu déployer pour faire en sorte que mes choix thématiques et méthodologiques ainsi que mes interprétations ne soient pas biaisés par cette subjectivité, la rigueur scientifique m'impose de faire preuve de transparence vis-à-vis de cet irréductible ancrage dans mon individualité, qui peut être à l'origine d'erreurs, de biais ou d'éléments arbitraires dans mes raisonnements, sans que j'en aie moi-même conscience. Afin d'assurer cette parfaite transparence, il me semble donc important, en amont de cette thèse, de fournir aux lecteur·rices quelques informations clefs pour juger au mieux la qualité de mon travail et de ma rigueur scientifique. Tout d'abord, je peux préciser que je suis un homme, hétérosexuel, cisgenre, blanc, athée avec une éducation judéo-chrétienne et issu d'une classe sociale "moyenne inférieure". Je peux ensuite ajouter que mes convictions politiques, bien que non partisans, sont en partie anarchistes, municipalistes libertaires et communistes libertaires, auxquelles viennent s'ajouter des convictions écologistes, alter-mondialistes, anti-spécistes, anti-sexistes et anti-racistes. À partir de ces informations, les lecteur·ices ont à leur disposition les éléments pour juger au mieux si les travaux et les conclusions présentées dans cette thèse ont souffert ou non de biais ou d'erreurs d'interprétations liées à ma subjectivité.

Du fait que les médecins avec lesquels j'ai travaillé pouvaient aussi bien être des femmes que des hommes, j'ai décidé d'employer certaines terminologies neutres issues de l'écriture inclusive. Ainsi, le point médian '.' est utilisé lorsqu'il est fait mention d'individus pouvant aussi bien être des femmes que des hommes, si bien qu'aucune distinction de genre n'est justifiée. Par exemple, nous employons les termes "*patient·e*", au lieu de "*patient ou patiente*", ou bien "*utilisateur·ices*", au lieu de "*utilisateurs et utilisatrices*". De même, les pronoms "*iel*" et "*iels*" sont utilisés à la place de "*elle ou il*" ou simplement "*ils*". Pour les déterminants dont la prononciation change en fonction du genre associé, tels que "*un ou une*", "*la ou le*" et "*celles et ceux*", "*du ou de la*" et "*sa ou son*", ceux-ci ont été remplacé par des formes contractées lorsqu'aucune distinction de genre n'est justifiée, tels que respectivement : "*un·e*", "*la·e*", "*celleux*", "*du/de la*" et "*sa/son*". Nous savons que ceci est un choix fort pouvant rendre la lecture de ce manuscrit difficile, en particulier pour les personnes atteintes

de troubles de la lecture comme la dyslexie. J'ai donc, autant que possible, limité l'usage de ces termes. Dans cette optique de lisibilité, les objets gardent le genre qui leur est généralement associé, par exemple "*un logiciel*" ou "*une application*", et les points médians ne sont pas appliqués aux accords de genre, par exemple "*un·e patient·e diagnostiqué*".

Il me paraît aussi important d'informer que la fin de cette thèse s'est déroulée durant l'épidémie du SARS-CoV-2, qui est toujours en cours au moment où j'écris ces lignes. Cela a, malheureusement, impacté mon travail. En particulier, les observations et les essais cliniques de l'outil informatique développé au cours de cette thèse ont été limités. Fort heureusement, les médecins avec lesquels j'ai travaillé ont su m'accorder de leur temps et ont grandement facilité mon travail dans un contexte sanitaire compliqué.

Enfin, je tiens à préciser que les citations utilisées pour habiller les différents chapitres de ce manuscrit sont issues d'auteur·ices dont j'ai apprécié la lecture durant cette thèse et dont les passages choisis m'avaient l'air plus ou moins liés au contenu de ces chapitres. Les lecteur·ices sont donc invités à ne pas surinterpréter les liens entre le contenu de ce manuscrit et les citations choisies pour l'habiller.

Remerciements

Tout d’abord, je tiens à remercier mes encadrants : Yves, Brice, François et Alexis. La thèse que je présente ici ne serait pas ce qu’elle est sans leurs conseils et leur exigence. Merci aussi à Sylvie Despres, Pascale Zaraté, Pavlos Delias, Daniela Grigori et Delphine Maucort-Boulch pour avoir accepté de faire partie de mon jury de thèse malgré des conditions de soutenance compliquées.

Merci aux membres du Lamsade pour leur accueil à chaque fois que je montais sur Paris. Un merci tout particulier à Georges et Antoine pour les nombreuses et passionnantes discussions que l’on a pu avoir. Merci aussi à tous les développeurs avec qui j’ai eu la chance de travailler (et d’essayer de nombreux jeux de sociétés) au sein de la DSII Gières des Hospices Civils de Lyon : Rémi, Khalil, Pajolma, Solenn, Mike, Samira, Eric, Florian, Aurélien, David, Nicolas et Laurent. Mes excuses à tous ceux que j’oublie, ça reste une grande famille qui a beaucoup évolué depuis mon arrivé.

Je souhaite ensuite remercier les médecins qui ont rendu possible le développement de l’outil d’aide à la décision présenté dans cette thèse : Professeur Raverot, Docteur Abeillon, Docteur Bouzehouane, Docteur Brac de la Perriere, Docteur Charriere, Docteur Lasolle, Docteur Moret, Docteur Renault et Docteur Villar-Fimbel. Avec un merci tout particulier au Professeur Philippe Moulin et au Docteur Christophe Riou qui ont suivi mon travail depuis le tout début.

Merci, bien sûr, à ma famille qui m’a toujours soutenu dans mes choix d’études et de carrière. Je tiens à remercier tout particulièrement ma tante Maryline et mon oncle Robert pour leur relecture de ce manuscrit.

Enfin, merci à mes amis pour me supporter au quotidien, dans tous les sens du terme. Merci à François, Raph, Kevin, Guin, Alex, Antoine Quentin, Jérem, Michael, Millian, Taqi, et bien sûr merci à Léo, Chtimy, Lucas et Dadou, mes compagnons de galère, pour les nombreuses discussions qui auront su rythmer nos thèses respectives. Merci aussi à celles et ceux que je ne cite pas ici, mais qui savent qu’ils et elles comptent tout autant. Je ne serais sûrement pas là, entrain d’écrire ces lignes, sans vous tous. Donc merci.

À Hervé,

À Philippe-Olivier,
mon cousin

They both became one with the force now.

Table des matières

Table des figures	xii
Liste des tableaux	xii
Liste des acronymes	xiii
1 Introduction	1
1.1 Contexte	2
1.2 Problématiques	3
1.3 Objectifs	5
I Comprendre les contraintes et les enjeux d'une aide à la décision en milieu médical	9
2 Revue des outils d'aide à la décision médicaux passés et actuels	11
2.1 Historique des outils d'aide aux diagnostics	14
2.1.1 1950-2000 : Les premiers essais de DDSS	14
2.1.2 2000-2010 : Une perte de confiance dans le savoir-faire des médecins	15
2.1.3 2010-2020 : L'essor de l'apprentissage automatique	19
2.2 Des preuves que les HIS, CDSS et DDSS sont bénéfiques	21
2.3 Une acceptabilité discutable	25
3 Étude des raisons de la non-acceptation des outils d'aide aux diagnostics	29
3.1 Un impact négatif des HIS, CDSS et DDSS	30
3.2 Des problèmes de responsabilité	33
3.3 Un écart entre conception et réalité pour l'aide en consultation	35
3.3.1 Les DDSS basés sur des "guidelines"	36
3.3.2 Les DDSS basés sur du "Machine Learning"	37

4 Proposition d'une nouvelle approche pour l'aide à la décision en consultations	41
4.1 Définition de la notion "d'information" dans un contexte médical . . .	42
4.2 Identification d'une approche adaptée aux contraintes du processus d'aide à la décision	44
4.3 Les promesses et les limites de l'approche "ajustive"	46
 II Comprendre le processus de décision et les besoins des médecins en consultations médicales	 49
5 Présentation du système informatique des HCL	51
5.1 Historique du système d'information hospitalier	52
5.2 Un ensemble d'outils pour le personnel soignant	54
5.3 Les fonctionnalités dédiées aux consultations médicales	57
 6 Observations et analyses du processus de travail des médecins en consultations médicales	 63
6.1 Observations de consultations	64
6.1.1 Protocole d'observation	64
6.1.2 Résultats	66
6.2 Analyses informatiques	68
6.2.1 Analyse générale	69
6.2.2 Heuristic Miner	73
6.2.3 Fuzzy Miner	75
6.3 Synthèse	78
 7 Modélisation du processus de décision des médecins en consultations médicales	 81
7.1 Formalisation du processus d'une consultation spécifique	82
7.2 Modélisation générique du processus de décision en consultation . .	85
7.3 Un problème de classification multi-labels	88
 III Proposer un outil d'aide à la décision dédié au soutien des consultations médicales coutumières	 91
 8 Définition de nos exigences en termes de "transparence" pour une meilleure acceptabilité	 93
8.1 Recherche d'une définition générale de la "transparence"	96

8.2	Définition de critères opérationnels pour évaluer le respect de nos exigences	99
8.2.1	L'exigence de compréhensibilité	100
8.2.2	L'exigence d'interprétabilité	101
8.2.3	L'exigence de retraçabilité	105
8.2.4	L'exigence de révisabilité	106
8.3	Évaluation et méthode d'agrégation	107
9	Sélection d'un système de classification répondant à nos exigences de "transparence"	111
9.1	Évaluation de la "transparence" de divers systèmes de classification .	112
9.1.1	Classification par réseau de neurones	112
9.1.2	Classification par plus proches voisins	115
9.1.3	Classification probabiliste	116
9.1.4	Classification par arbre de décision	118
9.1.5	Classification par système de règles	119
9.1.6	Classification par vecteurs de support	120
9.1.7	Résumé des résultats	122
9.2	Adaptation de Naive Bayes aux problèmes multi-labels	123
9.2.1	Discrétisation des variables numériques	124
9.2.2	Comptage des nombres d'occurrences des labels	127
9.2.3	Estimation des probabilités d'occurrences des labels	129
9.2.4	Anticipation des labels pour une nouvelle entrée	130
9.3	Estimation du coût potentiel de la "transparence" en termes de performance	132
9.3.1	Taux d'erreur : perte de Hamming	133
9.3.2	Précisions micro et macro	135
9.3.3	Sensibilité, ou rappel, micro et macro	137
9.3.4	Spécificité micro et macro	139
10	Conception d'un assistant virtuel pour les consultations médicales	143
10.1	Présentation de l'outil	144
10.1.1	Phase d'anticipation	145
10.1.2	Phase de recherche de données	147
10.1.3	Phase de présentation des données	148
10.2	Essais cliniques	151
10.2.1	Protocole	151
10.2.2	Résultats	154
10.3	Limites et perspectives	158

11 Conclusion	161
11.1 Contributions	161
11.1.1 Analyse critique des outils d'aide à la décision en milieux hospitaliers	162
11.1.2 Modélisation du processus de décision en consultation médicale	163
11.1.3 Proposition de critères opérationnels pour l'évaluation de la "transparence" de systèmes de classification	164
11.1.4 Développement d'un outil d'aide à la décision pour les consultations médicales	165
11.2 Limites	166
11.3 Perspectives	167
11.3.1 Adapter l'aide proposée à d'autres services hospitaliers	168
11.3.2 Repenser la place des outils informatiques en milieux hospitaliers	169
11.3.3 Investiguer l'intérêt d'une approche "ajustive" dans d'autres domaines	171
A Fiche d'observation	173
B Questionnaire d'appréciation	175
Bibliographie	177

Table des figures

1.1	Répartition des 14 hôpitaux composant les Hospices Civils de Lyon . . .	3
1.2	Synthèse des principales raisons des venues aux HCL de 2014 à 2019 .	4
2.1	Schématisation des sous-ensembles regroupant les différents types de systèmes d'information hospitaliers	12
2.2	Recommandations de l'HAS pour la prise en charge de patient·es atteint·es de TSA	17
5.1	Synthèse des moments clés de la genèse d'Easily®	52
5.2	Répartition des groupements hospitaliers français utilisant Easily® en 2019	53
5.3	Schématisation du fonctionnement du socle et des modules Easily® . .	55
5.4	Schématisation du fonctionnement de la passerelle Easily®	55
5.5	Schématisation des différentes couches du noyau Easily®	56
5.6	Interface utilisateur de CapMedecin en consultations médicales	58
5.7	Exemples de sous-modules dédiés à l'accès aux données patients	60
5.8	Exemples de sous-modules dédiés à la saisie de documents	61
5.9	Interface de l'application ePrescription pour les prescriptions informatisées	62
6.1	Disposition des agents dans la salle de consultation durant la phase d'observation	65
6.2	Degrés de présence, en pourcentage, des différents évènements dans les différentes traces d'activités collectées	69
6.3	Répartition des durées des traces d'activités collectées	71
6.4	Graphe de connexions entre les différents évènements collectés	72
6.5	Graphe de dépendance obtenu via <i>Heuristic Miner</i>	74
6.6	Graphe résultant de l'utilisation de <i>Fuzzy Miner</i>	77
7.1	Modélisation générique du processus de décision des médecins lors de consultations médicales coutumières	86
8.1	Schématisation d'un système de classification	94

8.2	Grille d'évaluation des différentes exigences en termes de "transparence" à respecter pour un système de classification	108
9.1	Architecture du réseau de neurones de BP-MLL	113
9.2	Évaluation de la "transparence" de différents systèmes de classification	122
9.3	Pertes de Hamming de différents systèmes de classification	134
9.4	Précisions micro et macro de différents systèmes de classification	136
9.5	Sensibilités, ou rappels, micro et macro de différents systèmes de classification	138
9.6	Spécificités micro et macro de différents systèmes de classification	140
10.1	Processus interne de l'outil d'aide à la décision	144
10.2	Interface utilisateur de l'outil d'aide à la décision	150
10.3	Disposition des agents lors des essais cliniques de l'outil d'aide à la décision	152
10.4	Matrice de corrélations entre les différents critères observés	156
10.5	Synthèse des réponses au questionnaire d'appréciation	157

Liste des tableaux

6.1	Synthèse des résultats des observations de consultations médicales	66
7.1	Exemple du processus d'une consultation spécifique	84
7.2	Formalisation en problème multi-labels de notre objet d'étude	89
8.1	Synthèse des différents éléments d'un système de classification et de leurs définitions	96
8.2	Synthèse de nos exigences en termes de "transparence"	98
9.1	Exemple simpliste de base d'apprentissage pour illustrer le fonctionnement de notre système de classification	125
10.1	Synthèse des résultats pour chaque médecin lors des essais cliniques	154

Liste des acronymes

ACC	<i>American College of Cardiology</i>	18
AHA	<i>American Heart Association</i>	18
BP-MLL	<i>Backpropagation for Multi-Label Learning</i>	112
CDSS	<i>Clinical Decision Support System</i>	12
CHICA	<i>Child Health Improvement through Computer Automation</i>	18
CHU	<i>Centre Hospitalier Universitaire</i>	52
CPOE	<i>Computer Physician Order Entry</i>	12
DDSS	<i>Diagnostic Decision Support System</i>	13
DPC	<i>Données Patient Communes</i>	59
DPI	<i>Dossier Patient Informatisé</i>	11
DSII	<i>Direction du Système d'Information et de l'Informatique</i>	2
EHR	<i>Electronic Health Record</i>	11
GHT	<i>Groupement Hospitalier de Territoire</i>	2
GIE	<i>Groupement d'Intérêt Économique</i>	1
HAS	<i>Haute Autorité de Santé</i>	16
HCL	<i>Hospices Civils de Lyon</i>	1
HIS	<i>Health Information System</i>	11
IA	<i>Intelligence Artificielle</i>	14
IMC	<i>Indice de Masse Corporelle</i>	87
IMCI	<i>Integrated Management of Childhood Illness</i>	18
IoT	<i>Internet of Things</i>	170
ML	<i>Machine Learning</i>	14
ML-kNN	<i>Multi-Label k-Nearest Neighbors</i>	115
NICE	<i>National Institute for health Care Excellence</i>	16
RAKEL	<i>Random k-Labelsets</i>	116
RGPD	<i>Règlement Général sur la Protection des Données</i>	35

RIPPER	<i>Repeated Incremental Pruning to Produce Error Reduction</i>	119
SMO	<i>Sequential Minimal Optimization</i>	120
SVM	<i>Support Vector Machine</i>	120

Introduction

L'aide à la décision est une activité consistant à aider un·e décideur·se à améliorer ses décisions, via une meilleure compréhension des enjeux de ces décisions, une examination attentive de données pertinentes, ou une utilisation plus rigoureuse de pratiques et théories pertinentes. Une aide à la décision est habituellement fournie sur demande, pour des situations spécifiques. Cependant, les demandeur·ses d'aide à la décision ont souvent, dans une certaine mesure, de bonnes connaissances concernant le domaine sur lequel va porter l'aide à la décision demandée.

De plus, les demandeur·ses d'aide à la décision sont souvent tenus, par de tierces parties ou le grand public, pour responsables des conséquences des décisions prises, qu'elles soient positives ou négatives. Dans de telles situations, la tâche de l'analyste chargé de fournir une aide à la décision est délicate dans le sens où iel risque d'empiéter sur l'expertise et la responsabilité du/de la décideur·se.

Le cas d'une aide à la décision visant à fournir un support à des médecins durant des consultations coutumières est paradigmatique. Les médecins sont des experts en termes de connaissances médicales et sont responsables des décisions qu'ils prennent. De nombreux outils ont été développés dans les dernières années, y compris des outils d'aide à la décision, pour leur apporter un soutien informatique durant leurs activités médicales. Cependant, comment s'assurer que les outils proposés aux médecins apportent un soutien pertinent sans empiéter sur leur expertise et leurs responsabilités ?

Les travaux présentés dans cette thèse ont été financés par le Groupement d'Intérêt Économique (GIE) "Hopsis"¹ et effectués en collaboration avec les employés des Hospices Civils de Lyon (HCL). Ils visent à répondre à cette question en se penchant plus particulièrement sur l'aide aux consultations médicales coutumières.

1. www.hopsis.org

1.1 Contexte

Créés en 1802, les HCL sont un Groupement Hospitalier de Territoire (GHT) de 14 hôpitaux répartis dans et aux alentours de la ville de Lyon (cf. Figure 1.1). Les HCL incluent aussi 3 départements informatiques, ou Direction du Système d'Information et de l'Informatique (DSII), dédiés au développement et au maintien des outils informatiques utilisés par les soignant·es. Plus spécifiquement, les informaticien·nes des DSII développent depuis 2015 un outil appelé Easily[®], permettant aux soignant·es d'accéder rapidement aux informations concernant les patient·es. Le GIE Hopsis est, quant à lui, chargé du financement et de la diffusion de Easily[®]. Nous donnons plus de précisions sur le fonctionnement et la genèse d'Easily[®] au Chapitre 5.

Depuis 2017, Easily[®] est utilisé quotidiennement par le personnel soignant des HCL et les DSII étudient les futurs outils pouvant être développés afin de soutenir les soignant·es dans leurs activités. C'est le cas, par exemple, des outils "O2", dédiés à simplifier la recherche clinique via la constitution de cohortes de patient·es et l'analyse de leurs données. DUCRAY *et al.*, (2020) fournissent un exemple récent de l'utilisation de tels outils.

Actuellement, près de 20.000 soignant·es travaillent aux HCL, dont près de 5.000 médecins. Comme nous le montre la Figure 1.2, les HCL recensent près de 1.250.000 venues tous les ans. Les consultations médicales composent une large partie de l'activité des médecins aux HCL avec près d'un million de consultations recensées tous les ans.

Les consultations médicales constituent donc un aspect central des processus de prises en charge des patient·es. Celles-ci servent autant au suivi des patient·es qu'à l'établissement de diagnostics servant de base à la prise en charge des patient·es. Cependant, durant ces consultations, les médecins prennent des décisions de manière répétée et des erreurs médicales peuvent se produire.

Les HCL projettent ainsi de développer des outils d'aide à la décision afin de minimiser la charge de travail du personnel soignant, ainsi que le risque d'erreurs médicales. Du fait de la centralité des consultations médicales dans l'activité des HCL, ainsi que dans les processus de prise en charge des patient·es, nous avons décidé de nous focaliser sur le soutien de ce type d'activités. Ainsi, nous cherchons à proposer un outil d'aide à la décision visant à apporter un soutien aux médecins durant leurs consultations médicales coutumières.

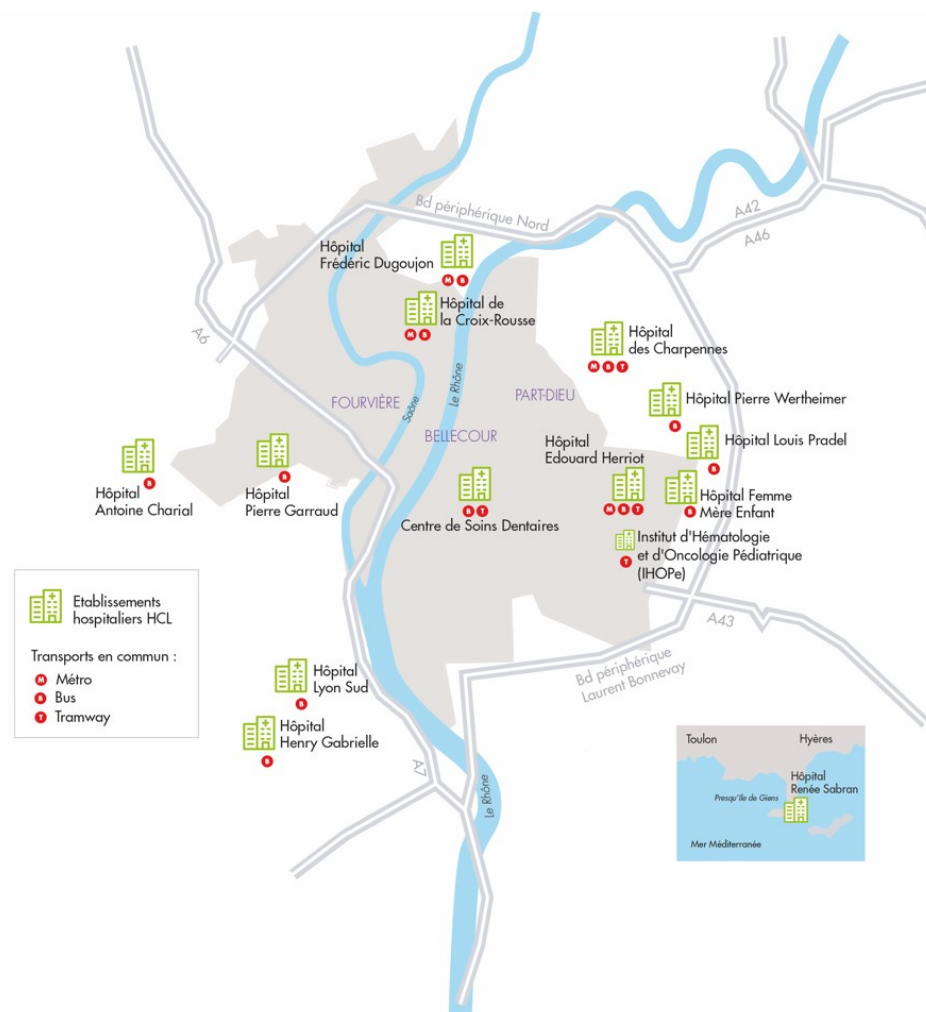


Figure 1.1.: Répartition des 14 hôpitaux composant les Hospices Civils de Lyon

1.2 Problématiques

La problématique centrale à laquelle nous avons tenté de répondre dans cette thèse est : "Comment aider les médecins durant leurs consultations médicales ?" Sachant que les consultations sont des activités pour lesquelles les médecins des HCL sont particulièrement coutumiers et compétents, tenter de répondre à cette problématique amène rapidement à devoir traiter d'autres questions tout aussi fondamentales.

Premièrement, nous pouvons nous demander : "Les médecins ont-ils besoin d'aide durant leurs consultations ?". En effet, comme mentionné précédemment, nous considérons que les médecins sont des experts médicaux et qu'ils sont compétents pour ce qui est du bon déroulé de leurs consultations. Ainsi, il est tout à fait possible

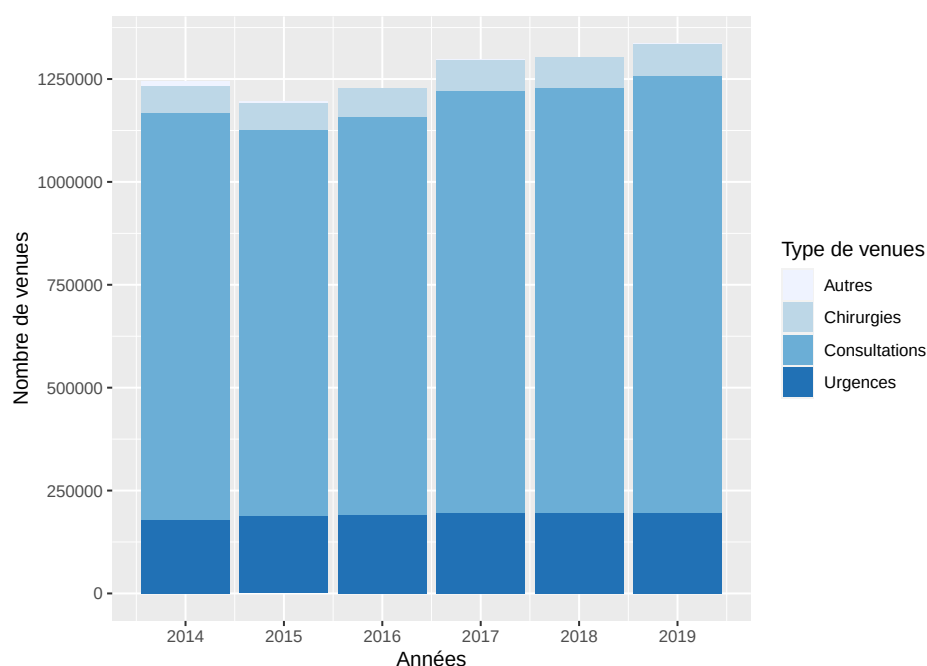


Figure 1.2.: Synthèse des principales raisons des venues aux HCL de 2014 à 2019

que les médecins n'aient besoin d'aucun support et qu'aucun outil d'aide à la décision ne soit pertinent dans de telles situations.

Cela nous amène à nous interroger sur les différentes approches employées pour aider les médecins dans leurs activités médicales, qu'elles soient coutumières ou non. En d'autres termes, nous tentons de répondre à la question : "Quels types d'outils sont actuellement fournis aux médecins afin de les soutenir dans leurs tâches?". Cette question en soulève trois autres : "Les outils actuellement fournis aux médecins ont-ils un impact positif sur les performances des médecins ou sur les soins apportés aux patient·es?", "Les outils actuellement fournis aux médecins sont-ils acceptés par ceux-ci et sont-ils bien intégrés dans leur processus de travail?" et "Les outils actuellement fournis aux médecins sont-ils adaptés à une aide aux consultations médicales coutumières?".

Enfin, ces interrogations nous amènent à devoir traiter une problématique d'autant plus fondamentale : "Que signifie réellement 'aider' un·e décideur·se compétent·e et responsable tel qu'un·e médecin?".

1.3 Objectifs

La thèse que nous soutenons dans ce manuscrit est qu'une aide à la décision adaptée et acceptable se doit de respecter le savoir-faire des médecins et de leur laisser la responsabilité des décisions prises en consultations, en se limitant à leur fournir les informations sur leurs patient·es qui leur sont nécessaires pour leurs prises de décision. Afin de défendre cette thèse, nous nous appuyons sur une analyse des différents aspects des consultations médicales et des actuels outils d'aide à la décision médicale, mais aussi sur les essais cliniques d'un nouvel outil conçu pour l'occasion. Le type d'outil d'aide à la décision que nous proposons de développer, fournissant aux médecins des informations ciblées concernant leurs patient·es, se base ainsi sur une meilleure compréhension des contraintes et des enjeux d'une aide à la décision lors de consultations médicales coutumières. Les travaux effectués durant cette thèse s'articulent autour de trois axes principaux.

Dans la Partie I, nous analysons les différentes approches employées pour aider les médecins, mais aussi le personnel soignant en général, dans leurs différentes activités, et à comprendre si ces approches sont applicables à notre cas. Cette partie reprend en grande partie le travail présenté dans RICHARD *et al.*, (2020b). Ainsi, au Chapitre 2, nous faisons un historique des différents outils d'aide à la décision médicale proposés au fil des années. Nous analysons ensuite leur impact sur les performances des médecins et sur la prise en charge des patient·es, avant de nous attarder sur l'acceptabilité de tels outils par les médecins. Grâce à cette étude, nous mettons en évidence que, malgré l'apport potentiel de tels outils, les médecins sont encore réticents à leur utilisation. Au Chapitre 3, nous explorons les diverses raisons pouvant expliquer cette non-acceptation des outils d'aide à la décision par les médecins. Nous mettons en évidence que les approches sur lesquelles se basent les outils d'aide à la décision sont le reflet de choix idéologiques forts dans le type d'aide à apporter aux médecins. Des choix idéologiques qui, sur de nombreux points, ne semblent pas correspondre aux réalités d'une activité médicale coutumière. Ceci peut expliquer, en partie, la non-acceptation des outils actuels dans de telles situations. Enfin, au Chapitre 4, nous tentons de déterminer quelle approche de l'aide à la décision serait la plus appropriée au support de médecins en situations coutumières. Nous y concluons qu'une approche "ajustive", cherchant à laisser un maximum d'autonomie au/à la décideur·se et s'adaptant à ses besoins, semble la plus appropriée compte tenu des contraintes sous-jacentes à une aide en consultations médicales coutumières.

Dans la Partie II, nous nous recentrons sur l'activité des médecins travaillant aux HCL et cherchons à analyser les spécificités de leurs processus de décision lors de leurs consultations. Notre objectif dans cette partie est d'identifier les points de ce processus pouvant être potentiellement coûteux en temps et en énergie pour les médecins et, ainsi, identifier leurs besoins en termes d'aide à la décision. Cette partie reprend ainsi les travaux présentés dans RICHARD *et al.*, (2018) en les approfondissant considérablement. Au Chapitre 5, nous nous concentrons sur la genèse et les spécificités d'Easily[®], le système informatique utilisé par les médecins travaillant aux HCL. Cela nous permettra d'avoir une meilleure connaissance générale des outils mis à la disposition des médecins durant leurs consultations, mais aussi du contexte et des objectifs qui ont présidé au développement de ces outils. Au Chapitre 6, nous analysons les données récoltées sur l'activité des médecins durant leurs consultations afin d'en dégager des processus récurrents. Ces analyses reposent sur des observations de consultations effectuées sur le terrain, ainsi que sur l'extraction informatique de processus à partir des traces d'activités des médecins collectées durant leurs consultations. Au Chapitre 7, nous proposons différentes modélisations du processus de décision des médecins durant une consultation médicale. Ces modèles nous ont permis de mettre en évidence le rôle central de la recherche d'information, sur la·e patient·e en cours, dans le processus de décision des médecins. Cette recherche d'informations est essentielle aux médecins pour leur permettre de prendre des décisions concernant la prise en charge de leurs patient·es, mais elle implique une charge de travail coûteuse en temps et en énergie.

À partir de cette observation, nous avons décidé d'orienter notre proposition vers un outil capable d'apprendre et d'anticiper les besoins en informations des médecins, en fonction de la pathologie du/de la patient·e en cours, et de fournir ces informations aux médecin afin de les décharger de cette charge de travail. Dans la Partie III, nous développons les différents aspects de l'outil d'aide à la décision que nous proposons. Les deux premiers chapitres de cette partie reprennent et approfondissent les travaux présentés dans RICHARD *et al.*, (2020a). Compte tenu du fait que nous utilisons un algorithme d'apprentissage afin de définir quel sous-ensemble d'informations les médecins ont besoin et pour quelle pathologie, cela peut générer une méfiance pouvant conduire à une non-acceptation de notre outil. Ainsi, au Chapitre 8, nous décrivons les différentes exigences en termes de "transparence" que nous nous sommes fixées afin de minimiser le risque de non-acceptation de notre outil. Au Chapitre 9, nous décrivons les différents algorithmes envisagés pour notre outil ainsi que la manière dont nous les avons évalués. Notre choix s'est finalement porté sur une variante de l'algorithme *Naive Bayes*, présentée dans ce même chapitre. Enfin, au Chapitre 10, nous présentons concrètement le fonctionnement de notre outil

d'aide à la décision, ainsi que les résultats des essais cliniques effectués en situations réelles aux HCL.

Partie I

Comprendre les contraintes et les enjeux
d'une aide à la décision en milieu médical

Revue des outils d'aide à la décision médicaux passés et actuels

” *To survive, let the past teach you. Past customs, struggles, leaders and thinkers. Let these help you. Let them inspire you, warn you, give you strength. But beware : God is change, past is past. What was cannot come again. To survive, know the past. Let it touch you. Then, let the past go.*

— **Octavia E. Butler**
Parable of the Talents

Dans ce chapitre, nous cherchons à présenter l'état actuel des outils d'aide à la décision dédiés au soutien des soignant·es, leur impact sur les performances des médecins, leur impact sur la prise en charge des patient·es et leur acceptabilité en pratique. Ce chapitre nous permettra aussi d'introduire les différentes notions concernant les outils d'aide à la décision qui nous seront utiles par la suite.

Comme schématisé en Figure 2.1, les outils d'aide à la décision que nous allons étudier ici font partie de l'ensemble, plus large, des systèmes d'information hospitaliers, ou *Health Information System* (HIS). Cet ensemble regroupe tout système informatique utilisé en milieux hospitaliers ou développé pour un contexte médical. De nombreux HIS ont été développés au cours des dernières décennies afin de soutenir la prise en charge des patient·es et d'améliorer divers processus hospitaliers. De nos jours, les HIS sont omniprésents dans les milieux médicaux et il est difficile de trouver un hôpital sans aucun système informatique.

Selon SHORTLIFFE et CIMINO, (2014), qui ont examiné les différentes applications des HIS, un des premiers types de HIS fut celui des systèmes informatiques permettant la saisie et la recherche de données médicales sur les patient·es. Plus communément appelé Dossier Patient Informatisé (DPI), ou *Electronic Health Record* (EHR), ceux-ci incluent les bases de données, les systèmes d'indexation et les systèmes de recherches utilisant des données médicales concernant les patient·es. De manière similaire, les

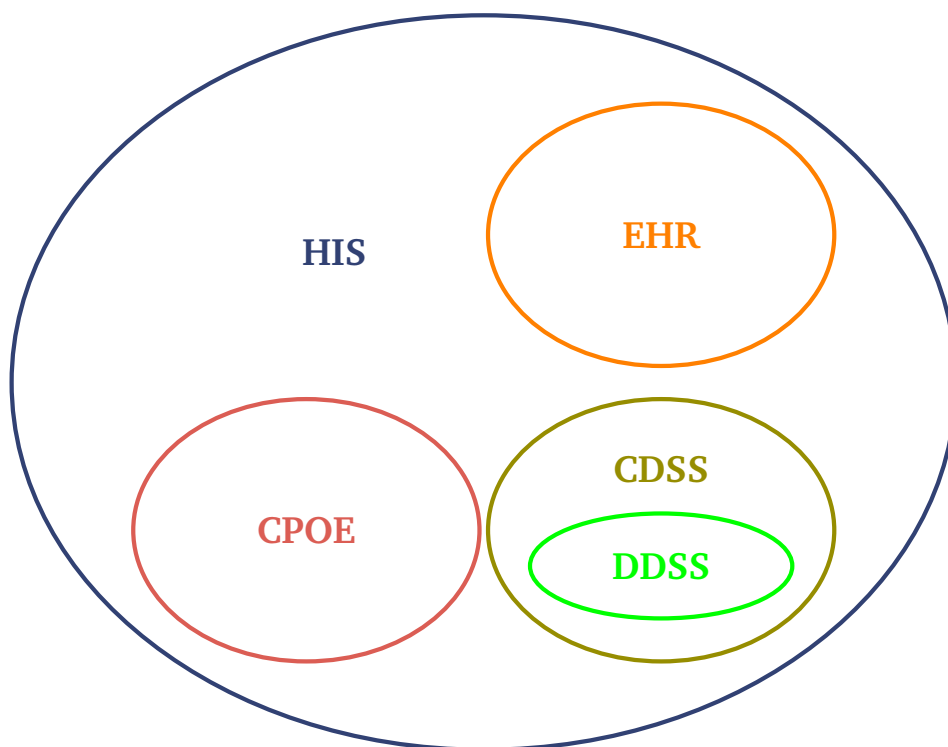


Figure 2.1.: Schématisation des sous-ensembles regroupant les différents types de systèmes d'information hospitaliers

systèmes de saisie d'ordonnances (KUPERMAN et GIBSON, 2003), ou *Computer Physician Order Entry* (CPOE), constituent l'ensemble des outils informatiques développés pour informatiser les ordonnances faites par les médecins.

Un autre sous-ensemble des HIS est celui des outils d'aide à la décision clinique, ou *Clinical Decision Support System* (CDSS). Comme développé par MUSEN *et al.*, (2014) et BERNER, (2016), les CDSS incluent tout système informatique destiné à transmettre des informations aux médecins, à un moment précis d'un processus médical, cela afin de les aider dans leurs prises de décision. L'objectif principal des CDSS est de minimiser le risque d'erreurs médicales. L'ensemble des CDSS est lui aussi composé de divers autres sous-ensembles.

Tout d'abord, les systèmes d'alerte, ou *Alert Systems*, sont conçus pour alerter rapidement le personnel soignant lors d'un événement indésirable. Le système HELP (WARNER *et al.*, 1972) offre un des plus anciens exemples de système d'alerte. Celui-ci émettait un signal lumineux et sonore si un·e patient·e venait à avoir un arrêt cardiaque. GermWatcher (DOHERTY *et al.*, 2006), quant à lui, était destiné à détecter et à traquer les infections bactériologiques chez les patient·es hospitalisés. Enfin, les systèmes d'alerte incluent aussi les systèmes intégrés dans des CPOE afin de

prévenir de potentielles allergies ou de mauvaises associations de médicaments lors de la saisie d'ordonnances. Dans le même ordre d'idée, les systèmes de rappels, ou *Reminders Systems*, forment un autre sous-ensemble des CDSS. Les systèmes de rappels incluent tout outil informatique utilisé par le personnel soignant pour les aider à se rappeler les actions à effectuer dans certaines conditions (GARG *et al.*, 2005). Par exemple, les interventions quotidiennes à effectuer pour le suivi d'un-e patient-e.

Enfin, nous arrivons au sous-ensemble que nous allons étudier dans ce chapitre, celui des outils d'aide aux diagnostics, ou *Diagnostic Decision Support System* (DDSS). Plus précisément, notre objectif est d'identifier les différentes manières de soutenir les médecins dans leurs décisions, d'examiner si l'impact de ces outils est bénéfique sur la prise en charge des patient-es et d'estimer si ces outils sont bien acceptés en pratique. Ceci devrait nous permettre d'évaluer si ces méthodes s'appliquent au contexte qui nous intéresse dans le cadre de cette thèse : l'aide à la décision en consultation médicale coutumière.

Pour ce faire, nous nous appuyons sur une analyse bibliographique effectuée sur PubMed¹ à l'aide de la requête suivante : ("*decision support system*" or "*computer-aided*" or "*machine learning*" or "*expert system*") and "*consultation*". 393 articles furent initialement trouvés via cette requête. Cet ensemble d'articles étant redondant au vu de nos objectifs car contenant de nombreuses revues et méta-analyses de DDSS, synthétisant les informations contenues dans d'autres articles de l'ensemble original, nous nous sommes concentrés sur ces revues et méta-analyses. Cependant, certains des articles les plus récents pouvant avoir été ignorés par ces revues et méta-analyses, nous avons donc conservé les articles présentant un DDSS spécifique publiés après 2017. Nous avons aussi conservé les articles présentant un DDSS spécifique incluant une étude de son impact ou de son acceptabilité, à cause du rôle central de ces notions dans notre étude.

En appliquant ces critères de sélection aux contenus des titres et résumés des articles, nous avons filtré 290 articles. Puis, en appliquant ces mêmes critères à l'intégralité du contenu des papiers restants, nous avons finalement obtenu 49 papiers. 12 d'entre eux (25%) sont des revues de DDSS utilisés dans des contextes médicaux généraux ou spécifiques, 27 (55%) sont des papiers présentant un DDSS spécifique (dont 20 incluant des essais cliniques ou des études de faisabilité) et 10 (20%) sont des études de l'impact des DDSS sur les performances des médecins, sur la santé des patient-es ou sur l'acceptabilité de ces systèmes. À ce corpus, nous ajoutons des articles issus

1. <https://www.ncbi.nlm.nih.gov/pmc/>

de différentes recherches sur GoogleScholar avec des requêtes similaires, ainsi que des livres, revues et études cités dans les articles sélectionnés.

Sur cette base, nous présentons en Section 2.1, l'historique des différents DDSS existants, depuis leurs ébauches dans les années 1950 jusqu'à leur utilisation actuelle, en nous concentrant sur les différentes méthodes utilisées. En Section 2.2, nous étudions l'impact des DDSS sur les performances des médecins et sur la santé des patient·es. Nous y montrons que l'utilisation d'un DDSS semble avoir un impact bénéfique sur ces deux points. Enfin, en Section 2.3, nous étudions l'appréciation générale de ces outils en situations pratiques. Nous y montrons que, malgré un impact potentiellement bénéfique des DDSS reconnu par les médecins, ceux-ci semblent peu enclins à utiliser de tels outils dans leurs pratiques quotidiennes.

2.1 Historique des outils d'aide aux diagnostics

Dans notre étude des différents outils d'aide à la décision développés au cours des dernières décennies, nous avons pu dégager trois grandes périodes détaillées dans les sous-sections ci-dessous. Les années situées entre 1950 et 2000 sont marquées par les premiers essais de systèmes experts visant à reproduire les capacités de diagnostic des médecins, mais aussi à initier des réflexions sur la place de l'informatique dans la prise en charge des patient·es. Aux alentours des années 2000, la mise en évidence des erreurs médicales comme l'une des principales causes de décès aux États-Unis a poussé les institutions de santé à financer le développement de guides de bonnes pratiques et d'outils fournissant des directives aux médecins. Enfin, les années 2010 sont particulièrement marquées par l'essor de l'Intelligence Artificielle (IA) et plus particulièrement des algorithmes dit de *Machine Learning* (ML).

2.1.1 1950-2000 : Les premiers essais de DDSS

Selon les revues de FIESCHI, (1986) et de MILLER, (1994), concernant les travaux sur les DDSS de 1954 à 1993, les premiers DDSS ont été développés dans le but de tenter de reproduire, via un ordinateur, le raisonnement d'un·e médecin faisant un diagnostic. Durant cette période, le terme consacré pour parler des DDSS était celui de "systèmes experts", ou *expert systems*. Durant ces premiers stades de l'histoire des DDSS, des années 1950 aux années 1970, la majorité des travaux ont porté sur la représentation du raisonnement des médecins, leurs connaissances et expertises, et

sur leurs usages possibles en situations pratiques. Selon MILLER, (1994), les premiers travaux sur le développement de systèmes experts datent des années 1970.

Dans les années 1980, le développement de systèmes experts dans différents contextes médicaux, comme la psychiatrie (MORELLI *et al.*, 1987) ou les consultations médicales (KULIKOWSKI, 1988), furent motivés par la prolifération des systèmes informatiques, mais aussi par les innovations dans les interfaces utilisateurs et les systèmes réseaux. Le système expert MYCIN (BUCHANAN et SHORTLIFFE, 1984 ; SHORTLIFFE *et al.*, 1973 ; SHORTLIFFE, 2012), dédié aux diagnostics d'infections bactériologiques, offre un premier exemple de DDSS produit à cette époque. Le système expert INTERNIST-I/QMR (MILLER *et al.*, 1986 ; MILLER, 2010), constitue un second exemple dédié aux diagnostics en médecine générale. Le système expert AI/RHEUM (KINGSLAND *et al.*, 1983), était quant à lui dédié aux diagnostics en rhumatologie. Enfin, le système expert DxPlain (BARNETT *et al.*, 1987 ; HOFFER *et al.*, 2005) était dédié à l'apprentissage des étudiant·es en médecine pour le diagnostic de pathologies courantes. Les exemples cités ici, ainsi que la plupart des systèmes experts produits durant cette période, étaient basés sur des ensembles de règles issues de l'expertise de différents spécialistes.

Selon MILLER et MASARIE JR, (1990), l'idée d'un DDSS utilisé comme un outil "magique" fournissant des recommandations que les médecins devaient suivre, poétiquement appelé le "*Greek Oracle*" model, a commencé à être considéré comme obsolète vers la fin des années 1980. En effet, les systèmes experts produits durant cette période avaient pour objectif d'expérimenter les potentielles applications de l'IA. À quelques exceptions près, comme DxPlain, la grande majorité de ces systèmes experts n'ont pas été utilisés en pratique, principalement du fait de leur coût de développement et de maintenance. Dans les années 1990, selon MILLER, (1994), la plupart des travaux sur les DDSS se sont orientés vers l'exploration d'autres méthodes issues de l'IA, telles que les réseaux de neurones artificiels (ZURADA, 1992) ou encore les systèmes à base de logique floue (ZADEH *et al.*, 1996), afin de proposer de nouvelles approches pour l'aide aux diagnostics.

2.1.2 2000-2010 : Une perte de confiance dans le savoir-faire des médecins

En 2000, l'institut national de médecine des États-Unis publia le rapport *To Err is Human : Building a Safer Health System* [L'erreur est humaine : construire un système de santé plus sûr]. Ce rapport, rédigé par DONALDSON *et al.*, (2000), examine de multiples études sur les erreurs médicales et conclut qu'entre 44.000 et 98.000 cas

de décès par an sont dus à des erreurs médicales évitables. À titre de comparaison, selon MOKDAD *et al.*, (2004) qui ont étudié les causes de décès aux USA pour l'année 2000, 43.000 décès sont dus à des accidents de véhicules motorisés, 75.000 décès sont dus à des microbes et 29.000 sont dus à des incidents impliquant des incendies. Le rapport *To Err Is Human* poussa les gouvernements et les politiques nationales de santé à considérer la sûreté des patient-es comme une priorité sanitaire.

Selon REIDER, (2016), durant les années 2000, de nombreuses politiques nationales furent mises en place pour favoriser la production de "guides de bonnes pratiques" (ou "*gold standard*" *guidelines*), l'éducation à la sûreté des patient-es et le développement de CDSS. D'autres études sur les erreurs médicales, telles que celle de LEAPE, (2000) ou celle de BERNER et GRABER, (2008), confortèrent les gouvernements dans leurs efforts et les incitèrent à continuer dans cette direction. Ainsi, de nombreux CDSS furent développés afin de prévenir de potentielles erreurs médicales, en particulier des systèmes d'alerte, des systèmes de rappels et des DDSS basés sur des "guides de bonnes pratiques" (ou *Guideline-based DDSS*). Durant cette période, de plus en plus de travaux sur les DDSS avaient pour objectif d'aider les médecins en leur fournissant des recommandations de diagnostics, ceci réinstallant petit à petit l'idée d'un DDSS utilisé tel un "oracle grec".

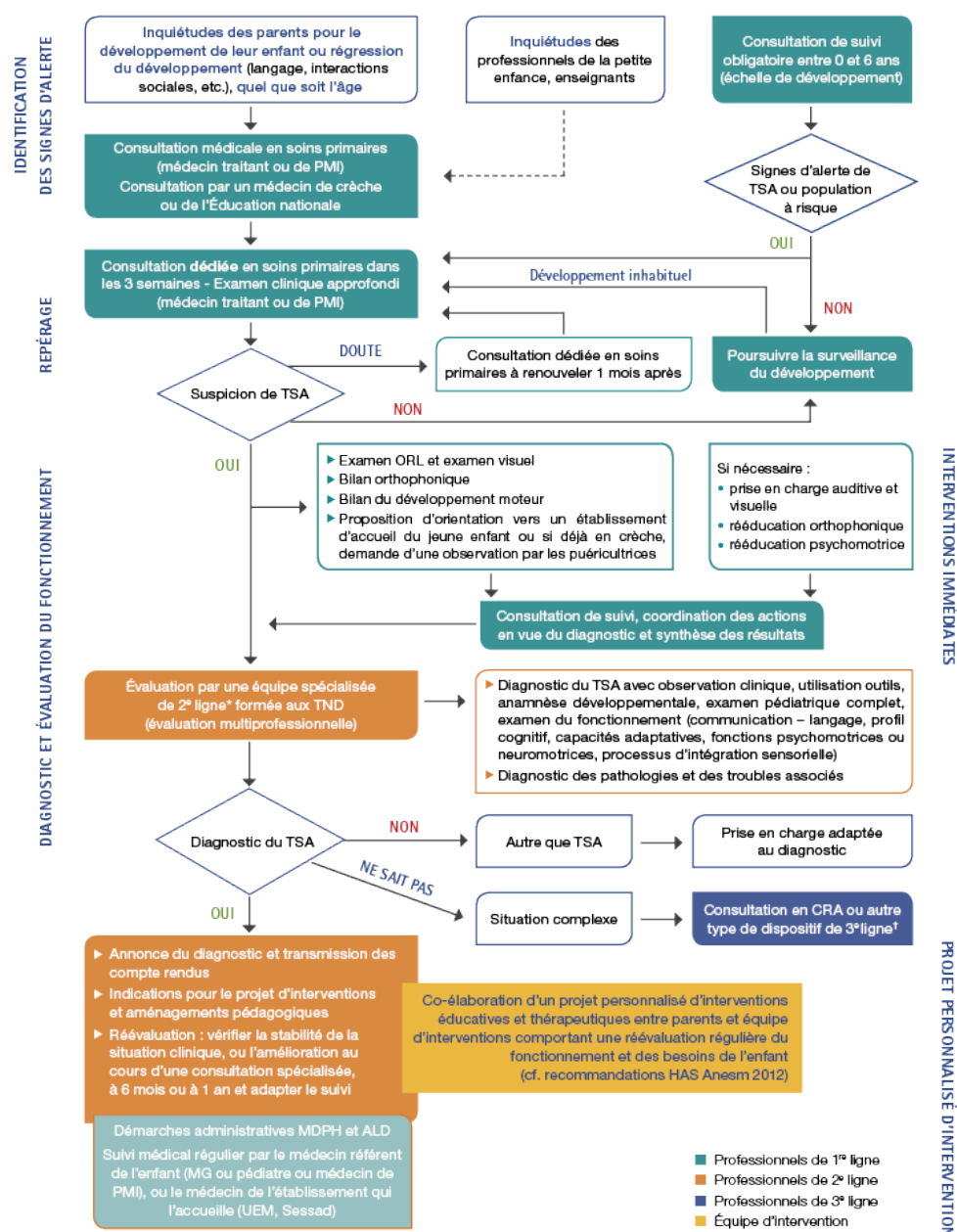
De même que les bases de connaissances utilisées par les systèmes experts et issues de l'expérience de médecins spécialistes, les "guides de bonnes pratiques" (ou *guidelines*) ont permis l'essor de ce que nous appelons des *Guideline-based DDSS*. Les *guidelines*, dont les "guides aux diagnostics", sont des listes d'instructions à suivre dans une situation spécifique. Ces *guidelines* sont généralement basées sur les bonnes pratiques à privilégier selon une autorité de santé (groupe de spécialistes, institution de santé, etc.) et peuvent être représentées sous la forme d'organigrammes. La Figure 2.2 montre un exemple d'organigramme fourni par la Haute Autorité de Santé (HAS), résumant les recommandations pour la prise en charge de patient-es atteint-es d'un trouble du spectre autistique. D'autres exemples de *guidelines* peuvent être trouvés sur le site web de la HAS², pour des *guidelines* françaises, ainsi que sur le site web de la *National Institute for health Care Excellence* (NICE)³, pour des *guidelines* anglaises.

Ainsi, les *Guideline-based DDSS* regroupent les systèmes experts se basant sur des "guides de bonnes pratiques", ou sur des connaissances de spécialistes, afin de fournir aux médecins les recommandations indiquées pour une situation spécifique (YANASE et TRIANTAPHYLLOU, 2019). Les *Guideline-based DDSS* incluent aussi les

2. www.has-sante.fr

3. www.nice.org.uk

4. www.has-sante.fr/jcms/c_468812



*Équipes de pédopsychiatrie (services de psychiatrie infanto-juvénile dont centres-médico-psychologiques - CMP), services de pédiatrie, centres d'action médico-sociale précoce (CAMSP), centres médico-psycho-pédagogiques (CMPP), réseaux de soins spécialisés sur le diagnostic et l'évaluation de l'autisme ou praticiens libéraux coordonnés entre eux par un médecin. †Professionnels exerçant en centre ressources autisme (CRA) ou en centre hospitalier pour des avis médicaux spécialisés complémentaires, notamment en neuropédiatrie, génétique clinique et imagerie médicale.

Figure 2.2.: Résumé des recommandations de la Haute Autorité de Santé (HAS) pour la détection et la prise en charge de patient·es atteint·es d'un trouble du spectre autistique (TSA) ⁴

systèmes fournissant directement aux médecins les "guides de bonnes pratiques" correspondants à une situation donnée, tels que le site web de la NICE ou bien l'outil

Quick Medical Reference (QMR) associé au système expert INTERNIST-1 (MILLER *et al.*, 1986 ; MILLER, 2010).

De nombreux *Guideline-based DDSS* furent donc conçus durant cette période, animés par une dynamique de prévention de potentielles erreurs médicales qui continue de persister. DxPlain (BARNETT *et al.*, 1987 ; HOFFER *et al.*, 2005), un des systèmes experts cités précédemment, fournissant des recommandations de médecine générale, est toujours en usage⁵. Le système *Child Health Improvement through Computer Automation* (CHICA) (ANAND *et al.*, 2004), utilisé à l'hôpital *Wishard Memorial* à Indianapolis, dédié à la prise en charge de jeunes patient·es en médecine générale, est un autre exemple de *Guideline-based DDSS*. CHICA se base sur les *guidelines* de l'état de l'Indiana pour générer des formulaires à l'attention des médecins et des parents. Ces formulaires sont utilisés pour récolter des informations sur les patient·es et pour recommander aux médecins des actions à effectuer durant leurs consultations. D'autres applications du système CHICA furent développées par la suite pour la prévention de la dépression maternelle (CARROLL *et al.*, 2013), la prévention des comportements suicidaires chez les adolescent·es (ETTER *et al.*, 2018) ou encore la prévention de l'apnée du sommeil chez le nourrisson (HONAKER *et al.*, 2018). Le DDSS proposé par GONZALVO *et al.*, (2017) est basé sur les *guidelines* du groupe de spécialistes CONSORT⁶ et fournit des recommandations de traitements pour des patient·es déjà polymédicamenteux. Le DDSS proposé par LÓPEZ *et al.*, (2017), nommé OphtalDDSS, est dédié à aider les médecins généralistes dans la détection de maladies oculaires, en accord avec les recommandations nationales. Basé sur les recommandations de l'*American College of Cardiology* (ACC) et de l'*American Heart Association* (AHA), le DDSS proposé par KIRBY *et al.*, (2018) est dédié à la détection préventive de maladies cardiaques sévères. YANG *et al.*, (2018) ont proposé un DDSS dédié à la prévention de réactions hypersensibles aux substances utilisées en radiographie, en se basant sur les recommandations des autorités de santé coréennes et les antécédents allergiques des patient·es. Les projets eIMCI (BESSAT *et al.*, 2019) et ALMANACH (BERNASCONI *et al.*, 2019) sont tous deux dédiés à la prise en charge de jeunes patient·es en médecine générale dans des pays en voie de développement, cela via l'utilisation de DDSS suggérant des actions ou des diagnostics basés sur les recommandations de l'*Integrated Management of Childhood Illness* (IMCI).

5. www.mghlcs.org/projects/dxplain

6. www.consort-statement.org

2.1.3 2010-2020 : L'essor de l'apprentissage automatique

Un récent développement dans cet historique des DDSS provient du fait que de plus en plus de données sont enregistrées sur les patient·es, du fait de la prolifération et à la généralisation des HIS. Il a rapidement été montré que ces larges volumes de données étaient difficilement analysables par un cerveau humain (YANASE et TRIANTAPHYLLOU, 2019) et noyaient le personnel soignant sous toutes sortes d'informations (PIVOVAROV et ELHADAD, 2015). Les algorithmes de fouille de données, ou *Data Mining*, et d'apprentissage automatique, ou *Machine Learning* (ML), se sont montrés bien plus adaptés et efficaces dans le traitement de ces larges volumes de données (DUA *et al.*, 2014 ; YANASE et TRIANTAPHYLLOU, 2019).

Les algorithmes de ML sont des méthodes conçues pour apprendre comment approximer des fonctions de classification à partir d'ensembles de données appelées bases d'apprentissage. Une fonction de classification peut être, par exemple, une fonction anticipant la valeur d'une variable y à partir de la valeur d'une variable x , ou bien une fonction capable de distinguer des images d'organes sains d'images d'organes malades à partir de l'analyse d'une matrice de pixels. Les problèmes de ML sont généralement divisés en trois catégories, en fonction du degré d'information fourni dans la base d'apprentissage : l'apprentissage supervisé, ou *supervised learning*, où toutes les données sont disponibles ; l'apprentissage semi-supervisé, ou *semi-supervised learning*, où une partie des données n'est pas disponible ; et l'apprentissage non supervisé, ou *unsupervised learning*, où aucune classification n'est prédéfinie. De nombreux algorithmes de ML ont été proposés afin de traiter ces problèmes, allant des algorithmes bayésiens aux réseaux de neurones artificiels, en passant par les arbres de décisions ou encore les machines à vecteurs de support. Nous revenons plus en détail sur ces algorithmes dans le Chapitre 9.

Ainsi, comme développé dans les revues de DUA *et al.*, (2014), OZAYDIN *et al.*, (2016), MIOTTO *et al.*, (2017), KULIKOWSKI, (2019) et YANASE et TRIANTAPHYLLOU, (2019), de nombreux travaux ont commencé à voir le jour vers la fin des années 2000 et le début des années 2010 sur le développement de DDSS basés sur tous types d'algorithmes de ML, ou *ML-based DDSS*. Ces travaux avaient pour objectif d'aider aux diagnostics de pathologies spécifiques et de minimiser le risque d'erreurs de diagnostic via le traitement de larges volumes, et ce dans plusieurs domaines médicaux. Par exemple, DEIG *et al.*, (2019) ont proposé une revue de plusieurs *ML-based DDSS* développés pour une utilisation en radio-oncologie, principalement dédiés à l'anticipation des risques de mauvaises réactions aux traitements donnés aux patient·es. De manière similaire, PEIFFER-SMADJA *et al.*, (2019) ont proposé une revue des *ML-based DDSS* dédiés au support des médecins pour des cas de

maladies infectieuses. Ces outils fournissent, entre autres, des recommandations de diagnostics ou de traitements, des détections préventives de risques de maladies ou des détections des risques de réactions aux traitements. Dans leur revue de différents HIS, KATARIA et RAVINDRAN, (2018) ont rapporté l'usage de *ML-based DDSS* pour anticiper les réponses possibles à un traitement ou pour anticiper la propagation de certaines maladies. La revue de GORDON *et al.*, (2018) recense différents usages d'algorithmes de ML en génétique médicale, que ce soit pour la recommandation de diagnostics comme pour l'évaluation du risque d'apparition d'une maladie génétique.

Comme exemples de *ML-based DDSS* en particulier, DE FAUW *et al.*, (2018) et ZHANG *et al.*, (2018) ont tous deux proposé des *ML-based DDSS* dédiés à la détection précoce de maladies oculaires via l'analyse d'images rétinienne. TITANO *et al.*, (2018) ont proposé un *ML-based DDSS* dédié à l'anticipation d'événements neurologiques via l'analyse de radiographies crâniennes. PEARCE *et al.*, (2019) ont proposé un *ML-based DDSS* pour évaluer le risque de prise en charge d'urgence d'un·e patient·e au moment de sa consultation. JAROSZEWSKI *et al.*, (2019) ont proposé un *ML-based DDSS* dédié à la détection précoce de maladies mentales. Enfin, de nombreux *ML-based DDSS* furent développés dans l'objectif de détecter précocement différents types de tumeurs, telles que des tumeurs mammaires (JOO *et al.*, 2004), des tumeurs cérébrales (HOLLON *et al.*, 2018) ou des tumeurs de la peau (ESTEVA *et al.*, 2017). Ces détections de tumeurs, en particulier celle de la peau, peuvent être utilisées via des outils de télémédecine, comme rapporté par ELSNER *et al.*, (2018) et PASQUALI *et al.*, (2020). Ces différents exemples de *ML-based DDSS* jouent le rôle d'extensions aux médecins, effectuant des tâches que des humains ne peuvent accomplir avec le même degré de précision.

D'après ce bref historique des DDSS, il apparaît que le rapport *To Err Is Human*, ainsi que les travaux l'ayant complété, ont mis en évidence certaines limitations concernant les compétences de diagnostic des médecins. En réponse, les autorités de santé de différents pays ont financé le développement de "guides de bonnes pratiques", ou *guidelines*, et de DDSS dédiés à améliorer l'adhérence des médecins à ces *guidelines*. Ainsi, les travaux sur les DDSS se sont concentrés, des années 1950 jusqu'aux années 2000, sur les *Guideline-based DDSS*. Les premiers essais de DDSS, des années 1950 jusqu'à la fin des années 1990, se basent principalement sur des recommandations issues de l'expertise de médecins spécialisés travaillant en collaboration avec les développeurs, puis sont suivis dans les années 2000 par des travaux basés sur les recommandations de hautes autorités de santé. Plus récemment, il est apparu que les algorithmes d'apprentissage automatique, ou ML, pouvaient être plus performants que les médecins pour effectuer certaines tâches (ex. identifier

des mélanomes microscopiques sur des images). Cela incita au développement de travaux sur les *ML-based DDSS* durant les deux dernières décennies. Bien que, comme nous l'avons vu, des *Guideline-based DDSS* sont toujours en développement actuellement, les *ML-based DDSS* ont commencé à dominer le domaine des DDSS depuis la fin des années 2000 et le début des années 2010⁷.

2.2 Des preuves que les HIS, CDSS et DDSS sont bénéfiques

Dans cette section, nous nous intéressons à l'impact que peuvent avoir les HIS en général sur les performances des médecins ou sur la sûreté des patient·es, avant de nous concentrer sur l'impact des CDSS, puis des DDSS. Comme l'a montré l'étude de PATEL *et al.*, (2000) sur l'impact des HIS et plus spécifiquement la représentation des connaissances dans les EHR, un simple système de saisie de données patient peut avoir un impact important sur le comportement des médecins et leurs processus de travail. Plus particulièrement, PATEL *et al.*, (2000) ont montré une standardisation des processus de travail des médecins au fur et à mesure de l'utilisation des outils, ceux-ci convergeant vers le processus induit par les EHR utilisés. Il est donc important de s'assurer que les outils mis à la disposition des médecins auront potentiellement un impact bénéfique sur leurs méthodes de travail, ainsi que sur les soins apportés aux patient·es.

La méta-analyse de 257 études proposée par CHAUDHRY *et al.*, (2006) concernant l'impact des HIS sur la qualité, l'efficacité et les coûts des soins médicaux a montré que les HIS peuvent avoir un impact bénéfique sur les performances du personnel

7. Notons qu'il existe certaines exceptions aux deux sous-ensembles prédominants des DDSS que sont les *Guideline-based DDSS* et les *ML-based DDSS*. Le DDSS proposé par GRÄSSER *et al.*, (2017) est dédié à la recommandation de traitements en se basant sur des mesures de similarités entre le cas courant et des cas précédemment rencontrés, cela pour chaque nouveau cas rencontré et sans processus d'apprentissage entre chaque cas. Bien que son utilisation rapproche cet outil des *ML-based DDSS*, le fait est qu'il n'utilise pas d'algorithme de ML à proprement parler. GIORDANENGO *et al.*, (2019) ont proposé, quant à eux, un DDSS dédié à la prise en charge de patient·es diabétiques, via la présentation de données collectées par les patient·es elleux-mêmes et de rappels d'actions à effectuer par les médecins durant leurs consultations. Cependant, GIORDANENGO *et al.*, (2019) ne se basent pas sur les recommandations d'une haute autorité de santé, mais sur des règles à appliquer sous certaines conditions proposées par des médecins impliqués dans le développement de l'outil. De plus, ces règles, établies par le consensus des médecins impliqués, ne sont pas destinées à d'autres médecins, mais au DDSS lui-même. Enfin, le *ML-based DDSS* proposé par SIMON *et al.*, (2019) utilise un algorithme de ML non pas pour faire des recommandations aux médecins, mais pour détecter et extraire des concepts complexes dans des documents médicaux numérisés. Avec ce DDSS, SIMON *et al.*, (2019) montrent qu'il est possible d'utiliser des algorithmes de ML pour autre chose que des recommandations, tout en apportant une aide pertinente en situations pratiques.

soignant. Selon LEAPE et BERWICK, (2005) et WACHTER, (2004), qui ont étudié les améliorations des systèmes de santé concernant la sûreté des patient·es cinq ans après le rapport *To Err Is Human*, mais aussi selon CLANCY, (2009) qui proposa une étude similaire dix ans après *To Err Is Human*, le premier impact de ce rapport fut l'automatisation des enregistrements d'erreurs médicales. Avec la généralisation des HIS, l'enregistrement d'actes médicaux devient plus régulé et contrôlé. En outre, avec le développement des systèmes d'alerte et de rappels, le risque d'erreurs potentielles se réduit année après année. De tels résultats pour les HIS, qui sont de plus en plus omniprésents dans le travail du personnel soignant et dans certains aspects de la prise en charge des patient·es, bien qu'ils n'aient pas été démontrés avant les années 2000, furent à n'en pas douter rapidement perçus par les médecins, les autorités de santé et les gouvernements. Dans ce contexte, la perte de confiance envers les médecins, incarnée par le rapport *To Err Is Human*, a fourni une opportunité historique aux CDSS de montrer leur utilité.

Anticipant l'appel pour le développement d'outils d'aide à la décision du rapport *To Err Is Human*, JOHNSTON *et al.*, (1994) ont étudié 28 essais contrôlés de différents types de CDSS (dosage assisté par ordinateur, *Guideline-based DDSS*, systèmes de rappels, etc.) afin d'évaluer l'impact potentiel des CDSS sur les performances des soignant·es. Dans cette étude, JOHNSTON *et al.*, (1994) définissent comme de "bonnes" performances les taux d'erreurs bas dans le dosage des médicaments et dans les diagnostics, mais aussi un respect élevé des directives sanitaires. Les auteur·ices rapportent que certains CDSS, plus spécifiquement les systèmes de dosage assisté par ordinateur, semblent avoir un impact bénéfique sur les performances des médecins. De même, HUNT *et al.*, (1998) ont étudié l'effet de plusieurs CDSS sur les performances du personnel soignant avec 68 essais contrôlés et mis à jour par GARG *et al.*, (2005) avec 97 autres essais contrôlés. Ces auteur·ices concluent que la plupart des CDSS peuvent améliorer les performances du personnel soignant. KAUSHAL *et al.*, (2003) ont proposé une étude des effets des CDSS, plus spécifiquement ceux inclus dans des CPOE, sur la sûreté des soins apportés aux patient·es. Les auteur·ices ont montré une potentielle réduction du taux d'erreur de traitement grâce à l'utilisation des CDSS. SLAIN *et al.*, (2014) ont rétrospectivement analysé un an d'utilisation d'un CDSS dédié au support des infirmier·ères en service d'urgence. Le CDSS était intégré dans le processus de travail du service d'urgence et proposait un pré-traitement des données des patient·es à leur arrivée. SLAIN *et al.*, (2014) ont rapporté une plus grande précision dans le placement des patient·es et dans le transfert d'informations entre les membres du personnel soignant grâce à l'utilisation du CDSS. ZIER *et al.*, (2017) ont analysé l'utilisation d'un CDSS durant un an, en comparaison avec trois ans sans CDSS. Ce CDSS était dédié au support de transferts d'organes par la

détection précoce de mort cérébrale. Les auteur·ices ont montré que ce CDSS avait un impact bénéfique sur les soins apportés aux patient·es en attente d'organes. Une exception cependant à tous ces résultats positifs : VERDOORN *et al.*, (2018) ont étudié une année d'utilisation d'un CDSS dédié à la prévention d'erreurs de traitement et ont obtenu de plus mauvais résultats avec le CDSS que sans. Les auteur·ices ont souligné la nécessité d'apporter des améliorations au processus de leur CDSS.

Dans le cas plus spécifique des DDSS, bien qu'il y ait quelques exceptions comme les études de ECCLES *et al.*, (2002) et POELS *et al.*, (2008), qui ont effectué des essais contrôlés de *Guideline-based DDSS* basés sur des scénarios issus de situations coutumières pour les médecins et ont rapporté que l'utilisation d'un DDSS ne présentait pas d'impact significatif (ni positif, ni négatif) sur les performances des médecins ou leur processus de travail, une majorité des études sur les DDSS montrent un impact bénéfique sur les performances des médecins. HECKERLING *et al.*, (1991), CHANG *et al.*, (1996), MURPHY *et al.*, (1996) et ELSTEIN *et al.*, (1996), qui ont effectué des essais contrôlés du système expert Iliad (WARNER *et al.*, 1988 ; WARNER JR, 1989), ont montré que le système expert améliorait la précision des diagnostics des médecins pour les cas complexes, particulièrement chez les étudiants en médecine (MURPHY *et al.*, 1996). TAYLOR *et al.*, (2008) ont effectué des essais contrôlés d'un *Guideline-based DDSS* dédié au support des médecins dans la prise en charge de patient·es asthmatiques. Les auteur·ices ont montré que le DDSS a aidé les médecins à améliorer leur processus de décision et à diminuer la durée de leurs consultations. WATROUS *et al.*, (2008) ont étudié, à l'aide d'essais contrôlés, l'impact d'un *Guideline-based DDSS* dédié à la détection de souffle cardiaque lors d'auscultations. Les auteur·ices ont rapporté une amélioration, due à l'utilisation du DDSS, de la sensibilité et de la spécificité de la classification, par les médecins, des types de souffles cardiaques. CARROLL *et al.*, (2013) ont effectué des essais cliniques de leur *Guideline-based DDSS* dédié à la prévention des dépressions maternelles, alertant les médecins lorsqu'une patiente remplit un certain nombre de critères. D'après les auteur·ices, le DDSS a montré un impact potentiellement bénéfique sur les soins apportés aux patientes. KOSTOPOULOU *et al.*, (2017) ont effectué des essais cliniques d'un *Guideline-based DDSS* dédié au support de médecins généralistes en leur fournissant une liste de diagnostics potentiels basés sur les données récoltées auprès des patient·es. Les auteur·ices ont montré une amélioration de la précision des diagnostics des médecins due à l'utilisation du DDSS. KOSTOPOULOU *et al.*, (2017) ont aussi rapporté le fait que les médecins utilisant le DDSS saisissaient bien plus de données concernant les patient·es. Enfin, KIRBY *et al.*, (2018) ont analysé un an d'utilisation, dans 13 hôpitaux différents, d'un *Guideline-based DDSS* dédié à la prévention de maladies cardiovasculaires. Les auteur·ices ont rapporté que l'utilisation

du DDSS a amélioré la précision des médecins, mais aussi la prise en charge des patient·es.

Les essais de DDSS dans des pays en voie de développement montrent aussi des résultats encourageants. DALABA *et al.*, (2014) ont étudié une année d'utilisation d'un *Guideline-based DDSS* dédié à la santé de l'enfant dans des centres de soins au Ghana. Les auteur·ices ont rapporté une diminution des complications médicales et des décès après l'introduction du DDSS. De manière similaire, BESSAT *et al.*, (2019) ont effectué des essais cliniques d'un *Guideline-based DDSS* dédié à la santé de l'enfant dans des établissements de soins généralistes au Burkina Faso. Les auteur·ices ont rapporté une amélioration de la sûreté des soins apportés aux patient·es due à l'utilisation du DDSS. Enfin, BERNASCONI *et al.*, (2019) ont analysé l'impact de l'introduction de *Guideline-based DDSS* dédié à la santé de l'enfant dans divers hôpitaux de différents pays en voie de développement. Les essais cliniques rapportés par les auteur·ices ont montrés une amélioration de la précision des médecins dans la prise en charge des patient·es.

Concernant les *ML-based DDSS*, bien que ceux-ci aient démontré leur capacité à surpasser les médecins en termes de précision durant différentes phases de tests (ESTEVA *et al.*, 2017), l'impact de leur utilisation en situations réelles reste encore peu étudié (YANASE et TRIANTAPHYLLOU, 2019). Prenons par exemple le cas des *ML-based DDSS* dédié à la détection de maladies infectieuses étudiés par PEIFFER-SMADJA *et al.*, (2019). Les auteur·ices rapportent que sur 60 *ML-based DDSS* étudiés, seuls trois présentaient des essais cliniques, ce qui est trop peu pour pouvoir tirer des conclusions. Actuellement, les *ML-based DDSS* représentent encore un domaine émergent et il reste hasardeux de déterminer si ils améliorent la sûreté des soins apportés aux patient·es ou modifient de manière bénéfique le processus de travail des médecins.

Si nous résumons, les différents essais cliniques de DDSS ont montré un impact théoriquement bénéfique sur les performances des médecins et sur la sûreté des soins apportés aux patient·es. Les *Guideline-based DDSS* se montrent assez performants lorsqu'ils sont utilisés en médecine générale, pour les soins primaires ou dans des pays en voie de développement. Ce sont là des situations pour lesquelles des "guides de bonnes pratiques" pour des cas spécifiques sont bienvenus. Pour ce qui est des *ML-based DDSS*, ceux-ci sont encore principalement évalués sur des critères de performances lors de phases de test (YANASE et TRIANTAPHYLLOU, 2019). Il reste délicat de déterminer si les actuels *ML-based DDSS* fournissent un soutien bénéfique aux médecins en situations pratiques.

2.3 Une acceptabilité discutable

Nous nous intéressons maintenant à l'acceptabilité des HIS, CDSS et DDSS dans leurs utilisations par le personnel soignant en situations pratiques. Selon SHORTLIFFE et CIMINO, (2014), les HIS les plus omniprésents sont les CPOE, les EHR, les systèmes de rappels et les systèmes d'alerte. Concernant ces derniers, CHADWICK *et al.*, (2017) ont étudié l'introduction, durant trois mois dans deux hôpitaux distincts, d'un système d'alerte dédié à la prévention du VIH et ont montré que, malgré une légère "fatigue" due aux alertes, le système d'alerte était généralement bien accepté par le personnel soignant. Cependant, l'introduction d'un système informatique dans un contexte hospitalier reste une tâche compliquée. Selon HEEKS *et al.*, (1999), qui ont étudié les causes potentielles de réussite ou d'échec des HIS, bien que certains HIS aient été introduits avec succès en situations pratiques, beaucoup ont essuyé des échecs. Un exemple que les auteur·ices explorent afin d'illustrer ce phénomène est celui de BEYNON-DAVIES et LLOYD-WILLIAMS, (1998), expliquant l'échec d'un CDSS dédié au support d'un service d'urgence londonien. Dans cet exemple, l'introduction du CDSS échoua car, selon BEYNON-DAVIES et LLOYD-WILLIAMS, (1998) : "*the speed and the depth of change were simply too aggressive for the circumstances*" [la vitesse et la profondeur des changements étaient simplement trop agressives au vu des circonstances]. L'annulation de ce projet de CDSS causa une perte estimée à 20 millions de livres sterling de l'époque. Un autre exemple utilisé par HEEKS *et al.*, (1999) est l'analyse proposée par GUAH, (1998) de l'échec de l'introduction d'un DDSS dédié au support de médecins lors de coloscopies, pour un hôpital universitaire du Royaume-Uni. D'après HEEKS *et al.*, (1999), le système produisait des informations statistiques qui n'étaient pas significantes pour les médecins et nécessitait d'apprendre un nouveau processus de travail totalement différent de leurs habitudes. Le développement du DDSS fut alors abandonné. Comme le montrent ces exemples, bien qu'un HIS se doit d'être performant et bénéfique aux médecins comme aux patient·es, il est tout aussi important que celui-ci soit accepté et utilisé en pratique par le personnel soignant, particulièrement au vu des sommes investies dans le développement d'un HIS.

SITTIG *et al.*, (2006) ont étudié les facteurs influençant l'acceptabilité des DDSS. Les auteur·ices rapportent dans leur étude qu'un fort pourcentage des recommandations des DDSS est ignoré par les médecins. Selon OVERHAGE *et al.*, (1997), TIERNEY *et al.*, (2003) et WEINGART *et al.*, (2003), ce pourcentage varie entre 54% et 91%. SITTIG *et al.*, (2006) ont aussi rapporté que les médecins étaient plus enclins à accepter les recommandations d'un DDSS pour les cas de patient·es âgés, avec des traitements multiples ou étant polypathologiques. ONEGA *et al.*, (2010) ont étudié l'acceptabilité

de différents DDSS utilisés par des radiologues, cela en comparant les préférences des médecins entre utiliser un DDSS et demander l'avis d'un autre médecin via une seconde lecture. Les auteur·ices ont interrogé 257 radiologues de différents hôpitaux à travers les USA. D'après leurs résultats, les radiologues étaient plus favorables à une seconde lecture par un pair, quand bien même une majorité d'entre eux reconnaissait que l'utilisation d'un DDSS apportait une meilleure précision qu'une seconde lecture. La méta-analyse proposée par MASUD *et al.*, (2019), sur l'utilisation de DDSS en services de radiologie, a montré des résultats similaires de faible acceptabilité, malgré l'amélioration des performances perçue et reconnue par les radiologues.

Seulement une poignée d'études a montré une bonne acceptabilité des *Guideline-based DDSS*, dans des situations vraiment spécifiques. C'est le cas des travaux de PORAT *et al.*, (2017), qui ont analysé l'acceptabilité par des médecins et des patient·es d'un *Guideline-based DDSS*. 34 médecins généralistes ont participé à l'étude en consultant 12 cas standard de patient·es durant des essais contrôlés. Les auteur·ices ont rapporté que 74% des médecins ont trouvé le DDSS utile, bien que l'utilisation du DDSS les obligeait à saisir plus de données sur les patient·es tout en interagissant avec ceux-ci. L'utilisation de DDSS dans des pays en voie de développement constitue aussi un cas particulier où les *Guideline-based DDSS* apparaissent comme largement acceptés par les médecins comme par les patient·es (DALABA *et al.*, 2014 ; BESSAT *et al.*, 2019 ; BERNASCONI *et al.*, 2019).

Concernant les *ML-based DDSS*, tout comme leur impact sur la sûreté des soins apportés ou les performances des médecins, leur acceptabilité en situations pratiques reste peu étudié (PEIFFER-SMADJA *et al.*, 2019). JAROSZEWSKI *et al.*, (2019) ont rapporté que, durant les essais cliniques de leur *ML-based DDSS* dédié à la prévention de maladies mentales, seulement 28% des participant·es ont répondu "very likely" [très probablement] lorsque le DDSS leur présentait la question : "Be honest, how likely are you to try the resources I just shared?" [Soyez honnête, à quel point allez-vous probablement essayer les ressources que je viens juste de partager?]. Dans leur étude sur l'acceptabilité de différents types de HIS dédiés à la prévention de maladies sexuellement transmissibles, NADARZYNSKI *et al.*, (2020) ont rapporté que, pour un premier contact, 70% des patient·es préféreraient une consultation en face à face. Les auteur·ices rapportent aussi que seulement 40% des patient·es trouvent les IA sous forme de *chatbot* acceptables.

Pour résumer, bien que l'on trouve des exceptions pour des situations spécifiques, les DDSS sont généralement peu acceptés en situations coutumières, où le support apporté semble redondant avec le savoir-faire et les capacités des médecins (MASUD

et al., 2019). Par conséquent, il apparaît que nous sommes actuellement face à une situation quelque peu paradoxale : les DDSS ont montré une capacité à améliorer les performances des médecins et la sûreté des soins apportés aux patient·es, une capacité reconnue par les médecins, cependant les DDSS restent en pratique peu utilisés et difficiles à intégrer dans le processus de travail des médecins. Il apparaît aussi que les capacités intrinsèques d'un DDSS ne peuvent être considérées comme le seul facteur déterminant de son utilité potentielle. Il y a, par conséquent, un besoin de mieux comprendre pourquoi les DDSS ne sont pas bien acceptés en situations pratiques et quelles sont les caractéristiques pouvant améliorer l'acceptabilité des DDSS dans ces situations. Cette compréhension doit nous permettre de proposer un outil plus adapté à notre cas d'aide à la décision en consultations médicales coutumières.

Étude des raisons de la non-acceptation des outils d'aide aux diagnostics

” *Il y a des objets que nous avons appris à utiliser. Nous les utilisons, mais pas comme le font les Visiteurs, c'est presque certain. Je suis totalement convaincu que dans la plupart des cas nous enfonçons des clous avec des microscopes.*

— **Arkadi & Boris Strougatski**

Stalker, pique-nique au bord du chemin

Nous avons vu au Chapitre 2 que les HIS, CDSS et DDSS ont potentiellement un impact bénéfique sur les performances des médecins et la sûreté des soins apportés aux patient·es. Cependant, nous avons aussi vu que l'introduction d'un CDSS dans un hôpital peut échouer et que les DDSS actuels sont généralement peu acceptés par les médecins, qui ignorent souvent les recommandations des DDSS en pratique.

Les premières explorations des barrières à l'usage de "guides de bonnes pratiques", ou *guidelines*, contiennent des indications utiles sur les raisons de la non-acceptation des outils d'aide à la décision par les médecins. CABANA *et al.*, (1999) ont effectué une méta-analyse de 76 études sur la non-acceptation des *guidelines* en situations cliniques et ont rapporté 7 barrières potentielles, classifiées en trois catégories :

1. Les barrières externes, telles que : la présence de *guidelines* contradictoires, l'incapacité de concilier les préférences de certain·es patient·es avec les recommandations de certaines *guidelines*, et d'autres facteurs environnementaux comme le manque de temps ou le manque de ressources ;
2. Les barrières affectant l'attitude des médecins envers les *guidelines*, telles que : les désaccords avec des *guidelines* spécifiques ou avec les *guidelines* en général, l'inertie des méthodes de travail précédentes, la croyance qu'ils ne peuvent effectuer les recommandations des *guidelines* ou la croyance que les recommandations des *guidelines* ne vont pas mener aux résultats désirés ;

3. Les barrières liées aux connaissances des médecins concernant les *guidelines*, telles que : les problèmes d'accès aux *guidelines*, le volume d'informations à emmagasiner, ainsi que le temps à consacrer pour se tenir informés de nouvelles directives.

Dans ce chapitre, nous cherchons à élargir et enrichir cette analyse des raisons potentielles de la non-acceptation des DDSS en l'appliquant plus largement aux différents aspects des HIS, CDSS et DDSS, avec une attention particulière sur l'usage de ces outils en situation de consultations médicales coutumières. Dans la Section 3.1, nous proposons un aperçu du potentiel impact négatif que peut avoir l'usage des DDSS en pratique, mais aussi plus généralement des CDSS et des HIS. Nous mettons en évidence que l'introduction de HIS dans les processus médicaux, loin de les alléger, entraîne une augmentation des tâches "administratives" demandées aux soignant·es et, ainsi, de leur charge de travail. En Section 3.2, nous explorons les problèmes de responsabilités éthiques soulevés par l'utilisation des DDSS en pratique. Il apparaît que l'influence que peuvent avoir de tels outils sur le processus de décision des médecins est à prendre en considération par les concepteurs des DDSS autant que par les médecins et les institutions de santé, en particulier si nous considérons que la bonne prise en charge d'un·e patient·e est de la responsabilité des médecins. Enfin, en Section 3.3, nous résumons les différentes raisons pour lesquelles les spécificités des DDSS, *Guideline-based* et *ML-based*, peuvent conduire à une non-acceptation par les médecins en situation de consultations médicales coutumières. Nous mettons ainsi en évidence que les DDSS actuels ne sont pas adaptés à l'aide à la décision en consultations coutumières et qu'il nous faudra envisager de nouvelles approches d'aide à la décision pour traiter le cas que nous étudions.

3.1 Un impact négatif des HIS, CDSS et DDSS

TSAI *et al.*, (2003) ont étudié l'impact, sur les performances de différents médecins, de mauvaises recommandations fournies par un DDSS. Les auteur·ices cherchaient ainsi à remettre en question un postulat généralement admis : si un DDSS fait une erreur ou fournit une mauvaise recommandation, la·e médecin sera à même de la détecter. Cette étude est basée sur 83 simulations adaptées de cas cliniques réels issus de services de cardiologie. Les sujets étaient 30 internes en médecine, en deuxième ou troisième année d'internat, et le DDSS était contrôlé de manière à pouvoir produire des recommandations ne correspondant pas aux recommandations officielles. TSAI *et al.*, (2003) ont rapporté que, lorsque le DDSS fournissait de bonnes recommandations, la précision des sujets s'améliorait. En revanche, la

précision des sujets diminuait grandement lorsque le DDSS fournissait de mauvaises recommandations. Les auteur·ices ont aussi rapporté que les sujets suivaient les recommandations du DDSS plus souvent lorsque les recommandations étaient fournies avec un degré de confiance élevé. POVYAKALO *et al.*, (2013) ont effectué une étude similaire concernant l'impact, sur les performances de 50 radiologues, d'un DDSS dédié à la détection de tumeurs mammaires. Dans cette étude, les auteur·ices ont évalué la capacité de discrimination des radiologues sur 180 mammographies, avec et sans DDSS. POVYAKALO *et al.*, (2013) ont rapporté que le DDSS aidait les radiologues les moins précis, mais pénalisait les radiologues les plus performants en réduisant leurs précisions. Plus généralement, BOWMAN, (2013) a étudié la sûreté dans l'usage des EHR. Il a ainsi rapporté que de mauvais choix de conceptions, des usages inadéquats des outils, des erreurs de programmation, ou des erreurs dans les données, pouvaient mener à des erreurs diminuant la sûreté des soins apportés aux patient·es. Ces types d'erreurs concernent, finalement, tout type de HIS, dont les CDSS et donc les DDSS. Une première barrière aux DDSS est ainsi la crainte de ne pas pouvoir détecter les erreurs potentielles d'un DDSS.

BERTILLOT, (2016) a étudié, via un ensemble d'entretiens avec des membres du personnel soignant de différents hôpitaux en France, les effets des HIS sur la rationalisation et la standardisation des méthodes de travail du personnel soignant. BERTILLOT, (2016) a ainsi montré que l'introduction, durant les dernières décennies, de différents HIS dans les hôpitaux a amélioré la traçabilité des patient·es hospitalisés et a permis une meilleure transmission des informations entre services. Cependant, cela a aussi permis la mise en place de systèmes d'évaluation dans ces mêmes hôpitaux, comparant les performances des différents services et facilitant l'introduction de pratiques managériales compétitives dans les hôpitaux publics. BERTILLOT, (2016) a aussi rapporté une augmentation de la charge administrative demandée au personnel soignant, ceux-ci devant saisir plus d'informations dans les systèmes d'information et ce plus régulièrement. Ce temps consacré aux tâches administratives, bien que nécessaire pour différentes raisons, n'est pas du temps consacré à la prise en charge des patient·es. Les résultats de l'étude effectuée par MITCHELL *et al.*, (2016) mettent en évidence les mêmes aspects de l'impact des HIS. Les auteur·ices se sont entretenus avec des experts dans la sûreté des soins apportés aux patient·es en les interrogeant sur leurs perceptions des travaux effectués sur le signalement des erreurs médicales. Dans cette étude qualitative, MITCHELL *et al.*, (2016) ont rapporté que les médecins percevaient, principalement à cause de leur manque de temps, le signalement systématique des erreurs médicales comme une charge de travail additionnelle.

HALL *et al.*, (2016) ont effectué une méta-analyse de 46 études sur le bien-être des médecins et la sûreté des soins apportés aux patient·es, afin de déterminer s'il existe une corrélation entre bien-être des médecins, surmenage et sûreté des soins. Les auteur·ices ont rapporté qu'un faible bien-être des médecins était significativement corrélé à un haut risque de surmenage, ainsi qu'à une mauvaise sûreté des soins apportés et donc à un haut risque d'erreurs médicales. WEST *et al.*, (2018), qui ont étudié les raisons et les conséquences du surmenage du personnel soignant, ont rapporté que l'utilisation de HIS comme l'un des facteurs menant au surmenage des médecins. Par conséquent, en essayant d'améliorer la sûreté des soins apportés aux patient·es, les HIS actuels augmentent la charge de travail du personnel soignant. Cette charge de travail supplémentaire réduit le bien-être des médecins et, par un effet collatéral, augmente le risque d'erreurs médicales en pratique. Ces travaux mettent ainsi en évidence la charge actuelle des médecins comme une seconde barrière à l'introduction des DDSS en situations pratiques, et ce pour deux raisons : soit les médecins manquent de temps pour se consacrer à la prise en main d'un nouvel outil, soit ils y sont forcés et la potentielle charge de travail supplémentaire en résultant risque d'impacter négativement la sûreté des soins apportés. De plus, si les médecins sont les seuls en charge de la détection des erreurs potentielles d'un outil supposé leur apporter une aide, ceci ne fera que générer plus de charge de travail, ce qui apparaît comme particulièrement contre-productif.

CABITZA *et al.*, (2017) ont étudié les conséquences imprévues de l'utilisation du *Machine Learning* (ML) en contextes médicaux. Les auteur·ices ont rapporté que les systèmes basés sur du ML, du fait de leurs performances mais aussi de leur opacité, peuvent amplifier la perte de compétences des médecins rapportée par TSAI *et al.*, (2003) et POVYAKALO *et al.*, (2013). Les auteur·ices ont aussi rapporté que l'incertitude intrinsèque aux contextes médicaux affectait les performances des systèmes de ML, réduisant leurs performances. De même, CHALLEN *et al.*, (2019) ont étudié l'impact potentiel de l'Intelligence Artificielle (IA) sur la sûreté des soins apportés. Les auteur·ices ont rapporté différentes causes d'erreurs médicales dues à l'utilisation d'outils IA. Par exemple, les systèmes basés sur du ML sont généralement entraînés sur des cas appartenant à un contexte bien précis et perdent leur précision lorsque ce contexte vient à changer. L'opacité de certains systèmes basés sur du ML et l'automatisation de certaines tâches sont aussi rapportés comme des facteurs augmentant le risque d'erreurs médicales. CHALLEN *et al.*, (2019) argumentent aussi que les systèmes basés sur de l'apprentissage par renforcement, apprenant au fur et à mesure de nombreux essais pratiques, sont potentiellement dangereux sur le long terme car ceux-ci peuvent faire des choix exploratoires non-sûrs ou renforcer seulement des comportements à court terme. Nous sommes faces ici à une troisième

barrière à l'utilisation de DDSS en pratique : le manque de connaissances des technologies utilisées, ainsi que l'incertitude en résultant. En effet, dans le cas des *ML-based DDSS*, les médecins sont supposés moins "compétents" que leur DDSS dans l'accomplissement de la même tâche, et il apparaît peu probable que ceux-ci soient en capacité de détecter si le DDSS a fait une erreur ou non¹.

3.2 Des problèmes de responsabilité

ITANI *et al.*, (2019), qui ont étudié l'utilisation d'algorithmes de fouille de données pour l'aide à la décision, ont montré que des facteurs sociétaux tels que les valeurs partagées par les patient-es et les médecins forment un aspect important à prendre en compte pour comprendre la non-acceptation des DDSS. Ces valeurs renvoient à des perceptions sociales et à des implications éthiques de l'utilisation des DDSS, mais aussi à des pressions sociales concernant la responsabilité des médecins par rapport aux conséquences de leurs décisions (TSOUKIÀS, 2020).

Selon GOODMAN, (2016), qui a étudié les problèmes éthiques et légaux entourant les DDSS, il y a un besoin de définir les responsabilités légales de l'usage des DDSS. En effet, si un-e médecin utilise un DDSS et que le DDSS fournit une mauvaise recommandation conduisant à une erreur médicale, qui doit être tenu pour responsable ? (DE DOMBAL, 1987). La réponse dépend clairement de la manière dont le DDSS a été développé puis utilisé.

Une erreur dans le programme du DDSS peut conduire à une mauvaise recommandation (BOWMAN, 2013). Dans ce cas, nous pouvons tenir pour responsables les concepteurs ou les développeurs du DDSS. Une erreur médicale peut aussi venir d'une mauvaise compréhension ou d'une mauvaise utilisation de l'outil par un-e médecin. La technologie utilisée par le DDSS peut aussi être une source d'erreur. Dans le cas, par exemple, d'un *Guideline-based DDSS* utilisant une base de règles définie par des spécialistes, ces "spécialistes" peuvent produire des règles pouvant être considérées comme "dangereuses" ou "insensées" par le reste de la communauté médicale. Le fait de passer par une haute autorité de santé et par des recommandations faisant consensus ne fait qu'amoindrir le problème puisqu'il existe toujours des cas où il n'y a pas consensus. C'est par exemple le cas lors de l'apparition de nouvelles pathologies. Nous retrouvons ici la barrière liée à la crainte de ne pas pouvoir détecter une erreur potentielle de l'outil, ainsi qu'une seconde barrière

1. Le domaine naissant des IA Explicables (DORAN *et al.*, 2017 ; GUNNING, 2017 ; RUDIN et RADIN, 2019), que nous développons plus en détail en Chapitre 8, offre des perspectives intéressantes pour atténuer ce problème.

possible : le désaccord avec certaines recommandations, voir avec le fait même d'avoir à suivre des recommandations d'un outil informatique.

Nous pouvons, à bon droit, défendre que les DDSS basés sur des algorithmes de ML sont plus dignes de confiance que les *Guideline-based DDSS*, principalement grâce aux hautes performances des algorithmes de ML, se basant sur un large volume de cas et pouvant même surpasser les performances des médecins (ESTEVA *et al.*, 2017). Cependant, les problèmes de responsabilités ne sont pas différents dans le cas des *ML-based DDSS* : si un *ML-based DDSS* fournit une mauvaise recommandation conduisant à une erreur médicale, qui est responsable ? Les algorithmes de ML sont entraînés et évalués sur des ensembles de données qui peuvent ne pas englober toute la variété des cas possibles. Ainsi, quand bien même un *ML-based DDSS* montrerait de bonnes performances lors de phases de tests, ces critères de performances apparaîtraient comme nécessaires mais insuffisants en termes de sûreté des soins apportés aux patient·es en conditions réelles.

De plus, les algorithmes de ML dit "supervisés" ne peuvent que reproduire les comportements qu'ils ont appris. Ainsi, tout comme pour les *Guideline-based DDSS*, si la base d'apprentissage utilisée est basée sur des comportements de médecins pouvant être considérés comme "dangereux" ou "insensés" par le reste de la communauté médicale, l'algorithme de ML entraîné sur cette base reproduira sûrement, voire amplifiera, ces comportements "dangereux" (GARCIA, 2016 ; SANDVIG *et al.*, 2016 ; ZOU et SCHIEBINGER, 2018). La principale différence avec les *Guideline-based DDSS* étant qu'il est bien plus difficile pour des médecins de détecter si un *ML-based DDSS* a un comportement non-voulu, en particulier si le processus du DDSS reste opaque pour les médecins. Ceci est d'autant plus préoccupant lorsque les médecins ont une grande confiance dans les recommandations du *ML-based DDSS* (TSAI *et al.*, 2003 ; POVYAKALO *et al.*, 2013). Nous retrouvons ici les barrières liées aux connaissances sur les technologies utilisées.

Les *Guideline-based DDSS* et les *ML-based DDSS* posent tous deux des problèmes lorsque les médecins sont considérés comme seuls responsables des soins apportés aux patient·es et de leurs conséquences. Ainsi, comme introduit précédemment, les médecins ne peuvent être considérés comme les seuls en charge de la prévention des erreurs médicales dues à l'utilisation d'un DDSS supposé les aider. Selon les *Asilomar AI Principles*², une liste de principes développés durant un séminaire organisé par l'institut *FutureOfLife* et dédiée à guider les institutions et les concepteurs dans le développement d'IA bénéfiques, les concepteurs d'IA et les institutions utilisant des IA se doivent de prendre leur part de responsabilité dans la prévention des erreurs

2. www.futureoflife.org/ai-principles

potentiellement dues à l'utilisation de systèmes d'IA. Ces principes concernent donc les *ML-based DDSS*, mais aussi les *Guideline-based DDSS* tels que les systèmes experts. L'émergence d'outils législatifs dédiés à la régulation de l'usage des systèmes informatiques, en particulier sur l'utilisation de données sensibles comme avec le Règlement Général sur la Protection des Données (RGPD) en Europe (VOIGT et BUSSCHE, 2017), témoigne d'une prise de conscience générale de ces problèmes. Une prise de conscience qui identifie les DDSS actuels comme n'étant pas encore à la hauteur des attentes exprimées. Nous avons ici une barrière externe à l'utilisation des DDSS : une méfiance de société envers les algorithmes en général.

3.3 Un écart entre conception et réalité pour l'aide en consultation

En plus des barrières mises en évidence dans les sections précédentes, la littérature suggère qu'il existe une raison supplémentaire pouvant expliquer la non-acceptation des DDSS par les médecins lors de situations coutumières : l'écart entre conception et réalité, ou *reality-design gap problem*.

Ce concept a été proposé par HEEKS *et al.*, (1999) et HEEKS, (2006) afin d'expliquer pourquoi certains HIS réussissent dans leurs objectifs, là où d'autres échouent. Les auteur·ices montrent que plus grand est l'écart entre la manière dont est conçu un HIS et les réalités du terrain, plus grand est le risque que le HIS échoue à remplir ses objectifs. Pour formaliser ce problème d'écart conception-réalité, HEEKS *et al.*, (1999) ont proposé le cadre méthodologique ITPOSMO, défini par sept dimensions pouvant chacune être sujette à un écart entre conception et réalité :

1. *Information* : les utilisateur·ices sont-ils familiers aux types d'informations fournis par le système ?
2. *Technologie* : l'hôpital possède-t-il les capacités techniques pour faire fonctionner le système ?
3. *Processus* : comment le système s'intègre-t-il dans le processus de travail des utilisateur·ices ?
4. *Objectifs et Valeurs* : les objectifs du système concordent-ils avec les objectifs et valeurs morales des utilisateur·ices ?
5. *Effectifs et Compétences* : le système nécessite-t-il de hautes compétences techniques pour être utilisé ?

6. *Systèmes de gestion* : le système nécessite-t-il des structures additionnelles afin de l'administrer ?
7. *Autres* : l'utilisation du système est-elle coûteuse en temps d'utilisation ? L'utilisation du système génère-t-il de la charge de travail supplémentaire ?

En détaillant, ci-après, le contenu des sept dimensions d'ITPOSMO pour le cas précis de l'utilisation de DDSS en consultations médicales coutumières, nous cherchons à mettre en évidence les contraintes expliquant la non-acceptation des DDSS dans ce contexte. Précisons tout de même qu'il existe certainement des exceptions aux caractéristiques générales que nous explorons dans notre analyse de l'écart conception-réalité des *Guideline-based DDSS* et *ML-based DDSS*. Notre objectif consiste simplement à mettre en évidence les tendances générales des approches actuellement utilisées dans la conception d'outil d'aide à décision clinique et à déterminer si ces approches sont applicables dans le cadre de consultations coutumières.

3.3.1 Les DDSS basés sur des "guidelines"

Les *Guideline-based DDSS* fournissent généralement des recommandations, provenant de "guides de bonnes pratiques", d'actions ou de traitements adaptés à une situation spécifique. Les médecins sont familiers aux directives provenant de hautes autorités de santé, une partie de leur travail étant de se tenir à jour des nouvelles directives pour les cas qu'ils peuvent être amenés à traiter. Ainsi, il y a peu d'écart conception-réalité sur le plan des informations fournies aux médecins par le système.

Sur le plan des capacités technologiques, les *Guideline-based DDSS* sont généralement intégrées dans des HIS déjà existants afin d'accéder aux données des patient·es et ne nécessitent pas de ressources supplémentaires de la part de l'hôpital pour sa mise en place. Il y a donc peu d'écart conception-réalité sur le plan technologique.

Les *Guideline-based DDSS* nécessitent généralement que les médecins saisissent les symptômes du/de la patient·e ou toute autre information supplémentaire demandée par le DDSS. En situations coutumières, ce type de processus peut être considéré comme redondant avec le processus d'un·e médecin durant une consultation (cf. Chapitre 7). Ainsi, sur le plan des processus il existe un certain écart conception-réalité.

L'objectif des *Guideline-based DDSS* est généralement d'améliorer l'adhérence des médecins aux bonnes pratiques définies par des autorités de santé. Dans des situations coutumières, cet objectif est vite confronté aux valeurs morales des médecins,

telles que leur libre arbitre ou leur désaccord avec certaines des bonnes pratiques recommandées (CABANA *et al.*, 1999). Les *Guideline-based DDSS* peuvent aussi automatiser trop de choses dans le processus des médecins, ce qui peut amener à une sensation de perte de contrôle pour les médecins (HEEKS, 2006). Il semble ainsi exister, en situations coutumières, un important écart entre l'objectif premier des *Guideline-based DDSS* et les valeurs des médecins tel que leur libre arbitre et leur volonté de garder le contrôle du processus de leurs consultations.

Sur le plan des compétences techniques, l'utilisation des *Guideline-based DDSS* ne nécessite pas de compétences supplémentaires de la part des médecins. Il n'y a donc pas d'écart conception-réalité sur cet aspect.

Sur le plan de l'administration du système, les bonnes pratiques recommandées par les autorités étant en constante évolution, l'utilisation de *Guideline-based DDSS* peut nécessiter des mises à jour régulières par un agent externe des recommandations à fournir. Cela génère un léger écart conception-réalité.

Enfin, l'utilisation de *Guideline-based DDSS* peut être coûteuse en temps pour les médecins, ceux-ci pouvant passer plus de temps à saisir des informations dans le système qu'à interagir avec leurs patient·es (PORAT *et al.*, 2017). De plus, les médecins se doivent de comprendre les raisons sous-tendant une recommandation pour éviter de potentielles erreurs médicales, ce qui génère une charge supplémentaire de travail pour les médecins.

Pour résumer, l'utilisation de *Guideline-based DDSS*, dans le cadre de consultations coutumières, fait face à deux contraintes majeures. Premièrement, les médecins sont souvent déjà au courant des bonnes pratiques à employer pour les cas dont ils sont coutumier·ères. Ainsi, l'utilisation d'un *Guideline-based DDSS* apparaît pour les médecins comme inutilement coûteuse en temps et redondante avec leur processus de travail. Deuxièmement, les médecins semblent avoir une forte volonté de garder le contrôle du déroulement de leurs consultations et de rester responsables de la sûreté des soins apportés aux patient·es, ce qui peut entrer en conflit avec l'objectif des *Guideline-based DDSS*.

3.3.2 Les DDSS basés sur du "Machine Learning"

Intéressons-nous maintenant au cas des *ML-based DDSS*. Ce cas est plus complexe à traiter que celui des *Guideline-based DDSS*, principalement car il constitue encore un domaine émergent et que ces outils sont encore peu utilisés en situations pratiques (PEIFFER-SMADJA *et al.*, 2019).

Les *ML-based DDSS* fournissent généralement des recommandations ou des degrés de risque. Les médecins sont plutôt familiers avec ces types d'informations. Cependant, les raisons sous-jacentes à une recommandation peuvent, selon l'algorithme de ML utilisé, ne pas être intelligibles pour les médecins. Ainsi, il peut y avoir un léger écart conception-réalité concernant le type d'informations fournies aux médecins.

Sur l'aspect de la technologie, l'utilisation de certains algorithmes de ML, tels que les réseaux de neurones, peut nécessiter de puissantes ressources technologiques lors de la phase d'entraînement de l'outil. Une fois entraîné, un *ML-based DDSS* ne nécessite généralement pas de puissantes ressources pour être utilisé. Cependant, dans le cas d'un apprentissage continu, de puissantes ressources technologiques peuvent être nécessaires en continu (KULIKOWSKI, 2019). Il y a donc, là aussi, un léger écart entre conception et réalité, dépendant principalement de l'algorithme de ML utilisé.

Les *ML-based DDSS* actuels utilisent généralement des données déjà saisies dans le HIS et ne nécessitent pas d'actions supplémentaires de la part des médecins. Ils peuvent ainsi être intégrés assez aisément dans le processus de travail des médecins, à l'instar d'une information supplémentaire concernant la·e patient·e en cours. Sur le plan des processus, il n'y a donc potentiellement que peu, voire pas, d'écart conception-réalité.

Concernant l'objectif principal des *ML-based DDSS*, celui-ci est de fournir un outil hautement performant pour guider les médecins dans des tâches qu'ils ne sont pas capables de réaliser avec la même précision sans un soutien informatique. Dans le cadre de cas coutumiers, pour lesquels les médecins peuvent être considérés comme compétents, cet objectif apparaît contradictoire et superflu, ce qui peut créer un sentiment de perte de contrôle et de méfiance envers le DDSS si celui-ci est imposé.

L'utilisation d'un *ML-based DDSS* peut nécessiter un apprentissage supplémentaire pour les médecins pour être à même d'interpréter les résultats du DDSS, en comprendre les forces et les limites, et ainsi prendre des décisions sans biais. Il y a donc un écart conception-réalité évident concernant les compétences requises par les médecins dans l'utilisation des *ML-based DDSS* en situations coutumières.

Sur le plan de l'administration système des *ML-based DDSS*, il est possible de mettre en place un système mettant régulièrement à jour la base d'apprentissage avec les nouveaux cas rencontrés et relançant l'apprentissage du système. Ceci nécessite d'évaluer en continu les performances du DDSS afin de prévenir des erreurs potentielles. Cependant, rien de tout cela ne nécessite l'intervention d'un

agent extérieur, chaque étape pouvant être automatisée. Il n'y a donc potentiellement pas d'écart conception-réalité sur le plan de l'administration du système.

Enfin, comprendre les raisons d'une recommandation fournie par un *ML-based DDSS*, quand cela est possible, peut générer une charge de travail additionnelle pour les médecins. Il y a donc un possible écart conception-réalité sur cet aspect.

Si nous résumons, une première barrière à l'utilisation de *ML-based DDSS* en situations coutumières provient de la technologie utilisée. Si les médecins ne peuvent pas comprendre le fonctionnement de l'outil qu'ils utilisent ou comment interpréter les résultats de cet outil, celui-ci leur apparaîtra opaque, générant de la méfiance ou un sentiment de perte de contrôle. Ce sentiment peut être renforcé par l'objectif des *ML-based DDSS*, qui est de surpasser les performances des médecins dans situations pour lesquelles ils pourraient se considérer comme compétents. Ce qui nous amène à la seconde barrière, commune avec les *Guideline-based DDSS*, qui est la volonté des médecins de garder le contrôle de leurs processus de travail en situations coutumières et de rester responsables de la sûreté des soins apportés aux patient·es.

Sur la base des points soulevés dans cette section, mais aussi dans l'ensemble de ce chapitre, nous pouvons avancer que les approches actuelles des DDSS pour aider à la décision ne semblent pas adaptées aux situations de consultations médicales coutumières. Il nous faut donc envisager de nouvelles approches pour aider les médecins dans leurs décisions en de telles situations.

Proposition d'une nouvelle approche pour l'aide à la décision en consultations

“ *'But you're supposed to help us !', said Masklin.
'I suggest you consider deeply the proper meaning
of the word "help", said the box.
'Either you are intelligent nomes, or just clever
animals. It's up to you to find out which.'*

— **Terry Pratchett**
The Bromeliad Trilogy

Dans ce chapitre nous cherchons à analyser, autant d'un point de vue pratique que philosophique, les différentes approches actuellement employées dans le développement d'outils d'aide à la décision en milieux médicaux. Comme nous l'avons développé en Chapitre 2, dans leurs applications pour aider aux décisions de diagnostics, les DDSS actuels sont dans une situation quelque peu paradoxale. D'un côté leur potentielle utilité et leurs apports apparaissent indiscutables, mais d'un autre côté les DDSS restent peu acceptés en pratique par les médecins. Cette situation paradoxale reflète, de notre point de vue, une difficulté plus générale à fournir une aide à la décision pertinente à l'attention de décideur·ses considérés comme compétent·es et responsables dans leurs décisions. En analysant en détail le cas spécifique des DDSS pour des consultations coutumières, nous cherchons à développer une approche nouvelle afin de nous attaquer à cette difficulté.

En ciblant ce problème, nous cherchons à contribuer à un large domaine de recherche dédié à l'analyse des défis faisant face aux différentes approches et méthodologies de l'aide à la décision, comme principalement développé en sciences de la décision et en recherche opérationnelle, quand elles sont employées dans la conception, l'implémentation et l'évaluation de politiques publiques (TSOUKIÀS *et al.*, 2013 ; DE MARCHI *et al.*, 2016). Ce domaine de recherche a produit, entre autres, des travaux sur l'évaluation de politiques environnementales (JEANMOUGIN *et al.*, 2017), la conception d'options politiques (FERRETTI *et al.*, 2019 ; PLUCHINOTTA *et al.*, 2018 ;

PLUCHINOTTA *et al.*, 2019) et le développement d'outils méthodologiques pour des politiques environnementales de grande envergure (CHOULAK *et al.*, 2019). Dans cet ensemble de contributions, nous nous appuyons sur l'approche méthodologique et épistémologique clarifiée dans TSOUKIÀS *et al.*, (2013), MEINARD et TSOUKIÀS, (2019) et MEINARD et CAILLOUX, (2020).

Nous commençons, en Section 4.1, par tenter d'établir ce qui peut être considéré comme une "information" dans un contexte médical. Cette réflexion nous permet de mettre en évidence que les recommandations d'actions ou de diagnostics ne sont qu'un type d'informations parmi d'autres et qu'il sera nécessaire d'évaluer quel type d'informations est le plus pertinent à fournir aux médecins pour le soutien de consultations médicales coutumières. En Section 4.2, nous étudions les spécificités des DDSS actuels par le prisme des travaux de MEINARD et TSOUKIÀS, (2019). Nous y mettons en évidence que le fait de fournir des recommandations de diagnostics dérive d'un choix fort dans la manière d'aider à la décision. Un choix qui, comme nous avons pu le constater, ne semble pas correspondre aux réalités d'une aide en situations de consultations coutumières. Nous proposons ainsi d'envisager l'approche "ajustive" de l'aide à la décision comme une alternative viable compte tenu des contraintes soulevées par notre cas applicatif. Enfin, en Section 4.3, nous détaillons les promesses d'une approche "ajustive" en montrant comment celle-ci peut permettre de dépasser les barrières bloquant actuellement l'usage des DDSS en pratiques coutumières. La principale contrainte de cette approche étant d'étudier finement le processus de décision à soutenir pour pouvoir s'y ajuster.

4.1 Définition de la notion "d'information" dans un contexte médical

Selon OSHEROFF *et al.*, (2012), l'objectif des CDSS est d'améliorer les décisions médicales, et donc la sûreté des soins apportés aux patient·es, en fournissant au personnel soignant la "bonne information, à la bonne personne, au bon moment du processus de décision". La définition de OSHEROFF *et al.*, (2012) s'est avérée efficace car elle fournit une formule synthétique apparaissant comme indiscutable. De plus, cette définition englobe convenablement l'immense diversité des CDSS. Cependant, le succès de cette formule réside, dans une large mesure, sur le caractère indéterminé de la phrase "la bonne information". Dans le cas des DDSS, cette "bonne information" est actuellement majoritairement incarnée par les *guidelines* et les recommandations de décision.

Dans cette section, nous explorons l'idée que la raison essentielle sous-jacente au manque d'acceptation des DDSS actuels par les médecins, en situations coutumières, est que cette "bonne information" n'est pas si adaptée après tout. Nous cherchons donc à identifier les différentes "informations" pouvant être fournies au personnel soignant dans leurs activités.

À première vue, nous pourrions penser que, dans notre contexte, la notion d'"information" est sans-équivoque : une information est une donnée brute formatée pour être lue par un·e médecin. L'interprétation de cette information par un·e médecin lui donne une connaissance sur la situation et lui permet de prendre une décision. Collectées par les EHR et les CPOE, les bases de données des hôpitaux sont riches de telles données brutes sur les patient·es. Pour ne citer que quelques exemples : poids, âges, symptômes, historiques des hospitalisations, traitements pris, allergies connues, etc. Cet ensemble de données brutes peut fournir aux médecins une première strate d'informations sur une situation donnée.

En suivant la même logique, l'évolution dans le temps de ces données brutes et leurs interconnexions dans le processus de prise en charge des patient·es, représente une seconde strate d'informations pour les médecins. La notion d'information apparaît donc comme finalement plus complexe que de premier abord, plusieurs strates d'informations pouvant exister.

Une troisième strate d'information peut être trouvée dans les *guidelines*. Comme mentionné en Chapitre 2, les *guidelines* sont généralement une liste d'actions préconisées par une autorité de santé pour une situation spécifique. Les *guidelines* peuvent être représentées sous différents formats : arbres de décisions, organigramme, listes d'éléments à vérifier, instructions sur comment interpréter certains résultats, etc. (FIELD, LOHR *et al.*, 1990). Cependant, en admettant qu'une "information" puisse référer à ce troisième type d'entité, nous pourrions admettre que les modèles permettant d'interpréter des données, telles que les théories ou les ensembles de savoir-faire, constituent aussi dans un certain sens un ensemble d'informations.

Nous pouvons voir que, dans des contextes médicaux, le terme d'information peut référer à une large diversité d'entités, allant des données brutes aux modèles interprétatifs, en passant par certaines représentations des données et de leurs évolutions.

Cette analyse de la notion "d'information" en contextes médicaux montre que l'approche actuelle des DDSS, consistant principalement à fournir des recommandations, constitue un type particulier d'approche pour l'aide à la décision, ancrée dans une compréhension très spécifique de la notion d'information. Ce type d'approche reflète une volonté de standardisation des pratiques, basée sur l'idée présupposée qu'une

telle standardisation amènera à une diminution des erreurs médicales. Cependant, comme mentionné par l'étude de WOOLF, (1993) concernant l'impact des *guidelines* sur les soins apportés, une telle standardisation peut être préjudiciable pour la santé des patient·es et peut interférer avec l'individualisation des soins apportés.

Cela suggère que, plutôt que de s'accrocher à la vision réduite de "l'information" visant à la standardisation des pratiques, nous devrions aspirer à l'identification du type d'information correspondant aux contraintes d'une aide en situations de consultations médicales coutumières, du type d'information dont les médecins ont besoin lorsqu'ils procèdent à un diagnostic.

4.2 Identification d'une approche adaptée aux contraintes du processus d'aide à la décision

Nous soutenons que, dans le but d'identifier ce qui peut constituer la "bonne information" à fournir aux médecins, nous avons besoin d'analyser, à un niveau méthodologique et épistémologique, la véritable signification de l'activité consistant à fournir une aide à la décision en situations de consultations médicales coutumières. Notre analyse a, jusqu'à présent, mis en évidence les nombreuses spécificités associées à ce contexte. Comme dans de nombreux contextes médicaux, le contexte spécifique que nous traitons soulève de nombreux problèmes de responsabilité en termes de prise en charge des patient·es. Cependant, une autre spécificité assez marquée dans notre cas est que les médecins sont considérés comme assez compétents pour ne pas être aisément surpassés par des outils sophistiqués. De plus, les médecins montrent aussi une forte volonté de garder le contrôle du processus de leurs consultations, ainsi que garder la responsabilité des soins apportés aux patient·es.

Ces spécificités reflètent des contraintes qui verrouillent les interactions entre les fournisseur·ses d'aide à la décision, incarnés dans notre cas par des outils d'aide à la décision, et les médecin·es, qui sont les preneur·ses de décision bénéficiant d'une aide à la décision. Les fournisseur·ses d'aide à la décision concernés par le fait de fournir une aide pertinente et acceptable n'ont pas d'autres choix que prendre en compte ces contraintes pour choisir le type d'approche permettant de déverrouiller leurs interactions avec les preneur·ses de décision, ici des médecins.

MEINARD et TSOUKIÀS, (2019) ont montré le rôle fondamental d'une analyse des contraintes verrouillant l'aide à la décision, qui est essentielle dans le choix d'une approche pertinente de l'aide à la décision et qui joue ensuite un rôle décisif pour

l'enracinement de la validité et la légitimité de l'aide à la décision fourni. Ce modèle permet un éclairage utile sur notre analyse des nombreux inconvénients des différents DDSS actuels.

Comme détaillé en Chapitre 2, les *Guideline-based DDSS* englobent les systèmes fournissant des recommandations basées sur des "guides de bonnes pratiques" ou sur l'expertise d'un ensemble de spécialistes, mais aussi les systèmes fournissant directement ces "guides de bonnes pratiques". Cette approche se montre pertinente lorsque les interactions entre les médecins et les outils d'aide à la décision sont contraintes par un besoin d'homogénéiser les processus de décisions et de les faire converger vers des "bonnes pratiques" collectivement reconnues par des institutions expertes faisant autorité. De telles interactions correspondent à ce que MEINARD et TSOUKIÀS, (2019) nomment comme des situations pour lesquelles "un schéma de gouvernance irrévocable" contraint le processus d'aide à la décision. Toujours selon MEINARD et TSOUKIÀS, (2019), dans de telles situations les fournisseur·ses d'aide à la décision doivent s'appuyer sur une approche "conformiste", visant à identifier les outils qui seront les plus acceptables pour les décideur·ses du schéma de gouvernance. Dans les situations où les compétences des médecins sont insuffisantes, dus à des problèmes de formation médicale comme nous avons pu par exemple l'observer en Chapitre 2 pour les travaux destinés aux pays en développement, et pour lesquels les institutions de santé jouent un rôle crucial dans un schéma de gouvernance puissant visant à renforcer la qualité et la sûreté des soins apportés aux patient·es, les *Guideline-based DDSS* sont pertinents.

Les *ML-based DDSS*, quant à eux, englobent les systèmes de recommandations basés sur des algorithmes de ML dit "supervisés". Ce type de DDSS présuppose qu'il existe une fonction liant les données de patient·es à un ensemble de classes spécifiques (ex. maladies, traitements, degrés de risque, etc.), et ce indépendamment des croyances et connaissances des preneur·ses de décision ou de leur contexte de décision. Il présuppose aussi que cette fonction peut être approximée via des algorithmes de ML, en particulier par des réseaux de neurones artificiels multi-couches qui sont connus pour être des approximateurs universels (HORNİK *et al.*, 1989), bien que d'autres algorithmes de ML peuvent être utilisés. L'objectif est alors de trouver la meilleure approximation possible de cette fonction, généralement évaluée sur la précision des classifications lors de phases de test (voir Chapitre 9). Une approche d'aide à la décision basée sur de tels outils se montre pertinente dans des situations où l'objectivité et la véracité sous-jacentes aux théories et aux algorithmes peuvent être considérées comme acquises et indiscutables. Pour ce type de situations, MEINARD et TSOUKIÀS, (2019) argumentent qu'il est plus pertinent d'avoir une approche "objectiviste" de l'aide à la décision. Dans des contextes médicaux, les situations

limitées par ces contraintes sont celles pour lesquelles le volume de données à collecter, la capacité des outils existants à les traiter et à surpasser les performances de toute autre forme d'expertises, sont clairement établis. C'est le cas, par exemple, pour la détection de maladies oculaires (ZHANG *et al.*, 2018), dans l'évaluation d'un risque d'infection (PEIFFER-SMADJA *et al.*, 2019) ou de mauvaise réaction à un traitement (DEIG *et al.*, 2019). Ce sont des tâches pour lesquelles nous pouvons supposer qu'une fonction de classification existe, mais ne pouvons supposer qu'il existe un·e décideur·se assez "compétent·e" pour l'approximer précisément.

Cependant, comme énoncé précédemment, le contexte d'aide à la décision auquel nous nous intéressons en priorité, les consultations médicales coutumières, n'est pas caractérisé par des contraintes similaires à celles présentées ci-dessus et pour lesquelles les actuels *Guideline-based DDSS* et *ML-based DDSS* apparaissent comme pertinents. Dans le cas de consultations coutumières, les médecins sont compétents : ils n'ont besoin ni d'être contrôlé par des autorités de santé sur leur conformité avec de "bonnes pratiques", ni d'être supplanté par des outils supposément plus performants. La contrainte principale de ce contexte, nous l'avons vu, est la volonté des médecins à rester en charge de leur processus de décision et à rester responsable de la sûreté des soins apportés aux patient·es. Ceci fait écho à ce que MEINARD et TSOUKIÀS, (2019) définissent comme des situations où la contrainte principale est le respect de "l'esprit d'initiative sanctifié" des décideur·ses. Nous pourrions argumenter que cette définition est vague, mais le cas des médecins effectuant des consultations coutumières peut être considéré comme un exemple permettant de la clarifier. Les deux caractéristiques distinctes des consultations coutumières, les compétences des médecins et leurs responsabilités, indique un besoin d'outils d'aide à la décision laissant les décideur·ses prendre leurs propres décisions et d'en être responsable. Le rôle d'un outil d'aide à la décision dans ce type de situations est de s'efforcer à faciliter, à alléger et à accélérer le processus privilégié par les décideur·ses et de s'ajuster à leurs besoins. Cette approche de l'aide à la décision est définie, par MEINARD et TSOUKIÀS, (2019), comme une approche "ajustive".

4.3 Les promesses et les limites de l'approche "ajustive"

Comme pour les approches "conformiste" et "objectiviste", l'approche "ajustive" de l'aide à la décision comporte des forces mais aussi des faiblesses et elle n'est adaptée qu'à certaines situations. Dans cette section nous cherchons à montrer, en reprenant

les barrières actuelles à l'utilisation des DDSS, en quoi une approche "ajustive" nous apparaît comme viable pour aider à la décision en consultations médicales coutumières.

Un premier avantage d'une approche "ajustive" est que celle-ci repose sur les compétences et le savoir-faire des médecins et l'expérience qu'ils ont accumulé. Nous avons vu que les approches actuellement utilisées pour soutenir les médecins supposent que les médecins manquent de compétences pour effectuer certaines tâches. Bien que de telles approches peuvent se révéler pertinentes dans des situations complexes, elles semblent inappropriées pour des situations coutumières et peuvent générer, pour les médecins, le sentiment d'être mis en arrière-plan de leurs propres processus de décisions et donc de la méfiance envers les outils d'aide à la décision qui leur sont proposés. Avec une approche "ajustive", les médecins gardent le contrôle de leurs processus de décision. De plus, avec une aide à la décision s'ajustant aux besoins des médecins, ceux-ci peuvent mettre en avant leurs compétences dans l'établissement de diagnostics et en interactions sociales.

Un second avantage important de l'approche que nous proposons d'expérimenter est que, comparée aux autres approches, celle-ci ne comporte que peu de risque de perte de savoir-faire pour les médecins. Comme introduit précédemment, l'usage des DDSS actuels peut tendre à la diminution des performances des médecins présentant de bonnes capacités de diagnostic (POVAKALO *et al.*, 2013). De même, nous avons vu que de mauvaises recommandations de diagnostics fournies par un DDSS ne sont pas nécessairement détectées par les médecins, en particulier les moins expérimentés (TSAI *et al.*, 2003). Ces craintes de perte de savoir-faire et de non détection des erreurs apparaissaient alors comme des barrières à l'utilisation des DDSS. Avec une approche "ajustive", les médecins continuent d'accumuler de l'expérience car ils ne peuvent pas se reposer sur un outil leur suggérant ce qu'ils sont censé faire. Ainsi, les risques d'impact négatif sur les performances des médecins sont minimisés.

Par corollaire, un troisième avantage à l'emploi d'une approche "ajustive" est qu'elle ne porte pas atteinte à la responsabilité des médecins dans la prise en charge des patient·es. En effet, les problèmes de responsabilités soulevés par les DDSS actuels, décrits en section 3.2, ne sont plus à considérer si nous laissons les médecins seuls responsables du déroulement de leur processus de décision et de la prise en charge des patient·es. Comme mentionné précédemment, une approche "ajustive" n'a pas pour vocation de suggérer aux médecins ce qu'ils peuvent ou sont censé faire, mais cherche plutôt à répondre aux besoins des médecins pour que ceux-ci puissent faire ce qu'ils sont censé faire. Ainsi, la responsabilité des consultations et de leurs conséquences est entièrement de la responsabilité des médecins.

Enfin, un dernier avantage de l'approche proposée est qu'elle n'est pas censée augmenter la charge de travail des médecins mais, au contraire, pourrait la diminuer. Nous avons vu précédemment que la charge de travail des médecins peut représenter une barrière à l'acceptation de nouveaux outils, en particulier si l'utilisation d'un DDSS nécessite un travail supplémentaire pour les médecins. Les DDSS actuels tendent à dire aux médecins comment ils sont censés travailler, sans prendre en compte leurs processus de décisions courants, ni l'impact du format de l'aide à décision employé. Avec une approche "ajustive", l'accent est mis sur la compréhension des processus de décision actuels des médecins. Cela afin d'établir sur quels points de leur processus de travail les médecins ont besoin en priorité d'un soutien informatique et quel type de système serait le plus pertinent de proposer.

Ce dernier point peut aussi être perçu comme une limite de l'approche "ajustive", car elle nécessite l'analyse et la compréhension du processus décision à soutenir via un outil d'aide à la décision. C'est pourtant ce à quoi nous allons nous atteler dans la prochaine partie de cette thèse afin de proposer un outil d'aide à la décision adapté aux contraintes d'une aide à la décision en consultations médicales coutumières et aux besoins des médecins des HCL dans de telles situations.

Partie II

Comprendre le processus de décision et
les besoins des médecins en consultations
médicales

Présentation du système informatique des HCL

” *For relaxation he could always engage Hal in a large number of semimathematical games, including checkers, chess, and polyominoes. If Hal went all out, he could win any one of them ; but that would be bad for morale. So he had been programmed to win only fifty percent of the time, and his human partners pretended not to know this.*

— **Arthur C. Clarke**
2001 : A Space Odyssey

Dans ce chapitre, nous présentons en détail l'outil informatique utilisé au quotidien par le personnel soignant des HCL : Easily[®]. Notre objectif est de cerner comment Easily[®] a été développé et quelles sont les fonctionnalités offertes au personnel soignant, en particulier aux médecins lors de leurs consultations. Pour ce faire, nous nous basons principalement sur la documentation interne des HCL et sur les communications faites par le GIE Hopsis concernant Easily[®]. Ce chapitre est aussi l'occasion d'ancrer dans un point de vue concret les motivations sous-jacentes au travail présenté dans cette thèse, visant à proposer de nouveaux outils aux médecins. Cela nous permettra, par la suite, de mieux analyser et comprendre l'usage d'Easily[®] par les médecins dans leurs processus de décision en consultations médicales.

En Section 5.1, nous nous attardons sur l'historique d'Easily[®], de sa genèse jusqu'à son utilisation actuelle dans de nombreux hôpitaux de France. En Section 5.2, nous détaillons l'architecture modulaire d'Easily[®], son système de portails métiers ainsi que les différentes couches de données accessibles via cet outil. Enfin, en Section 5.3, nous détaillons les fonctionnalités proposées aux médecins pour leurs consultations médicales.

5.1 Historique du système d'information hospitalier

Comme décrit en Figure 5.1, issue de la documentation interne des HCL et synthétisant les moments clés du développement d'Easily[®], la genèse d'Easily[®] commence en 2008, année où les HCL font face à un déficit de près de 90 millions d'Euros. En réponse à ce déficit budgétaire, les HCL et leur DSII lancent le projet "Cap2013" ayant pour objectif un retour à un équilibre budgétaire pour 2013. Cet équilibre sera finalement atteint en 2016. C'est dans ce contexte que le HIS Easily[®] fut pensé et développé, avec pour objectif l'optimisation des processus opérationnels des HCL, en offrant au personnel soignant un outil efficace et simple d'utilisation.

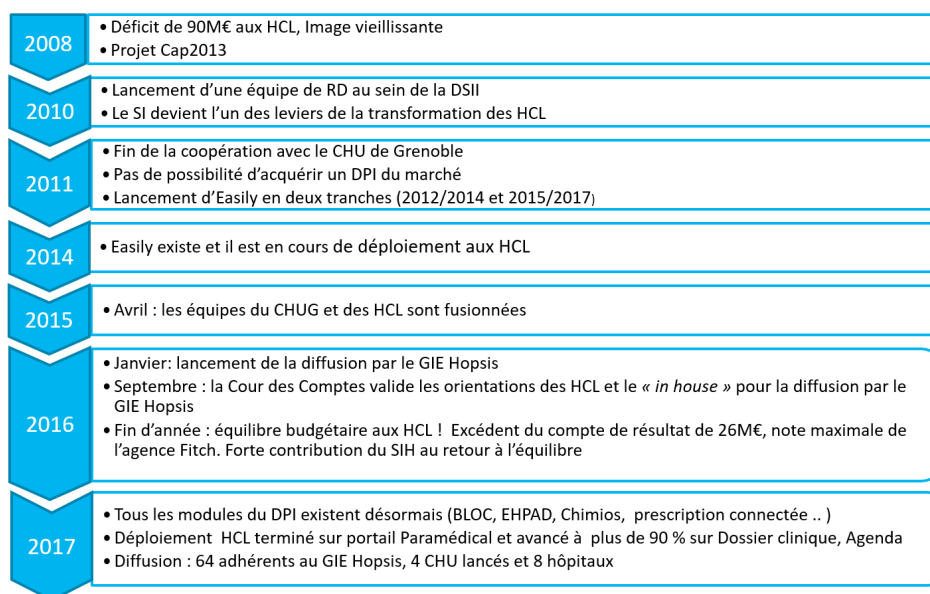


Figure 5.1.: Synthèse des moments clés de la genèse d'Easily[®]

En 2011, les HCL mettent fin à leur collaboration avec le Centre Hospitalier Universitaire (CHU) de Grenoble autour du développement du HIS Cristal-Net, alors utilisé depuis une dizaine d'années par le personnel des HCL. Riche de son expérience sur Cristal-Net, la DSII décide de développer son propre HIS : Easily[®], qui commencera à être déployé aux HCL courant 2014. De leur côté, les équipes informatiques du CHU de Grenoble continuent de travailler sur Cristal-Net, renommé Cristal-Link, avant de venir renforcer les équipes travaillant sur Easily[®] suite à une fusion des départements informatiques des deux groupes hospitaliers en avril 2015.

À partir de 2017, toutes les fonctionnalités essentielles d'Easily® sont opérationnelles et déployées au sein des HCL. Mais Easily® est aussi déployé dans 4 CHU et 8 Hôpitaux non HCL courant 2017, cela grâce à la diffusion d'Easily® par le GIE Hopsis depuis 2016. Cette diffusion d'Easily® continue aujourd'hui et, comme nous pouvons le voir en Figure 5.2, près d'une trentaine de GHT utilisent Easily® en 2019, soit plus d'une centaine d'établissements de santé dans toute la France.

Depuis 2017, la DSII se consacre à la maintenance des différentes fonctionnalités d'Easily®, que nous détaillerons en section 5.2, mais aussi à la création de nouveaux outils pour continuer d'améliorer le confort de travail du personnel soignant et la prise en charge des patient·es. Par exemple, les outils "O2" ont été récemment développés afin de faciliter la recherche clinique via la recherche de cohorte de patient·es et une extraction des données plus aisée.

C'est donc dans cette optique d'amélioration des fonctionnalités offertes aux médecins que s'ancre le travail présenté dans cette thèse. Ces nouvelles fonctionnalités seront, dans un premier temps, proposées et évaluées au sein des HCL mais, comme suggéré précédemment, celles-ci ont pour vocation d'être proposées à tout GHT utilisant Easily® si elles montrent de bons résultats.

5.2 Un ensemble d'outils pour le personnel soignant

Nous nous concentrons maintenant sur l'architecture générale d'Easily® et les outils proposés au personnel soignant. Afin de faciliter la maintenance et les futures évolutions d'Easily®, les développeurs de la DSII ont décidé de baser Easily® sur une architecture modulaire.

Comme schématisé en Figure 5.3, l'architecture d'Easily® se base sur un socle technique centralisant les fonctionnalités essentielles d'Easily®, telles que l'accès et l'ajout de données relatives aux patient·es et aux agendas des médecins. Ce socle sert alors de base au développement de modules/applications répondants aux différents besoins du personnel soignant : gestion des blocs et des lits, gestion des agendas, gestion des prescriptions, traçabilité des actions médicales, etc. Le socle Easily®, ainsi que les modules, sont développés en C#/.NET comme des applications web. Il est aussi possible d'intégrer des applications externes à Easily® à l'aide de modules servant d'interface entre le socle Easily® et ces applications externes. Cela est particulièrement utile lors de la mise en place d'Easily® dans des hôpitaux autres que ceux des HCL, ceux-ci pouvant vouloir garder certaines de leurs applications historiques.

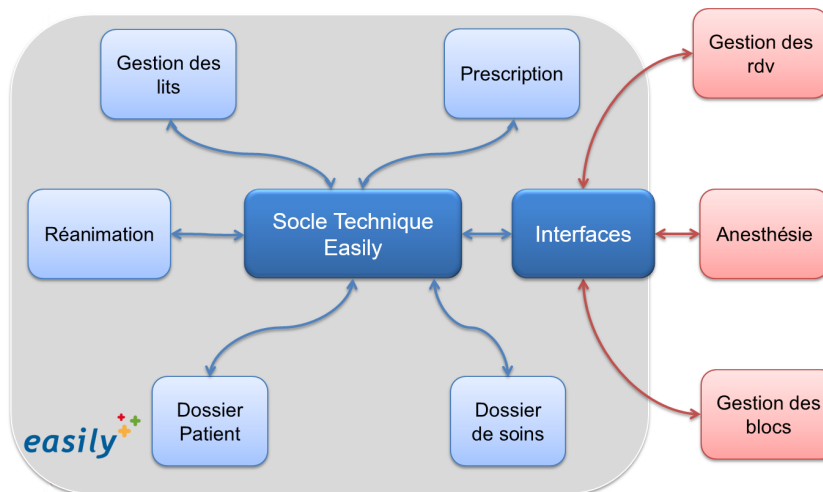


Figure 5.3.: Schématisation du fonctionnement du socle et des modules Easily®

De cette manière, les modules Easily® peuvent fonctionner et évoluer indépendamment. De même, l'ajout de nouvelles fonctionnalités, via de nouveaux modules, se voit grandement simplifié. Cependant, certains modules peuvent avoir besoin de communiquer entre eux ou d'être tenus au courant de l'activité d'autres modules. Pour ce faire, comme schématisé en Figure 5.4, Easily® utilise une passerelle permettant aux modules d'émettre et de recevoir des évènements relatifs à l'activité d'autres modules.

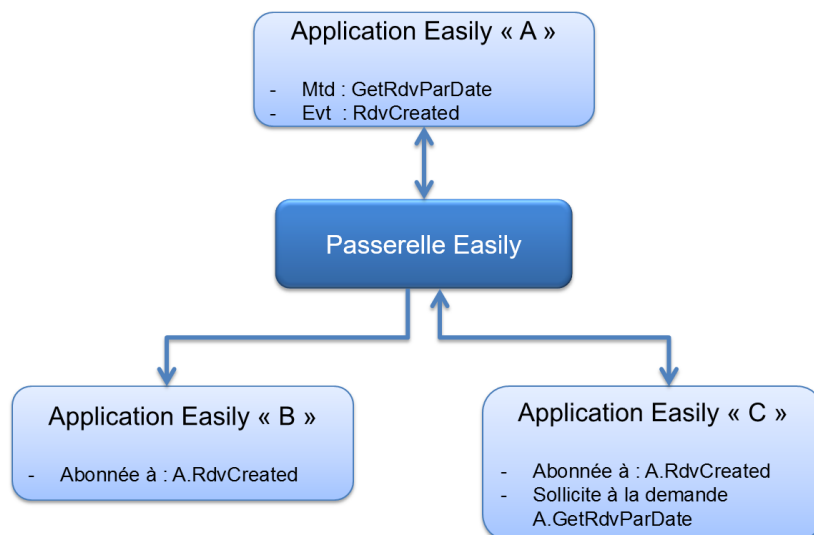


Figure 5.4.: Schématisation du fonctionnement de la passerelle Easily®

Concernant le fonctionnement du socle Easily®, celui-ci est divisé en 5 couches noyaux, schématisés en Figure 5.5. Le "cœur" noyau permet la gestion des données

essentielles propres aux établissements hospitaliers (structure hospitalière, intervenants, etc.). La couche noyau "patient" est dédiée à la gestion des données relatives aux patient·es (identités, coordonnées, venues, données biologiques, etc.). La couche noyau "Structure Hôtelière" est dédiée à la gestion des chambres, des lits, des boxes, etc. La couche noyau "Référentiel" permet la gestion des références utilisées par les différents modules/applications Easily[®]. Enfin, la couche noyau "Outils" est dédiée à la gestion des habilitations pour l'accès à certains modules et à la traçabilité des outils utilisés par le personnel soignant.

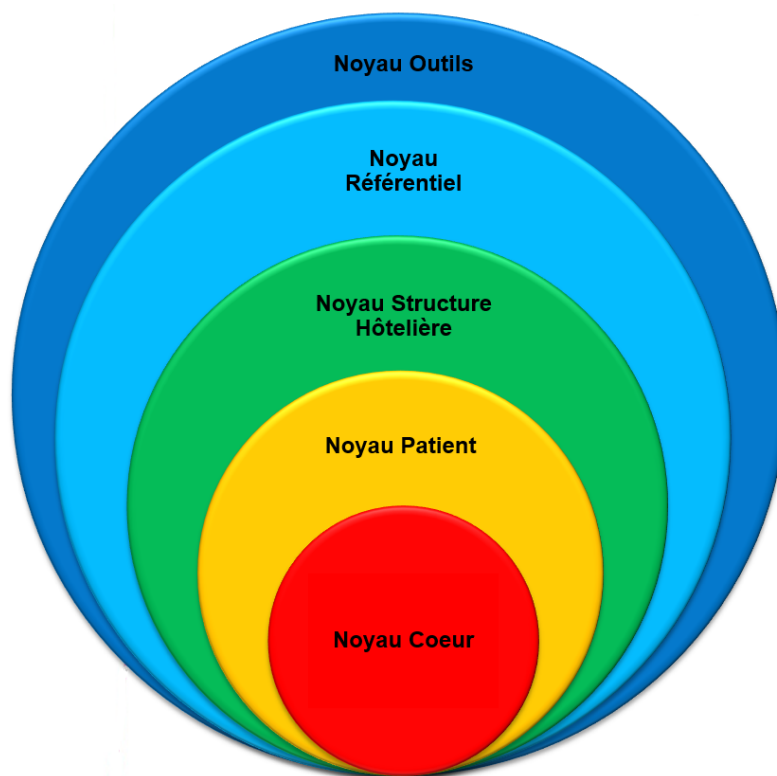


Figure 5.5.: Schématisation des différentes couches du noyau Easily[®]

Côté utilisateur·ices, Easily[®] se présente comme une application web, codée en HTML5 et JavaScript, avec des accès à différents portails "métiers". Chaque portail regroupe un ensemble de modules dédiés à chaque corps de métier d'un hôpital. Actuellement, Easily[®] propose, entre autres, les portails : Infirmier, Sage-Femme, animateur, Rééducateur, Nutrition, Hygiéniste, Secrétaire médical, Direction, Pharmacien, Parapharmacien, Biologie et Médecin.

5.3 Les fonctionnalités dédiées aux consultations médicales

Nous allons maintenant nous concentrer sur le portail Médecin d'Easily[®], et plus particulièrement sur les outils mis à la disposition des médecins pour leurs consultations. Outre l'agenda de rendez-vous pour consultations, permettant aux médecins d'avoir accès à la liste des patient·es attendus pour une journée donnée et permettant d'accéder directement au dossier de chacun·e de ces patient·es, l'outil principalement utilisé par les médecins lors de leurs consultations est "CapMedecin". Comme nous pouvons le voir en Figure 5.6, montrant un exemple de l'interface de CapMedecin pour une patiente fictive, cet outil sert d'accès à diverses fonctionnalités.

L'écran principal de CapMedecin est divisé en deux colonnes. Celle de gauche présente, par défaut, la liste des documents concernant la·e patient·e en cours de consultation. Ces documents sont triés du plus récent au plus ancien et sont regroupés en fonction du service médical associé. D'autres données sont accessibles sur cette colonne, via les différents onglets :

- Dossier de Spécialité : Accès rapide à certaines données du/de la patient·e par spécialité médicale ;
- Histoire : onglet par défaut, liste chronologique des documents médicaux liés au/à la patient·e ;
- Activité : Historique des actes médicaux du/de la patient·e ;
- Visionneuse doc : Affichage des documents médicaux du/de la patient·e, triés du plus récent au moins récent.

La colonne de droite est, quant à elle, dédiée au visionnage de résultats des examens médicaux du/de la patient·e, via différents onglets. Par défaut, la·e médecin a accès aux onglets suivants :

- Examens : Accès aux comptes-rendus des examens cliniques du/de la patient·e ;
- Visionneuse CR : Accès aux comptes-rendus des examens biologiques du/de la patient·e ;
- Biologie : Accès aux données biologiques du/de la patient·e, issues d'examens en laboratoire ;
- Anapath : Accès aux données anatomopathologiques du/de la patient·e ;
- Imagerie : Accès aux données d'imagerie médicale du/de la patient·e ;
- Photo : Accès aux photographies issues du dossier du/de la patient·e.

De plus, la colonne de droite permet d'afficher certaines données en fonction de l'onglet sélectionné dans la colonne de gauche. Par exemple, si l'onglet Histoire est sélectionné, la colonne de droite permet d'afficher un ou plusieurs documents sélectionnés par la·e médecin dans l'onglet Histoire.

À ces fonctionnalités principales, s'ajoutent des accès à divers sous-modules, dont nous pouvons voir une partie en Figure 5.7. Chacun de ces sous-modules permet aux médecins d'accéder à diverses informations concernant la·e patient·e en cours, ainsi qu'à diverses fonctionnalités utiles à l'activité d'une consultation. Tout d'abord, les médecins ont accès à l'ensemble des parcours médicaux dans lesquels la·e patient·e est inclus (cf. Figure 5.6, ligne du bas). En haut à droite de CapMedecin, les médecins ont accès divers sous-modules tels que :

- Post-it : Création d'une note concernant la·e patient·e ;
- Agenda : Accès à l'agenda du/de la patient·e ;
- Antécédents : Accès et saisie des antécédents médicaux du/de la patient·e (cf. Figure 5.7b) ;
- Rech. Clin. : Ajout du/de la patient·e à un essai clinique et accès aux essais cliniques auxquels la·e patient·e est affecté ;
- MedPhone : Accès au QR code du/de la patient·e (utile pour certaines applications mobiles liées à Easily[®]) ;
- Fiches de liaison : Accès aux messages liés à l'un des parcours du/de la patient·e ;
- Obligations Légales : Accès et production de documents administratifs ou législatifs (ex. consentement pour un examen) ;
- Motif et Diagnostic : Accès à l'historique des motifs de venues et des diagnostics du/de la patient·e ;
- DPC : Accès et saisie de Données Patient Communes (DPC), c'est-à-dire des données communes à différents services concernant la·e patient·e telles que sa taille ou son poids (cf. Figure 5.7a).

Données Patient communes

Poids	40 kg, le 19/05/2017 17:15 (poids mesuré)	
Taille	162 cm, le 28/06/2016 13:38 (taille mesurée)	
IMC	15,24 kg/m ² , le 19/05/2017 17:15 [Dénutrition]	
Surface corporelle	1,38 m ² , le 19/05/2017 17:15 (méthode Dubois)	
Grossesse	Non renseigné	Allaitement Non renseigné
Procréation	Non renseigné	

Modifier

Filtre ☒ Toutes les DPCs ☐ Favoris **Ajouter une DPC**

Libellé	Valeur	Auteur	Date
★ Evaluation de la douleur (0 à 10)	2	HCLIDE1, H	28/11/2017 08:00
★ Fréquence cardiaque	77	HCLIDE1, H	28/11/2017 16:00
★ Possibilité de procréation	Non renseigné	HCLMED1, H	28/06/2016 13:37
★ Température (°C)	38,8	HCLIDE1, H	28/11/2017 16:00
★ Tension artérielle maximale (mmHg)	120	HCLIDE1, H	28/11/2017 15:00
★ Tension artérielle minimale (mmHg)	80	HCLIDE1, H	28/11/2017 15:00
★ Trouble de la déglutition	Non	HCLMED1, H	15/09/2017 18:20

(a) Module DPC

Antécédents - GRANGER Hermione

Synthese

Données patient communes
Maj : 19/05/2017

40 Kg mesuré

162 cm mesuré

15,24 Kg/m²
Dénutrition

1,38 m²
Méthode Dubois

Allergies médicamenteuses

Libellé	Date
H2421 - HYPERSENSIBILITE URTICA URENS	02/08/2018

Transfusions

Libellé	Date
---------	------

Antécédents importants

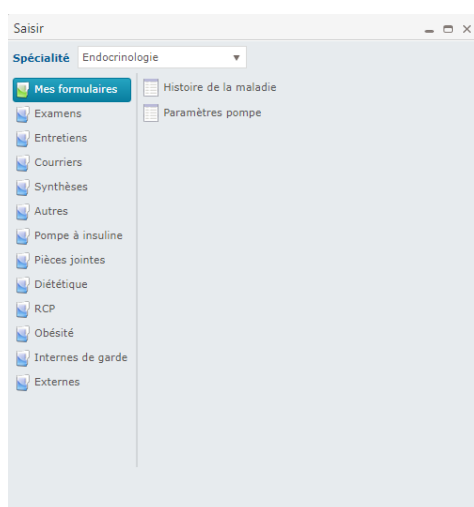
Type	Libellé	Date
MED	Y49.6 - Eff. indesir. psychodysléptiques (us. ther.)	02/08/2018

(b) Module Antécédents

Figure 5.7.: Exemples de sous-modules dédiés à l'accès aux données patients

Les médecins ont ainsi accès à toutes les informations concernant leurs patient·es, en quelques clics et à tout moment d'une consultation. Enfin, viennent les fonctionnalités liées à la génération de documents médicaux (cf. Figure 5.6, onglets à gauche de l'écran). Ceux-ci permettent, entre autres :

- La saisie de documents : comptes-rendus, observations, mots de suivi, etc. (cf. Figure 5.8) ;
- La dictée de divers documents, saisie ensuite par les secrétaires médicales ;
- L'enregistrement d'actes médicaux (Onglet PMSI) ;
- La création de parcours médicaux et/ou l'affiliation d'un·e patient·e à l'un de ces parcours ;
- L'accès à la prescription informatisée ou ePrescription (cf. Figure 5.9) ;
- L'accès et la saisie d'informations concernant l'hospitalisation du/de la patient·e ;
- Le téléchargement de documents ;
- La transmission d'informations sur la·e patient·e à d'autres services.



(a) Liste des documents pouvant être saisis

(b) Exemple d'un formulaire de saisie

Figure 5.8.: Exemples de sous-modules dédiés à la saisie de documents

Données Patient Communes

Poids : 40 kg (il y a 3 ans)

Taille : 162 cm (il y a 4 ans)

IMC : 15,24 kg/m² (il y a 3 ans)

SC : 1,38 m²

La venue du patient est fermée. Vous ne pouvez qu'arrêter la prescription.

La prescription sera initiée sur l'UF médicale : MEDECINE INTERNE PAV 03

Tout	Soins et Actes médicaux	Médicaments		
MEDICAMENTS				
	Prescription	Posologie	Durée	Alertes
☐	Paracétamol 500 mg gélule - voie ORALE	500 mg Matin, Soir soit 1000 mg/24h	pendant tout le séjour à partir du 06/06/2019 15:00	
Arrêté à partir du 26/02/2020 à 15:38				
☐	Aérosol - voie INHALEE Terbutaline sulfate 5 mg/2 ml solution pour inhalation	Matin, Midi, Soir à passer sur 15min 5 mg	pendant tout le séjour à partir du 06/06/2019 15:00	
Arrêté à partir du 26/02/2020 à 15:38				
☐	Salbutamol 1,25 mg/2,5 ml solution pour inhalation nstp - voie INHALEE	1,25 mg Soir	pendant tout le séjour à partir du 14/06/2016 10:00	
Arrêté à partir du 26/02/2020 à 15:38				
☐	Mométasone furoate 50 microgrammes/dose suspension pour pulvérisation nasale - voie NASALE	50 µg Soir	pendant tout le séjour à partir du 14/06/2016 10:00	
Arrêté à partir du 26/02/2020 à 15:38				
☐	Saccharomyces boulardii 50 mg gélule - voie ORALE	100 mg Soir	pendant tout le séjour à partir du 14/06/2016 10:00	
Arrêté à partir du 26/02/2020 à 15:38				
☐	Lopéramide chlorhydrate 2 mg gélule - voie ORALE	2 mg Soir	pendant tout le séjour à partir du 14/06/2016 10:00	
Arrêté à partir du 26/02/2020 à 15:38				
☐	Azithromycine 250 mg comprimé - voie ORALE	250 mg Soir	pendant tout le séjour à partir du 14/06/2016 10:00	
Arrêté à partir du 26/02/2020 à 15:38				
☐	Racécadotril 100 mg gélule - voie ORALE	100 mg Soir	pendant tout le séjour à partir du 14/06/2016 10:00	
Arrêté à partir du 26/02/2020 à 15:38				

Arrêter tout

Ordonnance arrêt de Pancarte 26/02/2020

Synthèse des alertes

Biologie

Ciclosporine 0,05 % collyre undecasil® HCL 1 récipient undose Mldi

pendant tout le séjour

Etat de la prescription

Arrêtée le 26/02/2020 15:38 par Docteur hclmed1

UF médicale 21124 MEDECINE INTERNE PAV 03

UF d'hébergement 21124 MEDECINE INTERNE PAV 03

Arrêt immédiat

Figure 5.9.: Interface de l'application ePrescription pour les prescriptions informatisées

Observations et analyses du processus de travail des médecins en consultations médicales

” *Le docteur ouvrit la fenêtre et le bruit de la ville s'enfla d'un coup. D'un atelier voisin montait le sifflement bref et répété d'une scie mécanique. Rieux se secoua. Là était la certitude, dans le travail de tous les jours. Le reste tenait à des fils et à des mouvements insignifiants, on ne pouvait s'y arrêter. L'essentiel était de bien faire son métier.*

— **Albert Camus**
La Peste

Dans ce chapitre, nous approfondissons les observations et analyses du processus de décisions des médecins durant leurs consultations présentées dans RICHARD *et al.*, (2018). Ces observations et analyses ont pour objectif de donner une meilleure compréhension de notre objet d'étude : les consultations coutumières des médecins aux HCL. Nous cherchons ici à comprendre différents aspects du processus de décision des médecins en consultations :

- Quelles sont les différentes actions élémentaires qui composent le processus d'une consultation ?
- Quelles sont les interconnexions entre ces différentes actions ?
- Quelles actions sont les plus fréquentes lors d'une consultation ?
- Quelles actions prennent le plus de temps aux médecins durant une consultation ?

Du fait de l'hétérogénéité des spécialités médicales, nous avons décidé de nous concentrer, dans le cadre de cette thèse, sur les consultations du service d'endocrinologie des HCL. Les consultations de cette spécialité sont généralement axées sur le suivi de patient·es atteint·es de maladies chroniques, ce qui correspond à l'aspect "coutumier" que nous cherchons à analyser.

Principalement du fait de la charge de travail des médecins, nous avons décidé de nous limiter à quelques observations de terrain, présentées en Section 6.1. Ces observations nous ont permis d'avoir un aperçu des différents types d'actions pouvant être effectuées par un·e médecin durant une consultation. Cependant, ces quelques observations ne sont pas suffisantes pour nous permettre de comprendre le processus de décision d'une consultation. Nous avons donc décidé d'employer des méthodes de fouille de données, plus précisément de *process mining* [minage de processus], sur les données relatives à l'activité des médecins durant leurs consultations. La Section 6.2 présente les méthodes de *process mining* employées ainsi que leurs résultats. Enfin, la Section 6.3 propose une discussion synthétique autour des résultats obtenus et de leurs interprétations.

6.1 Observations de consultations

La première étape de notre analyse fut d'observer, en conditions réelles, les processus de décision de médecins lors de consultations. Comme énoncé précédemment, du fait de la charge de travail des médecins, il n'était pas envisageable de mettre en place des observations de large ampleur. Aussi, seules les consultations de deux médecins ont été observées. Bien qu'elles n'aient pas de portée quantitative, ces observations nous ont permis d'avoir un aperçu des actions pouvant former le processus d'une consultation, ce qui nous donnera une meilleure compréhension des données recueillies par la suite.

6.1.1 Protocole d'observation

Nous considérons ici qu'une observation correspond à une et une seule consultation. Le début d'une observation est situé au moment de l'arrivée du/de la patient·e dans la salle de consultation, et la fin située au moment de sa sortie de la salle de consultation. Les médecins enchainent généralement plusieurs consultations durant une matinée ou un après-midi. Une séance d'observation correspond donc à plusieurs observations successives sans interruption.

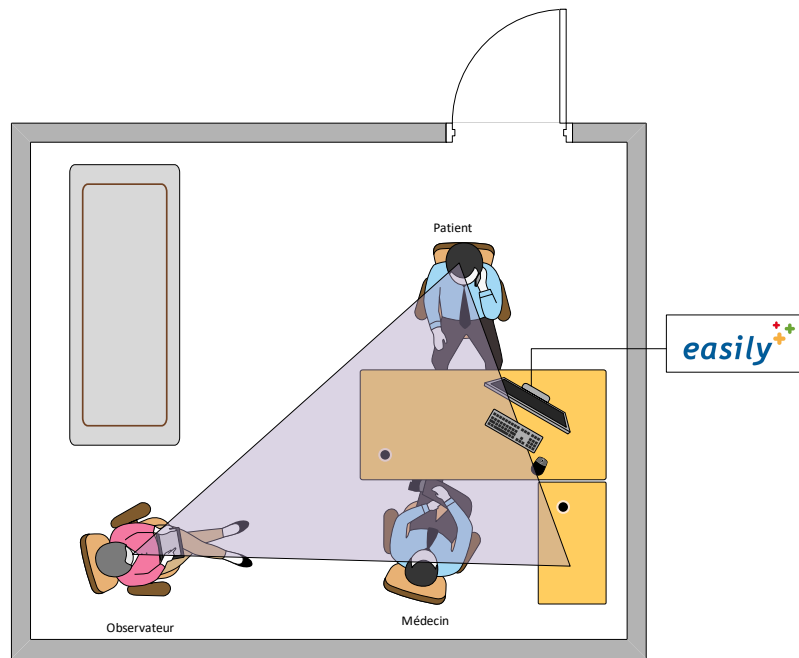


Figure 6.1.: Disposition des agents dans la salle de consultation durant la phase d'observation

Comme schématisé dans la Figure 6.1, l'observateur·ice se place derrière la·e médecin afin de ne pas être dans son champ visuel tout en pouvant observer ses interactions avec la·e patient·e et son utilisation d'Easily[®]. L'objectif est d'avoir un aperçu des interactions médecin-patient·e et médecin-Easily[®].

Les indicateurs suivants ont été collectés :

- Durée de la consultation (en minutes) ;
- Nombre d'actions effectuées par la·e médecin ;
- Temps d'utilisation d'Easily[®] par la·e médecin (en minutes).

Au début d'une observation, l'observateur·ice commence par noter le moment d'arrivée du/de la patient·e au format HH mm. Ensuite, durant l'observation, l'observateur·ice recense toutes les actions effectuées par la·e médecin, par exemple : consulter un document médical, prendre une mesure, prescrire un médicament, etc. Parallèlement, l'observateur·ice chronomètre le temps d'utilisation d'Easily[®] par la·e médecin, en considérant comme "utilisation d'Easily[®]" toute interaction avec le HIS même si la·e médecin interagit dans le même temps avec la·e patient·e, c'est-à-dire : recherche d'information, saisie d'information, saisie de prescription, etc. À la fin d'une observation, l'observateur·ice note le moment de départ du/de la patient·e, ainsi que le temps d'utilisation d'Easily[®] chronométré durant l'observation.

À noter que la·e médecin n'est pas au courant des données récoltées durant les observations.

6.1.2 Résultats

La Tableau 6.1 résume, pour chaque observation, le nombre d'informations recherchées par la·e médecin, le nombre de prescriptions effectuées, la durée de la consultation et le temps d'utilisation d'Easily® par la·e médecin.

Observation	Nombre d'informations recherchées	Nombre de prescriptions	Durée de la consultation	Temps d'utilisation d'Easily®
1	14	1	27'	10'
2	12	3	27'	13'
3	16	1	30'	13'
4	14	3	30'	16'
5	10	1	22'	12'
6	13	4	22'	13'
7	16	0	23'	6'
8	16	10	23'	11'
9	11	5	20'	10'
moyennes	13.6 ± 2.2	3.1 ± 3.1	24.9 ± 3.7	11.6 ± 2.8

(a) Médecin 1

Observation	Nombre d'informations recherchées	Nombre de prescriptions	Durée de la consultation	Temps d'utilisation d'Easily®
1	18	3	20'	10'
2	3	0	10'	7'
3	16	3	16'	8'
4	17	2	25'	8'
5	13	2	17'	9'
6	8	3	19'	9'
7	8	4	18'	12'
8	9	3	12'	7'
moyennes	11.5 ± 5.3	2.5 ± 1.2	17.1 ± 4.7	8.75 ± 1.7

(b) Médecin 2

Table 6.1.: Synthèse des résultats des observations de consultations médicales effectuées en conditions réelles au service d'endocrinologie des HCL

Ces observations nous ont permis de mieux comprendre certains points caractéristiques d'une consultation. Premièrement, nous avons pu recenser, deux types généraux d'actions possibles par les médecins durant une consultation :

- **Rechercher une information** : questions posées aux patient·es, informations recueillies lors de l'auscultation et informations dans Easily[®] (ce type d'action est généralement associé à la saisie de ladite information dans Easily[®]) ;
- **Faire une prescription** : prescriptions de médicaments, prescriptions d'analyses en laboratoire et recommandations de rendez-vous chez un·e autre médecin spécialiste.

Deuxièmement, l'action de rechercher une information concernant l'état du/de la patient·e apparaît comme bien plus fréquente que l'action de faire une prescription. Nous avons dans le premier cas une moyenne de 13.6 informations recherchées par consultation pour une moyenne de 3.1 prescriptions par consultation (cf. Tableau 6.1a). Dans le second cas (cf. Tableau 6.1b), nous observons une moyenne de 11.5 informations recherchées par consultation pour une moyenne de 2.5 prescriptions par consultation.

Troisièmement, nous pouvons observer que les médecins passent près de 50% de la durée de la consultation à utiliser leur HIS. Dans le premier cas (cf. Tableau 6.1a), nous observons une durée moyenne des consultations de 24.9 minutes avec une durée moyenne d'utilisation d'Easily[®] de 11.6 minutes. Et nous observons, dans le second cas (cf. Tableau 6.1b), une durée moyenne d'utilisation d'Easily[®] de 8.75 minutes pour une durée moyenne des consultations de 17.1 minutes.

Cependant, nous avons deux précisions à faire concernant ce troisième point :

1. Les médecins utilisent généralement Easily[®] tout en interagissant avec la·e patient·e et nous n'avions pas les moyens de faire la distinction entre temps d'utilisation avec et sans interactions médecin-patient·e ;
2. Bien que les médecins n'aient pas été tenus au courant des données récoltées, ils savaient que nous étions là pour observer leur utilisation d'Easily[®], ce qui a pu impacter leur processus de travail.

Enfin, il a été observé que les prescriptions étaient généralement rédigées à la fin des consultations, mais que les contenus de ces prescriptions étaient décidés au fur et à mesure de la consultation et énoncés de vive voix aux patient·es. Cette dernière observation fait écho aux résultats de l'étude menée par GIBSON *et al.*, (2006), qui ont montré le rôle central des "prescriptions verbales" dans les interactions médecin-patient·e.

6.2 Analyses informatiques

Comme énoncé précédemment, ces observations ne sont pas suffisantes pour comprendre et formaliser les processus de décision des médecins durant leurs consultations. Pour améliorer notre compréhension de ces processus, nous avons décidé de collecter les données relatives à l'activité de différents médecins durant leurs consultations afin de les analyser via des algorithmes de fouilles de données dédiées à l'extraction de processus, aussi nommés *process mining*.

Comme introduit et présenté par VAN DER AALST *et al.*, (2004), VAN DER AALST *et al.*, (2007), VAN DER AALST *et al.*, (2011) et VAN DER AALST, (2011), les algorithmes de *process mining* sont destinés à analyser les traces d'activités (ou *event logs*) récoltées par le système informatique d'une société, privée ou publique, afin d'extraire les processus sous-jacents au fonctionnement de cette société. Ainsi, le *process mining* forme un sous-domaine de la fouille de données, ou *data mining* (HAND, 2007), concentré sur l'extraction des processus. Précisons aussi que les données utilisées dans un contexte de *process mining* ne sont initialement pas destinées à être analysées, mais simplement à garder une trace de l'activité des utilisateur·ices pour diverses raisons.

Pour des exemples pratiques, ROJAS *et al.*, (2016) proposent une revue des utilisations d'algorithmes de *process mining* en contexte médical. Ceux-ci peuvent être utilisés pour comprendre l'utilisation d'un HIS au sein d'un hôpital (REBUGE *et al.*, 2013), découvrir les processus sous-jacents à l'activité d'un hôpital (MANS *et al.*, 2008), évaluer la conformité des activités d'un l'hôpital avec les directives en vigueur (KIRCHNER *et al.*, 2012) et améliorer les processus de travail courants via l'identification de goulots d'étranglement (BOSE et VAN DER AALST, 2011). D'après ROJAS *et al.*, (2016), les processus découverts par *process mining* peuvent aussi être employés afin d'aider au développement de nouveaux HIS, ce que nous cherchons à faire ici, en favorisant la compréhension de l'aide pouvant être apportée dans l'exécution d'un processus de travail.

Du fait des mises à jour régulières d'Easily[®], et donc des données enregistrées par le HIS, nous nous sommes concentrés sur l'analyse des données correspondant à un mois d'activité sans modification notable des types de données. Ainsi, nous avons récolté les données liées à l'activité de 72 médecins de 5 départements différents au sein des HCL : endocrinologie, dermatologie, urologie, gynécologie et rhumatologie. Près de 40.000 évènements ont été collectés, répartis dans près de 3.500 cas de consultations.

Les analyses présentées dans les sous-sections ci-dessous ont été effectuées à l'aide du langage R et la librairie bupaR (JANSSENSWILLEN, 2020). Le code source pour les reproduire est disponible sur le GitLab du LAMSADE ¹.

6.2.1 Analyse générale

Après une première étape de nettoyage des données générant du bruit et pouvant gêner nos analyses, nous obtenons 59 évènements distincts. En Figure 6.2, nous pouvons observer la présence, en pourcentage, de chacun de ces évènements dans les différentes traces d'activités collectées.

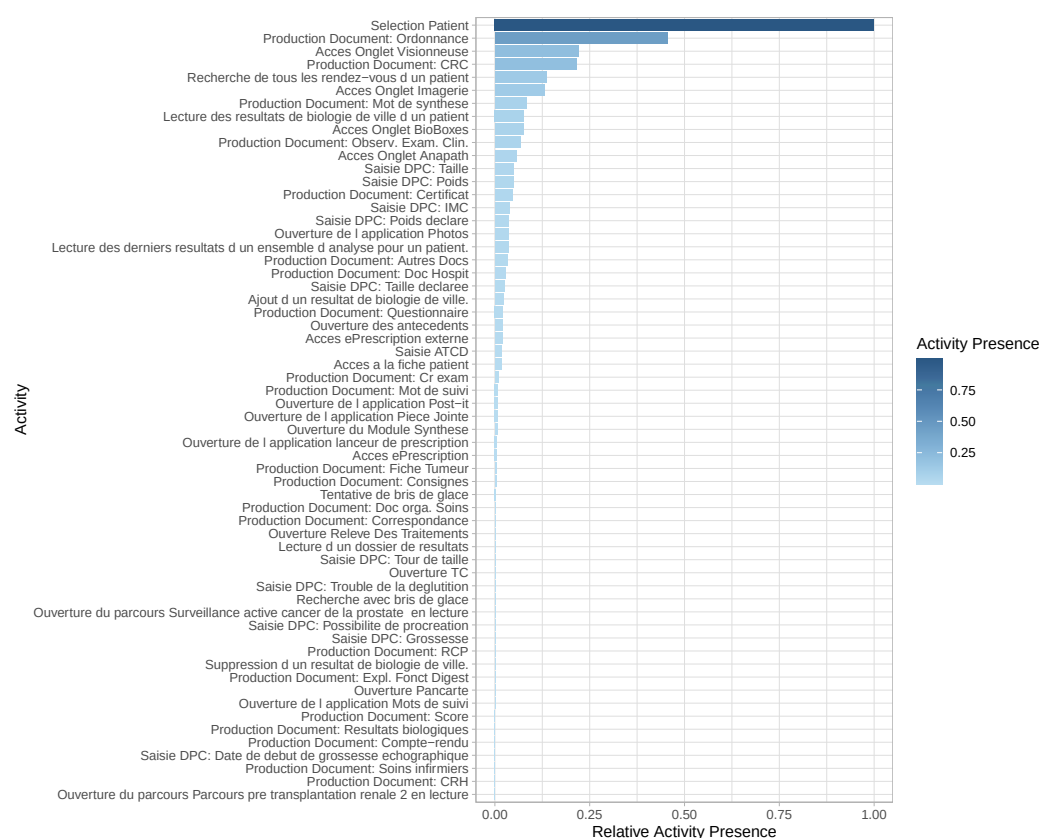


Figure 6.2.: Degrés de présence, en pourcentage, des différents évènements dans les différentes traces d'activités collectées

Nous retrouvons les deux types d'actions que nous avons pu observer en situation réelle : la recherche d'une information sur la·e patient·e en cours ("Accès Onglet Visionneuse", "Accès Onglet Anapath", "Lecture des résultats de biologie de ville d'un patient", etc.) et la production de prescriptions ("Production Document Ordonnance",

1. https://git.lamsade.fr/a_richard/consultation-process-analysis

"Accès ePrescription externe", "Recherche de tous les rendez-vous d'un patient", etc.). Nous pouvons aussi ajouter à la catégorie "rechercher une information" les événements de type "Saisie", qui correspondent à la saisie par la·e médecin d'une information collectée auprès du/de la patient·e. Enfin, nous observons aussi des événements liés à la production de documents autres que des ordonnances tels que : "CRC" (le compte-rendu de la consultation), "Mot de synthèse" (message envoyé au médecin traitant résumant la consultation), "Observ. Exam. Clin." (compte-rendu d'examen clinique), etc. Ces documents aident au suivi des patient·es car ils fournissent aux médecins un résumé des décisions prises lors de précédentes consultations. Ce dernier groupe d'événements est donc plus proche d'une recherche d'information sur la·e patient·e que d'une production d'ordonnance.

Concernant la présence de ces événements dans les différentes traces d'activités collectées, nous observons que l'événement "Sélection Patient" est présent dans 100% des traces. Cet événement, qui correspond à l'ouverture du dossier patient dans Easily[®], débute plus de 99% des traces collectées. Il s'agit aussi de l'événement terminant le plus des traces collectées, près de 38%, ce qui suggère que les médecins n'ont fait qu'ouvrir le dossier de leur patient·e pour ces consultations.

Viennent ensuite, pour les activités présentes dans plus de 12.5% des traces collectées :

- "Production Document : Ordonnance", présent dans presque la moitié des traces collectées ;
- "Accès Onglet Visionneuse", indiquant que la·e médecin a visionné au moins un document médical concernant la·e patient·e ;
- "Production Document : CRC", correspondant à la production d'un compte-rendu de la consultation ;
- "Recherche de tous les rendez-vous d'un patient", indiquant que la·e médecin a accédé à l'agenda du/de la patient·e, suggérant l'ajout d'un nouveau rendez-vous pour un suivi ;
- "Accès Onglet Imagerie", indiquant que la·e médecin a visionné au moins un résultat d'imagerie médicale concernant la·e patient·e.

La production d'ordonnances paraît ici plus fréquente que la recherche d'informations sur la·e patient·e en cours. Cependant, si nous combinons les activités liées à la recherche d'informations, telles que "Accès Onglet Visionneuse", "Accès Onglet Imagerie" ou "Accès Onglet BioBoxes", nous obtenons une fréquence relative similaire à celle de l'événement "Production Document : Ordonnance". De plus, les traces d'activités récoltées ne permettent pas de connaître avec précision quelles informations ont été regardées par les médecins dans ces onglets, ni en quelle quantité.

Nous pouvons donc supposer que la présence de ce type d'activités est bien plus importante qu'observée.

Nous pouvons aussi approximer la durée moyenne d'une consultation à partir de la durée séparant le premier évènement et le dernier évènement d'une même trace d'activité. Cependant, un certain nombre de traces présentent une durée de plus d'une heure, ce qui suggère que la·e médecin a accédé au dossier de sa/son patient·e plusieurs heures après le moment de la consultation. En Figure 6.3, est représentée la répartition des durées des traces d'activités inférieures à 60 minutes.

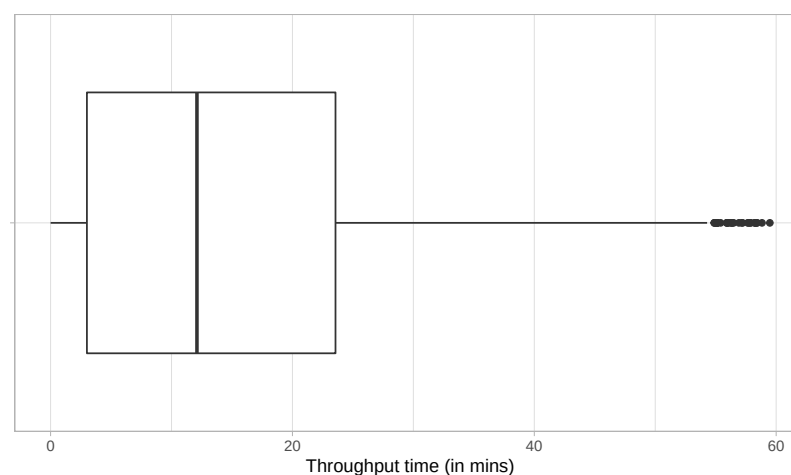


Figure 6.3.: Répartition des durées inférieures à 60 minutes pour les traces d'activités collectées

Nous pouvons ainsi estimer, à partir de cette figure, la durée moyenne d'une consultation entre 10 et 25 minutes, en supposant une marge d'au moins 5 minutes entre la durée d'une trace d'utilisation d'Easily[®] et la durée totale d'une consultation. La fin d'une consultation étant définie comme le moment où la·e patient·e sort de la salle de consultation et non pas comme le moment de la dernière action du/de la médecin dans Easily[®].

Enfin, nous pouvons calculer la matrice de connexions entre les différents évènements collectés afin de représenter le processus d'une consultation. Une matrice de connexions est construite comme suit : chaque évènement se voit affecter à une ligne et à une colonne de la matrice, et pour chaque cellule de la matrice est comptabilisé le nombre de fois où l'évènement correspondant à la ligne est suivi par l'évènement correspondant à la colonne. Cette matrice permet ainsi de savoir combien de fois un évènement A a été suivi par un évènement B dans les traces collectées. Cela nous permet ensuite de créer un graphe de connexions entre les différents évènements

observés. La Figure 6.4 montre le graphe de connexions obtenu pour les traces d'activités que nous avons collectées aux HCL.

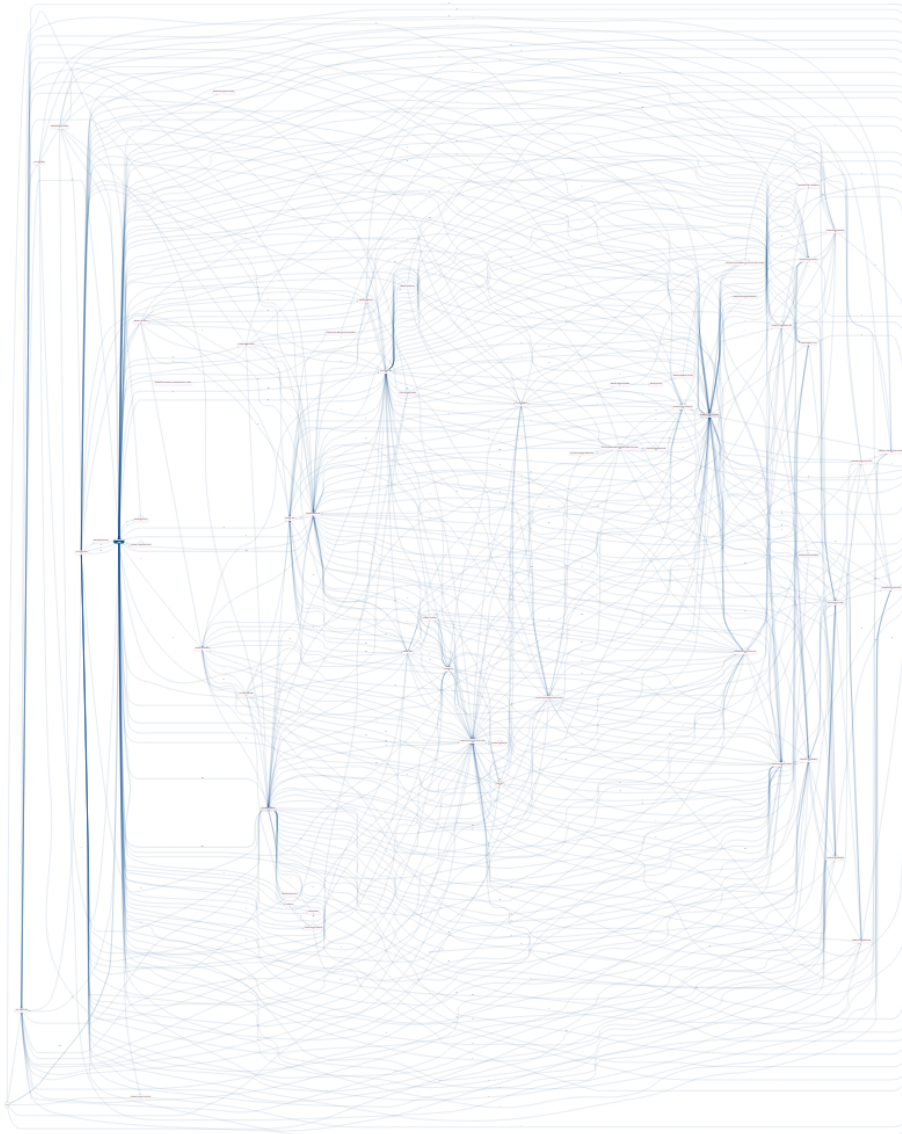


Figure 6.4.: Graphe de connexions entre les différents évènements collectés

Nous pouvons observer que le graphe obtenu contient de très nombreuses connexions s'entremêlant entre elles. Ce type de résultat correspond à ce que VAN DER AALST, (2011) nomme des processus de type "spaghetti". L'étude de ce type de processus nécessite l'usage d'algorithmes de *process mining* afin d'extraire les processus les plus significatifs.

6.2.2 Heuristic Miner

L'algorithme *Heuristic Miner*, proposé par WEIJTERS *et al.*, (2006), se base sur l'heuristique que les événements les plus co-dépendants forment les processus sous-jacents les plus révélateurs d'un processus global. Ainsi, l'algorithme *Heuristic Miner* calcule le degré de dépendance $a \Rightarrow_W b$, pour chaque couple d'événements a et b d'un ensemble de traces W , via l'Équation (6.1) (avec $|a >_W b|$ le nombre de fois où a est suivi de b dans W).

$$a \Rightarrow_W b = \frac{|a >_W b| - |b >_W a|}{|a >_W b| + |b >_W a| + 1} \quad (6.1)$$

Une fois ces degrés de dépendance calculés, nous obtenons une matrice de dépendances connectant les différents événements entre eux et permettant ainsi de créer un graphe de dépendance. Enfin, un seuil λ est utilisé afin que seuls soient visibles les liens associés à un degré de dépendance supérieur à ce seuil. La Figure 6.5 montre le graphe de dépendance obtenu par l'algorithme *Heuristic Miner* à partir des traces d'activités collectées aux HCL, avec $\lambda = 0.9$ et en supprimant les événements n'appartenant pas à un chemin entre les nœuds "start" et "end".

À partir de ce graphe, nous pouvons observer que le déroulé d'une consultation se divise en deux phases bien distinctes :

1. Une phase de recherche d'informations sur la·e patient·e en cours, illustrée par les événements "Accès Onglet BioBoxes", "Lecture des résultats de biologie de ville d'un patient", "Saisie DPC : Poids déclaré" et "Ouverture des antécédents" ;
2. Une phase de production de documents, tels que des ordonnances et comptes-rendus de consultation, qui clôturent généralement le processus de consultation.

Notons tout de même la présence, dans la seconde phase, entre l'accès à l'outil "ePrescription" et la production d'ordonnances, de l'événement "Saisie ATCD", qui correspond à la saisie d'antécédents médicaux concernant la·e patient·e en cours. Cela peut correspondre, comme nous avons pu l'observer en situations réelles, au fait de demander aux patient·es si ceux-ci ont déjà suivi un traitement ou si ils ont des contre-indications pour suivre un traitement particulier, des allergies par exemple. La présence de cet événement en fin de processus peut ainsi refléter une dernière recherche d'informations sur la patient·e juste avant la saisie, et donc la décision finale, du traitement à prescrire.

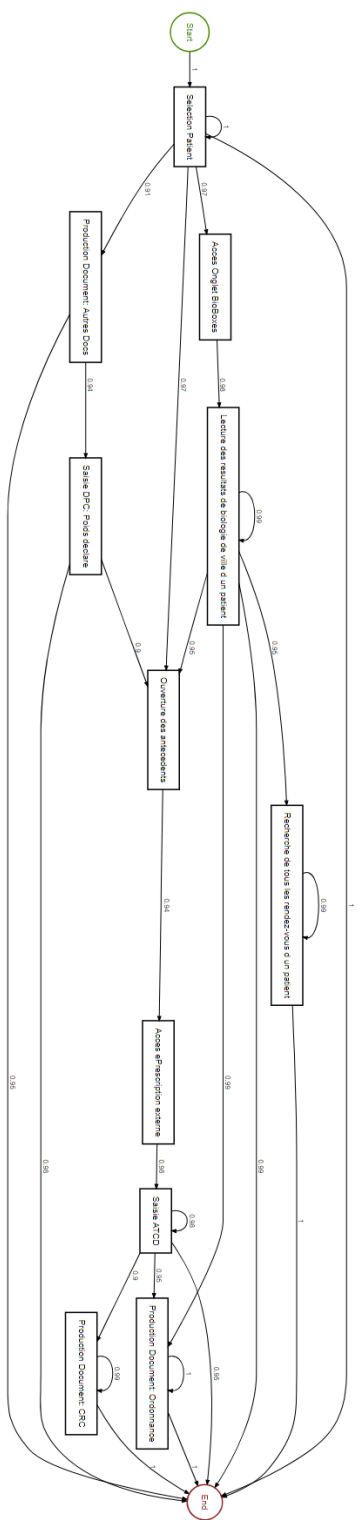


Figure 6.5.: Graphe de dépendance obtenu via l'utilisation de *Heuristic Miner* sur les traces d'activité correspondant à 3439 consultations effectuées sur un mois par 75 médecins de 5 départements différents.

Nous observons aussi que certains événements sont "co-dépendants" d'eux-mêmes, ce qui suggère que les actions liées à ces événements sont répétées plusieurs fois d'affilée. C'est le cas pour les événements : "Lecture des résultats de biologie de ville d'un patient", "Recherche de tous les rendez-vous d'un patient", "Saisie ATCD", "Production Document : Ordonnance", "Production Document : CRC".

Enfin, plusieurs types d'événements conduisent à la fin du processus. Si nous nous concentrons sur les événements dont le degré de dépendance avec le nœud "end" est égal à 1, nous avons :

- "Sélection Patient", ce qui suggère, comme énoncé en début de section 6.2.1, que les médecins ont simplement ouvert puis refermé le dossier de leur patient·e ;
- "Recherche de tous les rendez-vous d'un patient", ce qui suggère que les médecins ont finalisé leur consultation en proposant un nouveau rendez-vous au/à la patient·e en cours ;
- "Production Document : Ordonnance", ce qui suggère la clôture de la consultation avec la production des ordonnances pour les traitements prescrits par la·e médecin ;
- "Production Document : CRC", ce qui suggère que les médecins ont finalisé leur consultation par la rédaction d'un compte-rendu synthétisant les points abordés durant la consultation et les décisions prises concernant les prescriptions faites.

Si la clôture d'une consultation par l'événement "Sélection Patient" n'apporte que peu d'informations sur le processus d'une consultation, ce n'est pas le cas des trois autres situations. Qu'une consultation se clôture régulièrement par la planification d'un prochain rendez-vous ou la production de documents de synthèse, sans signe que les médecins aient recherché plus d'informations sur leur patient·e en cours après ces événements, laisse suggérer que les médecins décident de clôturer leur consultation lorsque ceux-ci pensent avoir fait le tour de toutes les décisions possibles pour le cas en cours. En d'autres termes, les médecins mettent fin à leur consultation quand l'accumulation d'informations sur la·e patient·e n'est plus possible ou n'est plus utile au choix d'une prescription à faire.

6.2.3 Fuzzy Miner

Afin d'analyser avec un autre point de vue les traces d'activités collectées lors de consultations médicales, nous avons utilisé un autre algorithme de *process mining*

bien connu : *Fuzzy Miner*. Pour ce faire, nous nous sommes appuyés sur la librairie *fuzzymineR*². Cette librairie se base sur l'outil ProM 6.4.1 (VAN DONGEN *et al.*, 2005 ; VERBEEK *et al.*, 2010) et permet l'utilisation de l'algorithme *Fuzzy Miner* dans le langage R.

L'algorithme *Fuzzy Miner*, proposé par GÜNTHER et VAN DER AALST, (2007), fut initialement conçu pour analyser les processus de type "spaghetti", comme celui qui nous étudions. Pour ce faire, *Fuzzy Miner* utilise plusieurs métriques afin de mesurer le degré de "signifiante" de chaque nœud et de chaque arc du graphe, ainsi que le degré de connexion entre les différents nœuds. Ensuite, en se basant sur différents seuils de significances, l'algorithme *Fuzzy Miner* va regrouper ou supprimer des nœuds et des arcs suivant la logique suivante :

- Les nœuds et arcs hautement "signifiants" sont préservés ;
- Les nœuds et arcs peu "signifiants" mais fortement "corrélés" sont regroupés dans un nouveau nœud, ou "*cluster*" ;
- Les nœuds et arcs peu "signifiants" et peu "corrélés" sont supprimés du graphe.

Ainsi, plus les seuils de significances sont élevés plus les nœuds et arcs sont regroupés ou supprimés, et donc plus le graphe sera simplifié. Malheureusement, le package *fuzzymineR* ne permet pas de savoir quels nœuds ont été regroupés entre eux, ce qui limitera fortement notre analyse et les conclusions que nous pourrions en tirer.

La Figure 6.6 montre le graphe obtenu par l'utilisation de l'algorithme *Fuzzy Miner* à partir des traces d'activités collectées, avec des seuils de signifiante des nœuds et des arcs fixés à 0.75.

Nous pouvons voir que, du fait d'un choix de seuils élevés, de nombreux nœuds et arcs ont été regroupés dans des nœuds *cluster*. Cependant, nous pouvons observer quelques similitudes et différences avec les résultats obtenus avec l'algorithme *Heuristic Miner*. Premièrement, nous observons que le processus de consultation commence par la sélection du/de la patient·e et se termine par la production d'un compte-rendu de la consultation. Ceci suggère, comme énoncé en fin de section 6.2.2, que les médecins mettent fin à leurs consultations lorsque ceux-ci ont fait le tour des prescriptions à faire pour la·e patient·e en cours.

2. <https://github.com/nirmalpatel/fuzzymineR>

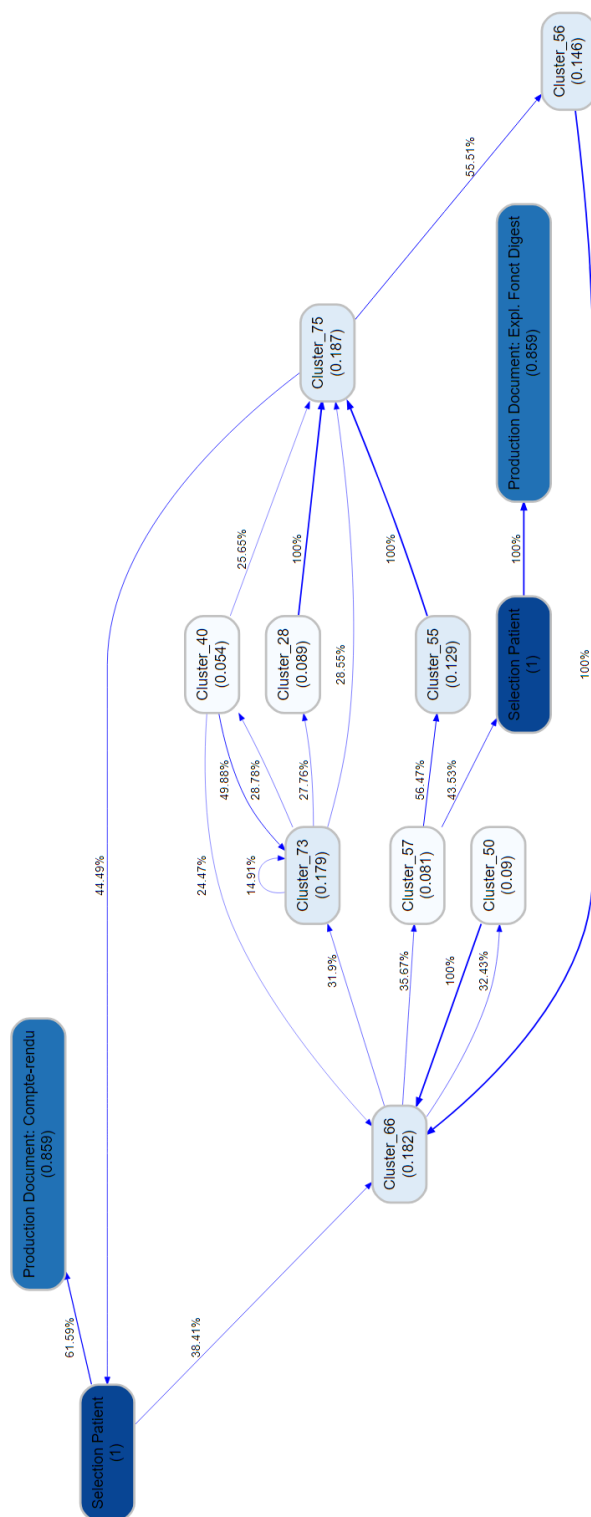


Figure 6.6.: Graphe résultant de l'utilisation de l'algorithme Fuzzy Miner, avec un seuil de signifiacnce des nœuds et des arcs fixé à 0.75

Deuxièmement, à l'inverse du graphe obtenu via *Heuristic Miner*, l'algorithme *Fuzzy Miner* met en évidence de nombreux sous-processus au processus général d'une consultation. La production d'ordonnance étant un évènement unique et la production de compte-rendu étant bien identifiable en fin de processus, nous pouvons supposer que les différents "*clusters*" regroupent des évènements liés à la recherche d'informations sur la·e patient·e en cours. Ceci suggère que le processus d'une consultation se base sur de nombreux sous-processus de recherche d'informations, qui diffèrent en fonction du type des informations recherchées. Par exemple, la recherche d'une donnée biologique ne suit pas le même processus que la recherche d'un antécédent familial. Ces différents sous-processus sont liés entre eux, parfois de manière cyclique, ce qui suggère que les médecins ne recherchent nécessairement pas toutes les informations d'un certain type en même temps, mais font des recherches répondant à leurs besoins en information au fur et à mesure de la consultation. Par exemple, un·e médecin peut rechercher une donnée biologique, puis vérifier si un antécédent familial est présent dans le dossier de la·e patient·e, puis rechercher une autre donnée biologique, puis rechercher les résultats d'un scanner, puis vérifier les antécédents allergiques du/de la patient·e, et ainsi de suite.

6.3 Synthèse

À partir des observations et des analyses effectuées, nous pouvons maintenant apporter une réponse aux questions soulevées en début de ce chapitre. Cette synthèse de nos résultats nous permettra de proposer, par la suite, des modèles formalisant le processus de décision des médecins lors de leurs consultations.

Nous avons commencé par mettre en évidence, autant dans nos observations que dans nos analyses, les différentes actions élémentaires composant le processus d'une consultation. Nous avons proposé de regrouper ces différentes actions en deux grands types d'actions : la recherche d'informations et la production de prescriptions. La recherche d'information est, d'après nos observations et analyses, le type d'action le plus récurrent de la part des médecins lors de leurs consultations.

Comme nous avons pu l'analyser dans les traces d'activités des médecins, la production de prescriptions arrive généralement en fin de consultation. Au contraire, la recherche d'informations semble constituer la partie initiale (cf. Figure 6.5) et centrale (cf. Figure 6.6) d'une consultation. Cela nous permet de supposer que la recherche d'informations sur la·e patient·e en cours est nécessaire aux médecins afin de déterminer des prescriptions à lui fournir en fin de consultations.

La recherche d'informations est généralement liée à la saisie de ces informations et à la production de documents synthétisant la consultation effectuée (ex. un compte-rendu de consultation). Ce type de documents est nécessaire aux médecins afin de se remémorer ce qui a été décidé, et pour quelles raisons, au cours de précédentes consultations. La production de documents de synthèse clôture les consultations tout autant que la production d'ordonnances. Cette observation nous permet de supposer que les médecins mettent fin à leurs consultations lorsque ceux-ci pensent que toutes les décisions de prescriptions pouvant être prises ont été prises. En d'autres termes, les médecins mettent fin à leurs consultations lorsque la recherche de nouvelles informations n'est plus nécessaire pour décider des prescriptions à faire au/à la patient·e en cours. Cela met d'autant plus en évidence l'aspect essentiel de la collecte d'informations des médecins au cours d'une consultation.

Le processus de travail des médecins, lors de leurs consultations, peut ainsi se résumer à un travail d'enquête pour déterminer les décisions à prendre concernant le traitement de leurs patient·es. Les médecins accumulent alors des informations, élaguent le domaine des prescriptions pouvant être faites aux patient·es et s'arrêtent lorsqu'il ne reste plus aucune ambiguïté sur les prescriptions à faire pour la consultation en cours.

Modélisation du processus de décision des médecins en consultations médicales

” *In the shooter hypothesis, a good marksman shoots at a target, creating a hole every ten centimeters. Now suppose the surface of the target is inhabited by intelligent, two-dimensional creatures. Their scientists, after observing the universe, discover a great law : "There exists a hole in the universe every ten centimeters". They have mistaken the result of the marksman's momentary whim for an unalterable law of the universe.*

— Liu Cixin

The Three-Body Problem

Dans ce chapitre, nous cherchons à formaliser et modéliser le processus de décision des médecins durant leurs consultations médicales. Le but de ces modélisations est de mettre en évidence les points clés du processus de décision des médecins en consultation afin de déterminer quel type de support serait le plus adapté à leurs besoins. Ce chapitre reprend ainsi le travail proposé dans RICHARD *et al.*, (2018).

Nous pouvons trouver, dans la littérature, de nombreuses propositions de modèles visant à expliquer le processus de raisonnement des médecins lorsqu'ils établissent un diagnostic (CROSKERRY, 2009 ; PELACCIA *et al.*, 2011 ; STAUSBERG et PERSON, 1999). Ces études mettent en évidence l'impact de l'expérience des médecins sur leurs décisions. Cependant, comme précisé au Chapitre 6, nous ne cherchons pas ici à comprendre les processus internes des médecins, mais simplement les décisions/actions pouvant être effectuées par un·e médecin durant une consultation et la manière dont ces décisions/actions sont connectées entre elles.

TAYLOR *et al.*, (1971) ont proposé une comparaison entre les processus de diagnostics de différents médecins et le processus de diagnostic d'un système expert. Dans

cette étude, les processus de 6 médecins ont été observés pour la résolution de 20 cas de goitre non toxique. TAYLOR *et al.*, (1971), en schématisant les points potentiels de divergence entre les processus des médecins et celui du système expert, ont mis en évidence un processus de décision cyclique. Les médecins, tout comme le système expert, mettent à niveau leurs estimations au fur et à mesure qu'ils accumulent de l'information sur le cas traité. LEAPER *et al.*, (1973) proposent, quant à eux, une étude de 1307 diagnostics effectués par un total de 28 médecins, pour des cas de patient-es souffrant de douleurs abdominales. Un des principaux résultats de cette étude est que, pour un·e même patient·e et un même diagnostic final, les médecins ne suivent pas nécessairement le même "chemin" dans leurs raisonnements, et ce bien qu'ils passent généralement par les mêmes étapes. Ces études offrent des modélisations, parfois sommaires, du processus de décision pouvant s'appliquer au cas des consultations médicales. Cependant, ces études datent d'avant la généralisation des HIS et ne permettent donc pas de rendre compte de l'impact de ces outils sur les processus de décision des médecins. Si certains travaux, plus récents, proposent une analyse de l'usage d'un HIS en consultations médicales (GIBSON *et al.*, 2006 ; REBUGE *et al.*, 2013), ceux-ci ne proposent pas une formalisation générique des processus analysés.

Les modèles proposés dans ce chapitre sont principalement basés sur les observations et analyses effectuées en Chapitre 6. En Section 7.1, nous proposons une formalisation permettant de détailler les actions effectuées par un·e médecin durant une consultation spécifique. En Section 7.2, nous proposons une abstraction du type de décisions que peuvent faire les médecins durant une consultation afin d'établir un modèle générique du processus de décision des médecins lors de consultations médicales. Nous y montrons que le principal besoin des médecins lors de leurs consultations réside dans l'accumulation d'informations concernant la·e patient·e en cours. Enfin, en Section 7.3, nous montrons que le besoin en information des médecins peut être modélisé comme un problème de classification multi-labels et peut donc être résolu par des algorithmes d'apprentissage.

7.1 Formalisation du processus d'une consultation spécifique

Nous cherchons ici à établir un modèle permettant de formaliser finement le déroulement d'une consultation spécifique. Définissons $\mathcal{A}$ comme l'ensemble des alternatives

possibles, en termes d'actions/décisions, pour un·e médecin durant une consultation. Comme observé en Chapitre 6, deux sous-ensembles d'actions/décision se distinguent : "Rechercher" une information spécifique concernant la·e patient·e en cours de consultation et "Prescrire" quelque chose au/à la patient·e. Nous ajoutons une action supplémentaire : "Fin de consultation", représentant la décision du/de la médecin de mettre fin à la consultation en cours, de rédiger les différentes prescriptions décidées, de rédiger un document de synthèse et de planifier un nouveau rendez-vous avec la·e patient·e. Comme nous avons pu l'analyser en Chapitre 6, ces actions sont celles qui clôturent généralement une consultation. Nous définissons aussi Φ comme l'ensemble des médecins à même de faire des consultations médicales, $\mathcal{P}$ comme l'ensemble des patient·es et $\mathcal{T}$ comme l'ensemble des moments où une action/décision $A \in \mathcal{A}$ a été prise par un·e médecin durant une consultation.

Soit $\mathcal{C}$ l'ensemble des consultations, avec $c \in \mathcal{C}$ une consultation spécifique. Chaque consultation $c \in \mathcal{C}$ est identifiée par un triptyque $\langle p, \phi, t_0 \rangle$, avec $p \in \mathcal{P}$ un·e patient·e, $\phi \in \Phi$ un·e médecin et $t_0 \in \mathcal{T}$ le début de la consultation. Soit aussi $T_c \subset \mathcal{T}$, le sous-ensemble des moments $t \in \mathcal{T}$ constituant une consultation $c \in \mathcal{C}$, tel que $\forall t \in T_c, \exists A_t \in \mathcal{A}$.

Nous pouvons maintenant définir $\mathcal{X}$ comme l'ensemble des données sur les patient·es pouvant être "recherchées" par un·e médecin au cours d'une consultation et $X_t^{\phi,p} \in \mathcal{X}$ comme le sous-ensemble des données connues par un·e médecin $\phi \in \Phi$ à propos d'un·e patient·e $p \in \mathcal{P}$ à un moment $t \in T_c$ d'une consultation $c \in \mathcal{C}$. En nous reposant, encore une fois, sur nos analyses du Chapitre 6, nous supposons que, à tout moment $t \in T_c$, un·e médecin décide de faire une action $A_t \in \mathcal{A}$ en se basant sur les données connues $X_t^{\phi,p}$. C'est-à-dire, soit décider d'une donnée $X \in \mathcal{X}$ telle que $X \notin X_t^{\phi,p}$, soit décider d'une prescription à faire (sans nécessairement rédiger cette prescription).

Comme introduit par CROSKERRY, (2009), PELACCIA *et al.*, (2011) et STAUSBERG et PERSON, (1999), les médecins prennent ces décisions en se reposant des règles idiosyncratiques. Nous supposons que ces règles suivent un format tel que défini en Équation (7.1), où la conjonction des informations $X_t^{\phi,p}$ connues du/de la médecin ϕ sur un·e patient·e p , à un moment t , implique la décision d'une action $A_t \in \mathcal{A}$.

$$\bigwedge_{X \in X_t^{\phi,p}} X = x \implies A_t = a \quad (7.1)$$

Cependant, ces règles sont vouées à évoluer en fonction de différents facteurs tels que l'apparition de nouveaux traitements, de nouveaux types d'analyses médicales,

ou simplement de l'état de fatigue du/de la médecin. Il est ainsi difficile d'établir un ensemble fini et définitif de ces règles. Toutefois, savoir que les médecins suivent un système de règles se basant sur les informations connues sur un·e patient·e à un moment t d'une consultation nous permet de détailler le déroulé d'une consultation spécifique à l'aide d'un tableau.

T_c	$\mathcal{X}$							$\mathcal{A}$
	Sexe	Âge	IMC	Pathologie	HDL	LDL	TG	
t_0	♂	55	∅	HChol	∅	∅	∅	Rechercher HDL
t_1	♂	55	∅	HChol	1.1	∅	∅	Rechercher LDL
t_2	♂	55	∅	HChol	1.1	5.53	∅	Rechercher TG
t_3	♂	55	∅	HChol	1.1	5.53	1.98	Prescrire Ezetrol
t_4	♂	55	∅	HChol	1.1	5.53	1.98	Rechercher IMC
t_5	♂	55	24.43	HChol	1.1	5.53	1.98	Fin de Consultation

Table 7.1.: Exemple du processus de décision d'un·e médecin durant une consultation médicale spécifique, pour le suivi d'un·e patient·e hypercholestérolémique

En Tableau 7.1, nous pouvons voir un exemple factice et simplifié, inspiré de nos observations, détaillant le processus d'une consultation pour un·e patient·e hypercholestérolémique. L'ensemble $\mathcal{X}$ correspond ici à l'ensemble des données patient : {Sex, Age, IMC, Pathologie, HDL, LDL, TG}, représenté par les colonnes centrales du tableau. Dans cet exemple, la·e médecin commence une consultation avec $X_{t_0}^{\phi,p} \neq \emptyset$. En effet, l'âge, le sexe et la pathologie pour laquelle un·e patient·e est suivi sont généralement connus des médecins en début de consultation. Ainsi, pour chacun des moments $t \in T_c$ de la consultation, référencés dans la colonne de gauche, la·e médecin décide d'une action $A_t \in \mathcal{A}$, référencée dans la colonne de droite. Ces décisions, comme énoncé précédemment, se basent sur l'ensemble des règles d'usage que s'est créé la·e médecin au cours de ses études et de son activité.

Cette formalisation est particulièrement utile pour rendre compte du déroulement d'une consultation spécifique. Elle permet de mettre en évidence, étape par étape, l'accumulation d'informations par les médecins, ainsi que l'ensemble des données $X_t^{\phi,p}$ nécessaires à un·e médecin pour décider d'une prescription spécifique. Cependant, cette formalisation ne permet pas de mettre en évidence les liens génériques

entre les décisions visant à "Rechercher" une information sur un·e patient·e et celles visant à "Prescrire" quelque chose à ce·tte patient·e.

7.2 Modélisation générique du processus de décision en consultation

Afin de formaliser et de mettre en évidence les interactions entre les différents types de décisions pouvant être prises par les médecins en consultations, nous proposons ici un modèle plus générique du processus de décision d'une consultation médicale. L'objectif de ce modèle est de mettre en évidence le processus de décision commun à toute consultation médicale, indépendamment des données recueillies sur un·e patient·e ou de la spécialité du/de la médecin. Alors que le modèle présenté en Section 7.1 donne l'image d'un processus de décision linéaire, se déroulant étape par étape, le modèle que nous présentons dans cette section met en lumière un processus de décision cyclique. Cette observation fait écho aux analyses faites dans le Chapitre 6, mais aussi à des études antérieures du processus d'une consultation médicale (TAYLOR *et al.*, 1971 ; LEAPER *et al.*, 1973).

Le modèle que nous proposons est représenté graphiquement en Figure 10.1 à l'aide d'un diagramme d'activité. Ce diagramme est composé de 9 nœuds : 1 nœud "départ" (disque noir), 6 nœuds "action" (blocs arrondis) et 2 nœuds "conditionnel" (losanges). Nous y retrouvons les deux types de décisions précédemment définis : le choix de "rechercher" une information encore inconnue sur la·e patient·e (nœud 3) et le choix de "prescrire" quelque chose au/à la patient·e (nœud 5). Nous y retrouvons aussi l'action spéciale "Fin de consultation" (nœud 6), représentant le choix du/de la médecin de mettre fin à la consultation en cours.

Premièrement, comme nous l'avons relevé dans nos analyses du Chapitre 6, les médecins mettent fin à leurs consultations lorsque ceux-ci considèrent qu'ils ne peuvent pas faire plus de prescriptions à leur patient·e pour la consultation en cours. Nous avons ici notre condition d'arrêt du processus d'une consultation, représentée par la question en nœud (1), suivi par le premier "embranchement" du processus : s'il n'y a plus de prescription à faire pour la consultation en cours, alors mettre fin à la consultation (nœud 6), sinon continuer la consultation (nœud 2).

Comme nous l'avons vu au Chapitre 6 et formalisé en section 7.1, les médecins décident des prescriptions à faire à leurs patient·es à partir de l'ensemble $X_t^{\phi,p}$ des informations connues à leurs propos à un instant t . Ainsi, si à un instant t

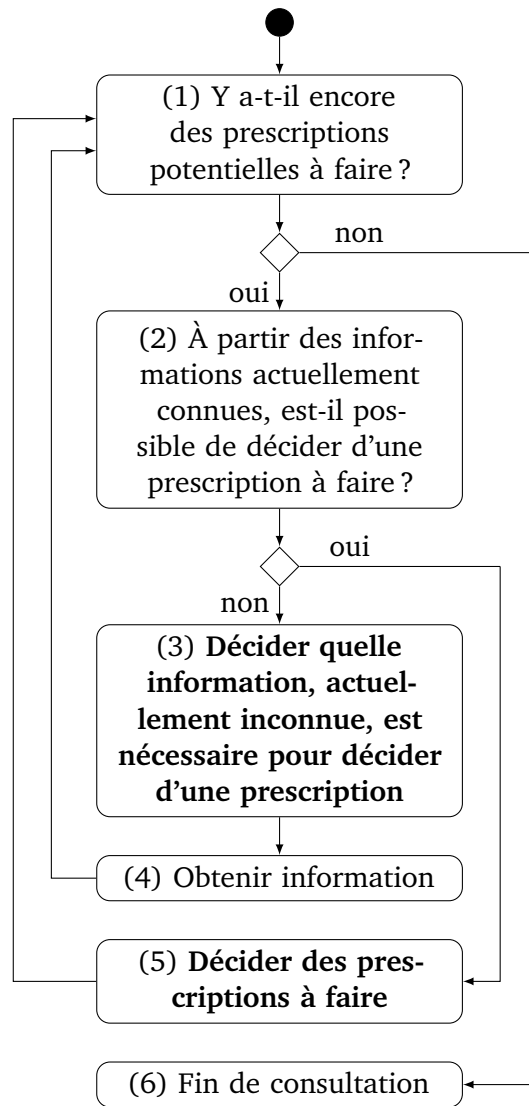


Figure 7.1.: Modélisation générique du processus de décision des médecins lors de consultations médicales coutumières

un·e médecin considère qu'il reste des ambiguïtés concernant des prescriptions pouvant être fournies à sa/son patient·e, iel va tenter de déterminer si une nouvelle prescription peut être décidée à partir des informations actuellement connues sur la·e patient·e (nœud 2). Nous arrivons ainsi au second "embranchement" du processus : si une prescription peut être décidée à partir des informations actuellement connues sur la·e patient·e, alors décider de la prescription à faire (nœud 5), sinon continuer à accumuler de l'information sur la·e patient·e (nœud 3).

Le nœud (3) correspond à la décision de "rechercher" une information concernant la·e patient·e p en cours de consultation. Plus précisément, et comme formalisé en section 7.1, cette décision porte sur le choix à un instant t d'une information

$X \in \mathcal{X}$ alors inconnue du/de la médecin ϕ , c'est-à-dire $X \notin X_t^{\phi,p}$, parmi toutes les informations alors inconnues du/de la médecin à cet instant t .

Une fois l'information à "rechercher" choisie, la-e médecin tente d'obtenir cette information, soit en la demandant au/à la patient-e ou bien en effectuant une recherche dans son HIS (nœud 4). Précisons que, si un-e médecin tente d'obtenir une information sur la-e patient-e mais n'y parvient pas, cela est aussi à considérer comme une connaissance sur la-e patient-e, car pouvant amener la-e médecin à prescrire une analyse en laboratoire. De plus, il peut arriver que l'obtention d'une information annule la potentialité d'une prescription envisagée et puisse, parfois, amener à la fin d'une consultation. C'est le cas dans l'exemple présenté en Tableau 7.1 pour illustrer notre premier modèle. Dans cet exemple, la dernière décision du/de la médecin, avant de mettre fin à la consultation, est de rechercher l'Indice de Masse Corporelle (IMC) de sa/son patient-e, afin de vérifier que sa valeur ne dépasse pas un certain seuil limite. Un IMC de 24.43 est relativement "normal", ainsi aucune prescription supplémentaire n'est envisagée par la-e médecin. La-e médecin ayant, à ce moment de la consultation, fait le tour des prescriptions possibles, iel décide de mettre fin à la consultation. C'est la raison pour laquelle le nœud (4) est relié au nœud (1), et non au nœud (2).

Le nœud (5), quant à lui, correspond au choix de la prescription à fournir au/à la patient-e. Comme formalisé en section 7.1, cette décision repose essentiellement sur l'ensemble $X_t^{\phi,p}$ des informations connues à un instant t par la-e médecin ϕ à propos de sa/son patient-e p . Ces prescriptions peuvent être rédigées au moment où elles sont décidées (nœud 5), ou bien à la fin de la consultation (nœud 6), comme nous avons pu l'observer dans le Chapitre 6. Une fois la prescription décidée, la-e médecin continue d'essayer de déterminer s'il reste d'autres prescriptions pouvant être faites (nœud 1).

Nous obtenons ainsi un processus cyclique avec deux sous-processus internes. Le premier correspond à l'accumulation, par la-e médecin, d'informations concernant la-e patient-e en cours de consultation. Cette accumulation peut alors, pour un cas spécifique, se décomposer en une suite d'actions comme formalisé en section 7.1. Ce premier sous-processus apparaît essentiel au déclenchement du second sous-processus, qui correspond au choix des prescriptions à faire au/à la patient-e en cours. Tout comme pour le premier sous-processus, il est possible, pour un cas spécifique, de décomposer ce second sous-processus en plusieurs étapes successives, ce qui nous ramène à notre premier modèle décrit en section 7.1. Ce second modèle, plus générique que le premier, permet de mettre en évidence l'aspect central de l'accumulation d'information dans le processus d'une consultation médicale. Plus pré-

cisement, le choix de quelle information inconnue est nécessaire afin de déterminer rapidement quelle prescription fournir au/à la patient·e apparaît comme l'élément déterminant d'une consultation.

7.3 Un problème de classification multi-labels

À partir des modèles proposés dans ce chapitre, ainsi que des analyses présentées dans le Chapitre 6, nous avons pu mettre en évidence le rôle central de la recherche de données patients dans le déroulé du processus de décisions des médecins en consultation. Cependant, nous avons aussi pu voir au Chapitre 2 que l'émergence des HIS a entraîné l'accumulation d'un très grand nombre de données patients, noyant le personnel soignant sous des flots de données médicales (PIVOVAROV et ELHADAD, 2015). Et bien qu'une grande majorité des données patients dont les médecins ont besoin durant les consultations est présente en base de données, l'accès à ces données n'est pas toujours aisé. Les médecins passent alors plus de temps à rechercher de l'information sur leurs patient·es qu'à analyser ces données et à établir des diagnostics (voir Chapitre 6). Nous en concluons qu'un outil capable d'anticiper, de retrouver et de présenter de manière synthétique les données patients dont les médecins ont besoin durant leurs consultations, serait un outil d'aide à la décision pertinent dans notre contexte.

Ce problème peut aisément se formaliser comme un problème classification multi-labels. Comme défini par TSOUMAKAS et KATAKIS, (2007), un problème multi-labels est une extension des problèmes de classification "classiques", ou mono-label. Chaque entrée est définie par un ensemble de variables endogènes $\mathcal{X}$, mais aussi par un ensemble $\mathcal{Y}$ de labels pouvant être associés à $\mathcal{X}$. En d'autres termes, dans un problème de classification mono-label, chaque entrée appartient à une seule et unique classe, alors que dans un problème de classification multi-labels, chaque entrée peut appartenir à plusieurs classes en même temps. Un exemple classique de problème multi-labels est la labellisation automatique d'images.

Dans notre cas, en reprenant le modèle micro présenté en section 7.1 et en ne gardant que les données connues sur la·e patient·e en début de consultation (c-à-d. son âge, son sexe, son IMC et la pathologie pour laquelle iel est suivi) ainsi que l'ensemble des informations recherchées par la·e médecin, nous obtenons une formalisation multi-labels de notre problème. La Tableau 7.2 montre un exemple simplifié, avec des patient·es factices, de notre problème au format multi-labels.

$\mathcal{X}$: données connues sur la·e patient·e				$\mathcal{Y}$: données sur la·e patient·e nécessaires au médecin					
Sexe	Age	IMC	Pathologie	HbA1c	Glucose	HDL	LDL	Créatinine	Microalbumine
♀	42	34.23	DT2	1	1	0	0	0	0
♂	52	27.15	HChol	0	0	1	1	0	0
♂	24	21.12	DT1	1	1	0	0	1	1
♀	67	26.22	HChol	0	0	1	1	0	0

Table 7.2.: Exemple, avec des données factices, de la formalisation de notre cas d'étude en problème de classification multi-labels (DT1 : Diabète Type 1, DT2 : Diabète Type 2, HChol : Hypercholestérolémie)

Comme mentionné précédemment, l'ensemble $\mathcal{X}$ des variables endogènes de notre problème représente les données connues du médecin sur la·e patient·e en début de consultation. En d'autres termes, son âge, son sexe et la pathologie pour laquelle il est suivi, auxquels nous ajoutons son IMC, qui est une information calculée quasiment à chaque consultation.

L'ensemble $\mathcal{Y}$ des variables exogènes (les labels) représente l'ensemble des données sur la·e patient·e généralement recherchées par la·e médecin et pouvant donc l'intéresser au cours d'une consultation. Chaque label $Y \in \mathcal{Y}$ peut prendre soit la valeur 1, correspondant à "la·e médecin a recherché cette information", soit la valeur 0, correspondant à "la·e médecin n'a pas recherché cette information". Dans l'exemple figurant en Tableau 7.2, l'ensemble $\mathcal{Y}$ est limité aux données : HbA1c (hémoglobine glyquée), Glucose, cholestérol HDL, cholestérol LDL, Créatinine et Microalbumine. Cependant, en pratique, l'ensemble $\mathcal{Y}$ peut compter plus d'une centaine de données potentiellement nécessaires aux médecins au cours d'une consultation.

Comme abordé plus en détail dans le Chapitre 9 de cette thèse, certain nombre d'algorithmes d'apprentissage ont été proposés afin de résoudre les problèmes de classification multi-labels. L'algorithme d'apprentissage BP-MLL, proposé par ZHANG et ZHOU, (2006), se basant sur un réseau de neurones adapté au traitement de problèmes multi-labels. ML-kNN, proposé par ZHANG et ZHOU, (2007), qui est une adaptation de l'algorithme des k plus proches voisins (kNN) pour le traitement des problèmes multi-labels. Mais aussi, le méta-algorithme RAKEL, proposé par TSOUMAKAS *et al.*, (2011b), qui consiste à ramener un problème multi-labels à plusieurs problèmes mono-label. Ce méta-algorithme permet alors d'utiliser sur un problème multi-labels des algorithmes d'apprentissage destinés à traiter des problèmes mono-label.

Plusieurs méthodes sont donc applicables à notre problème multi-labels. Cependant toutes ne sont pas nécessairement adaptées aux contraintes de notre contexte applicatif. Il est tout à fait possible qu'aucun des algorithmes existants ne convienne.

Si cela s'avère être le cas, il nous faudra alors envisager de proposer un nouvel algorithme. De plus, nous avons vu dans le Chapitre 3 que l'utilisation d'algorithmes ML peut générer de la méfiance de la part des médecins et donc une non-acceptation de l'outil proposé. Il nous faut donc tenir compte de cet aspect dans la sélection de l'algorithme d'apprentissage à utiliser dans notre outil.

Partie III

Proposer un outil d'aide à la décision dédié
au soutien des consultations médicales
coutumières

Définition de nos exigences en termes de "transparence" pour une meilleure acceptabilité

” *It was a technical advance worth communicating, worth explaining, worth demystifying. Whatever the social implications of the HealthGuard implant, they could no more be presented in a vacuum, divorced from the technology which made the device possible, than vice versa. Once people ceased to understand how the machines around them actually functioned, the world they inhabited began to dissolve into an incomprehensible dreamscape.*

— **Greg Egan**
Distress

Nous avons défendu que proposer un outil fournissant, en début de consultation, des données ciblées sur des patient·es pouvait être une aide adaptée aux besoins des médecins. Cependant, l'utilisation d'algorithmes d'apprentissage peut être un frein à l'acceptation d'un tel outil (cf. Chapitre 3). Il a été montré que la "transparence" d'un outil, et des algorithmes qu'il utilise, était un facteur pouvant améliorer son acceptabilité (HOLZINGER *et al.*, 2017 ; SINHA et SWEARINGEN, 2002). Nous proposons donc dans ce chapitre d'explorer comment développer un outil "transparent", afin d'améliorer l'acceptabilité de l'aide que notre outil pourra offrir. En explorant le concept de "transparence" des outils d'aide à la décision, nous nous focalisons sur un critère bien particulier d'évaluation de ces outils. Il existe bien d'autres critères au moyen desquels il est possible et légitime d'évaluer des outils d'aide à la décision, comme par exemple la consistance logique (FRANCO et MARTIN, 2009 ; QU et VERES, 2016) ou la défendabilité (CASTILHO *et al.*, 2018). Nous ne cherchons pas ici à dévaloriser ces critères. Nous nous concentrons simplement sur la "transparence" car

elle est reconnue comme étant vectrice d'acceptabilité, or c'est l'acceptabilité que nous ciblons en premier lieu.

Bien que de nombreuses définitions de la "transparence" se côtoient dans la littérature, peu arrivent à englober tous les aspects qui peuvent être associés au concept de "transparence des algorithmes". De plus, il y a peu de propositions de critères opérationnels permettant d'établir si un algorithme peut être considéré "transparent". Dans ce chapitre, notre objectif est de reprendre et d'étoffer nos travaux présentés dans RICHARD *et al.*, (2020a) sur les critères opérationnels permettant l'évaluation de la "transparence" de différents algorithmes, en nous focalisant sur ceux dédiés à la résolution de problèmes de classification.

Comme développé au chapitre 7, notre cas d'usage peut être formalisé comme un problème de classification multi-labels que nous souhaitons résoudre à l'aide d'un système de classification basé sur un algorithme d'apprentissage. Avant d'explorer le concept de "transparence" dans ce contexte, il nous faut définir certains termes que nous utiliserons par la suite quand nous parlerons de classification.

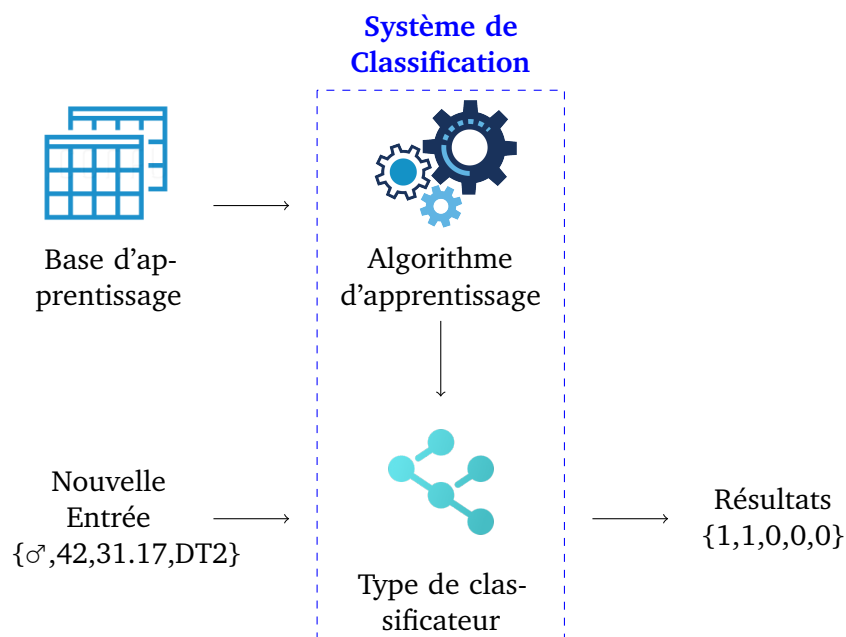


Figure 8.1.: Schématisation d'un système de classification

Nous appelons *algorithme d'apprentissage* un algorithme dédié à l'approximation d'une fonction de classification à partir d'une base de données, aussi appelée *base d'apprentissage*. Un algorithme d'apprentissage prend donc en entrée un ensemble d'entrées, chacune composée d'un même ensemble de *variables*. Les variables en question se divisent en deux catégories : les variables endogènes X , dont les valeurs

caractérisent une entrée, et les variables exogènes Y , dont les valeurs définissent la classe à laquelle est associée l'entrée. Dans le cadre d'un apprentissage dit supervisé, les valeurs des variables endogènes X et exogènes Y sont connues pour chaque entrée composant la base d'apprentissage. Par exemple, chaque entrée de la base d'apprentissage peut correspondre à un·e patient·e avec comme variables endogènes X la liste de ses symptômes et comme variable exogène Y son diagnostic.

Une fois que l'ensemble des entrées de la base d'apprentissage est analysé, l'algorithme d'apprentissage génère un *classificateur* dont le processus permet de déterminer la valeur, alors inconnue, des variables exogènes Y pour toute nouvelle entrée associée à un ensemble de variables endogènes X connues. Si nous reprenons notre exemple, pour un·e nouveau·elle patient·e, le classificateur tentera d'anticiper la valeur de la variable Y , son diagnostic, à partir de la liste de symptômes X de la·e patient·e. Précisons qu'un classificateur peut prendre différentes formes selon l'algorithme d'apprentissage utilisé : arbres de décision (QUINLAN, 1986), réseaux de neurones entraînés (GOODFELLOW *et al.*, 2016), ensembles de règles (COHEN, 1995), etc.

Cependant, un même algorithme d'apprentissage, pour différentes bases d'apprentissage, générera différents classificateurs. Par exemple, un algorithme d'apprentissage générant des arbres de décision va, pour deux bases d'apprentissage différentes, générer deux arbres de décision totalement différents. De plus, la complexité de ces arbres peut varier en fonction de la complexité de la fonction de classification que tente d'approximer l'algorithme d'apprentissage. Ainsi, l'un des arbres de décision générés peut apparaître simple et parfaitement lisible pour l'utilisateur·ice, tandis que l'autre peut présenter de trop nombreux embranchements et donc apparaître totalement incompréhensible. Le seul invariant sur lequel nous pouvons alors travailler est le *type de classificateur* généré par l'algorithme d'apprentissage.

Enfin, comme schématisé en Figure 8.1, nous appelons *système de classification* l'ensemble composé d'un algorithme d'apprentissage et du type de classificateurs produit par cet algorithme d'apprentissage. Par exemple, YAO *et al.*, (2005) proposent un système de classification utilisant des arbres de décision générés par une variante de l'algorithme d'apprentissage C4.5 (QUINLAN, 1993) afin de classifier des données médicales. Dans cet exemple, le type de classificateur utilisé est un arbre de décision, l'algorithme d'apprentissage est C4.5, les entrées sont des données médicales concernant un·e patient·e et la sortie du classificateur utilisé est la durée d'hospitalisation potentielle d'un·e patient·e.

Ces termes maintenant définis, et synthétisés dans la Tableau 8.1, nous pouvons étudier le concept "transparence" pour les systèmes de classification afin d'en extraire

Nom	Définition
Algorithme d'apprentissage	Processus générant des classificateurs à partir de bases d'apprentissage
Type de classificateur	Structure commune à tous les classificateurs générés par un même algorithme d'apprentissage
Classificateur	Processus permettant de classifier des entrées pour un problème de classification spécifique
Base d'apprentissage	Ensemble d'entrées définissant un problème de classification spécifique
Entrée	Ensemble de données à classifier
Résultat	Classification proposée par un classificateur pour une entrée donnée

Table 8.1.: Synthèse des différents éléments d'un système de classification et de leurs définitions

un ensemble d'exigences à respecter pour maximiser l'acceptabilité de ces systèmes et tenter de définir des critères opérationnels permettant d'évaluer le respect d'un système de classification envers ces exigences. En section 8.1, nous explorons les différentes définitions de la "transparence" proposées dans la littérature afin d'en dégager une définition plus générale regroupant les différentes exigences que doit respecter un système de classification pour être considéré "transparent". En section 8.2, nous proposons différents critères opérationnels, chacun lié à l'une de nos exigences en termes de "transparence", afin de pouvoir évaluer concrètement si un système de classification spécifique peut être considéré "transparent". En section 8.3, nous proposons une méthode d'évaluation et une méthode d'agrégation applicables à nos critères, compte tenu de leurs propriétés, afin de proposer une évaluation générale de la "transparence" d'un système de classification.

8.1 Recherche d'une définition générale de la "transparence"

Bien que le concept de "transparence des algorithmes" soit aussi ancien que les tout premiers systèmes de recommandations, l'omniprésence actuelle d'algorithmes d'apprentissage que l'on peut à bon droit considérer comme "opaques", tels que les réseaux de neurones artificiels, a remis la question de la "transparence" des algorithmes sur le devant de la scène avec l'émergence d'un domaine de recherche

sur les IA explicables (GUNNING, 2017 ; GOEBEL *et al.*, 2018). Cette question de la "transparence" des algorithmes, et par extension des systèmes de classification, est d'autant plus importante dans un contexte médical comme le nôtre (HOLZINGER *et al.*, 2017). En effet, les DDSS peuvent influencer les médecins vers de mauvaises déductions si ceux-ci ne comprennent pas comment le système fonctionne ou comment interpréter ses résultats (cf. Chapitre 4).

La littérature contient différentes définitions du concept de "transparence des algorithmes" et de notions connexes. Généralement, le concept de "transparence" est défini comme le fait de fournir le code source d'un système, des schémas de son fonctionnement ou des explications à ses résultats (CASTELLUCCIA et LE MÉTAYER, 2019 ; CRAMER *et al.*, 2008 ; DATTA *et al.*, 2016 ; GÖRITZLEHNER *et al.*, 2014 ; HERLOCKER *et al.*, 2000 ; KIM et HINDS, 2006 ; LYONS, 2013 ; MONTAVON *et al.*, 2018 ; SINHA et SWEARINGEN, 2002 ; THELISSON, 2017 ; TINTAREV et MASTHOFF, 2007 ; TSOUKIÀS, 2020). Cependant, certain·es auteur·ices précisent que ces explications doivent être compréhensibles par les utilisateur·ices (KIM et HINDS, 2006 ; SINHA et SWEARINGEN, 2002). D'autres définissent qu'un classificateur est "transparent" si celui-ci possède un processus de raisonnement compréhensible et interprétable par l'utilisateur·ice (KARSENTY et BOTHEREL, 2005 ; DINKA *et al.*, 2006 ; MICHENER et BERSCH, 2013 ; SPAGNOLLI *et al.*, 2017 ; DORAN *et al.*, 2017 ; ABDOLLAHI et NASRAOUI, 2018 ; PÉGNY et IBNOUHSEIN, 2018). D'autres encore définissent la "transparence" comme la capacité de retracer toutes les données utilisées durant le processus d'utilisation (AKKERMANS *et al.*, 2004 ; HOLZNER et HOLZNER, 2006 ; HEDBOM, 2008 ; HELBIG *et al.*, 2010 ; ZARSKY, 2013 ; GROTH, 2013 ; AMIRIBESHELI *et al.*, 2016 ; CHASE et MEIKLEJOHN, 2016 ; ANANNY et CRAWFORD, 2018). Enfin, certain·es auteur·ices mentionnent la capacité de prendre en compte les retours utilisateurs comme un aspect permettant à un système d'apparaître plus "transparent" aux yeux des utilisateur·ices (ZARSKY, 2013 ; CARUANA *et al.*, 2015). Notons que, si certaines définitions proposées regroupent plusieurs des aspects cités ci-dessus (ITANI *et al.*, 2019 ; PÉGNY et IBNOUHSEIN, 2018), aucune ne semble capturer la totalité des aspects du concept de "transparence".

L'ensemble de ces contributions se retrouve autour de quelques idées clés, que les différent·es auteur·ices cités ci-dessus mettent en avant de manière plus ou moins claire et exhaustive, et que nous proposons de synthétiser en clarifiant une série d'exigences fondamentales. L'exigence de *compréhensibilité* met en avant l'importance que tout utilisateur·ice puisse cerner rapidement le fonctionnement et les résultats d'un système de classification. L'*interprétabilité* correspond à l'exigence que tout utilisateur·ice puisse, sans biais, tirer des conclusions et décider d'actions à entreprendre à partir des résultats présentés par le système. L'exigence de *retraçabilité*

met en avant l'importance de pouvoir retracer le déroulé des actions effectuées par un système à partir d'un résultat, ainsi que les données utilisées par ledit système. Cette exigence est d'autant plus importante dans un contexte applicatif tel que le nôtre, où les données sont extrêmement sensibles. Enfin, l'exigence de *révisabilité* met en avant l'importance de prendre en compte les retours des utilisateur·ices dans l'amélioration d'un système, afin que ceux-ci puissent mieux s'approprier le fonctionnement dudit système.

Cependant, toutes ces exigences ne sont pas associées à tous les éléments d'un système de classification, et elles n'appellent pas toutes à la même interprétation selon les éléments auxquels elles sont associées. De plus, nous cherchons ici à définir des critères de "transparence" invariants, applicables à toute utilisation pratique d'un système de classification. Ainsi, les éléments dépendant fortement de l'usage d'un système de classification, tel que la base d'apprentissage utilisée ou le classificateur produit, ne sont pas pris en compte dans notre définition de la "transparence". Nous nous concentrons donc principalement sur la "transparence" de l'algorithme d'apprentissage et du type de classificateur, car ces deux éléments permettent de définir un système de classification invariablement de son usage. Le Tableau 8.2 synthétise les différentes exigences en termes de "transparence" que nous nous sommes fixées, ainsi que les éléments d'un système de classification auxquels elles s'appliquent.

Exigences		S'applique à	
Nom	Définition	Algorithme d'apprentissage	Type de classificateur
Compréhensibilité	Nécessité d'avoir un fonctionnement rapidement cernable par tout utilisateur·ice	X	X
Interprétabilité	Nécessité de s'assurer que les conclusions des utilisateur·ices ne seront pas biaisées	X	X
Retraçabilité	Nécessité de connaître les actions effectuées par le système et les données utilisées	X	X
Révisabilité	Nécessité de pouvoir prendre en compte les retours utilisateurs dans l'amélioration du système		X

Table 8.2.: Synthèse de nos exigences en termes de "transparence", de leurs définitions et des éléments d'un système de classification auxquelles elles sont associées

Comme énoncé précédemment, l'exigence de *compréhensibilité* s'applique autant au processus d'un système de classification qu'à ses résultats. Ainsi, cette exigence s'applique à tous les éléments d'un système de classification, c'est-à-dire, à l'algorithme d'apprentissage et au type de classificateur utilisés. L'exigence d'*interprétabilité* s'applique aux différents processus impliqués dans la génération et la proposition de résultats aux utilisateur·ices. En effet, pour interpréter un résultat sans biais, un·e

utilisateur·ice doit être au courant de la manière dont ce résultat a été produit. Cette exigence s'applique donc aussi bien à l'algorithme d'apprentissage utilisé qu'au type de classificateur généré par celui-ci. Cependant, comme nous le verrons en section 8.2.2, l'*interprétabilité* doit être évaluée différemment pour ces deux éléments, principalement car leurs fonctions respectives au sein d'un système de classification diffèrent fortement. L'exigence de *retraçabilité* met en avant l'importance de pouvoir retracer le processus suivi et de pouvoir retrouver les données utilisées par un système de classification. Cette exigence s'applique donc autant à l'algorithme d'apprentissage utilisé qu'au type de classificateurs générés. Enfin, l'exigence de *révisabilité* met en avant le fait qu'un·e utilisateur·ice impliqué dans l'amélioration d'un système de classification, et comprenant comment ses retours sont pris en compte, percevra ledit système comme étant plus "transparent". L'utilisateur·ice n'ayant potentiellement accès qu'au classificateur, nous considérons que ce principe n'est applicable qu'au type de classificateur utilisé par un système de classification.

En nous basant sur les différentes exigences que nous nous sommes fixées à partir des différents aspects de la "transparence" provenant de la littérature, nous proposons la définition suivante.

Définition 8.1. *Un système de classification est considéré "transparent" si, et seulement si :*

- *l'algorithme d'apprentissage et le type de classificateur utilisés sont **compréhensibles** ;*
- *l'algorithme d'apprentissage et le type de classificateur utilisés sont **interprétables** ;*
- *l'algorithme d'apprentissage et le type de classificateur utilisés sont **retraçables** ;*
- *et le type de classificateur utilisé est **révisable**.*

8.2 Définition de critères opérationnels pour évaluer le respect de nos exigences

La définition d'exigences associées au concept de "transparence" permet la clarification de concepts abstraits, mais n'est pas suffisante pour déterminer si un système de classification peut être considéré comme "transparent" ou non dans la pratique. Pour cela, nous avons besoin de critères opérationnels, c'est-à-dire de propriétés concrètes dont nous pouvons vérifier si un système donné les satisfait ou non. Malheureusement, à notre connaissance, la littérature sur la "transparence" des

algorithmes manque de propositions de tels critères opérationnels. Cette section est consacrée à définir, pour chacune de nos exigences, les critères opérationnels qui nous permettront, par la suite, d'évaluer la "transparence" de divers systèmes de classification.

8.2.1 L'exigence de compréhensibilité

Comme énoncé en section 8.1, le concept de "transparence" peut souvent être résumé par "*donner des explications à l'utilisateur·ice*". Cependant, plusieurs auteur·ices soulignent aussi le fait que les explications données à l'utilisateur·ice doivent être *compréhensibles* par cette même utilisateur·ice (DORAN *et al.*, 2017 ; KIM et HINDS, 2006 ; SINHA et SWEARINGEN, 2002). Quand bien même un système de classification *C* fournirait des explications "compréhensibles" de ses résultats, si tout le fonctionnement de *C* n'est pas "compréhensible", il sera compliqué de qualifier *C* de "transparent". La *compréhensibilité* est donc une exigence importante dans la définition de la "transparence" d'un système de classification, et elle doit s'appliquer à tous les éléments composant ledit système.

Il nous faut donc déterminer un ou plusieurs critères opérationnels permettant de déterminer si un système de classification sera "compréhensible" par toutes ses utilisateur·ices potentiels. Pour ce faire, nous nous basons sur la définition proposée par MONTAVON *et al.*, (2018) : est "compréhensible" toute chose appartenant à un domaine auquel l'utilisateur·ice peut donner un sens. En d'autres termes, est "compréhensible" un système se basant sur des notions déjà connues des utilisateur·ices. Précisons que nous ne parlons pas de notions "faciles à comprendre", car cela supposerait une étape supplémentaire dans la compréhension du système. Les utilisateur·ices doivent connaître les notions sur lesquelles se base un système afin de pouvoir comprendre directement ledit système.

Dans notre cas, les utilisateur·ices étant des médecins, nous considérons que ceux-ci pourront donner un sens à toute notion étudiée durant leur scolarité et leurs études médicales. Afin de prendre en compte le fait que les notions vues par les médecins durant leurs études peuvent varier, nous proposons de nous assurer que les notions sur lesquelles repose un système sont bien déjà connues par les utilisateur·ices en nous focalisant sur le programme du tronc commun du concours de première année de médecine¹. Plus généralement, nous considérons comme "compréhensible" toute chose reposant sur des notions dont nous pouvons être sûrs qu'elles sont incluses dans le programme du cursus scolaire de toutes les utilisateur·ices potentiels.

1. www.enseignementsup-recherche.gouv.fr

Nous arrivons donc à la définition suivante d'un critère opérationnel pour évaluer la "compréhensibilité" d'un système de classification.

Définition 8.2. *Un algorithme d'apprentissage et un type de classificateur sont considérés comme **compréhensibles** si, et seulement si, ceux-ci se basent sur des notions incluses dans le programme du cursus scolaire de toutes les utilisateur·ices potentiels.*

Ainsi, un système de classification est basé sur un ensemble E de notions et les utilisateur·ices potentiels de ce système possèdent tous un ensemble S de notions étudiées durant leur cursus universitaire, tel que $E \cap S$ peut être vide. En nous basant sur notre Définition 8.2, nous voyons que la "compréhensibilité" d'un système de classification peut être évaluée sur un continuum allant de $E \cap S = \emptyset$ à $E \cap S = E$. Notons aussi que l'ensemble S peut évoluer et que, par conséquent, un système jugé "non compréhensible" pour un groupe d'utilisateur·ices peut être jugé "compréhensible" pour un autre groupe lors d'une évaluation ultérieure, et inversement.

8.2.2 L'exigence d'interprétabilité

Selon SPAGNOLLI *et al.*, (2017), l'objectif d'un système "transparent" est de s'assurer que les utilisateur·ices sont en capacité de prendre des décisions informées, sans biais, à partir des résultats fournis par le système. Néanmoins, être simplement "compréhensible" n'est pas suffisant pour prévenir toutes les mésinterprétations possibles des résultats par un·e utilisateur·ice. La "compréhensibilité" seule, définie en section 8.2.1, ne permet donc pas à un système de classification d'être considéré comme "transparent". Un système de classification, son fonctionnement, ses résultats, se doivent aussi d'être "interprétables" par l'utilisateur·ice. Il nous faut donc déterminer les critères opérationnels permettant d'évaluer si un système de classification est "interprétable" ou non par ses utilisateur·ices, au sein de l'exigence posée par SPAGNOLLI *et al.*, (2017).

Prenons l'exemple standard d'un système de classification dédié à la classification d'images (GUNNING, 2017). En pratique, l'utilisateur·ice utilise un classificateur produit par un algorithme d'apprentissage, et non directement un algorithme d'apprentissage. Ainsi, si un·e utilisateur·ice fournit au classificateur l'image d'un animal et que le résultat du classificateur est "c'est une image d'humain", l'utilisateur·ice peut légitimement demander au classificateur : "Pourquoi m'as-tu fourni ce résultat?". Nous avons alors deux possibilités :

1. Le classificateur a fourni un résultat correct ou attendu : l'utilisateur·ice souhaite alors en savoir plus sur les raisons sous-jacentes à ce résultat ;
2. Le classificateur a fourni un résultat incorrect ou inattendu : l'utilisateur·ice souhaite alors savoir pourquoi le classificateur ne lui a pas fourni le "bon" résultat.

Dans le premier cas, l'utilisateur·ice s'attend à des explications des étapes du raisonnement ayant conduit à un résultat spécifique. En fonction du type de classificateur, ces explications pourront prendre différentes formes, telles que : "parce que l'animal présenté a des vêtements, des cheveux et n'a pas de griffes" ou "parce que l'image présentée est similaire à ces autres images d'humains". On voit ici que l'explication se doit avant tout de fournir un raisonnement *compréhensible*, au sens de la Définition 8.2.

De plus, afin de prévenir de potentielles mésinterprétations, l'utilisateur·ice peut aussi se demander : "À quel point puis-je me fier à cette classification ?" et s'attendre à ce que le classificateur lui fournisse un risque d'erreur du résultat fourni, ce risque d'erreur pouvant varier d'un type de classificateur à un autre.

Dans le second cas, l'utilisateur·ice s'attend à avoir accès à l'ensemble du processus permettant au classificateur de classifier, cela afin de retracer le cheminement du classificateur et de détecter sur quel(s) point(s) le classificateur a pu, et peut encore, faire "fausse route". Cependant, le processus d'un classificateur peut être compliqué à appréhender. Le classificateur se doit donc d'être capable de fournir une version, ou une représentation, *compréhensible* de son processus. Enfin, face à une erreur du système de classification, l'utilisateur·ice peut aussi légitimement se demander : "Dans quelle mesure puis-je me fier à cet outil en général ?". Pour répondre à cette interrogation, le classificateur se doit de pouvoir fournir différentes métriques telles que : son taux d'erreurs, sa précision, sa sensibilité et sa spécificité. Ces métriques, expliquées plus précisément en Chapitre 9, sont généralement évaluées lors de la phase de test d'un classificateur, mais peuvent être mises à jour au fur et à mesure de l'utilisation du classificateur.

En nous basant sur l'ensemble des points abordés dans les paragraphes précédents, nous proposons la définition suivante des critères opérationnels permettant d'évaluer si un type de classificateur est "interprétable" ou non.

Définition 8.3. *Le type de classificateur utilisé par un système de classification est considéré comme **interprétable** si, et seulement si, il peut fournir aux utilisateur·ices :*

- des explications compréhensibles de ses résultats, incluant :

- le raisonnement, ou le processus, ayant conduit aux résultats fournis ;
- le risque d'erreur des résultats fournis ;
- une version compréhensible de son processus général ;
- des métriques générales telles que :
 - son taux d'erreur ;
 - sa précision ;
 - sa sensibilité ;
 - sa spécificité.

Cependant, une fois face au processus général d'un classificateur, ou simplement face aux explications d'un résultat, l'utilisateur·ice peut légitimement se demander : "Comment ce processus a-t-il été construit ? D'où vient-il ?". Un simple classificateur n'est pas en mesure de répondre à de telles interrogations. Son rôle est d'appliquer son processus pour classer des *entrées* fournies, pas de savoir comment son processus a été construit. À ces interrogations, seul l'algorithme d'apprentissage employé dans le système de classification peut y répondre, car son rôle est justement de construire des classificateurs. Cette distinction entre *interprétabilité* du classificateur et *interprétabilité* de l'algorithme d'apprentissage rejoint les réflexions de PÉGNY et IBNOUHSEIN, (2018) sur la transparence des algorithmes d'apprentissage, qui proposent la distinction entre "intelligibilité" de la procédure et "intelligibilité" de la sortie.

Ainsi, si un système de classification se doit d'utiliser un classificateur "interprétable", comme défini en Définition 8.3, il se doit aussi d'utiliser un algorithme d'apprentissage "interprétable". Le fait d'utiliser un algorithme d'apprentissage "interprétable" permet à l'utilisateur·ice de mieux comprendre l'outil qu'il utilise, ses forces et limitations, et ainsi prendre des décisions plus éclairées. Un premier moyen de déterminer si un algorithme d'apprentissage est "interprétable", serait d'évaluer si les utilisateur·ices peuvent aisément reproduire le processus de cet algorithme. Cependant, évaluer l'*interprétabilité* d'un algorithme de cette manière impliquerait une charge de travail supplémentaire pour les médecins, ce que nous souhaitons éviter. Nous avons donc recherché des critères opérationnels maximisant la l'*interprétabilité* d'un algorithme d'apprentissage.

Premièrement, il est difficile de considérer qu'un algorithme est reproductible si toutes ses étapes ne sont pas *compréhensibles* par les utilisateur·ices. Les étapes d'un algorithme "interprétable", ainsi que ses embranchements et sa condition d'arrêt, se doivent donc d'être *compréhensibles* dans le sens de la Définition 8.2. Deuxièmement, un algorithme peut être amené à utiliser diverses structures (ex. tableau,

matrices, etc.). Ces structures se doivent donc aussi d'être *compréhensibles* pour que l'on puisse qualifier l'algorithme "d'interprétable". Troisièmement, comme énoncé précédemment, nous considérons que plus un algorithme est aisément reproductible par un être humain, plus cet algorithme peut être considéré "interprétable". Par exemple, si l'algorithme exécute un ensemble de sous-processus en parallèle, cela rendra plus difficile sa reproductibilité par un être humain. Ainsi, l'algorithme ne doit présenter aucun sous-processus à exécuter en parallèle. Dans le même ordre d'idée, plus un algorithme possède d'embranchements et de sous-embranchements, plus sa reproductibilité par un être humain sera compliquée. Ainsi, pour être considéré "interprétable", un algorithme doit donc avoir un nombre d'embranchements aussi réduit que possible. Enfin, si l'exécution de certaines étapes de l'algorithme dépend de la valeur d'au moins une variable globale mise à jour au fur et à mesure de l'algorithme, cela implique de tenir à jour la valeur de cette variable tout au long de l'exécution de l'algorithme, ce qui complique encore une fois la reproductibilité dudit algorithme par un être humain. Ainsi, l'emploi de variables globales pour déterminer l'exécution, ou non, de certaines étapes de l'algorithme se doit d'être aussi réduit que possible.

Nous arrivons donc à la définition suivante des critères opérationnels nécessaires et suffisants pour évaluer si un algorithme d'apprentissage est "interprétable" ou non :

Définition 8.4. *L'algorithme d'apprentissage utilisé par un système de classification est considéré comme **interprétable** si, et seulement si :*

- *chacune de ses étapes est compréhensible ;*
- *les structures utilisées sont compréhensibles ;*
- *son processus est aussi aisément reproductible que possible par un être humain.*

Notons que l'emploi d'une expression telle que *aussi ... que possible* implique que nous ne pouvons dire d'un algorithme d'apprentissage qu'il est absolument "interprétable". Par corollaire, l'évaluation de l'*interprétabilité* d'un système de classification est, dans une certaine mesure, subjective et dépendante de ce que l'on peut considérer comme "possible" en termes de "reproductibilité" pour un algorithme d'apprentissage. Nous pouvons aussi noter que, définie ainsi, l'*interprétabilité* se base sur le concept de *compréhensibilité*, les informations à fournir imposées par l'*interprétabilité* se devant d'être compréhensibles. Ces deux notions ne sont donc pas totalement indépendantes.

8.2.3 L'exigence de retraçabilité

Une exigence supplémentaire à prendre en compte est la capacité d'un système de classification à retracer, à partir des résultats fournis, les données et opérations utilisées pour produire ces résultats. Notons que nous parlons ici de la capacité à pouvoir remonter les étapes de l'exécution d'un système, avec ou sans mise en place d'un système de collecte de traces d'activités dudit système. Comme avancé par HEDBOM, (2008), un·e utilisateur·ice (direct ou indirect) peut légitimement demander à savoir quelles données personnelles ont été utilisées par un système de classification, mais aussi comment ces données ont été utilisées et pourquoi. Ce point est d'autant plus important dans des contextes médicaux, où les données utilisées sont particulièrement sensibles.

Les critères de *compréhensibilité* et d'*interprétabilité* seuls ne suffisent pas à assurer la capacité à retracer, à partir des résultats, les données et opérations effectuées par un système de classification. Prenons, par exemple, un système de classification totalement *compréhensible* et *interprétable*. Si ce système a un fonctionnement dit "stochastique", c'est-à-dire si certaines des opérations qu'il effectue dépendent de variables aléatoires, il devient difficile de retracer les opérations effectuées à partir des résultats du système.

Le déterminisme, ou la non-stochastocité, de l'algorithme d'apprentissage et du type de classificateur utilisé apparaît comme un critère nécessaire et suffisant pour assurer la *retraçabilité* du processus d'un système de classification à partir d'un résultat. Nous proposons donc la définition suivante d'un critère opérationnel permettant d'évaluer la *retraçabilité* d'un système de classification :

Définition 8.5. *L'algorithme d'apprentissage et le type de classificateur définissant un système de classification sont considérés comme étant **retraçables** si, et seulement si, leurs processus ne sont pas stochastiques.*

Notons que, défini ainsi, l'exigence de *retraçabilité* est soit rempli, soit non rempli. L'utilisation d'un système de classification seulement *retraçable*, c'est-à-dire déterministe mais aucunement *compréhensible* ou *interprétable*, peut rendre compliqué, mais pas impossible, la tâche de retracer les opérations et données utilisées à partir des résultats. Ainsi, un système de classification *C* ne peut être considéré que "partiellement" *retraçable* si *C* ne remplit pas les critères de *compréhensibilité* et d'*interprétabilité*. Cependant, l'introduction de cette nuance rend l'exigence de *retraçabilité* fortement dépendante des exigences de *compréhensibilité* et d'*interprétabilité*.

8.2.4 L'exigence de révisabilité

Enfin, le concept de "transparence" d'un système de classification peut être associé à la capacité de ce système à prendre en compte les retours utilisateurs pour améliorer ses futurs résultats (ZARSKY, 2013). Lorsqu'un système de classification permet à ses utilisateur·ices de faire des retours sur les résultats du système, et que ces retours sont pris en compte par le système, ledit système apparaît généralement de moins en moins "opaque" à ses utilisateur·ices au fur et à mesure de son utilisation.

Prenons l'exemple de CARUANA *et al.*, (2015), qui ont étudié l'acceptabilité de différents types de classificateurs dans un contexte médical. Les auteur·ices ont rapporté qu'un classificateur basé sur un ensemble de règles était bien plus apprécié des utilisateur·ices qu'un classificateur basé sur un réseau de neurones. Et ce, malgré les meilleures performances du second lors des phases de test. Selon CARUANA *et al.*, (2015), la possibilité de pouvoir directement modifier les "mauvaises" règles apprises par le système de classification, est un élément décisif dans cette préférence des utilisateur·ices.

Cependant, tous les types de classificateurs ne sont pas directement modifiables par les utilisateur·ices. Un autre moyen de prendre en compte les retours utilisateurs est d'employer un algorithme d'apprentissage dit continu, ou *online*. Les algorithmes d'apprentissage sont dans leur majorité non continus, ou *offline*, mais tous peuvent être modifiés de manière plus ou moins aisée pour devenir des algorithmes d'apprentissage continus.

Notons que l'utilisation d'un apprentissage continu peut être considérée comme nettement moins "transparente" pour les utilisateur·ices qu'une modification directe du classificateur. Les classificateurs ne pouvant être modifiés que via un algorithme d'apprentissage continu ne sont alors considérés que comme "partiellement" *révisable*.

Nous proposons donc la définition suivante d'un critère opérationnel permettant d'évaluer la *révisabilité* d'un système de classification.

Définition 8.6. *Le type de classificateur utilisé par un système de classification est considéré comme étant **totale**ment révisable si, et seulement si :*

- *les utilisateur·ices peuvent directement modifier le processus du classificateur ;*

*et comme **partiel**lement révisable si, et seulement si :*

- *les retours utilisateurs peuvent être pris en compte par l'algorithme d'apprentissage pour modifier le processus du classificateur.*

Notons que, la révisabilité d'un système de classification dépend aussi de son contexte d'utilisation. De plus, nous pourrions argumenter qu'un système de classification n'étant ni *compréhensible*, ni *interprétable*, pourra difficilement être considéré comme "totalement", voire "partiellement", *révisable*. La *révisabilité* d'un type de classificateur n'est alors pas indépendante des critères de *compréhensibilité* et d'*interprétabilité* d'un système de classification.

8.3 Évaluation et méthode d'agrégation

Nous avons, dans la section 8.2, établi un ensemble E d'exigences que doit respecter un système de classification pour être considéré "transparent". Pour chaque exigence $e \in E$, nous avons défini un ensemble $C_e \subset C$ de critères opérationnels permettant d'établir si un système de classification respecte ou non l'exigence e , avec C l'ensemble de tous les critères opérationnels définis. De plus, pour toute paire d'exigences distinctes e_0 et e_1 de E , les deux ensembles de critères opérationnels associés C_{e_0} et C_{e_1} sont indépendants, en d'autres termes $\forall e_0, e_1 \in A : C_{e_0} \cap C_{e_1} = \emptyset$. Un système de classification s est alors considéré "transparent" si, et seulement si, la totalité des exigences en termes de "transparence" est respectée par s . Notre objectif est maintenant d'utiliser ces critères opérationnels pour évaluer la "transparence" de différents systèmes de classification existants afin de déterminer lequel, ou lesquels, utiliser dans notre outil d'aide à la décision. Nous sommes donc face à un problème de décision multi-critères (ROY et BOUYSSOU, 1993), où chaque système de classification représente une alternative $s \in S$ à évaluer à partir de chacune des exigences $e \in E$, avec S l'ensemble des systèmes à évaluer, et où chaque exigence $e \in E$ est à évaluer à partir de chacun des critères opérationnels associés $c \in C_e$.

Comme relevé par GUITOUNI et MARTEL, (1998) dans leur guide sur le choix d'une méthode d'aide à la décision multi-critères, une fois établis l'approche et les critères, deux points sont importants à prendre en compte pour évaluer des alternatives. Premièrement, la manière dont est évaluée indépendamment chaque exigence $e \in E$, qui va varier en fonction des critères opérationnels associés $c \in C_e$. Deuxièmement, la manière dont sont agrégées les évaluations des différentes exigences $e \in E$ pour établir une évaluation globale de chaque système $s \in S$.

Une exigence $e \in E$ peut, comme développé en section 8.2, aussi bien dépendre d'un seul critère que de plusieurs. Nous avons déjà mis en évidence que certaines exigences, comme celle de *compréhensibilité* ou celle de *révisabilité*, peuvent être considérées comme "totalement" respectées ou "partiellement" respectées. De plus,

pour l'exigence d'*interprétabilité*, nous avons établi plusieurs critères opérationnels à respecter, par le type de classificateur ou par l'algorithme d'apprentissage utilisés. Dans ce cas, il se peut que seul un sous-ensemble de critères soit validé et donc que l'exigence associée ne soit que "partiellement" respectée.

Ainsi, nous considérons que l'évaluation d'une exigence $e \in E$ peut se résumer à répondre à la question : est-ce que le système de classification $s \in S$ remplit l'ensemble des critères $c \in C_e$ nécessaires au respect de l'exigence $e \in E$? En fonction du nombre de critères opérationnels $c \in C_e$ validés pour une exigence $e \in E$, nous considérons que l'évaluation de ladite exigence e varie entre "pas du tout" respectée, si aucun critère n'est validé, et "totalement" respectée, si tous les critères sont validés. Dans le cas où une majorité de critères opérationnels est validée, mais pas la totalité, nous considérons que l'exigence e n'est que "partiellement" respectée. Enfin, dans le cas où une minorité de critères est validée, nous considérons que l'exigence e n'est "pas vraiment" respectée. Compte tenu du caractère subjectif de certains critères opérationnels que nous avons définis, et du concept de "transparence" en général, nous considérons important de garder des valeurs sémantiques à nos évaluations plutôt que des valeurs numériques. La Figure 8.2 présente la grille d'évaluation que nous utilisons par la suite pour évaluer le respect de chaque exigence $e \in E$ par différents systèmes de classification.

	<i>Pas du tout</i>	<i>Pas vraiment</i>	<i>Partiellement</i>	<i>Totalement</i>
Système Compréhensible ?				
Classificateur Interprétable ?				
Algorithme Interprétable ?				
Système Traçable ?				
Classificateur Révisable ?				

Figure 8.2.: Grille d'évaluation des différentes exigences en termes de "transparence" à respecter pour un système de classification

Maintenant que nous avons défini la manière dont nous allons évaluer si les différentes exigences en termes de "transparence" que nous nous sommes fixées sont

respectées par un système de classification donné, il nous faut définir la manière dont nous allons agréger ces évaluations pour établir si ledit système de classification peut être considéré "transparent". Les méthodes d'agrégation peuvent se distinguer en trois catégories (GUITOUNI et MARTEL, 1998) :

- Compensatoires : méthodes supposant qu'une bonne évaluation de n'importe quelle exigence peut compenser une mauvaise évaluation de n'importe quelle autre exigence ;
- Partiellement compensatoires : méthodes supposant qu'une bonne évaluation de certaines exigences peut compenser une mauvaise évaluation de certaines autres exigences ;
- Non compensatoires : méthodes supposant qu'aucune bonne évaluation de n'importe quelle exigence ne peut compenser une mauvaise évaluation de n'importe quelle autre exigence.

Un exemple classique de méthode d'agrégation compensatoire est la moyenne pondérée des évaluations effectuées. La méthode d'agrégation proposée par BÉRANGER, (2015), dans son modèle d'évaluation de l'éthique des projets informatiques en contexte médical, peut aussi être considérée comme une méthode compensatoire.

Dans notre cas, nous cherchons un système de classification respectant la totalité de nos exigences en termes de "transparence". Nous ne pouvons pas considérer qu'une bonne évaluation pour une de nos exigences puisse compenser une mauvaise évaluation pour n'importe quelle autre de nos exigences. Il nous faut donc nous orienter vers une méthode d'agrégation non compensatoire telle que, par exemple, Electre Tri (ROY et BOUYSSOU, 1993) ou MR-Sort (BOUYSSOU et MARCHANT, 2007).

Toutefois, il nous paraît possible de nous passer d'une méthode d'agrégation sophistiquée, car nous recherchons simplement un système respectant au moins "partiellement" la totalité de nos exigences. De plus, nos exigences ne sont pas totalement indépendantes, une bonne évaluation sur une exigence pouvant améliorer l'évaluation d'une autre exigence, et inversement. Nous proposons donc une méthode d'évaluation générale consistant simplement à considérer comme "transparent" tout système de classification respectant une majorité, si ce n'est la totalité, des exigences que nous nous sommes fixés en termes de "transparence". Cependant, si une seule de nos exigences est évaluée comme "pas vraiment" ou "pas du tout" respectée, alors le système ne sera pas considéré "transparent". De cette manière, seuls les systèmes respectant une majorité de nos exigences en termes de "transparence", au moins "partiellement", seront envisagés pour être utilisés dans notre outil d'aide à la décision.

Sélection d'un système de classification répondant à nos exigences de "transparence"

” *Legends of prediction are common throughout the whole Household of Man. Gods speak, spirits speak, computers speak. Oracular ambiguity or statistical probability provides loopholes, and discrepancies are expunged by Faith.*

— Ursula K. Le Guin
The Left Hand of Darkness

Nous avons vu au Chapitre 8 les différentes exigences que se doit de satisfaire, selon nous, un système de classification pour être considéré "transparent". Associée à l'approche "ajustive" que nous proposons pour aider les médecins dans leurs consultations (cf. Chapitre 4 et Chapitre 7), cette "transparence" permettrait de maximiser l'acceptation de l'outil que nous souhaitons développer. Dans ce chapitre, nous cherchons donc à identifier s'il existe au moins un système de classification qui remplit nos exigences en termes de "transparence". Si tel est le cas, nous pourrions l'utiliser dans notre outil d'aide à la décision en consultation.

Dans la Section 9.1, nous évaluons la "transparence" de différents systèmes de classification pouvant être utilisés pour résoudre des problèmes de classification multi-labels. Ces évaluations nous amènent à la conclusion que l'algorithme Naive Bayes, de par l'utilisation de notions connues des médecins et de par sa simplicité, correspond au type d'algorithme que nous recherchons. Cependant, l'algorithme Naive Bayes est historiquement dédié à la résolution de problèmes mono-label. Dans la Section 9.2, nous présentons une variante de l'algorithme de Naive Bayes adaptée à la résolution de problèmes de classification multi-labels. Enfin, bien qu'il nous faille proposer un outil "transparent", il nous faut aussi nous assurer que cette "transparence" n'implique pas de faibles performances en classification. Dans la Section 9.3, nous analysons les performances des divers systèmes de classification évalués, ainsi

que la variante de l'algorithme Naive Bayes proposée, et ce sur différentes bases de données multi-labels.

9.1 Évaluation de la "transparence" de divers systèmes de classification

Dans cette section, nous utilisons les critères opérationnels proposés au Chapitre 8 afin d'évaluer la "transparence" des systèmes de classification applicables à notre cas. Plusieurs systèmes de classification pouvant se baser sur une même approche, nous ne nous concentrons donc sur un système de classification reconnu dans la littérature par approche. Ainsi, nous commençons par évaluer les systèmes de classification BP-MLL (ZHANG et ZHOU, 2006) et ML-kNN (ZHANG et ZHOU, 2007), tous deux dédiés au traitement de problèmes multi-labels. Nous évaluons ensuite les systèmes Naive Bayes (JOHN et LANGLEY, 1995), basé sur la théorie des probabilités, C4.5 (QUINLAN, 1993), générant des arbres de décision, RIPPER (COHEN, 1995), générant des ensembles de règles, et SMO (PLATT, 1998 ; KEERTHI *et al.*, 2001), générant des machines à vecteur de support.

9.1.1 Classification par réseau de neurones

Commençons par l'évaluation du système de classification *Backpropagation for Multi-Label Learning* (BP-MLL) (ZHANG et ZHOU, 2006). Ce système de classification est basé sur un réseau de neurones artificiels adapté au traitement de problèmes de classification multi-labels. Un réseau de neurones artificiels, comme introduit par KOHONEN, (1988) et ZURADA, (1992), est un ensemble de nœuds connectés entre eux selon une certaine architecture, généralement en plusieurs "couches", pouvant varier selon le type de problème à résoudre. Chaque nœud, ou "neurone", du graphe est associé à une fonction d'activation et chaque lien, ou "synapse", est associé à un poids w . Si la somme des poids des synapses actives arrivant sur un neurone permet l'activation dudit neurone, alors les synapses partant de ce neurone seront activées et ainsi de suite pour chaque neurone du réseau. L'objectif de l'algorithme d'apprentissage de ce type de système de classification est de trouver, par un processus dit de *back-propagation*, l'ensemble des poids permettant de minimiser l'erreur du réseau sur une tâche donnée. Le classificateur issu de cet apprentissage est donc un réseau de neurones dit "entraîné", avec un ensemble de poids optimisés pour accomplir la tâche pour laquelle le réseau a été entraîné.

La Figure 9.1 montre l'architecture du réseau de neurones utilisée par BP-MLL. Celui-ci reprend une architecture assez classique, avec en entrée une couche de d neurones, correspondant aux attributs $\{a_1, a_2, \dots, a_d\}$ présents dans chaque instance de la base d'apprentissage, et en sortie une couche de Q neurones, correspondant aux labels $\{c_1, c_2, \dots, c_Q\}$ dont nous cherchons à anticiper les valeurs. Le réseau de neurones de BP-MLL ne contient qu'une seule couche intermédiaire, aussi appelée couche "cachée", composée de M neurones. À la couche d'entrée, ainsi qu'à la couche intermédiaire, est ajouté un neurone de "biais" permettant de générer du bruit pour éviter un potentiel sur-apprentissage.

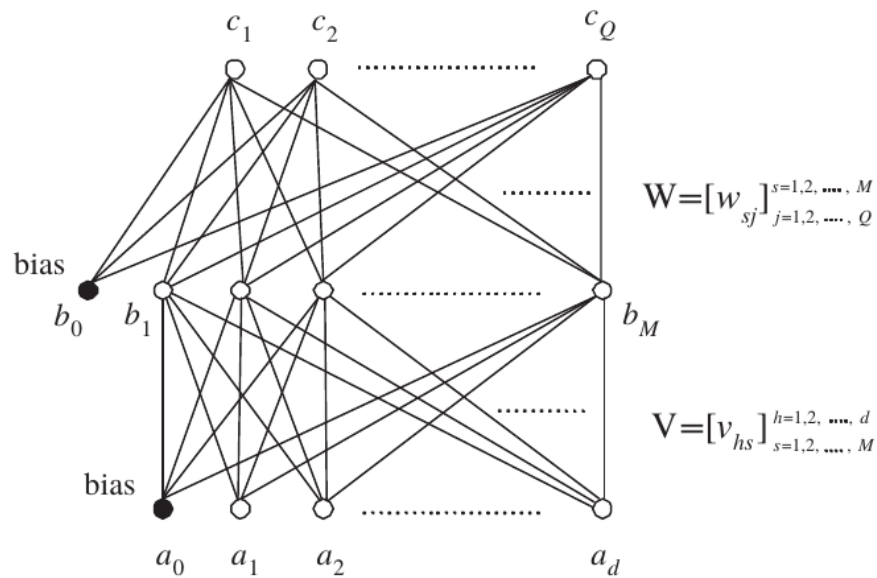


Figure 9.1.: Architecture du réseau de neurones BP-MLL, représentation issue de ZHANG et ZHOU, (2006)

Si nous nous référons au programme du tronc commun des universités de médecine¹, les notions sur lesquelles se basent les réseaux de neurones, en particulier la théorie des graphes, ne font pas partie du domaine de connaissances de toutes nos utilisatrices potentielles. Ainsi, en accord avec les critères énoncés au Chapitre 8, nous considérons le système de classification BP-MLL comme n'étant *pas du tout compréhensible*.

Concernant l'*interprétabilité* de BP-MLL, il nous faut faire la distinction entre son algorithme d'apprentissage, qui est un algorithme dit de *backpropagation* mettant à jour les poids d'un réseau "non entraîné", et le type de classificateur produit par cet algorithme, qui est un réseau de neurones "entraîné".

1. www.enseignementsup-recherche.gouv.fr

Pour ce qui est de l'*interprétabilité* du classificateur, bien que des travaux sont en cours pour extraire le processus de classification d'un réseau de neurones (DOŠILOVIĆ *et al.*, 2018 ; SAMEK *et al.*, 2017), ces recherches sont encore à un stade précoce et nécessitent de plus amples investigations. Nous ne préférons donc pas nous appuyer sur ces travaux. Pour ce qui est de BP-MLL, les réseaux de neurones produits ne permettent pas de fournir des explications *compréhensibles*, que ce soit pour un résultat spécifique ou pour leur processus général. Cependant, il est possible de fournir le risque d'erreur d'un résultat, le risque d'erreur général des réseaux de neurones produits, ainsi que leur précision, sensibilité et spécificité. Seule une minorité des critères définis comme nécessaires à l'*interprétabilité* d'un classificateur sont donc remplis. Ainsi, nous considérons le type de classificateur utilisé par BP-MLL comme n'étant *pas du tout interprétable*.

L'algorithme d'apprentissage utilisé par BP-MLL se base sur une mise à jour des poids du réseau de neurones par *backpropagation*. BP-MLL ne présente pas d'embranchements ni de sous-processus à exécuter en parallèle, mais nécessite la mise à jour d'une variable d'erreur globale afin de déterminer quand arrêter l'algorithme. Enfin, BP-MLL est basé sur succession de calculs matriciels ainsi que d'autres notions mathématiques difficilement qualifiables de *compréhensibles*, dans le sens défini en Définition 8.2. Il en est de même pour les structures utilisées : un réseau généralement représenté par une matrice de poids peut être difficile à interpréter. Ainsi, une majorité des critères de la Définition 8.4 n'étant rempli, nous considérons l'algorithme d'apprentissage utilisé par BP-MLL comme n'étant *pas du tout interprétable*.

Pour ce qui est de la *traçabilité* des résultats, l'algorithme d'apprentissage et le type de classificateur utilisé par BP-MLL sont tous deux totalement déterministes. Cependant, du fait de leur faible *compréhensibilité* et *interprétabilité*, nous ne pouvons considérer BP-MLL que comme *partiellement retraçable*.

Enfin, il est difficile de considérer un réseau de neurones entraîné comme directement modifiable par un·e utilisateur·ice, principalement à cause de sa faible *interprétabilité*. De plus, s'il est possible d'entraîner un réseau de neurones de manière continue, cette approche est très peu préconisée car les réseaux de neurones ont une tendance à facilement "oublier" leurs acquis. Ce problème est connu sous le nom de *vanishing gradient problem* (HOCHREITER, 1998). Ainsi, nous considérons le type de classificateur utilisé par BP-MLL comme n'étant *pas du tout révisable*.

9.1.2 Classification par plus proches voisins

Considérons maintenant le système de classification *Multi-Label k-Nearest Neighbors* (ML-kNN) (ZHANG et ZHOU, 2007). ML-kNN est une extension de l'algorithme des k plus proches voisins dédié au traitement de problèmes multi-labels. L'algorithme des k plus proches voisins, sur lequel se base ML-kNN, utilise principalement des calculs de distance entre les différentes entrées. Nous pouvons supposer que la notion de "distance" fait partie du champ de connaissances de tout·e utilisateur·ice potentiel. ML-kNN se base en plus sur des calculs avancés de probabilités de co-occurrence entre chaque label, afin de fournir une classification plus précise. Bien que la théorie des probabilités fasse partie du cursus universitaire des médecins¹, les notions avancées utilisées par ML-kNN, comme les calculs des probabilités de co-occurrence des labels, n'appartiennent pas nécessairement au champ de connaissances de tou·tes les utilisateur·ices potentiels. Ainsi, en accord avec la Définition 8.2, nous considérons le système de classification ML-kNN comme n'étant que *partiellement compréhensible*.

Le type de classificateur produit par ML-kNN est un ensemble de *centroids* et un ensemble de probabilités. Une nouvelle entrée sera affectée au *centroid* le plus "proche" et les labels seront déterminés selon les probabilités associées. Ainsi, le classificateur peut potentiellement fournir une explication telle que : "La·e patient·e X est similaire à ces autres exemples de patient·es. Je vous fournis donc des informations similaires concernant X". Pour ce qui est de son processus général, il existe des manières d'expliquer graphiquement les distances et les probabilités utilisées. Cependant, comme dit précédemment, les calculs de probabilités que présenterais ML-kNN ne sont pas nécessairement *compréhensibles* pour tout·e utilisateur·ice potentiel. Ceci étant dit, un classificateur produit par ML-kNN peut fournir le risque d'erreur d'un résultat, ainsi que son risque d'erreur global, sa précision, sa sensibilité et sa spécificité. Ainsi, nous considérons le type de classificateur utilisé par ML-kNN comme n'étant *partiellement interprétable* par les utilisateur·ices.

L'algorithme d'apprentissage présente peu d'embranchements et de sous-processus parallèles. ML-kNN ne nécessite pas non plus l'emploi d'une variable globale pour déterminer l'exécution d'une ou plusieurs actions. Les entrées de la base d'apprentissage sont traitées de la même manière les unes après les autres. Les conditions d'arrêt sont *compréhensibles*, et il en est de même pour les structures utilisées qui sont, dans le pire des cas, des ensembles de probabilités. Cependant, ML-kNN se base à chaque étape sur des calculs avancés de probabilités, rendant les différentes étapes de l'algorithme peu *compréhensibles*. Ainsi, nous considérons l'algorithme d'apprentissage utilisé par ML-kNN comme n'étant *partiellement interprétable*.

Concernant la *traçabilité* des résultats de ML-kNN, il est à noter que l'algorithme des plus proches voisins est généralement initialisé de manière aléatoire. Les *centroids* initiaux étant déterminés en choisissant aléatoirement des entrées dans la base d'apprentissage. Il est possible d'éviter cela en choisissant les extrema de la base d'apprentissage, mais ceci n'est pas applicable à tous les cas. Mise à part l'initialisation, tout le processus de l'algorithme d'apprentissage est déterministe et il en est de même pour le type classificateur. À notre connaissance, l'algorithme ML-kNN se base sur une version standard de l'algorithme des plus proches voisins. Ainsi, nous considérons le système de classification ML-kNN comme n'étant *pas vraiment retraceable*.

Enfin, bien qu'il soit difficile de considérer le type de classificateur utilisé par ML-kNN comme modifiable directement par les utilisateur-ices, l'utilisation d'un apprentissage continu est parfaitement envisageable. L'algorithme d'apprentissage de ML-kNN peut aisément mettre à jour le classificateur au fur et à mesure de l'acquisition de nouvelles entrées. Ainsi, nous considérons le type de classificateur utilisé par ML-kNN comme *partiellement révisable*.

9.1.3 Classification probabiliste

Maintenant que nous avons évalué les systèmes de classification spécifiquement dédiés aux problèmes de classification multi-labels, nous pouvons évaluer les systèmes de classification mono-label. Il est possible d'utiliser des systèmes de classification mono-label sur de problèmes multi-labels en utilisant des méta-algorithmes tels que *Random k-Labelsets* (RAkEL) (TSOUMAKAS *et al.*, 2011b).

Commençons par le système de classification Naïve Bayes (JOHN et LANGLEY, 1995). Le fonctionnement de son algorithme d'apprentissage consiste à parcourir la base d'apprentissage pour estimer un ensemble de probabilités. On note les probabilités $P(C = c)$, qu'une instance de la base d'apprentissage soit de la classe c , et les probabilités $P(X = x|C = c)$ qu'une variable X ait pour valeur x sachant que l'instance est de la classe c . Pour les variables "numériques", c'est-à-dire prenant leurs valeurs dans un domaine continu, l'algorithme suppose que celles-ci suivent une loi normale et tente de l'estimer à partir de la moyenne et de l'écart-type des valeurs observées. Une fois cet ensemble de probabilités calculé, il sert de base au classificateur pour estimer la probabilité $P(C = c|X')$ qu'une nouvelle entrée X' soit de la classe c , cela en appliquant une version "naïve" du théorème de Bayes qui suppose chaque variable X comme indépendante (cf. Équation (9.5)).

Si $P(C = c|X') > P(C \neq c|X')$, alors le classificateur déterminera que l'entrée X' appartient à la classe c .

Le système de classification Naive Bayes se base ainsi exclusivement sur des notions de base de la théorie des probabilités : le calcul de fréquences et le théorème de Bayes. Ces notions mathématiques font partie du programme commun des facultés de médecine¹. De plus, ce sont des notions régulièrement utilisées par les médecins, lors de recherches cliniques pour un nouveau traitement par exemple. Ainsi, nous considérons le système de classification Naive Bayes comme étant *totale-ment compréhensible* par les utilisateur·ices.

Le type de classificateur utilisé par Naive Bayes est composé de plusieurs ensembles de fréquences calculées à partir de la base d'apprentissage, donc des approximations de probabilités. Le processus général du classificateur peut alors être représenté à l'aide d'un réseau bayésien. De plus, ce type de classificateur peut fournir, avec un résultat, une explication telle que : "Sachant que la·e patient·e est suivi pour Hypercholestérolémie, la probabilité que vous ayez besoin de telle information est de X%". Un calcul plus ou moins détaillé peut aussi faire office d'explication *compréhensible* pour des médecins. Ainsi, nous considérons le type de classificateur utilisé par Naive Bayes comme étant *totale-ment interprétable* par les utilisateur·ices.

Concernant l'algorithme d'apprentissage utilisé par Naive Bayes, celui-ci consiste simplement à parcourir la base d'apprentissage en une fois afin de mettre à jour des fréquences d'occurrences utiles. Il n'y a donc aucun embranchement, ni sous-processus à exécuter en parallèle, ni variable globale à tenir à jour pour savoir si une étape doit être exécutée ou non. Les différentes étapes ainsi que les structures utilisées sont parfaitement *compréhensibles* par des médecins. Ainsi, nous considérons l'algorithme d'apprentissage utilisé par le système de classification Naive Bayes comme étant *totale-ment interprétable*.

Bien que se basant sur la théorie des probabilités, l'algorithme d'apprentissage et le type de classificateur utilisé par Naive Bayes sont tous deux parfaitement déterministes. Ainsi, nous considérons Naive Bayes comme étant *totale-ment retraçable*.

Enfin, il paraît compliqué d'envisager qu'un·e utilisateur·ice puisse mettre directement à jour des fréquences calculées par un algorithme. Cependant, nous pouvons parfaitement envisager la mise à jour de ces fréquences, et donc du classificateur, via l'algorithme d'apprentissage au fur et à mesure de l'acquisition de nouvelles entrées. Ainsi, nous considérons le type de classificateur utilisé par Naive Bayes comme *partiellement révisable*.

9.1.4 Classification par arbre de décision

Évaluons maintenant la "transparence" du système de classification C4.5 (QUINLAN, 1993). C4.5 est un système basé sur l'induction et l'utilisation d'arbres de décision. Pour ce faire, l'algorithme d'apprentissage de C4.5 parcourt les entrées de la base d'apprentissage et tente de déterminer les conditions permettant de subdiviser ces entrées en différents sous-ensembles pour que toutes les entrées associées appartiennent à la même classe. Ces sous-ensembles deviennent alors les "feuilles" de l'arbre de décision final. Jusqu'à obtenir ce résultat, l'algorithme d'apprentissage de C4.5 calcule, pour chaque "feuille" f transitoire, le gain d'information normalisé de chaque attribut de la base d'apprentissage. Un nouveau nœud est alors créé en divisant, généralement en deux, le domaine de l'attribut avec le plus grand gain d'information normalisé. Les entrées du sous-ensemble de f sont alors réparties selon cette nouvelle condition, et ainsi de suite. L'arbre de décision ainsi obtenu est utilisé comme classificateur.

Le système de classification C4.5 se base principalement sur deux notions : les arbres de décision et le gain d'information, notion connexe à l'entropie de Shannon (SHANNON, 1948). Comme nous l'avons vu et développé au Chapitre 2, les médecins sont habitués à consulter des "guides de bonnes pratiques" souvent représentés sous forme d'arbres de décision. Les arbres de décisions font donc partie du champ de connaissances de nos utilisateur·ices potentiels. L'entropie de Shannon est, quant à elle, une notion issue du domaine des théories de l'information et de la communication. C'est une notion qui ne fait donc pas partie du cursus universitaire de toutes nos utilisateur·ices potentiels¹. Ainsi, nous ne considérons le système de classification C4.5 que comme *partiellement compréhensible*.

Le type de classificateur utilisé par C4.5 est composé d'arbres de décision, qui sont, comme mentionné précédemment, totalement compréhensibles par toutes les utilisateur·ices. Lors de la proposition d'un résultat, un arbre de décision peut fournir le cheminement et les embranchements l'ayant mené à ce résultat, ainsi que le risque d'erreur de ce résultat. De même, l'entièreté d'un arbre de décision, en d'autres termes le processus général du classificateur, peut être fournie aux utilisateur·ices. Enfin, le risque d'erreur général du classificateur, sa précision, sensibilité et spécificité sont calculés lors de la phase de test du classificateur et peuvent donc être fournis à l'utilisateur·ice. Ainsi, nous considérons le type de classificateur utilisé par C4.5 comme *totalement interprétable* par les utilisateur·ices.

L'algorithme d'apprentissage utilisé par C4.5 ne présente pas d'embranchements, ni de sous-processus à exécuter en parallèle. Cependant, certaines étapes et conditions

sont basées sur l'entropie de Shannon qui peut difficilement être considérées comme *compréhensibles*. De plus, l'exécution de l'algorithme, en particulier sa condition de fin, nécessite la mise à jour d'une variable globale. Les structures d'arbres binaires utilisées par l'algorithme sont aisément compréhensibles, même pour un non-informaticien. Ainsi, nous considérons l'algorithme d'apprentissage utilisé par C4.5 comme n'étant que *partiellement interprétable*.

Pour ce qui est de la traçabilité des résultats fournis par C4.5, le type de classificateur et l'algorithme d'apprentissage utilisés sont tous deux totalement déterministes. Du fait de la relativement bonne *compréhensibilité* et *interprétabilité* de C4.5, nous considérons C4.5 comme *totalement retraçable*.

Enfin, les arbres de décision produits et utilisés par C4.5 sont des classificateurs directement modifiables par les utilisateur·ices. Nous considérons donc le type de classificateur utilisé par C4.5 comme *totalement révisable*.

9.1.5 Classification par système de règles

Le système de classification *Repeated Incremental Pruning to Produce Error Reduction* (RIPPER), est un système proposé par COHEN, (1995) permettant la génération d'ensembles de règles. L'algorithme d'apprentissage de RIPPER fonctionne en trois étapes. Premièrement, l'algorithme tente de faire "croître" les règles déjà existantes en y ajoutant une condition sur l'attribut avec le plus fort gain d'information, générant de nouvelles règles par la même occasion. Deuxièmement, l'algorithme "élague" les règles existantes des sous-conditions qui n'influent pas le gain d'information de la règle. En d'autres termes, si une règle a les mêmes performances avec ou sans une sous-condition, alors cette sous-condition est enlevée de la règle. Enfin, l'algorithme "optimise" les sous-conditions des règles existantes pour minimiser l'erreur de l'ensemble des règles. Ainsi, l'algorithme d'apprentissage de RIPPER génère un ensemble de règles utilisé en tant que classificateur. Le fait de faire "croître" puis "élaguer" les règles de manière répétée permet, selon COHEN, (1995), d'obtenir un ensemble restreint de règles courtes.

Le système de classification RIPPER se base donc sur la notion de règles de type SI ... ALORS Nous pouvons supposer que, tout comme le concept d'arbre de décision, le concept de règles fait partie du champ de connaissances de tout·e utilisateur·ice potentiel. Cependant, tout comme C4.5, RIPPER se base aussi sur les notions d'entropie de Shannon et de gain d'information pour déterminer comment modifier les règles existantes. Nous avons vu que ces notions, issues de la théorie

de l'information, ne font pas partie du cursus universitaire de tout·e utilisateur·ice potentiel¹. Ainsi, nous considérons le système de classification RIPPER comme n'étant que *partiellement compréhensible*.

Le type de classificateur utilisé par RIPPER est constitué d'un ensemble de règles. Comme énoncé précédemment, la notion de règles est supposée *compréhensible*, mais un large ensemble de règles peut être compliqué à analyser. Cependant, RIPPER utilise plusieurs heuristiques pour minimiser la complexité des règles produites tout en maximisant les performances. RIPPER peut donc fournir les règles impliquées dans la production d'un résultat spécifique, tout comme l'ensemble des règles composant le classificateur. Le classificateur peut aussi fournir le risque d'erreur d'un résultat, son risque d'erreur global, ainsi que sa précision, sa sensibilité et sa spécificité. Ainsi, nous considérons le type de classificateur utilisé par le système de classification RIPPER comme étant *totalement interprétable* par les utilisateur·ices.

L'algorithme d'apprentissage utilisé par RIPPER présente peu d'embranchements et pas de sous-processus à exécuter en parallèle. Certaines étapes de l'algorithme se basent sur la notion d'entropie et sont donc peu *compréhensibles* par des utilisateur·ices, de même que la condition d'arrêt utilisée qui dépend aussi d'une variable globale. Les structures utilisées, telles que les ensembles de règles, sont considérées comme assez *compréhensibles*, même pour des non-informaticiens. Ainsi, nous considérons l'algorithme d'apprentissage utilisé par le système de classification RIPPER comme n'étant que *partiellement interprétable*.

Concernant la traçabilité des résultats produits par RIPPER, l'algorithme d'apprentissage et le type de classificateur sont tous deux totalement déterministes. Du fait de la relativement bonne *compréhensibilité* et *interprétabilité* de RIPPER, nous pouvons considérer RIPPER comme étant *totalement retraçable*.

Enfin, un système de règles est parfaitement modifiable directement par les utilisateur·ices. Ainsi, nous considérons le type de classificateur utilisé par RIPPER comme étant *totalement révisable*.

9.1.6 Classification par vecteurs de support

Pour finir, évaluons le système de classification *Sequential Minimal Optimization* (SMO) (PLATT, 1998 ; KEERTHI *et al.*, 2001). SMO se base sur la génération de machines à vecteurs de support, ou *Support Vector Machine* (SVM), permettant la classification de nouvelles entrées. Les SVM sont un ensemble de fonctions mathématiques, aussi appelé "*kernel*", permettant de séparer le domaine d'une ou plusieurs

variables. Utilisées ensemble, ces fonctions permettent d'estimer à quelle classe appartient une nouvelle entrée. L'algorithme d'apprentissage de SMO se base sur l'approximation des multiplicateurs de Lagrange (ROCKAFELLAR, 1993) permettant d'optimiser le *kernel* de fonctions. Pour ce faire, l'algorithme sélectionne les multiplicateurs de Lagrange ne respectant pas les conditions Karush-Kuhn-Tucker (KARUSH, 2014) afin de les optimiser, et ainsi de suite jusqu'à obtenir une solution satisfaisant toutes les conditions Karush-Kuhn-Tucker.

Le système de classification SMO est donc basé sur des notions mathématiques assez avancées ne faisant pas nécessairement partie du cursus universitaire de tout·e utilisateur·ice potentiel. Ainsi, nous considérons que le système de classification SMO n'est *pas vraiment compréhensible*.

Le type de classificateur utilisé par SMO est un ensemble de fonctions permettant de séparer les entrées en fonction de leur classe. SMO peut alors, pour un résultat spécifique, fournir une explication de type "Résultat fourni car $X > f(x)$ " ou fournir les fonctions impliquées dans le résultat. SMO peut aussi fournir l'ensemble des fonctions et zones utilisées par le classificateur. Cependant, l'agrégation de fonctions peut être difficilement *compréhensible* par des utilisateur·ices non informaticien·nes ou non mathématicien·nes. Le classificateur utilisé par SMO est tout de même capable de fournir le risque d'erreur d'un résultat, son risque d'erreur global, sa précision, sa sensibilité et sa spécificité. Ainsi, nous considérons le type de classificateur utilisé par SMO comme n'étant *pas vraiment interprétable* par les utilisateur·ices.

Concernant l'algorithme d'apprentissage utilisé par SMO, celui-ci ne présente pas de sous-processus à exécuter en parallèle, mais contient certains embranchements dont l'exécution dépend de l'état des variables utilisées. Les étapes de l'algorithme font appel à des notions mathématiques avancées telles que les multiplicateurs de Lagrange et les conditions Karush-Kuhn-Tucker. Les structures utilisées restent assez simples : variables et fonctions. Ainsi, nous considérons l'algorithme d'apprentissage utilisé par SMO comme n'étant *pas vraiment interprétable*.

L'algorithme d'apprentissage et le classificateur utilisés par SMO sont tous deux totalement déterministes. Cependant, du fait de la faible *compréhensibilité* et *interprétabilité* de SMO, nous considérons SMO comme n'étant que *partiellement retraçable*.

Enfin, bien que les fonctions permettant le processus du classificateur soient difficilement modifiables directement par des utilisateur·ices, il est possible de rendre continu l'algorithme d'apprentissage utilisé par SMO, comme le montrent TAX et LASKOV, (2003) par exemple. Cette modification semble cependant plus compliquée

que pour ML-kNN ou Naive Bayes. Ainsi, nous considérons le type de classificateur utilisé par SMO comme n'étant *pas vraiment révisable*.

9.1.7 Résumé des résultats

Résumons en Figure 9.2 les résultats de nos évaluations. Comme détaillé au Chapitre 8, un système de classification sera considéré "transparent" si une majorité d'exigences en termes de "transparence" sont *totalement* respectées. Cependant, si une seule de nos exigences n'est *pas vraiment* ou *pas du tout* respectée, alors le système de classification ne sera pas considéré "transparent". Si aucun des algorithmes étudiés ne remplit nos exigences, il nous faudra trouver une autre solution. Notons que C4.5 et RIPPER ayant des résultats similaires, nous les avons regroupés en une seule entité.

Systèmes de Classification :

○ BP-MLL ◇ ML-kNN □ Naive Bayes △ C4.5 ou RIPPER ☆ SMO

	<i>Pas du tout</i>	<i>Pas vraiment</i>	<i>Partiellement</i>	<i>Totalement</i>
Système Compréhensible ?	○	☆	◇ △	□
Classificateur Interprétable ?	○	☆	◇	△ □
Algorithme Interprétable ?	○	☆	◇ △	□
Système Retraçable ?		◇	○ ☆	△ □
Classificateur Révisable ?	○	☆	◇ □	△

Figure 9.2.: Évaluation de la "transparence" de différents systèmes de classification

La Figure 9.2 montre que le système de classification Naive Bayes respecte *totalement* une majorité des exigences que nous nous sommes fixées, avec seulement le type de classificateur utilisé considéré comme n'étant que *partiellement révisable*. Les systèmes de classification C4.5 et RIPPER remplissent aussi *totalement* une majorité de nos exigences. Cependant, leurs algorithmes d'apprentissage respectifs

sont considérés comme moins *compréhensibles* et *interprétables* que celui de Naive Bayes.

Concernant les systèmes de classification ne remplissant pas toutes nos exigences en termes de "transparence", nous avons ML-kNN avec 4 exigences *partiellement* respectées et une exigence *pas vraiment* respectée. SMO avec une exigence *partiellement* respectée et 4 exigences *pas vraiment* respectées. Enfin BP-MLL, avec une exigence *partiellement* respectée et 4 exigences *pas du tout* respectées.

Ainsi, notre choix se porte sur le système de classification Naive Bayes. Notons que ce choix aurait sûrement été différent si nos utilisateur·ices potentiels n'avaient pas été des médecins habitués à l'utilisation de probabilités.

9.2 Adaptation de Naive Bayes aux problèmes multi-labels

Le système de classification Naive Bayes, bien que remplissant nos exigences en termes de "transparence", est dédié à la résolution de problèmes mono-label. Il nous faut donc l'adapter à la résolution de problèmes multi-labels. Une manière commune d'utiliser un système de classification mono-label sur un problème de classification multi-labels est l'emploi du méta-algorithme RAKEL (TSOUMAKAS *et al.*, 2011b). Celui-ci génère différents sous-ensembles aléatoires de labels, à partir d'une base d'apprentissage multi-labels, afin de la rendre utilisable par un algorithme d'apprentissage mono-label. Plusieurs classificateurs sont ainsi générés et utilisés en même temps lors de la phase de classification. Si une majorité de classificateurs indique une même valeur 0 ou 1 pour un label, alors ce label se verra affecté de cette valeur.

Cependant, l'usage de RAKEL, du fait de son caractère stochastique et de la combinaison de plusieurs classificateurs, amènera à un système de classification moins "interprétable" et moins "traçable". C'est pourquoi nous proposons une variante de l'algorithme Naive Bayes, détaillée en Algorithme 1, traitant les labels indépendamment les uns des autres. Le principal apport de cette proposition est de pouvoir traiter directement des problèmes de classification multi-labels tout en restant le plus "transparent" possible.

L'algorithme d'apprentissage de notre système de classification fonctionne en trois étapes : la discrétisation des variables numériques, le comptage du nombre d'occurrences de chaque label et l'estimation des probabilités d'occurrence de chaque

Algorithme 1 : Variante de l'algorithme d'apprentissage Naive Bayes adaptée à la résolution de problèmes multi-labels

Données : Une base d'apprentissage $\mathcal{I}$, un ensemble de variables $\mathcal{X}$ et un ensemble de labels $\mathcal{Y}$

Résultat : Les ensembles de probabilités approximées $P_{\mathcal{Y}}$ et $P_{\mathcal{X}|\mathcal{Y}}$

```

// Calcul des sous-ensembles des variables numériques
1 pour chaque variable  $X \in \mathcal{X}$  faire
2   | Discrétiser le domaine de  $X$  à partir de ses valeurs dans  $\mathcal{I}$ 

// Compter les occurrences  $N(Y = y)$  et  $N(Y = y \cap X = x)$ 
3 pour chaque entrée  $I \in \mathcal{I}$  faire
4   | pour chaque label  $Y \in \mathcal{Y}$  faire
5     |  $y_I \leftarrow$  valeur de  $Y$  pour l'entrée  $I$ 
6     | Incrémenter de 1 le nombre d'occurrences  $N(Y = y_I)$ 
7     | pour chaque variable  $X \in \mathcal{X}$  faire
8       |  $t_X^I \leftarrow$  le sous-ensemble de  $X$  correspondant à sa valeur dans l'entrée  $I$ 
9       | Incrémenter de 1 le nombre d'occurrences  $N(Y = y_I \cap X = t_X^I)$ 

// Estimer les probabilités  $P_{\mathcal{Y}}$  and  $P_{\mathcal{X}|\mathcal{Y}}$ 
10 pour chaque valeur  $y$  possible de chaque label  $Y \in \mathcal{Y}$  faire
11   |  $P(Y = y) \leftarrow \frac{N(Y=y)}{|\mathcal{I}|}$ 
12   | Ajouter  $P(Y = y)$  à l'ensemble  $P_{\mathcal{Y}}$ 
13   | pour chaque valeur  $t_X$  possible de chaque variable  $X \in \mathcal{X}$  faire
14     |  $P(X = t_X | Y = y) \leftarrow \frac{N(Y=y \cap X=t_X)}{N(Y=y)}$ 
15     | Ajouter  $P(X = t_X | Y = y)$  à l'ensemble  $P_{\mathcal{X}|\mathcal{Y}}$ 

16 retourner  $P_{\mathcal{Y}}$  et  $P_{\mathcal{X}|\mathcal{Y}}$ 

```

label. Un ensemble de probabilités est ainsi généré par l'algorithme d'apprentissage et utilisé par le système de classification pour déterminer les labels à associer à une nouvelle entrée.

Dans la suite de ce chapitre, nous allons utiliser la Tableau 9.1 comme base d'apprentissage, afin d'illustrer le fonctionnement de l'Algorithme 1.

9.2.1 Discrétisation des variables numériques

La version de base de Naive Bayes, présentée en Section 9.1.3, part de l'hypothèse que les variables prenant leurs valeurs dans un domaine continu (ex. $\mathbb{R}$), que nous appellerons par la suite des "variables numériques", suivent une loi normale de répartition. Cependant, il se peut que cette hypothèse ne soit pas applicable

$\mathcal{X}$		$\mathcal{Y}$	
Age	Pathologie	HbA1c	HDL
42	DT2	1	0
52	HChol	0	1
24	DT1	1	0
67	HChol	1	1

Table 9.1.: Exemple simpliste de base d'apprentissage pour illustrer le fonctionnement de notre système de classification

dans notre cas. En nous basant sur nos observations (cf. Chapitre 6), les médecins travaillent avant tout avec des notions de "seuil" (ex. si $\text{HbA1c} > 7$ alors ...), ou avec des concepts "flous", tels que l'idée de "poids élevé" ou de "tension basse". Partir de l'hypothèse qu'une variable numérique X suit une loi normale ne permet que de dire si une valeur $x \in X$ est plus ou moins "probable" selon la loi normale estimée, ce qui ne correspond pas à une notion de seuil. Ainsi, il apparaît pertinent, afin de nous ajuster au mieux aux manières de raisonner des médecins, de discrétiser le domaine des variables numériques en différents sous-ensembles, passant alors d'un domaine continu à un domaine discret. La première étape de notre algorithme d'apprentissage consiste donc à discrétiser le domaine de ces variables en différents sous-ensembles (Algorithme 1, ligne 2).

Dans un premier temps, nous avons envisagé d'utiliser l'algorithme de regroupement, ou *clustering*, *fuzzy c-means* (BEZDEK, 1981). Cet algorithme permet de déterminer un ensemble de sous-ensembles flous T_x à partir des valeurs observées d'une variable X . L'intérêt des sous-ensembles flous, tels que définis par ZADEH, (1965), est de permettre la représentation mathématique de notions "interprétables" par un·e utilisateur·ice, telles que "élevée", "moyenne" ou "faible". De plus, une valeur $x \in X$ fait partie de tous les sous-ensembles flous $t \in T_X$ selon un certain degré d'appartenance $\mu_t(x) \in [0..1]$. Ainsi, lorsque nous récupérons le sous-ensemble t_X^I correspondant à la valeur de X pour l'instance I , afin de calculer le nombre d'occurrences de $X = t_x \cap Y = y$ (Algorithme 1, ligne 8), nous considérons le sous-ensemble $t \in T_X$ avec le plus haut degré d'appartenance (Équation (9.1)).

$$t \leftarrow \arg \max_{t \in T_X} \mu_t(x) \quad (9.1)$$

Cependant, l'utilisation de *fuzzy c-means* nécessite l'introduction dans le système de classification de notions ne faisant pas partie du programme universitaire des

médecins¹. L'exigence de "compréhensibilité" n'est alors plus respectée par le système de classification et nous ne pouvons plus le considérer "transparent".

C'est pourquoi nous nous sommes orientés, dans un second temps, vers une méthode de discrétisation inspirée des histogrammes et consistant simplement à diviser le domaine de chaque variable numérique X en n sous-ensembles de taille h , telle que $h = \frac{\max(X) - \min(X)}{n}$. Le domaine d'une variable numérique X est alors défini à partir de ses valeurs minimum et maximum observées. Ainsi, le sous-ensemble t correspondant à une valeur $x \in X$ est le sous-ensemble $t \in T_X$ tel que $\min(t) \leq x < \max(t)$. Précisons que le choix arbitraire d'un nombre n de sous-ensembles va nécessairement impacter les résultats de l'outil. Plus n sera grand, plus l'algorithme sera précis, mais plus il lui faudra de données d'apprentissage pour avoir des résultats significatifs. Afin de nous rapprocher de la manière de faire des médecins, qui divisent généralement le domaine des données biologiques à l'aide de seuils minimum et maximum, nous pouvons préconiser d'utiliser $n = 3$, ou $n = 5$ pour plus de précision. Cependant, cette méthode ne garantit aucunement de retomber sur les seuils que définiraient des médecins.

Si nous reprenons la Tableau 9.1 comme exemple de base d'apprentissage, seule la variable *Age* est une variable numérique. Supposons que nous souhaitons diviser le domaine de cette variable en, par exemple, trois sous-ensembles. Les valeurs minimum et maximum observées sont respectivement 24 et 67. La taille h des sous-ensembles est donc égale à $(67 - 24)/3$, soit 14.3. Nous obtenons ainsi les trois sous-ensembles $[24, 38.3[$, $[38.3, 52.6[$ et $[52.6, 67[$.

Notons que cette méthode est particulièrement sujette à des effets de seuils, c'est-à-dire des résultats totalement différents pour des valeurs gravitant autour d'un seuil. Par exemple, deux entrées, avec respectivement $age = 38$ et $age = 39$, ne donneront pas les mêmes résultats car elles n'appartiennent pas au même sous-ensemble, bien que ces deux valeurs soient très proches. Ces effets de seuil potentiels peuvent être, cependant, considérablement amoindris si la base d'apprentissage utilisée contient un large panel d'entrées. Nous pouvons aussi envisager d'alerter l'utilisateur·ice lorsqu'une donnée numérique est proche d'un seuil. Enfin, afin d'éviter des erreurs au niveau des optima, nous considérons que si $x < \min(X)$ alors x appartient au sous-ensemble $t \in T_X$ tel que $\min(t) = \min(X)$, et si $x > \max(X)$ nous considérons que x appartient au sous-ensemble $t \in T_X$ tel que $\max(t) = \max(X)$. Si nous reprenons notre exemple, les sous-ensembles calculés précédemment correspondent alors aux sous-ensembles $] - \infty, 38.3[$, $[38.3, 52.6[$ et $[52.6, +\infty[$.

Cette méthode de discrétisation, bien que limitée et sujette à des effets de seuil, est bien plus "compréhensible" et "interprétable" que toute autre méthode de discrétisation.

sation que nous aurions pu envisager, car elle repose sur des notions parfaitement *compréhensibles* part tout-es nos utilisateur-ices potentiels¹. Nous souhaitons tout de même présenter la première méthode de discrétisation envisagée car dans la section 9.3 nous comparons les performances de classification obtenues pour chacune des deux méthodes.

9.2.2 Comptage des nombres d'occurrences des labels

Une fois les variables numériques discrétisées, la seconde étape de notre algorithme consiste à parcourir l'ensemble des instances I de la base d'apprentissage $\mathcal{I}$, afin de compter le nombre d'occurrences de $Y = 1$ et $Y = 0$ pour chaque label $Y \in \mathcal{Y}$ (Algorithme 1, ligne 3). Dans le même temps, l'algorithme récupère le sous-ensemble t_X^I correspondant à la valeur de $x \in X$ pour l'instance I , afin de calculer le nombre d'occurrences de $X = t_X \cap Y = y$ dans la base d'apprentissage $\mathcal{I}$. Précisons que si X est une variable nominale, alors $t_X = x$.

Nous obtenons ainsi un ensemble de taille $2 \times |\mathcal{Y}|$, contenant le nombre d'occurrences de $Y = 0$ et $Y = 1$ de chaque label $Y \in \mathcal{Y}$, et un ensemble de taille $2 \times |\mathcal{Y}| \times |\mathcal{T}| \times |\mathcal{X}|$, contenant le nombre d'occurrences de chaque $X = t_X \cap Y = y$, avec $\mathcal{T}$ l'ensemble de tous les T_X possibles. Le nombre d'occurrences d'un label $Y \in \mathcal{Y}$ égale à une valeur y est notée $N(Y = y)$, et le nombre d'occurrences d'une variable $X \in \mathcal{X}$ égale à un sous-ensemble t_X alors que $Y = y$ est notée $N(Y = y \cap X = t_X)$. Si nous reprenons notre exemple défini par la tableau 9.1, nous obtenons les nombres d'occurrences présentés en Équation (9.2). Afin d'alléger la présentation, les nombres occurrences égaux à zéro ne sont pas présentés.

$$N(\text{HbA1c} = 0) = 1 \quad (9.2a)$$

$$N(\text{HbA1c} = 1) = 3 \quad (9.2b)$$

$$N(\text{HDL} = 0) = 2 \quad (9.2c)$$

$$N(\text{HDL} = 1) = 2 \quad (9.2d)$$

$$N(\text{HbA1c} = 1 \cap \text{Age} \in]-\infty, 38.3]) = 1 \quad (9.2e)$$

$$N(\text{HDL} = 0 \cap \text{Age} \in]-\infty, 38.3]) = 1 \quad (9.2f)$$

$$N(\text{HbA1c} = 1 \cap \text{Age} \in [38.3, 52.6]) = 1 \quad (9.2g)$$

$$N(\text{HbA1c} = 0 \cap \text{Age} \in [38.3, 52.6]) = 1 \quad (9.2h)$$

$$N(\text{HDL} = 1 \cap \text{Age} \in [38.3, 52.6]) = 1 \quad (9.2i)$$

$$N(\text{HDL} = 0 \cap \text{Age} \in [38.3, 52.6]) = 1 \quad (9.2j)$$

$$N(\text{HbA1c} = 1 \cap \text{Age} \in [52.6, +\infty]) = 1 \quad (9.2k)$$

$$N(\text{HDL} = 1 \cap \text{Age} \in [52.6, +\infty]) = 1 \quad (9.2l)$$

$$N(\text{HbA1c} = 1 \cap \text{Pathologie} = \text{DT2}) = 1 \quad (9.2m)$$

$$N(\text{HDL} = 0 \cap \text{Pathologie} = \text{DT2}) = 1 \quad (9.2n)$$

$$N(\text{HbA1c} = 0 \cap \text{Pathologie} = \text{HChol}) = 1 \quad (9.2o)$$

$$N(\text{HDL} = 1 \cap \text{Pathologie} = \text{HChol}) = 2 \quad (9.2p)$$

$$N(\text{HbA1c} = 1 \cap \text{Pathologie} = \text{DT1}) = 1 \quad (9.2q)$$

$$N(\text{HDL} = 0 \cap \text{Pathologie} = \text{DT1}) = 1 \quad (9.2r)$$

$$N(\text{HbA1c} = 1 \cap \text{Pathologie} = \text{HChol}) = 1 \quad (9.2s)$$

9.2.3 Estimation des probabilités d'occurrences des labels

La troisième et dernière étape de notre algorithme consiste à calculer, pour chaque valeur possible y de chaque label Y , et pour chaque valeur/sous-ensemble possible x de chaque variable X , leurs fréquences d'occurrences dans la base d'apprentissage à partir de leurs nombres d'occurrences calculés précédemment (Algorithme 1, ligne 10). Ces fréquences d'occurrences nous permettent ainsi d'obtenir une estimation des probabilités $P(Y = y)$, à partir du nombre $|\mathcal{I}|$ d'entrées dans la base d'apprentissage et du nombre d'occurrences $N(Y = y)$ (cf. Équation (9.3a)), et des probabilités $P(X = x|Y = y)$, à partir des nombres d'occurrences $N(Y = y \cap X = x)$ et $N(Y = y)$ (cf. Équation (9.3b)).

$$P(Y = y) = \frac{N(Y = y)}{|\mathcal{I}|} \quad (9.3a)$$

$$P(X = x|Y = y) = \frac{N(Y = y \cap X = x)}{N(Y = y)} \quad (9.3b)$$

Si nous reprenons notre exemple, nous obtenons l'ensemble des probabilités décrites en Équation (9.4). Afin d'alléger la présentation, nous n'avons pas reporté les probabilités égales à zéro.

$$P(\text{HbA1c} = 0) = \frac{1}{4} = 0.25 \quad (9.4a)$$

$$P(\text{HbA1c} = 1) = \frac{3}{4} = 0.75 \quad (9.4b)$$

$$P(\text{HDL} = 0) = \frac{2}{4} = 0.5 \quad (9.4c)$$

$$P(\text{HDL} = 1) = \frac{2}{4} = 0.5 \quad (9.4d)$$

$$P(\text{Age} \in]-\infty, 38.3[| \text{HbA1c} = 1) = \frac{1}{3} = 0.33 \quad (9.4e)$$

$$P(\text{Age} \in]-\infty, 38.3[| \text{HDL} = 0) = \frac{1}{2} = 0.5 \quad (9.4f)$$

$$P(\text{Age} \in [38.3, 52.6[| \text{HbA1c} = 1) = \frac{1}{3} = 0.33 \quad (9.4g)$$

$$P(\text{Age} \in [38.3, 52.6[| \text{HbA1c} = 0) = \frac{1}{1} = 1.0 \quad (9.4h)$$

$$P(\text{Age} \in [38.3, 52.6[| \text{HDL} = 1) = \frac{1}{2} = 0.5 \quad (9.4i)$$

$$P(\text{Age} \in [38.3, 52.6[| \text{HDL} = 0) = \frac{1}{2} = 0.5 \quad (9.4j)$$

$$P(\text{Age} \in [52.6, +\infty[| \text{HbA1c} = 1) = \frac{1}{3} = 0.33 \quad (9.4k)$$

$$P(\text{Age} \in [52.6, +\infty[| \text{HDL} = 1) = \frac{1}{2} = 0.5 \quad (9.4l)$$

$$P(\text{Pathologie} = \text{DT2} | \text{HbA1c} = 1) = \frac{1}{3} = 0.33 \quad (9.4m)$$

$$P(\text{Pathologie} = \text{DT2} | \text{HDL} = 0) = \frac{1}{2} = 0.5 \quad (9.4n)$$

$$P(\text{Pathologie} = \text{HChol} | \text{HbA1c} = 0) = \frac{1}{1} = 1.0 \quad (9.4o)$$

$$P(\text{Pathologie} = \text{HChol} | \text{HDL} = 1) = \frac{2}{2} = 1.0 \quad (9.4p)$$

$$P(\text{Pathologie} = \text{DT1} | \text{HbA1c} = 1) = \frac{1}{3} = 0.33 \quad (9.4q)$$

$$P(\text{Pathologie} = \text{DT1} | \text{HDL} = 0) = \frac{1}{2} = 0.5 \quad (9.4r)$$

$$P(\text{Pathologie} = \text{HChol} | \text{HbA1c} = 1) = \frac{1}{3} = 0.33 \quad (9.4s)$$

9.2.4 Anticipation des labels pour une nouvelle entrée

L'ensemble des probabilités calculées par l'algorithme d'apprentissage de notre système de classification sert de base à notre classificateur lors de la labellisation de toute une nouvelle entrée $I' = \{X_1 = x_1, X_2 = x_2, \dots, X_{|\mathcal{X}|} = x_{|\mathcal{X}|}\}$. Pour ce faire, le classificateur estime, pour chaque label $Y \in \mathcal{Y}$, la probabilité $P(Y = 1 | I')$ en appliquant une version "naïve" du théorème de Bayes, qui suppose que chaque variable $X \in \mathcal{X}$ est indépendante des autres, simplifiant ainsi le calcul de $P(Y = 1 | I')$

(Équation (9.5)). Notons que, pour éviter toute division par zéro, toute valeur de probabilité nulle est remplacée par une valeur $\epsilon > 0$, par exemple $\epsilon = 0.001$.

$$P(Y = 1|I') = \frac{P(Y = 1) \times \prod_{i=1}^{|\mathcal{X}|} P(X_i = x_i|Y = 1)}{\sum_{y=0}^1 \left(P(Y = y) \times \prod_{i=1}^{|\mathcal{X}|} P(X_i = x_i|Y = y) \right)} \quad (9.5)$$

Ainsi, le classificateur détermine qu'un label $Y \in \mathcal{Y}$ vaut 1 si $P(Y = 1|I') > \lambda$. Le seuil λ étant, par défaut, égal à 0.5, car si $P(Y = 1|I') > 0.5$ alors $P(Y = 1|I') > P(Y = 0|I')$. Cependant, en fonction de l'usage, il peut être intéressant de modifier ce seuil pour être plus ou moins permissif au moment de la labellisation d'une nouvelle entrée.

Pour illustrer ce processus, reprenons les probabilités calculées pour notre exemple et supposons une nouvelle entrée $I' = \text{Age} = 34, \text{Pathologie} = \text{DT2}$. Le classificateur va, dans un premier temps, récupérer les probabilités correspondant aux données de I' , c'est à dire : $P(\text{HbA1c} = 0)$, $P(\text{HbA1c} = 1)$, $P(\text{HDL} = 0)$, $P(\text{HDL} = 1)$, $P(\text{Age} \in]-\infty, 38.3[| \text{HbA1c} = 0)$, $P(\text{Age} \in]-\infty, 38.3[| \text{HbA1c} = 1)$, $P(\text{Age} \in [24, 38.3[| \text{HDL} = 0)$, $P(\text{Age} \in]-\infty, 38.3[| \text{HDL} = 1)$, $P(\text{Pathologie} = \text{DT2} | \text{HbA1c} = 0)$, $P(\text{Pathologie} = \text{DT2} | \text{HbA1c} = 1)$, $P(\text{Pathologie} = \text{DT2} | \text{HDL} = 0)$ et $P(\text{Pathologie} = \text{DT2} | \text{HDL} = 1)$. À partir de cet ensemble de probabilités, le classificateur calcule une estimation des probabilités $P(\text{HbA1c} = 1|I')$ (cf. Équation (9.6a)) et $P(\text{HDL} = 1|I')$ (cf. Équation (9.6b)), avec $\epsilon = 0.001$.

$$\begin{aligned} P(\text{HbA1c} = 1 | I') &= \frac{0.75 \times 0.33 \times 0.33}{0.75 \times 0.33 \times 0.33 + 0.25 \times 0.001 \times 0.001} \\ &= \frac{0.08}{0.08000025} \\ &\simeq 0.99 \end{aligned} \quad (9.6a)$$

$$\begin{aligned} P(\text{HDL} = 1 | I') &= \frac{0.5 \times 0.001 \times 0.001}{0.5 \times 0.001 \times 0.001 + 0.5 \times 0.5 \times 0.5} \\ &= \frac{0.0000005}{0.1250005} \\ &\simeq 0.000004 \end{aligned} \quad (9.6b)$$

Ainsi, avec un seuil $\lambda = 0.5$, nous obtenons $P(\text{HbA1c}) > \lambda$ et $P(\text{HDL}) < \lambda$. Le classificateur détermine donc, pour l'entrée I' , que le label HbA1c est égal à 1 et que le label HDL est égal à 0. Nous notons aussi avec notre exemple que le processus

de notre système de classification est aisément reproductible par un être humain, facilitant ainsi l'interprétabilité de ses résultats.

9.3 Estimation du coût potentiel de la "transparence" en termes de performance

Nous avons, en section 9.1, évalué la "transparence" de divers systèmes de classification et nous avons déterminé que Naive Bayes était celui qui remplissait le mieux les exigences que nous nous étions fixées. Cependant, bien que nous cherchions un système de classification qui soit le plus "transparent" possible vis-à-vis de nos utilisateur·ices, nous ne voulons pas payer cette "transparence" par une perte trop importante de performance. C'est pourquoi, dans cette section, nous tentons d'évaluer si la "transparence" de Naive Bayes implique une baisse significative de performance. Si une baisse trop importante de performance est identifiée, il nous faudra nous replier vers un système moins "transparent", afin de trouver un bon équilibre entre "transparence" et performance.

Pour ce faire, nous avons évalué les performances des systèmes précédemment cités à l'aide de la librairie Java Mulan (TSOUMAKAS *et al.*, 2011a). La librairie Mulan, dédiée aux problèmes de classification multi-labels, donne accès à un ensemble de systèmes de classification ainsi qu'à des systèmes d'évaluation telle que la validation croisée. Dans ces expérimentations, les systèmes de classification mono-label ont été associés au méta-algorithme RAKEL (TSOUMAKAS *et al.*, 2011b). Nous avons aussi évalué les performances de deux autres systèmes de classification issus de la variante de Naive Bayes proposée en section 9.2 : FuzzyBayes, basé sur une discrétisation via des sous-ensembles flous, et HistBayes, basé sur une discrétisation inspirée des histogrammes.

Enfin, notons que les métriques utilisées pour évaluer les systèmes de classification multi-labels diffèrent quelque peu de celles utilisées pour les classifications mono-label. Nous allons donc détailler chacune d'entre elles afin de rendre plus compréhensible et interprétable les résultats de nos expérimentations. Pour ces expérimentations nous avons utilisé des bases d'apprentissage multi-labels reconnues, telles que : *Birds* (BRIGGS *et al.*, 2013), *CAL500* (TURNBULL *et al.*, 2008), *Emotions* (TROHIDIS *et al.*, 2008), *Genbase* (DIPLARIS *et al.*, 2005), *Medical* (PESTIAN *et al.*, 2007) et *Scene* (BOUTELL *et al.*, 2004). Nous y avons ajouté la base *Consultations*, composée de 50 instances, 4 variables et 18 labels, et correspondant à notre cas

pratique. Le code pour reproduire nos résultats est disponible sur le GitLab de LAMSADE².

9.3.1 Taux d'erreur : perte de Hamming

Commençons par l'évaluation du risque d'erreurs global. Pour les problèmes de classification mono-label, le taux d'erreur moyen se calcule en divisant le nombre d'erreurs par la taille de la base de test. Cependant, dans le cas d'un problème multi-labels, il nous faut une métrique prenant en compte le taux d'erreur pour chaque label. Ainsi, comme introduit par SCHAPIRE et SINGER, (2000), la perte de Hamming est une métrique répondant à nos besoins.

Soit C un classificateur et D une base de test. D est composé de $|\mathcal{X}|$ variables endogènes (ou attributs) et de $|\mathcal{Y}|$ variables exogènes (ou labels). Pour chaque entrée i allant de 0 à $|D| - 1$, nous nommons Y_i l'ensemble des valeurs réelles des labels et Z_i l'ensemble des valeurs des labels déterminées par C .

La perte de Hamming se calcule via la fonction $HammingLoss : C \times D \mapsto [0..1]$, définie en Équation (9.7).

$$HammingLoss(C, D) = \frac{1}{|D|} \sum_{i=0}^{|D|-1} \frac{Y_i \Delta Z_i}{|\mathcal{Y}|} \quad (9.7)$$

Avec Δ la différence symétrique entre deux ensembles de labels (équivalente à l'opérateur logique XOR). Par exemple, si nous supposons $Y_i = \{0, 1, 1, 0, 1\}$ et $Z_i = \{0, 0, 1, 0, 1\}$, alors $Y_i \Delta Z_i = 1$, seul le deuxième label étant différent.

L'utilisation de la perte de Hamming permet donc d'obtenir la moyenne du taux d'erreur pour l'ensemble des labels de chaque entrée. La Figure 9.3 montre la distribution des pertes de Hamming, obtenue par validation croisée sur chaque base d'apprentissage, des différents systèmes de classification étudiés. Pour mettre en évidence la "transparence" des systèmes, nous les avons ordonnés du moins "transparent" (BP-MLL, en rouge) au plus "transparent" (HistBayes, en vert). Notre objectif est de savoir si l'utilisation d'un système de classification "transparent" implique une perte de Hamming plus importante.

Nous pouvons tout d'abord observer que les systèmes de classification les plus "transparent" ne sont pas nécessairement ceux présentant les taux d'erreurs les plus élevés. Au contraire, mis à part pour la base *Emotions*, le taux d'erreur du

2. https://git.lamsade.fr/a_richard/transparent-performances

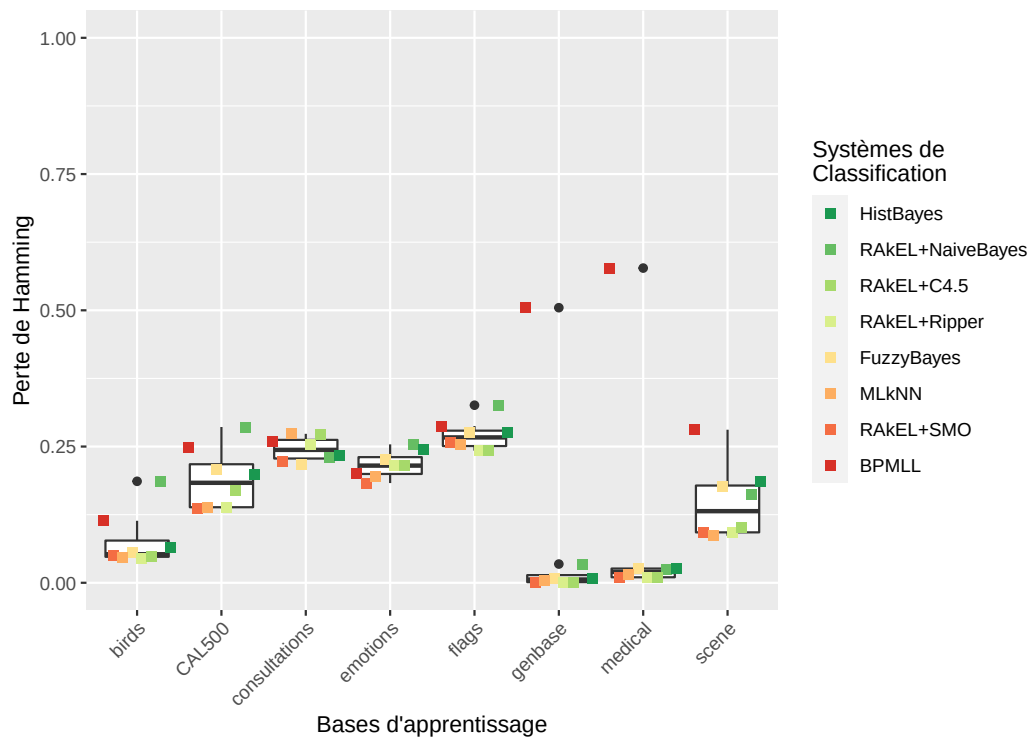


Figure 9.3.: Distribution des pertes de Hamming de différents systèmes de classification évaluées pour différentes bases d'apprentissage, résultats obtenus par validation croisée

système BP-MLL (en rouge) est généralement un des plus élevés, si ce n'est le plus élevé. Nous observons aussi que, mise à part pour la base *Emotions*, une plus grande "transparence" ne semble pas corrélée à un plus haut taux d'erreur.

Enfin, nous observons que, mis à part pour les bases d'apprentissage *CAL500* et *Scene*, les taux d'erreurs mesurés pour chaque système de classification sont assez proches les uns des autres. De plus, pour notre cas pratique représenté par la base *Consultations*, le taux d'erreurs de HistBayes (système que nous souhaitons utiliser pour sa "transparence") fait partie des plus basses mesurées.

9.3.2 Précisions micro et macro

En classification mono-label, une métrique souvent utilisée pour évaluer la performance d'un système de classification est la précision. Celle-ci est calculée via la fonction $Prec : C \times D \times \mathcal{Y} \mapsto [0..1]$ définie en Équation (9.8).

$$Prec(C, D, Y) = \frac{VPs(C, D, Y)}{VPs(C, D, Y) + FPs(C, D, Y)} \quad (9.8)$$

Avec la fonction $VPs : C \times D \times \mathcal{Y} \mapsto \mathbb{N}^+$, qui calcule le nombre de "vrais positifs" déterminé par C pour un label $Y \in \mathcal{Y}$ d'une base de test D . C'est-à-dire, le nombre de fois que C a déterminé que $Y = 1$ et que, effectivement $Y = 1$. Et la fonction $FPs : C \times D \times \mathcal{Y} \mapsto \mathbb{N}^+$, qui calcule le nombre de "faux positifs" de C pour le label $Y \in \mathcal{Y}$. C'est-à-dire, le nombre de fois que C a déterminé que $Y = 1$, mais qu'en réalité $Y = 0$. La fonction $Prec$ calcule ainsi le taux de "réponses correctes" de C sur toutes les fois où il a déterminé que $Y = 1$.

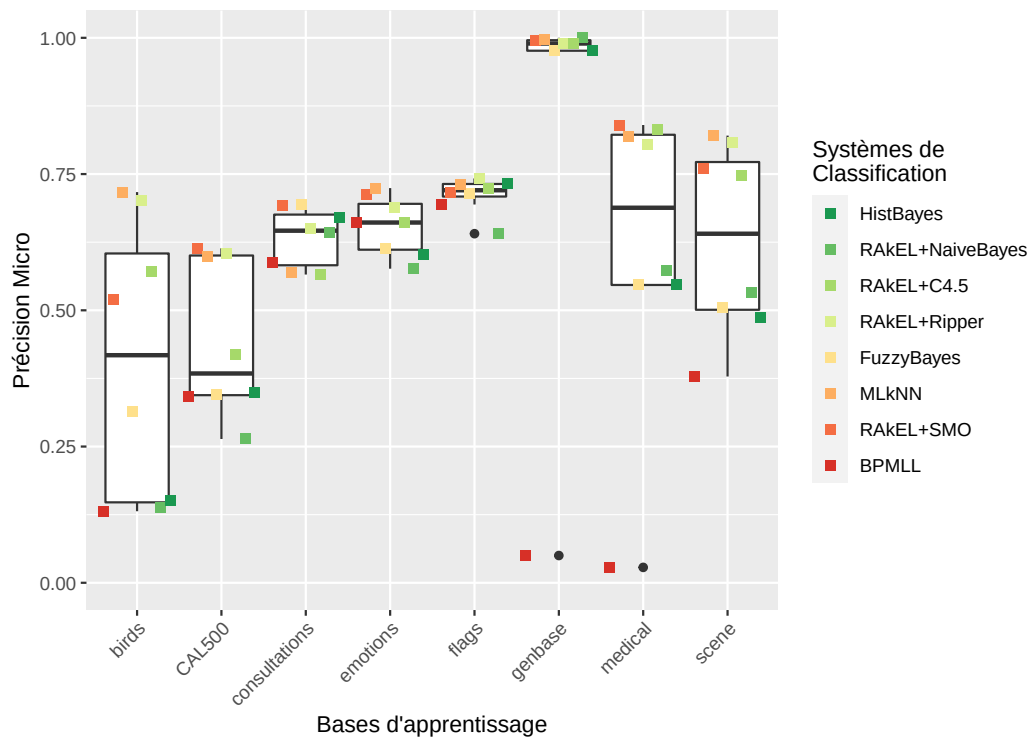
Cependant, dans le cadre d'un problème multi-labels, il nous faut mesurer la précision globale du système de classification C en prenant en compte la précision moyenne de C pour chaque label. Nous avons alors deux possibilités, chacune ayant ses avantages et ses inconvénients.

Premièrement, nous pouvons calculer la précision totale de C , indépendamment des labels, ce qui est appelé précision micro. La précision micro se calcule via la fonction $Prec_{micro} : C \times D \mapsto [0..1]$ définie en Équation (9.9).

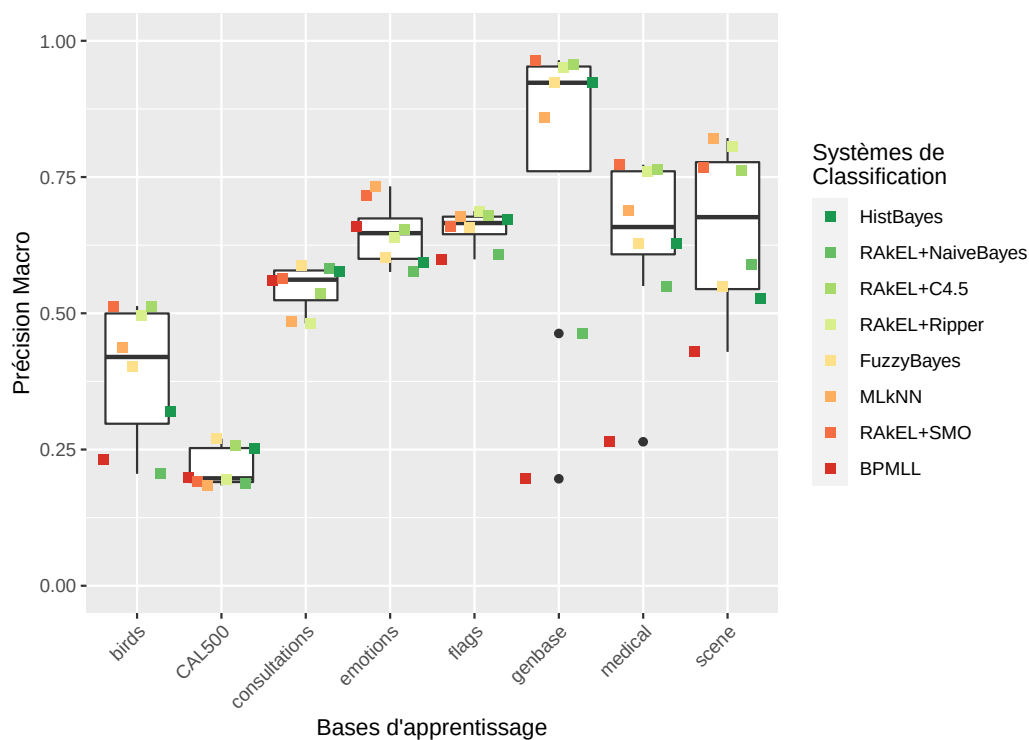
$$Prec_{micro}(C, D) = \frac{\sum_{Y \in \mathcal{Y}} VPs(C, D, Y)}{\sum_{Y \in \mathcal{Y}} VPs(C, D, Y) + FPs(C, D, Y)} \quad (9.9)$$

Deuxièmement, nous pouvons calculer la moyenne de la précision de C pour chaque label $Y \in \mathcal{Y}$. Cette seconde approche, nommée précision macro, sera plus sensible aux *extrema* que la première (ex. si C est très précis pour un label spécifique). La précision macro se calcule via la fonction $Prec^{macro} : C \times D \mapsto [0..1]$ définie en Équation (9.10).

$$Prec^{macro}(C, D) = \frac{\sum_{Y \in \mathcal{Y}} Prec(C, D, Y)}{|\mathcal{Y}|} \quad (9.10)$$



(a) Précisions micro



(b) Précisions macro

Figure 9.4.: Distributions des précisions micro et macro de différents systèmes de classification évaluées pour différentes bases d'apprentissage, résultats obtenus par validation croisée

Comme pour le taux d'erreur, nous cherchons à évaluer si la "transparence" d'un système de classification influe sur sa précision. La Figure 9.4 montre, pour les différents systèmes évalués, les distributions de leurs précisions micro (Figure 9.4a) et de leurs précisions macro (Figure 9.4b). Ces métriques ont été obtenues par validation croisée sur les différentes bases d'apprentissage citées précédemment.

Nous pouvons observer que les systèmes de classification le plus "transparents" ne sont pas nécessairement les moins précis. De même, il ne semble pas y avoir de corrélation entre "transparence" et précision.

Nous observons aussi que, pour la moitié des cas, il n'y a pas une forte disparité entre les précisions macro et micro des différents systèmes de classifications évalués. Et pour l'autre moitié, il y a une disparité plus nette entre les différentes précisions mesurées. Dans notre cas, la précision de HistBayes fait partie des systèmes les plus précis.

9.3.3 Sensibilité, ou rappel, micro et macro

Lors de l'évaluation des performances d'un système de classification mono-label, sa précision est généralement associée à son rappel. Cette métrique est aussi appelée sensibilité lorsqu'elle est associée à la spécificité d'un système de classification, métrique que nous définirons sous peu. La sensibilité, ou le rappel, se calcule via la fonction $Sens : C \times D \times \mathcal{Y} \mapsto [0..1]$ définie en Équation (9.11).

$$Sens(C, D, Y) = \frac{VPs(C, D, Y)}{VPs(C, D, Y) + FNs(C, D, Y)} \quad (9.11)$$

Avec la fonction $FNs : C \times D \times \mathcal{Y} \mapsto \mathbb{N}^+$ qui calcule le nombre de "faux négatifs" déterminé par C pour un label $Y \in \mathcal{Y}$ de la base de test D . C'est-à-dire, le nombre de fois que C a déterminé que $Y = 0$ alors que, en réalité, $Y = 1$. Ainsi, la sensibilité (ou rappel) détermine le taux de "bonnes réponses" sur toutes les fois où C aurait dû déterminer que $Y = 1$.

Cependant, dans le cadre de problèmes multi-labels, il nous faut prendre en compte la sensibilité de C pour chaque label. Ainsi, tout comme pour le calcul de la précision, nous calculons la sensibilité micro via la fonction $Sens_{micro} : C \times D \mapsto [0..1]$ définie en Équation (9.12) et la sensibilité macro via la fonction $Sens^{macro} : C \times D \mapsto [0..1]$ définie en Équation (9.13).

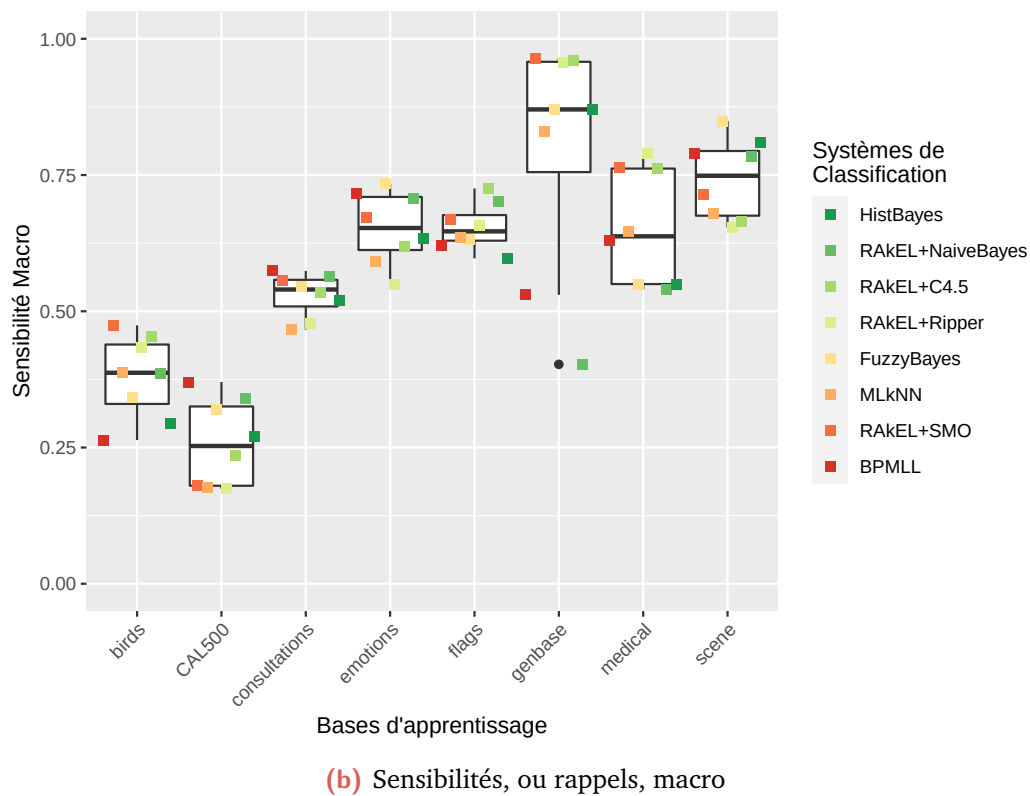
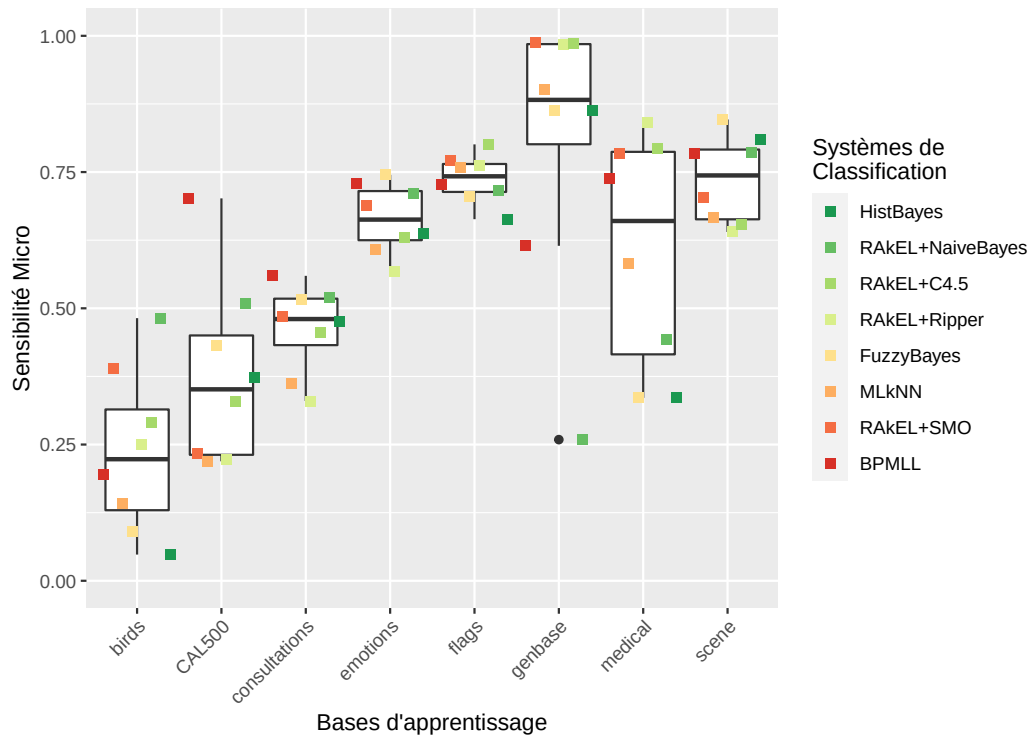


Figure 9.5.: Distributions des sensibilités, ou rappels, micro et macro de différents systèmes de classification évaluées pour différentes bases d'apprentissage, résultats obtenus par validation croisée

$$Sens_{micro}(C, D) = \frac{\sum_{Y \in \mathcal{Y}} VPs(C, D, Y)}{\sum_{Y \in \mathcal{Y}} VPs(C, D, Y) + FNs(C, D, Y)} \quad (9.12)$$

$$Sens^{macro}(C, D) = \frac{\sum_{Y \in \mathcal{Y}} Sens(C, D, Y)}{|\mathcal{Y}|} \quad (9.13)$$

La Figure 9.5 montre les sensibilités, ou rappels, micro et macro mesurées pour chaque système de classification étudié. Ces mesures ont été obtenues par validation croisée sur les différentes bases d'apprentissage citées précédemment.

Nous pouvons observer en Figure 9.5a une forte disparité des sensibilités micro mesurées pour au moins 5 bases d'apprentissage sur 8. Cette disparité se retrouve, dans une moindre mesure, pour les sensibilités macro (cf. Figure 9.5b). Cependant, tout comme pour le taux d'erreur et la précision, il ne semble pas y avoir de corrélation entre "transparence" et sensibilité. Les systèmes de classification les plus "transparents" n'étant pas nécessairement les moins sensibles, et inversement.

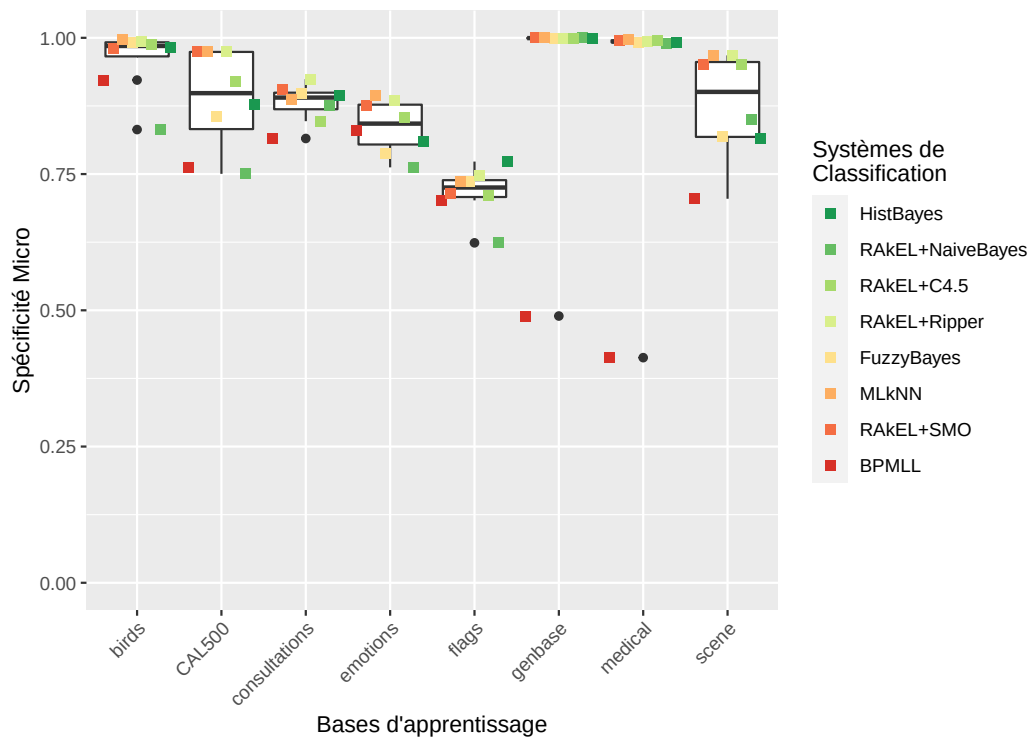
Pour notre cas, illustré par la base *Consultations*, les résultats des différents systèmes de classification sont assez proches. HistBayes, le système que nous souhaitons utiliser, ne semble pas faire partie des plus performants en termes de sensibilité, mais reste proche de la médiane pour notre cas pratique.

9.3.4 Spécificité micro et macro

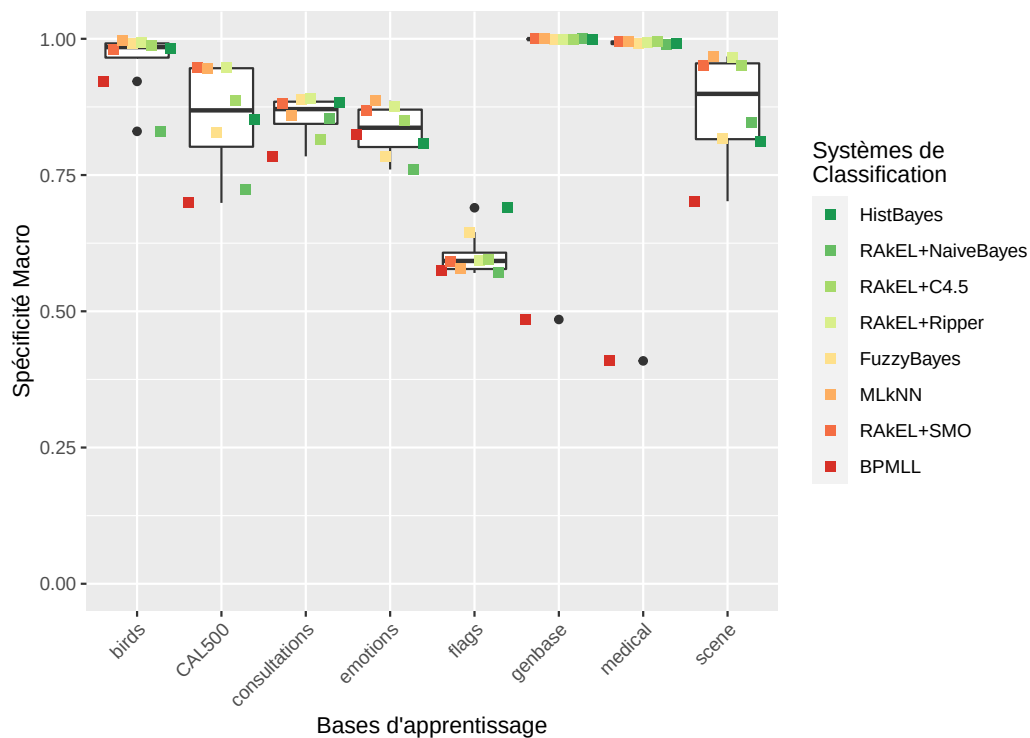
Enfin, une troisième métrique est généralement utilisée, en association avec la sensibilité, pour évaluer la performance d'un système de classification. Cette métrique, nommée spécificité, se calcule via la fonction $Spec : C \times D \times \mathcal{Y} \mapsto [0..1]$ définie en Équation (9.14).

$$Spec(C, D, Y) = \frac{VN_s(C, D, Y)}{VN_s(C, D, Y) + FPs(C, D, Y)} \quad (9.14)$$

Avec la fonction $VN_s : C \times D \times \mathcal{Y} \mapsto \mathbb{N}^+$, qui calcule le nombre de "vrais négatifs" déterminé par C pour le label $Y \in \mathcal{Y}$ de la base de test D . En d'autres termes, VN_s calcule le nombre de fois où C a déterminé que $Y = 0$ et que, effectivement, $Y = 0$. Ainsi, là où la sensibilité calcule le taux de $Y = 1$ trouvé par C , la spécificité calcule le taux de $Y = 0$ trouvé par C .



(a) Spécificités micro



(b) Spécificités macro

Figure 9.6.: Distributions des spécificités micro et macro de différents systèmes de classification évaluées pour différentes bases d'apprentissage, résultats obtenus par validation croisée

Cependant, dans le cadre de problème multi-labels, il nous faut prendre en compte la spécificité de C pour chaque label. Nous utilisons donc, comme pour les métriques précédentes, une spécificité micro calculée via la fonction $Spec_{micro} : C \times D \mapsto [0..1]$ définie en Équation (9.15), et une spécificité macro calculée via la fonction $Spec^{macro} : C \times D \mapsto [0..1]$ définie en Équation (9.16).

$$Spec_{micro}(C, D) = \frac{\sum_{Y \in \mathcal{Y}} VNs(C, D, Y)}{\sum_{Y \in \mathcal{Y}} VNs(C, D, Y) + FPs(C)} \quad (9.15)$$

$$Spec^{macro}(C, D) = \frac{\sum_{Y \in \mathcal{Y}} Spec(C, D, Y)}{|\mathcal{Y}|} \quad (9.16)$$

La Figure 9.6 montre la distribution des spécificités, calculées par validation croisée sur différentes bases d'apprentissage, des systèmes de classification étudiés.

Nous observons, mis à part pour les bases d'apprentissage *CAL500* et *Scene*, une grande similarité des spécificités micro et macro mesurées pour les différents systèmes de classification étudiés. De plus, tout comme pour les métriques précédentes, aucune corrélation n'est visible entre "transparence" et spécificité. Dans notre cas, illustré par la base *Consultations*, l'algorithme HistBayes montre une des spécificités les plus élevées.

Si nous résumons les résultats obtenus pour les différentes métriques de performances étudiées, l'emploi d'un algorithme "transparent" n'implique pas nécessairement une baisse des performances. Dans notre cas, l'emploi du système de classification HistBayes, ne montre pas une baisse significative des performances en comparaison avec des systèmes que nous considérons moins "transparents". C'est pourquoi notre choix se porte sur ce système de classification pour le développement d'un outil d'aide à la décision en consultations médicales.

Notons tout de même que ces résultats sont limités par notre définition de la "transparence" et des critères opérationnels que nous utilisons pour l'évaluer. Même si nous avons tenté d'être le plus générique possible, ces critères n'ont pas pour vocation d'être universel, mais de synthétiser ce qui nous semble pertinent compte tenu de notre contexte applicatif.

Conception d'un assistant virtuel pour les consultations médicales

” *In a properly automated and educated world, then, machines may prove to be the true humanizing influence. It may be that machines will do the work that makes life possible and that human beings will do all the other things that make life pleasant and worthwhile.*

— **Isaac Asimov**
Robot Visions

Dans ce dernier chapitre, nous présentons l'outil développé durant cette thèse. Nommé "CoBoy", celui-ci est destiné, comme expliqué au Chapitre 7, à anticiper les informations nécessaires aux médecins pour leurs consultations et à leur fournir ces informations en début de consultations. Précisons toutefois que le prototype actuel de CoBoy a été développé en collaboration avec les médecins du service d'endocrinologie et diabétologie des HCL et n'est donc calibré que pour apporter un soutien durant les consultations de ce service.

Cette proposition s'ancre dans l'approche "ajustive" de l'aide à la décision que nous voulions développer, comme expliqué en Chapitre 4, car l'outil ajuste l'aide fournie aux besoins des médecins pour chaque consultation, tout en les laissant maîtres du déroulé de leurs processus de décision et donc de leurs consultations. À ce stade, il nous faut vérifier la viabilité et la faisabilité d'une telle proposition en situation réelle.

La Section 10.1 présente le fonctionnement de l'outil CoBoy : la manière dont il anticipe les besoins des médecins, comment il récupère les données à fournir aux médecins et comment il présente ces données aux médecins. La Section 10.2 présente les résultats des essais cliniques de l'outil CoBoy effectués au sein des HCL afin d'évaluer la faisabilité de son utilisation en situations réelles. Enfin, la Section 10.3

propose une discussion autour des résultats obtenus, des limites rencontrées dans la mise en place de l'outil et des perspectives offertes par ce type d'outils.

10.1 Présentation de l'outil

Nous détaillons dans cette section le processus interne de l'outil d'aide à la décision développé dans le cadre de cette thèse lors de l'arrivée d'un-e patient-e pour une consultation. Pour illustrer ce processus, nous supposons l'arrivée du/de la patient-e $X = \{sex = \varnothing, age = 47, imc = 23.11, pathologie = DNID\}$, avec *DNID* le sigle correspondant à la pathologie "Diabète Non Insulino-Dépendant", ou "Diabète de type 2" (cf. Figure 10.2).

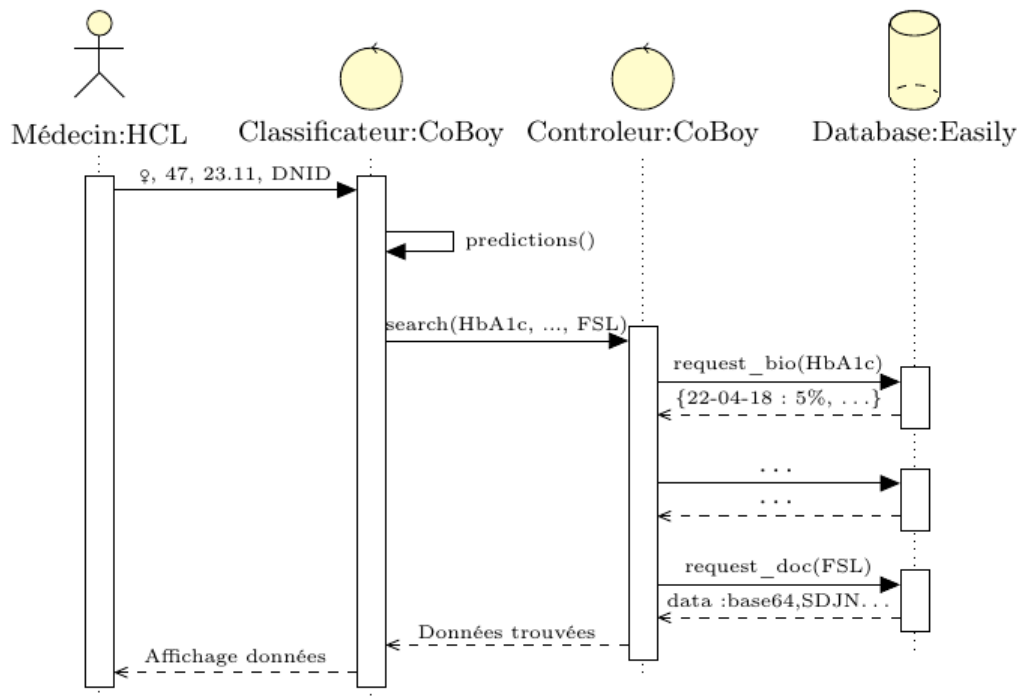


Figure 10.1.: Processus interne de CoBoy pour anticiper et fournir les données sur un-e patient-e nécessaires à un-e médecin pour une consultation médicale

La Figure 10.1 illustre les différentes étapes du processus interne de CoBoy lors de l'arrivée d'un-e patient-e en consultation. Ce processus se divise en trois phases : une phase d'anticipation des données nécessaires au/à la médecin, à partir des données connues sur X , suivie d'une phase de recherche des données identifiées comme étant nécessaires au/à la médecin, puis d'une phase de présentation des données récupérées.

10.1.1 Phase d'anticipation

La première phase du processus interne consiste à identifier le sous-ensemble d'informations concernant la·e patient·e pouvant être nécessaires au/à la médecin pour sa consultation. Pour ce faire, CoBoy reçoit en entrée l'âge, le sexe, l'IMC et la pathologie d'un·e patient·e X . Cet ensemble de données est alors utilisé par le classificateur de Naive Bayes, décrit en Section 9.2, pour calculer la probabilité $P(Y = 1|X)$ pour chaque information Y pouvant être récupérée par l'outil ($Y = 1$ signifiant que la·e médecin a besoin de l'information Y pour sa consultation). Chaque information dont la probabilité est supérieure à un seuil λ est identifiée comme probablement utile au/à la médecin pour la·e patient·e en cours.

Reprenons notre exemple $X = \{sex = \varphi, age = 47, imc = 23.11, pathologie = DNID\}$ qui, après discrétisation des variables numériques, correspond à $X' = \{sex = \varphi, age \in [44, 55[, imc \in [19.14, 23.14[, pathologie = DNID\}$ (cf. Chapitre 9), et supposons que seules deux informations sont récupérable par CoBoy : l'hémoglobine glyquée (HbA1c), et le cholestérol HDL. Supposons aussi les probabilités suivantes :

$$P(HbA1c = 1) = 0.6 \quad (10.1a)$$

$$P(HbA1c = 0) = 0.4 \quad (10.1b)$$

$$P(HDL = 1) = 0.4 \quad (10.1c)$$

$$P(HDL = 0) = 0.6 \quad (10.1d)$$

$$P(sex = \varphi | HbA1c = 1) = 0.45 \quad (10.1e)$$

$$P(sex = \varphi | HbA1c = 0) = 0.55 \quad (10.1f)$$

$$P(age \in [44, 55[| HbA1c = 1) = 0.6 \quad (10.1g)$$

$$P(age \in [44, 55[| HbA1c = 0) = 0.4 \quad (10.1h)$$

$$P(imc \in [19.14, 23.14[|HbA1c = 1) = 0.5 \quad (10.1i)$$

$$P(imc \in [19.14, 23.14[|HbA1c = 0) = 0.5 \quad (10.1j)$$

$$P(pathologie = DNID|HbA1c = 1) = 0.9 \quad (10.1k)$$

$$P(pathologie = DNID|HbA1c = 0) = 0.1 \quad (10.1l)$$

$$P(sex = \varphi|HDL = 1) = 0.55 \quad (10.1m)$$

$$P(sex = \varphi|HDL = 0) = 0.45 \quad (10.1n)$$

$$P(age \in [44, 55[|HDL = 1) = 0.4 \quad (10.1o)$$

$$P(age \in [44, 55[|HDL = 0) = 0.6 \quad (10.1p)$$

$$P(imc \in [19.14, 23.14[|HDL = 1) = 0.7 \quad (10.1q)$$

$$P(imc \in [19.14, 23.14[|HDL = 0) = 0.3 \quad (10.1r)$$

$$P(pathologie = DNID|HDL = 1) = 0.1 \quad (10.1s)$$

$$P(pathologie = DNID|HDL = 0) = 0.9 \quad (10.1t)$$

CoBoy calcule alors les probabilités $P(HbA1c = 1|X')$ et $P(HDL = 1|X')$ en appliquant le théorème de Bayes (cf. Équation (9.5)). Nous obtenons ainsi les probabilités : $P(HbA1c = 1|X') \simeq 0.94$ et $P(HDL = 1|X') \simeq 0.12$. Avec un seuil $\lambda = 0.5$, seule l'information HbA1c est considérée comme probablement nécessaire au/à la médecin pour la patiente en cours. CoBoy passe alors en phase de recherche de données pour tenter de récupérer cette information et la transmettre au/à la médecin.

Durant nos premiers essais, nous avons pu observer que se reposer uniquement sur un simple apprentissage n'était pas suffisant pour déterminer toutes les informations à fournir aux médecins. La limite majeure vient du fait qu'il n'est actuellement pas possible, comme abordé en Chapitre 6, d'avoir une trace fine des informations regardées par les médecins. La base d'apprentissage utilisée par CoBoy a été construite à partir de comptes-rendus de consultations, mais ceux-ci ne contiennent pas nécessairement toutes les informations consultées par les médecins, seulement une synthèse. Il a donc été décidé de rajouter un système de règles définies par des médecins lors d'entretiens pour les pathologies qu'ils traitent le plus couramment. Comme nous avons pu le développer dans le Chapitre 9, les systèmes de règles sont particulièrement "transparents" pour les utilisateur·ices, d'autant plus si ce sont ces utilisateur·ices qui créent les règles à appliquer et non un algorithme d'apprentissage. Nous avons ainsi pu construire une base d'une quinzaine de règles "SI ... ALORS ..." correspondant aux besoins des médecins, par exemple :

SI pathologie = diabète ALORS RECHERCHER dernier test d'effort

SI pathologie = hypercholestérolémie ALORS RECHERCHER score calcique

SI pathologie = phénylcétonurie ALORS RECHERCHER phénylalanine

Nous obtenons un système mixte qui, dans un premier temps, détermine à partir des règles formulées par les médecins un sous-ensemble d'informations à fournir puis, dans un second temps, calcule les probabilités des informations restantes comme vu précédemment.

Actuellement, l'outil CoBoy est capable de récupérer, selon les besoins, près de 50 types de données biologiques et biométriques, ainsi que près de 30 types de documents médicaux. Concernant le nombre de pathologies connues par CoBoy, celui-ci en recense actuellement près d'une vingtaine.

10.1.2 Phase de recherche de données

Une fois la phase d'anticipation passée, CoBoy a un sous-ensemble d'informations à rechercher en base de données. Pour chaque information de ce sous-ensemble, CoBoy exécute une requête qui dépend principalement du type de l'information recherchée.

Pour les informations de type "biologie", issues d'examens biologiques effectués aux laboratoires des HCL, c'est l'historique des différents résultats qui est recherché, ainsi que l'unité de mesure et les seuils minimum et maximum associés à ces données. Pour

des raisons de performance, seuls les cents résultats les plus récents sont retournés. C'est le cas, entre autres, de l'hémoglobine glyquée (HbA1c) et du cholestérol HDL vus précédemment.

Précisons que, pour des raisons d'infrastructures internes à la DSII, la version actuelle de l'outil n'est capable de fournir que les données biologiques d'un·e patient·e datant au maximum de la veille de la consultation. Cela ne pose problème que si les patient·es font leurs analyses le jour de la consultation, ce qui, en pratique, n'arrive que rarement. Ajoutons aussi que, si la·e patient·e effectue une analyse dans un laboratoire extérieur aux HCL et si les résultats de cette analyse ne sont pas saisis dans Easily[®], CoBoy sera dans l'incapacité technique de fournir ces données aux médecins. Ce deuxième cas de figure est, comme nous le développerons en Section 10.2, bien moins rare que le premier.

Pour les informations de type "document médical" telles que : les comptes-rendus d'IRM ou de scanner, les comptes rendus d'autres spécialités (ophtalmologie, diététique, cancérologie, etc.), les ordonnances, etc., ce sont les documents les plus récents qui sont recherchés par CoBoy.

Enfin, pour les informations "générales" telles que l'historique des traitements, les antécédents familiaux, les allergies, etc., c'est l'ensemble des données concernant ces informations qui est recherché.

Ainsi, CoBoy obtient pour chaque information pouvant probablement être nécessaire au/à la médecin l'ensemble des données associées concernant la·e patient·e en cours. CoBoy peut alors passer à la phase de présentation des données. Il est cependant possible qu'aucune donnée ne soit trouvée en base pour une ou plusieurs des informations recherchées. Dans ce cas, CoBoy indiquera simplement que l'information a bien été recherchée en base mais qu'aucune donnée n'a été trouvée pour la·e patient·e en cours.

10.1.3 Phase de présentation des données

Une fois que toutes les informations probablement nécessaires au/à la médecin ont été recherchées et récupérées en base, CoBoy peut les présenter au/à la médecin. La manière dont CoBoy présente les données récupérées sur la·e patient·e dépend fortement du type d'information auquel elles sont associées. La Figure 10.2 présente un exemple de l'interface utilisateur principale proposée par l'outil CoBoy. Celle-ci s'affiche lorsque la·e médecin sélectionne un·e patient·e dans sa liste de rendez-vous pour une consultation.

La colonne de gauche est dédiée à l’affichage des informations de type "biologie" et des informations "générales". Les informations de type "biologie" sont représentées à l’aide d’un graphique montrant l’évolution des données associées au cours du temps. La·e médecin·e peut connaître la valeur de chaque donnée en survolant les points formant le graphique. De plus, la donnée la plus récente est affichée au-dessus du graphique pour être directement visible par la·e médecin·e. Concernant les données liées aux informations "générales", celles-ci sont simplement affichées au format texte. En haut de la colonne de gauche, est mise à disposition une liste d’accès rapides vers les informations retrouvées en base par l’outil CoBoy.

La colonne de droite est dédiée, quant à elle, à l’affichage des informations de type "document médical", telles que : les comptes-rendus d’examen·s en laboratoire (prise de sang, IRM, radios, etc.), la dernière ordonnance faite par la·e médecin·e, les comptes rendus d’autres spécialités médicales (ex. ophtalmologie), etc. Ces documents sont affichés via un lecteur PDF et chaque document est accessible via un système d’onglets.

Chaque information fournie par l’outil est associée à une explication, accessible au/à la médecin·e en survolant un bouton "information". L’explication fournie dépend de la manière dont CoBoy a déterminé qu’il fallait fournir cette information pour la consultation en cours. Les informations fournies par défaut, comme les documents "Synthèse et Suivi" et "Dernière Ordonnance", sont associées à l’explication "Information par défaut". Les informations déterminées à partir des règles établies par les médecins sont associées à une explication détaillant les prémices de la règle utilisée, par exemple : "L’information HbA1c a été recherchée car : pathologie = Diabète Type 2". Enfin, les informations déterminées par le classificateur Naive Bayes (cf. Section 9.2), sont associées à une explication du type "En me basant sur vos précédentes consultations et connaissant les informations suivantes sur votre patient : $\text{age} \in [44.0, 55.0[$, $\text{sexe} = \text{Female}$, $\text{IMC} \in [19.14, 23.14[$ et suivi pour=DNID, la probabilité que vous ayez besoin de l’information 'Biochimie Sanguine' est de 72.29%". La·e médecin·e peut ensuite avoir accès au détail des calculs de $P(Y = 1|X)$ si iel le souhaite.

Comme introduit précédemment, il se peut qu’une information ait été recherchée par l’outil mais que celui-ci n’ait trouvé aucune donnée pour cette information concernant la·e patient·e en cours. Cependant, comme mentionnée en Chapitre 6, l’absence de données peut aussi servir aux médecins dans leurs processus de décision. Ainsi, l’outil CoBoy informera la·e médecin·e que l’information a été recherchée, mais qu’aucune donnée n’a été trouvée en base pour la·e patient·e en cours.



Enfin, chaque information fournie est associée à un bouton permettant aux médecins d'indiquer si iels ont effectivement besoin de cette information pour ce type de patient·e (voir Figure 10.2, bouton bleu en dessous du nom de chaque information). L'outil étant encore à l'état de prototype, ce bouton nous sert à évaluer la précision des informations fournies en pratique. Cependant, *in fine*, l'objectif est d'enrichir la base d'apprentissage de l'outil, via des retours utilisateurs, enfin d'améliorer ses performances mais aussi pour inclure les médecins dans le processus d'apprentissage de l'outil.

10.2 Essais cliniques

Afin d'évaluer la viabilité et la faisabilité de notre outil d'aide à la décision en situations réelles, nous avons effectué une série d'essais cliniques avec des médecins travaillant aux HCL. Nous allons, dans cette section, détailler le protocole d'observation utilisé ainsi que les résultats obtenus.

10.2.1 Protocole

L'étude de faisabilité que nous avons effectuée reprend en grande partie le protocole d'observation détaillé au Chapitre 6. Nous en reprenons ainsi le terme "observation", désignant l'étude d'une consultation débutant à l'arrivée d'un·e patient·e en salle de consultation et se terminant sa sortie de la salle, et le terme "séance d'observations", correspondant à une succession d'observations. La Figure 10.3 schématise la disposition des agents et des outils utilisés lors des observations. Un second écran a été mis en place afin que les médecins puissent utiliser CoBoy tout en ayant accès à leur environnement Easily[®] habituel (cf. Chapitre 5).

La première partie de notre étude a pour objectif d'évaluer si le nombre d'informations fournies par l'outil a eu un impact positif ou négatif sur le processus de décision des médecins. Les critères observés étaient les suivants (cf. Annexe A) :

- Le nombre d'informations fournies par l'outil CoBoy, pour quantifier l'aide apportée par l'outil ;
- Le nombre d'informations recherchées par la·e médecin, explicitées oralement au/à la patient·e, pour quantifier les besoins en informations du/de la médecin ;

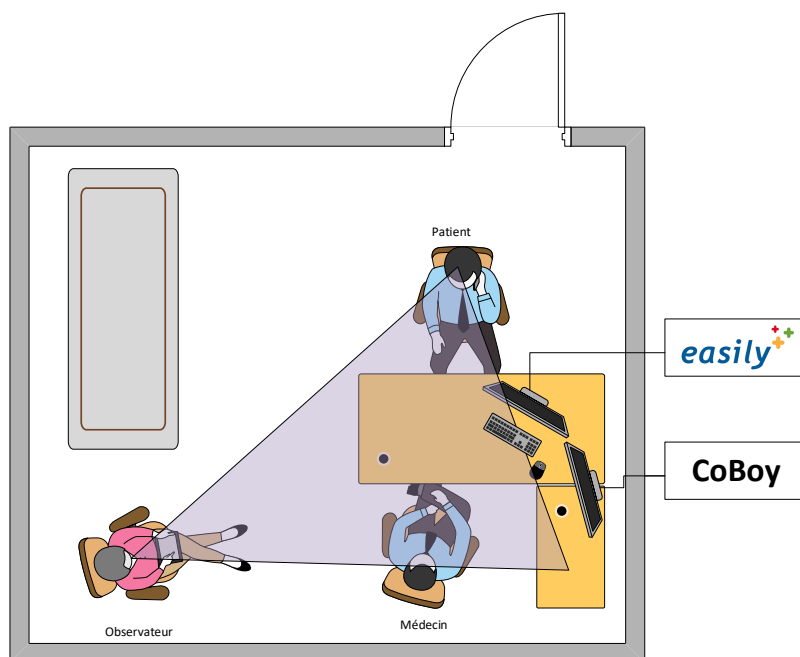


Figure 10.3.: Disposition des agents lors des essais cliniques de l'outil d'aide à la décision

- Le nombre de clics de souris effectués par la·e médecin dans Easily[®], pour quantifier les interactions médecin-HIS ;
- La durée de la consultation.

Notre hypothèse de base est que, si les informations fournies par l'outil d'aide à la décision sont utiles aux médecins, ceux-ci auront moins besoin de rechercher d'autres informations. En d'autres termes, plus le nombre d'informations fournies par l'outil augmente, plus le nombre d'informations recherchées par les médecins devrait diminuer. De même, nous avons choisi le nombre de clics de souris effectués par les médecins dans Easily[®] comme indicateur de leurs interactions avec leur HIS lorsqu'ils recherchent ou saisissent des informations. Ainsi, plus le nombre d'informations fournies par l'outil augmente, plus le nombre de clics effectués par les médecins devrait diminuer. Concernant la durée de la consultation, notre hypothèse de base est que l'utilisation de notre outil ne devrait pas l'impacter, ni positivement ni négativement, car la durée d'une consultation est fixée à l'avance par les médecins.

Cependant, comme nous avons pu le détailler dans le Chapitre 3, l'introduction d'un nouvel outil dans le processus de travail des médecins ne se fait pas sans un certain coût d'entrée. Ainsi, nous devrions observer, dans tous les cas, une augmentation du nombre d'informations recherchées, du nombre de clics et de la durée de la consultation du fait de l'introduction de CoBoy dans le processus des médecins.

Toutefois, si nous n’observons qu’une faible corrélation entre l’augmentation de ces trois critères et l’augmentation du nombre d’informations fournies par l’outil, nous pourrions avancer que l’aide apportée par l’outil a pu compenser cet inévitable coût d’entrée.

La seconde partie de notre étude portait sur l’évaluation de l’acceptabilité potentielle de l’outil proposé. Pour ce faire, nous avons demandé aux médecins observés de remplir un court questionnaire (cf. Annexe B) à la fin d’une séance d’observation, c’est-à-dire une fois toutes leurs consultations prévues terminées. Afin d’essayer de minimiser l’impact de la présence de l’observateur·ice sur les réponses des médecins, les questionnaires ont été remplis anonymement puis mélangés. Les deux parties de notre étude ne peuvent donc pas être reliées entre elles. Ce questionnaire comportait trois questions principales visant à évaluer respectivement la facilité d’utilisation de l’outil, l’utilité perçue de l’outil et son acceptabilité en pratique :

1. *Diriez-vous que la prise en main de CoBoy est :*
 - *Très Facile*
 - *Plutôt Facile*
 - *Ni Facile / Ni Difficile*
 - *Plutôt Difficile*
 - *Très Difficile*
2. *Diriez-vous que, en consultation, CoBoy est :*
 - *Vraiment Utile*
 - *Plutôt Utile*
 - *Ni Utile / Ni Inutile*
 - *Plutôt Inutile*
 - *Totalement Inutile*
3. *Si les HCL intégraient CoBoy dans Easily[®], utiliseriez-vous CoBoy durant vos consultations ?*
 - *Certainement*
 - *Plutôt Oui*
 - *Ne sais pas*
 - *Plutôt Non*
 - *Pas du tout*

Pour chacune de ces questions, il a été proposé aux médecins qui le souhaitent de compléter leurs réponses via un champ "commentaires additionnels". Enfin, deux questions plus générales ont été posées aux médecins pour les amener à donner leur avis général sur l'outil et à proposer des suggestions d'amélioration :

- Avez-vous des commentaires généraux à propos de CoBoy ?
- Avez-vous des suggestions d'amélioration ?

10.2.2 Résultats

Au total, nous avons pu observer 49 consultations effectuées par 7 médecins travaillant aux HCL, spécialisés en endocrinologie et utilisant quotidiennement Easily[®]. Entre 4 et 11 consultations ont pu être observées pour chaque médecin, avec une moyenne de 7 observations par médecin. La Tableau 10.1 résume, pour chaque médecin observé, les moyennes et les écarts-types des résultats obtenus pour chaque critère observé.

Médecin	Nombre d'observations	Nombre moyen d'informations fournies	Nombre moyen d'informations recherchées	Nombre moyen de clics	Durée moyenne d'une consultation (en minutes)
1	8	14 ± 3	14 ± 2	82 ± 26	23 ± 2
2	6	11 ± 5	14 ± 5	126 ± 44	26 ± 7
3	4	14 ± 1	15 ± 2	188 ± 78	35 ± 11
4	11	8 ± 6	8 ± 2	96 ± 32	18 ± 4
5	8	12 ± 6	19 ± 5	97 ± 39	23 ± 6
6	6	15 ± 1	16 ± 1	208 ± 69	36 ± 10
7	6	4 ± 4	15 ± 4	100 ± 57	23 ± 8

Table 10.1.: Synthèse des résultats pour chaque médecin lors des essais cliniques

Pour chaque médecin, à l'exception du 7ème, le nombre moyen d'informations fournies et le nombre moyen d'informations recherchées sont assez proches. Cela peut s'expliquer par plusieurs phénomènes que nous avons pu observer lors des essais cliniques :

- Dans certains cas, les informations fournies correspondent aux besoins des médecins, mais ceux-ci demandent aux patient·es de leur fournir cette même information afin de la valider ;
- Dans d'autres cas, les informations fournies correspondent aux besoins des médecins, mais ces informations ne sont pas à jour et les médecins doivent les rechercher eux-mêmes. Ce cas de figure s'est régulièrement produit durant nos observations, en particulier en ce qui concerne les informations biologiques provenant de laboratoires externes aux HCL et dont les données n'étaient pas accessibles via Easily[®] ;

- Dans d'autres cas encore, les informations fournies correspondent aux besoins des médecins, mais ceux-ci ont besoin d'informations supplémentaires ;
- Enfin, dans certains autres cas, les informations fournies ne correspondent que peu, voire pas du tout, aux besoins des médecins, qui se retrouvent alors dans une situation de consultation classique.

Nous observons aussi, dans la Tableau 10.1, que les interactions avec Easily[®], mesurées par le nombre moyen de clics de souris, varient fortement d'un-e médecin à un-e autre. Selon les cas de figure observés en pratique, certains médecins ont développé un processus de travail optimisé par rapport à leur besoins, tandis que d'autres vont explorer plusieurs pistes avant de clore leurs consultations. Les médecins qui travaillent de cette dernière manière sont ceux qui interagissent le plus avec Easily[®]. Concernant la durée des consultations observées, sauf exception, celles-ci apparaissent assez stables d'une consultation à l'autre, les médecins cherchant à s'en tenir autant que possible à leur agenda.

Afin de relever une tendance générale dans l'impact de notre outil sur le processus de travail des médecins observés, nous avons calculé les coefficients de corrélations, via la méthode de Pearson, entre les différents critères cités précédemment. Le coefficient de corrélation de Pearson $\rho_{X,Y} \in [-1, 1[$ entre deux variables numériques X et Y se calcule, comme détaillé en Équation (10.2), via l'espérance $\mathbb{E}$, la moyenne μ et l'écart-type σ de ces deux variables. Plus ce coefficient est proche de 1, plus les variables X et Y sont corrélées positivement, et corrélées négativement si $\rho_{X,Y}$ est proche de -1 . À l'inverse, plus ce coefficient est proche de 0, moins les deux variables sont corrélées.

$$\rho_{X,Y} = \frac{\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (10.2)$$

En calculant ces coefficients de corrélation pour chaque paire de critères observés, nous obtenons la matrice de corrélation présentée en Figure 10.4. Nous pouvons tout d'abord y observer que la durée d'une consultation et le nombre de clics de souris effectués durant cette consultation sont très fortement corrélés, avec un coefficient de corrélation à 0.88. Dans une moindre mesure, la durée d'une consultation est aussi corrélée au nombre d'informations recherchées par la-e médecin durant cette consultation. Ces deux points font sens, puisque l'on peut aisément supposer que plus un-e médecin recherche des informations ou plus iel a besoin d'interagir avec Easily[®], plus la durée de la consultation est rallongée. En revanche, le nombre d'informations fournies par l'outil ne semble que peu corrélé à la durée des consultations observées,

ce qui valide notre hypothèse en la matière. La raison en est que les médecins tentent autant que possible de s'en tenir à des durées de consultation fixes.

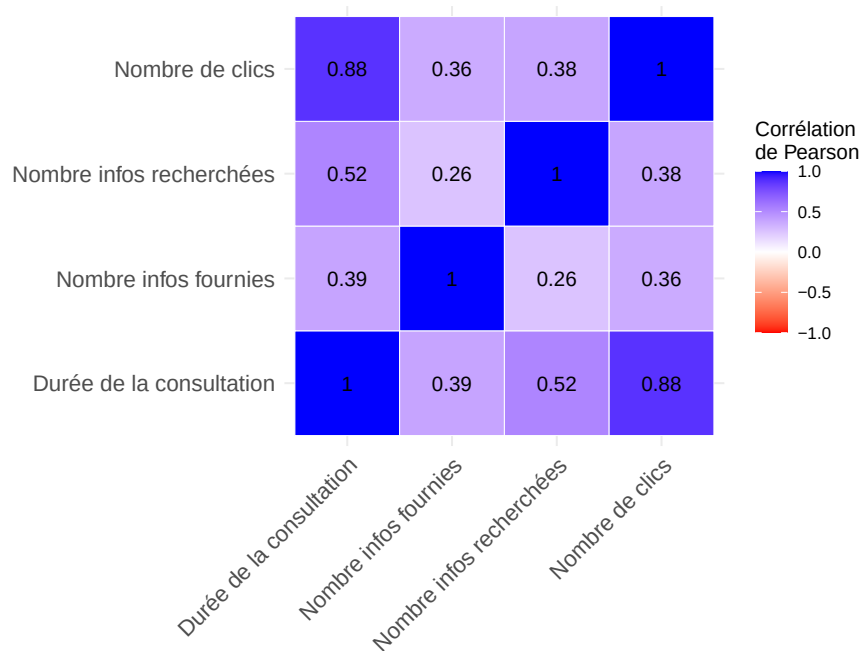


Figure 10.4.: Matrice de corrélations entre les différents critères observés

Nous observons aussi que le nombre d'informations fournies par l'outil n'est que faiblement corrélé au nombre d'informations recherchées par les médecins et au nombre de clics effectués par les médecins. Bien que les coefficients calculés pour ces trois critères ne soient pas négatifs, ceux-ci sont relativement proches de 0. Ces résultats vont ainsi dans le sens de nos hypothèses de départ, car l'augmentation du nombre d'informations n'apparaît pas corrélée à une augmentation du nombre de clics, ni à une augmentation du nombre d'informations recherchées par les médecins. Au contraire, l'aide apportée par notre outil semble avoir permis de compenser son coût d'entrée.

Concernant les résultats de notre questionnaire d'appréciation de notre outil, ceux-ci apparaissent bien plus positifs et sans équivoque. La Figure 10.5 montre la répartition des réponses des médecins aux trois questions principales portant sur la facilité de prise en main de l'outil, son utilité perçue et son acceptabilité potentielle en pratique. Comme les cinq réponses possibles diffèrent d'une question à une autre, nous avons choisi d'établir une échelle allant de -2 à +2, -2 et -1 correspondant aux réponses négatives, 0 à la réponse neutre et +1 et +2 aux réponses positives. Nous avons aussi décidé d'accepter les réponses intermédiaires, certains médecins n'ayant pas pu se décider entre deux propositions adjacentes.

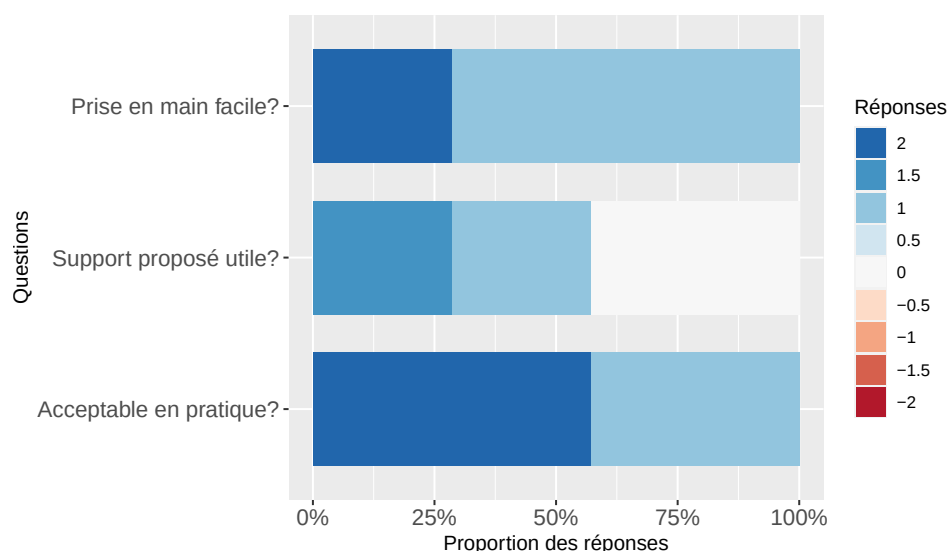


Figure 10.5.: Synthèse des réponses au questionnaire d'appréciation

Nous observons que 75% des médecins, soit 5 médecins sur 7, ont trouvé la prise en main de CoBoy "plutôt facile". Les deux médecins restants ont considéré la prise en main de CoBoy "très facile". La gêne occasionnée par le fait d'avoir à utiliser un second écran, ainsi que par une interface utilisateur assez sommaire, ont été les principales remarques concernant la prise en main de CoBoy.

Concernant l'utilité ressentie de l'outil, les résultats sont plus partagés, bien que globalement positifs. Plus de 50% des médecins interrogés, soit 4 médecins sur 7, ont trouvé CoBoy "plutôt utile", voire presque "vraiment utile". L'autre moitié des médecins interrogés n'a trouvé CoBoy "Ni utile, ni inutile". D'après les remarques faites par ce second groupe de médecins, ceux-ci ne pouvaient estimer l'utilité pratique de l'outil du fait de certaines limites rencontrées lors des essais, mais ne trouvaient pas l'approche proposée "inutile".

Enfin, concernant l'acceptabilité potentielle de l'outil en pratique, plus de la moitié des médecins interrogés a répondu qu'ils utiliseraient "certainement" CoBoy durant leurs consultations, si celui-ci était intégré dans Easily® et si les limites rencontrées étaient surmontées. L'autre moitié a répondu "plutôt oui" à cette même question et sous les mêmes conditions. Cette forte acceptabilité, en comparaison de la faible acceptabilité des DDSS actuels (cf. Chapitre 2), étaye le bien-fondé de notre choix d'une approche "ajustive" pour l'aide à la décision en consultation. Malgré les limites actuelles de notre outil, les médecins observés se sont montrés particulièrement enthousiastes vis-à-vis des possibilités offertes par l'approche "ajustive" de l'aide à la décision sur laquelle nous nous basons. Notons que l'intégration des médecins dans le processus de conception de l'outil a aussi pu favoriser cette forte acceptabilité.

Les commentaires généraux les plus récurrents portaient, pour les plus négatifs d'entre eux, sur les limites techniques et ergonomiques actuelles de l'outil, et pour les plus positifs d'entre eux, sur le confort de travail offert par le fait de disposer d'une synthèse des informations à propos de leur patient·e pour toute la durée de la consultation. Les suggestions d'amélioration portaient principalement sur des propositions pour pouvoir accéder aux données biologiques manquantes, ainsi que sur diverses améliorations possibles de l'interface utilisateur.

10.3 Limites et perspectives

Comme nous l'avons vu, une des principales limites à la viabilité de l'aide apportée était l'incapacité de fournir des données à jour pour certaines informations biologiques issues de laboratoires externes aux HCL. Cependant, les raisons de ces limites sont externes à notre outil d'aide à la décision et proviennent du contexte d'utilisation de celui-ci. En effet, bien que de nombreuses données sur les patient·es soient disponibles et accessibles via les bases de données des HCL, si les données nécessaires à la prise de décision ne sont pas disponibles, alors l'aide apportée par l'outil est nulle. Ainsi, il apparaît évident qu'un outil comme CoBoy nécessite de se reposer sur une structure solide et viable. Si cette base est absente, l'aide apportée par un outil dit "ajustif" sera faible, voire inexistante, puisque l'outil n'aura pas de structure de base à "ajuster".

Notre outil étant encore à l'état de prototype lors de nos essais, et n'étant pas intégré dans le processus de travail des médecins, il reste difficile d'estimer son impact réel en situation pratique. Nous avons, de plus, relevé que la simple intégration d'un second écran dans l'espace de travail des médecins n'était pas sans impact. La première raison en est que les médecins sont actuellement habitués à travailler avec un seul écran et le fait de devoir naviguer entre deux écrans, ou simplement de penser à regarder le second écran, peut générer une gêne dans leurs processus de travail habituels. La seconde raison est la place physique que prend un deuxième écran dans une salle de consultation qui, en fonction de sa position, peut créer une barrière supplémentaire entre la·e médecin et la·e patient·e. Cependant, certains médecins ont relevé le confort de travail permis par le fait d'utiliser de deux écrans, en particulier le fait de pouvoir saisir des comptes rendus sur un écran tout en affichant des informations concernant la·e patient·e sur l'autre écran. Une réflexion autour de la place physique des outils informatiques dans une salle de consultation pourrait donc être pertinente afin de trouver un équilibre entre confort de travail et interaction médecin-patient·e.

Malgré ces limites, l'outil CoBoy, ainsi que l'approche "ajustive" que nous proposons, ont suscité un certain intérêt auprès des médecins ayant participé aux essais. Un certain nombre de propositions ont été faites afin d'améliorer l'outil et de pallier les problèmes rencontrés lors des essais, les médecins ressentant l'utilité potentielle de l'outil si ces problèmes pouvaient être surmontés. Concernant les données biologiques externes, par exemple, il a été proposé de saisir ces données dans Easily[®], soit par les médecins durant la consultation, soit par leurs secrétaires médicales en amont de la consultation, ou bien de les saisir directement via l'interface de CoBoy. Cependant, cela pourrait générer une charge de travail trop importante en comparaison de l'aide apportée, un équilibre sera à trouver. L'idéal serait, sans doute, de pouvoir extraire sans risque d'erreurs les données biologiques des comptes rendus de laboratoire, ou bien d'avoir une structure informatique commune entre les HCL et les laboratoires externes aux HCL et que cette structure soit utilisée par les laboratoires pour transmettre aux médecins les résultats de leurs analyses.

De nombreuses suggestions ont aussi été faites concernant l'interface utilisateur de CoBoy. L'idée principale, ressortant des discussions avec les médecins, est de pouvoir avoir, en seul coup d'œil, une synthèse des informations importantes concernant la·e patient·e en cours, avec un accès rapide aux informations "secondaires". L'interface proposée par GIORDANENGO *et al.*, (2019), pour le suivi de patient·es diabétiques, offre un aperçu de ce vers quoi pourrait tendre l'interface utilisateur de CoBoy, mais pour tout type de pathologies. Il est aussi envisagé de donner aux médecins la possibilité d'indiquer à l'outil d'afficher une information de telle ou telle manière, via un système de configuration de l'interface, et sous telles ou telles conditions, via un système de création de règles. Le fait de pouvoir configurer leur environnement de travail pourrait améliorer le confort de travail des médecins, mais aussi leur appréciation de l'outil. Cependant, cela peut aussi générer une charge de travail supplémentaire. Encore une fois, un équilibre devra être trouvé entre modularité de l'outil et configuration de base.

En résumé, bien qu'il soit actuellement difficile d'estimer si l'outil impacte positivement ou négativement le processus de travail des médecins, l'approche proposée est majoritairement appréciée et acceptée par les médecins ayant participé aux essais. Les nombreuses suggestions faites par les médecins pour une meilleure intégration de CoBoy dans leur processus de travail mettent en évidence un intérêt certain pour l'outil proposé, mais aussi pour l'approche "ajustive" sur laquelle il se base. L'intégration d'un outil tel que CoBoy, cherchant à fournir aux médecins un sous-ensemble d'informations ciblées, apparaît adapté aux besoins des médecins durant leurs consultations et semble bien plus acceptable pour les médecins que les approches actuelles des DDSS (cf. Chapitre 3). Il pourra cependant être pertinent de

réitérer de tels essais une fois l'outil bien intégré dans le processus de travail des médecins, afin d'estimer son impact et son acceptabilité réels.

Conclusion

Les travaux présentés dans cette thèse se sont concentrés sur la question de l'aide à la décision en situations coutumières en milieux hospitaliers. Plus spécifiquement, nous avons étudié comment aider à la décision des médecins, considérés experts et responsables, durant leurs consultations médicales. À travers le traitement de ce sujet, nous avons été amenés à travailler sur un certain nombre d'aspects de l'aide à la décision en consultations médicales : les différentes approches envisageables pour aider les médecins dans leurs décisions, le processus de décision et les besoins des médecins en consultation, ou encore l'impact de la transparence des outils utilisés sur leur acceptabilité par les médecins.

Afin d'apporter une conclusion à ce manuscrit, la Section 11.1 propose une synthèse de nos différentes contributions. La Section 11.2, quant à elle, synthétise les différentes limites de nos travaux, au regard desquelles nos résultats doivent être interprétés. Enfin, la Section 11.3 propose plusieurs pistes de réflexion pour de futurs travaux sur des thématiques proches de celles traitées dans cette thèse.

11.1 Contributions

Au cours de ce manuscrit nous avons défendu la thèse qu'il est plus adapté et acceptable de respecter le savoir-faire des médecins et de leur laisser la responsabilité des décisions prises en consultations, en se limitant à leur fournir les informations sur leurs patient-es qui leur sont nécessaires pour leur prises de décision. Nous avons ainsi été amenés à traiter un certain nombre de problématiques relatives à l'aide à la décision en consultations médicales.

- Quels sont les outils qui sont actuellement utilisés pour aider la prise de décision en milieux hospitaliers ?
- Comment travaillent les médecins durant leurs consultations et quels sont leurs besoins ?
- Comment proposer une aide adaptée et acceptable pour les médecins ?

Nous proposons dans cette section une synthèse de nos différentes contributions sur ces différentes problématiques.

11.1.1 Analyse critique des outils d'aide à la décision en milieux hospitaliers

Dans la première partie de cette thèse, dont les points abordés ont aussi été publiés dans RICHARD *et al.*, (2020b), nous nous sommes intéressés aux différentes approches et outils développés afin d'apporter une aide à la décision en milieux médicaux. Notre objectif était de déterminer si les outils d'aide à la décision actuels, ou passés, pouvaient s'adapter aux contraintes d'une aide à la décision dans des situations coutumières comme le sont les consultations médicales. Pour cela, nous avons retracé l'historique des différents courants qui ont présidé au développement d'outils d'aide à la décision en milieux médicaux, en s'intéressant plus particulièrement aux outils d'aide aux diagnostics, ou DDSS. Nous nous sommes ensuite intéressés à l'impact de ces outils sur les performances des médecins, ainsi qu'à leur acceptabilité par les médecins en situations pratiques. Nous avons ainsi pu mettre en évidence que, malgré un impact potentiellement bénéfique sur les performances des médecins, perçu par ceux-ci, les outils d'aide à la décision actuels sont très peu acceptés en pratiques.

Afin de comprendre ce constat quelque peu paradoxal, nous avons investigué les différentes barrières pouvant expliquer la non-acceptation des outils d'aide à la décision actuels. Nous avons montré que, en pratique, l'introduction d'outils informatiques en milieux hospitaliers n'est pas nécessairement à l'avantage du personnel soignant utilisant ces outils. En effet, ceux-ci doivent s'adapter à l'utilisation d'un nouvel outil, ce qui peut générer une charge supplémentaire de travail, du stress et des erreurs dans d'autres aspects de leur travail. Nous avons aussi relevé que l'utilisation des outils d'aide à la décision actuels soulève de nombreux problèmes en termes de responsabilité. En particulier, en cas d'erreurs médicales, si la responsabilité doit reposer sur les médecins, alors ceux-ci se doivent de rester en alerte des erreurs pouvant venir de l'outil d'aide à la décision, générant là aussi une surcharge de travail pouvant amener à plus d'erreurs médicales. Ce constat est d'autant plus déplorable lorsque l'outil proposé est supposé apporter une aide à ses utilisateur·ices. Il est donc important, lors de l'élaboration d'un outil d'aide à la décision, de bien évaluer l'effort d'adaptation qui sera demandé aux médecins au regard de l'apport proposé par l'outil.

Enfin, nous avons mis en évidence que la principale barrière à l'utilisation des outils d'aide à la décision actuels en situations coutumières venait d'un choix d'approches inadaptées aux contraintes du contexte qui nous intéresse. En effet, les outils actuels reposent sur des approches que nous avons qualifiées de "conformiste" ou "objectiviste", selon la taxonomie proposée par MEINARD et TSOUKIÀS, (2019), ce qui reflète des choix idéologiques forts, parfois non assumés, sur ce qu'est l'aide à la décision. Comme nous l'avons montré, ces choix idéologiques ne correspondent ni aux objectifs, ni aux besoins, ni aux contraintes des médecins durant leurs consultations médicales coutumières. Dans de telles situations, l'esprit d'initiative des médecins, ainsi que leur responsabilité, sont hautement engagés. Ce sont les contraintes les plus importantes à prendre en compte dans la proposition d'un outil d'aide à la décision en consultations médicales coutumières. Ainsi, nous avons proposé que développer un outil basé sur une approche "ajustive" de l'aide à la décision, c'est-à-dire ajustant l'aide apportée aux besoins des décideur·ses, serait plus adaptée aux contraintes d'une aide en consultations coutumières.

11.1.2 Modélisation du processus de décision en consultation médicale

En parallèle de notre analyse des outils d'aide à la décision actuels, nous avons observé et analysé le processus de travail des médecins lors de leurs consultations. Ces analyses reprennent et approfondissent des travaux présentés dans RICHARD *et al.*, (2018). Nous avons mis en évidence, via ces analyses, les différents types d'actions possibles pour les médecins durant une consultation. La recherche d'informations concernant la·e patient·e en cours a été identifiée comme une activité indispensable aux médecins pour décider des prescriptions à faire à leurs patient·es. Fort de ces analyses, nous avons pu proposer deux modélisations du processus de décision en consultations afin de mieux en comprendre les rouages.

Le premier modèle proposé est dédié à la formalisation de consultations spécifiques, afin de détailler les décisions prises par les médecins au fur et à mesure de leurs consultations. Ce premier modèle nous a permis de mettre en évidence que les décisions des médecins reposent principalement sur les informations collectées sur la·e patient·e au cours de la consultation. Les médecins peuvent alors décider de rechercher une nouvelle information encore inconnue ou de faire une prescription à leur patient·e. La consultation prend fin lorsque la·e médecin estime que plus aucune prescription ne peut être faite.

Nous avons ensuite proposé un second modèle se voulant plus générique et permettant de mettre en évidence les liens entre les différentes actions/décisions possibles pour les médecins lors de leurs consultations. Ce second modèle nous a permis de mettre en évidence le processus cyclique d'une consultation. Dans ce processus, l'accumulation d'informations sur la·e patient·e en cours constitue un sous-processus central et essentiel au bon déroulé du processus de décision général.

À partir de ces deux modèles, nous avons mis en évidence le besoin pour les médecins d'accumuler des informations sur leurs patient·es. Cependant, l'accès à ces informations peut être fastidieux pour les médecins, quand bien même ces informations seraient accessibles via leur HIS. Nous avons donc émis l'hypothèse qu'un outil adapté aux besoins des médecins en consultations coutumières serait un outil capable d'apprendre et d'anticiper de quel sous-ensemble d'informations ont besoin les médecins en fonction de la pathologie de leurs patient·es. Ce type d'outil correspondrait ainsi à l'approche "ajustive" que nous voulons expérimenter.

Enfin, nous avons montré que le problème qu'un tel outil devrait résoudre peut se formaliser comme un problème de classification multi-labels, où chaque label correspond à une information pouvant potentiellement être utile aux médecins. Ce type de problèmes peut alors être traité informatiquement via des systèmes de classification reposant sur des algorithmes d'apprentissage.

11.1.3 Proposition de critères opérationnels pour l'évaluation de la "transparence" de systèmes de classification

Dans la première partie de cette thèse, nous avons mis en évidence une non-acceptation des outils d'aide à la décision, en partie du fait de la méfiance générée envers les algorithmes utilisés, qui sont parfois obscurs pour les utilisateur·ices. Dans l'objectif de proposer un outil d'aide à la décision acceptable pour les médecins, quand bien même celui-ci serait basé sur un système de classification, nous avons investigué quels pourraient être les critères maximisant la "transparence" du système à développer. Cette direction a été prise car un certain nombre d'études a montré l'impact bénéfique que peut avoir la "transparence" d'un algorithme sur son acceptabilité par les utilisateur·ices. Nos travaux sur ces questions ont été présentés dans RICHARD *et al.*, (2020a).

Dans un premier temps, nous avons proposé une définition générale regroupant les différentes exigences que nous nous sommes imposées en termes de "transparence". Parmi ces exigences, nous avons répertorié : la compréhensibilité du système,

l'interprétabilité de l'algorithme et des résultats, la retraçabilité du système et la révisabilité du système. Pour chacune de ces exigences, nous avons proposé une série de critères opérationnels qui nous permettraient d'évaluer la "transparence" d'un système de classification.

Nous avons ensuite étudié différents systèmes de classifications existants, reposant sur des approches bien distinctes, et nous avons évalué leur "transparence" à partir de nos critères opérationnels. Notre objectif était de sélectionner le système de classification répondant à la totalité des exigences que nous nous étions fixées. Si aucun des systèmes étudiés n'avait respecté ces exigences, il nous aurait fallu envisager d'en proposer un nouveau. Finalement, le système de classification *Naive Bayes* a été évalué comme répondant à toutes nos exigences en termes de transparence, en grande partie du fait de la "simplicité" de son algorithme d'apprentissage et du fait que la théorie des probabilités sur laquelle il s'appuie est connue des médecins.

Enfin, nous avons effectué une série d'expérimentations afin d'évaluer si le choix d'un système "transparent" engendrerait une perte de performances significative. Si tel avait été le cas, il nous aurait fallu modifier le système sélectionné ou envisager d'en utiliser un moins transparent mais plus performant. En sortie de nos expérimentations, il s'est avéré que la "transparence" n'était pas nécessairement corrélée à une baisse de performances. Et, bien que le système *Naive Bayes* ne soit pas des plus performants, pour les cas applicatifs simples comme celui que nous traitons ses performances restent convenables.

11.1.4 Développement d'un outil d'aide à la décision pour les consultations médicales

Après nous être positionnés sur l'approche qui nous paraissait la plus pertinente en termes d'aide à la décision en situations de consultations médicales coutumières, après avoir étudié le processus de travail des médecins et leurs besoins en termes de support, et après avoir sélectionné un système de classification répondant à nos critères de transparence afin de maximiser son acceptation par les médecins, nous avons pu développer un outil d'aide à la décision pour les consultations médicales coutumières.

Cet outil, nommé CoBoy, prend la forme d'un assistant personnel qui, à partir de l'âge, du sexe et de l'IMC des patient·es, ainsi que la pathologie pour laquelle iels sont suivis, va anticiper et fournir les informations nécessaires aux médecins pour leurs consultations. L'outil a été développé en collaboration avec les médecins du service

endocrinologie et diabétologie des HCL, il est donc calibré pour les pathologies liées à ce service. Actuellement, CoBoy prend en compte une vingtaine de pathologies et est capable de rechercher parmi une centaine d'informations patient en fonction des besoins des médecins.

Enfin, nous avons effectué une série d'essais cliniques de l'outil proposé afin d'évaluer la faisabilité de son introduction dans le processus de travail des médecins. Bien que l'impact sur le processus de travail des médecins ne soit pas parfaitement évaluable, les premiers résultats sont positifs. De plus, les médecins ayant essayé l'outil durant ces essais ont montré un grand intérêt pour celui-ci et se sont montrés impliqués dans leurs suggestions d'améliorations. Bien que beaucoup de travail reste à faire, l'outil proposé semble correspondre aux besoins des médecins en consultations médicales et aux contraintes induites par une aide à la décision en de telles situations.

11.2 Limites

Afin d'avoir une meilleure appréciation de nos différentes contributions, il est important de prendre en compte les limites de nos travaux. Dans cette section, nous proposons de revenir sur les différentes limites que nous avons pu relever au cours des différents chapitres de cette thèse. Celles-ci se répartissent entre les limites venant de notre cas applicatif, les limites techniques venant de notre contexte applicatif et les limites liées aux cadres d'analyse et aux paradigmes mis à contribution.

Comme nous avons pu le détailler tout au long de cette thèse, et plus particulièrement en Partie I, nous nous sommes concentrés sur la proposition d'un outil d'aide à la décision adapté aux consultations médicales coutumières. Dans ces situations, les médecins sont considérés comme des experts et leur responsabilité est hautement engagée. Ce n'est, bien entendu, pas le cas pour toutes les situations où des décisions médicales sont prises. La majeure partie de nos conclusions, en particulier celles concernant l'approche choisie, n'est donc pas nécessairement pertinente pour des cas applicatifs autres que les consultations médicales coutumières, ou plus généralement, pour des cas où la responsabilité des décideur·ses n'est pas aussi fortement engagée.

De manière similaire, nous avons étudié différents systèmes informatiques à travers le prisme des contraintes et des enjeux liés au contexte médical dans lequel nous avons travaillé. Bien que cela nous ait permis de mettre en évidence un certain nombre d'exigences qui s'imposent aux outils que nous pouvions utiliser, ces exigences sont potentiellement spécifiques à un contexte médical. De plus, nous nous

sommes concentrés sur l'étude de certains types de systèmes, tels que les DDSS et les systèmes de classification, appliqués au cas des consultations médicales. Nos conclusions concernant l'utilisation de ces systèmes ne sont donc pas nécessairement applicables à des systèmes de classification utilisés dans d'autres conditions que les consultations médicales, ou à d'autres types de systèmes informatiques utilisés durant des consultations médicales.

Le contexte médical dans lequel s'est déroulée cette thèse n'a pas non plus été sans contraintes. L'une des barrières majeures avec lesquelles nous avons dû travailler a été la disponibilité limitée des médecins. En effet, ceux-ci ont un emploi du temps chargé avec lequel il nous a fallu composer. Cette contrainte nous a particulièrement limités lors de nos analyses du processus de travail des médecins, présentées au Chapitre 6 et au Chapitre 10, pour lesquelles un panel d'observations plus conséquent nous aurait permis de tirer des conclusions plus robustes. Nous avons tout de même pu relever des tendances générales, comme les modèles présentés au Chapitre 7, mais qui gagneraient à être vérifiées avec des observations de plus grandes ampleurs.

Enfin, lors du développement de notre outil, nous avons principalement travaillé avec des médecins spécialisés en endocrinologie. L'outil d'aide à la décision que nous proposons est donc, par construction, optimisé pour les consultations de cette spécialité. C'est le cas en particulier du système de classification utilisé, car celui-ci n'a appris qu'à partir des consultations d'endocrinologie. Cependant, nous avons construit notre outil dans une optique de généricité, afin de pouvoir l'utiliser pour des consultations dans d'autres spécialités médicales. De plus, nos analyses ont mis en évidence certaines similitudes dans le processus de travail des médecins, indépendamment de leur spécialité médicale (cf. Chapitre 6), ce qui laisse supposer que notre approche pourrait être facilement adaptable pour des spécialités médicales autres que l'endocrinologie.

11.3 Perspectives

Les travaux présentés dans cette thèse nous ont permis de mettre en évidence des axes de recherche intéressants pour de futurs travaux. Dans cette section, nous proposons d'approfondir les perspectives qui ont pu être émises dans les différents chapitres de cette thèse.

11.3.1 Adapter l'aide proposée à d'autres services hospitaliers

Dans le Chapitre 10, nous avons détaillé un ensemble d'améliorations possibles pour l'outil d'aide à la décision développé au cours de cette thèse. Mises à part des limitations du fait de la communication entre les différentes structures du réseau médical, qui n'est pas vraiment du ressort de la recherche scientifique, nous avons relevé qu'un gros travail serait à faire sur l'interface utilisateur. De plus, l'outil que nous avons proposé est spécifiquement dédié aux consultations médicales et particulièrement à celles concernant des pathologies chroniques suivies en endocrinologie, puisque c'est avec des médecins de ce type de service que nous avons travaillé. Bien que nous ayons tenté de proposer une aide la plus générique possible, il est tout à fait possible que l'outil proposé ne corresponde pas totalement aux besoins de médecins d'autres services durant leurs consultations. Ainsi, il pourrait être intéressant d'étudier, via d'autres essais cliniques, l'applicabilité de l'outil proposé dans d'autres types de services hospitaliers.

Pour aller plus loin, il est tout à fait possible que les contraintes d'une aide à la décision en consultations médicales coutumières, c'est-à-dire l'importance de l'autonomie et de la responsabilité du/de la décideur·se, se retrouvent dans d'autres situations hospitalières. De plus, nous pouvons supposer que les besoins du personnel soignant dans ces autres situations ne sont pas nécessairement un besoin en informations comme pour les consultations médicales. Ainsi, il serait intéressant d'étudier l'applicabilité d'une approche "ajustive" de l'aide à la décision dans des situations médicales autres que celles étudiées dans cette thèse. Plus précisément, il serait intéressant d'étudier les contraintes qui s'imposent à une aide à la décision dans diverses situations médicales pour estimer quelle serait l'approche la plus adaptée à chaque situation.

Une situation médicale qu'il serait intéressant de traiter dans de futurs travaux de recherche est celle de la prise en charge des patient·es en services d'urgences. Dans de telles situations, le personnel soignant est amené à prendre rapidement de nombreuses décisions, autant dans l'ordre de priorité des patient·es à traiter que dans l'affectation des ressources médicales. De plus, nous pouvons supposer que le transfert des informations patient au sein du service d'urgence, entre le service d'urgence et les urgentistes, ainsi qu'entre le service d'urgence et les autres services de l'hôpital, est d'une importance cruciale pour la bonne prise en charge des patient·es. Si des algorithmes d'ordonnancement ou de gestion des ressources sont envisageables, il y a fort à parier que ceux-ci génèreraient de la méfiance de la part du personnel soignant et seraient peu utilisés en pratique. Ainsi, étudier les

différents processus de travail et de décision de ce type de service peut être un axe de recherche extrêmement exigeant et stimulant.

11.3.2 Repenser la place des outils informatiques en milieux hospitaliers

Dans la section précédente, nous avons développé différentes perspectives en lien avec la création de nouveaux outils d'aide à la décision en milieux médicaux. Cependant, il est également important de réfléchir à l'amélioration de l'ensemble des outils informatiques utilisés par le personnel soignant et à la place qu'occupent ces outils dans les processus de prise en charge des patient·es.

Dans le Chapitre 2, nous avons évoqué l'intérêt d'impliquer les utilisateur·ices dans le processus de conception d'outils informatiques pour améliorer leur acceptabilité et anticiper de potentielles barrières à leur utilisation en pratique. L'implication des médecins pour l'amélioration de l'outil que nous avons développé, sur laquelle nous avons insisté au Chapitre 10, semble montrer que les médecins apprécient le fait d'être intégrés dans la conception des outils qu'ils pourraient être amenés à utiliser au quotidien. GIORDANENGO *et al.*, (2019) offrent une autre illustration de l'intérêt d'intégrer les médecins dans le processus de conception, ceux-ci mettant en évidence de nombreux points d'améliorations et potentielles barrières. Un axe de recherche intéressant serait donc d'investiguer les méthodes de conception pouvant faciliter l'intégration des membres du personnel soignant dans la conception de nouveaux outils ou l'amélioration d'outils existants. Par exemple, proposer des interfaces utilisateur librement configurables par les médecins peut permettre un meilleur confort de travail. Il peut aussi être intéressant d'analyser, avec les utilisateur·ices, la manière dont ils utilisent leurs outils afin de dégager les pratiques communes ou les points d'améliorations potentiels.

Durant nos essais cliniques (cf. Chapitre 10), nous avons pu voir que la simple introduction d'un second écran peut générer autant une gêne qu'un confort de travail pour les médecins. En effet, l'introduction d'un second écran crée une barrière physique entre les médecins et leurs patient·es, et gêne ainsi leurs interactions. Ce constat peut faire écho à ce que nous avons relevé dans le Chapitre 3 concernant les barrières à l'acceptation d'un outil informatique par les médecins. Si l'introduction d'un nouvel outil informatique génère plus de contraintes que d'apports, celui-ci risque fort de ne pas être utilisé en pratique. Dans ce chapitre, nous nous étions concentrés sur l'aspect "logiciel" (ou *software*) des outils informatiques mais, comme nous avons pu le constater durant nos essais cliniques, l'aspect "physique" (ou

hardware) est aussi à prendre en compte. Nous pensons donc que l'étude de la place physique que prennent les outils informatiques dans les divers processus hospitaliers serait un axe de recherche intéressant pour de futurs travaux. Pour reprendre l'exemple des consultations médicales, la conception de supports informatiques spécifiques, comme des écrans modulables, pourrait être intéressante à investiguer.

Pour aller plus loin, la communication d'informations entre les différents outils informatiques utilisés, "logiciels" ou "physiques", peut constituer un autre axe de recherche intéressant. Nous pouvons par exemple penser au domaine de l'internet des objets, ou *Internet of Things* (IoT), qui s'intéresse aux objets connectés et à leurs protocoles de communication. L'utilisation de tels objets peut permettre de créer une "intelligence ambiante" qui fonctionnerait en synergie avec le personnel soignant. Nous pourrions par exemple imaginer une salle de consultation munie de divers objets connectés (balance, tensiomètre, etc.) qui enverraient directement les données collectées au système d'information utilisé par les médecins et à l'outil d'aide à la consultation qui leur afficherait automatiquement ces informations. Bien entendu, l'utilisation de tels outils est à encadrer d'un point de vue éthique, puisque les données utilisées sont extrêmement sensibles et puisque, d'un point de vue pratique, nous ne pouvons anticiper l'impact qu'auraient ces outils sur la qualité des soins apportés aux patient·es.

Enfin, comme nous avons pu le détailler au Chapitre 3, l'impact des outils informatiques, et de leur omniprésence dans les services hospitaliers, sur la qualité des soins apportés aux patient·es, reste sous-étudié. L'introduction de nouveaux outils informatiques dans le processus de travail des médecins peut être à double tranchant si cela impliquent une charge de travail supplémentaire pour les soignant·es. La principale vocation d'un outil informatique hospitalier est de faciliter la prise en charge des patient·es et d'améliorer la qualité des soins. Comme nous l'avons montré et soutenu tout au long de cette thèse, un moyen d'y parvenir est de réduire la charge de travail "administrative" des soignant·es, via des outils innovants, pour que ceux-ci puissent se concentrer sur la prise en charge de leurs patient·es. De même, il nous paraît pertinent d'étudier et d'améliorer les outils informatiques actuellement utilisés par les soignant·es, toujours dans cette optique de leur permettre de se concentrer sur leur domaine d'expertise : la bonne prise en charge des patient·es et la qualité des soins qui leurs sont apportés.

11.3.3 Investiguer l'intérêt d'une approche "ajustive" dans d'autres domaines

Nous pouvons supposer que des contraintes similaires à celles impliquées par une aide à la décision en consultation peuvent se retrouver dans des situations autres que médicales. Des situations où le processus de décision est proche de celui d'un diagnostic, ou des situations où la responsabilité des décideur·ses est fortement impliquée, seraient un parfait terrain d'expérimentation pour une approche "ajustive" de l'aide à la décision.

Dans le contexte de la mise en œuvre de politiques environnementales, par exemple, MEINARD et THÉBAUD, (2019) posent le constat que les plans de gestion environnementaux en France sont actuellement handicapés par le manque d'une large base de données sur les types de végétations, tandis que les institutions environnementales se focalisent sur la production coûteuse de directives non adaptées et inutilisables par les experts sur le terrain. Dans ce type de contexte, un outil d'aide à la décision basé sur une approche "ajustive" et sur l'étude des besoins des experts serait largement bénéfique.

Plus généralement, nous pourrions envisager d'étudier l'intérêt d'une aide à la décision "ajustive" pour les décisions d'intérêt général, pour lesquelles la responsabilité des décideur·ses est hautement engagée. Un axe de recherche intéressant serait le développement d'outils dédiés aux prises de décisions collectives. Nous pourrions par exemple imaginer une "ville intelligente" avec des outils permettant de collecter et de présenter des informations pouvant servir de support lors de conventions citoyennes et des outils permettant de faciliter les délibérations en synthétisant le débat en cours.

Enfin, les questions éthiques soulevées par l'utilisation de méthodes d'aide à la décision dans de telles situations, en particulier de méthodes informatiques, peuvent être particulièrement intéressantes à traiter. Les différentes exigences que nous nous sommes imposées en termes de "transparence", détaillées au Chapitre 8, peuvent servir de base à d'autres travaux du même type pour des cas applicatifs autres que médicaux. De plus, si nous nous sommes intéressés en particulier au concept de "transparence" du fait de notre contexte applicatif, il nous paraît important d'étudier d'autres concepts, tout aussi fondamentaux, liés à l'éthique des algorithmes et de proposer des critères opérationnels afin d'évaluer des algorithmes par le prisme de ces concepts. Nous pouvons par exemple penser aux concepts de "justice" et d'"équité", formalisés entre autres par BEAUCHAMP, CHILDRESS *et al.*, 2009 ; CLÉMENT *et al.*,

2008, qui peuvent être particulièrement adaptés à l'évaluation d'outils utilisés dans la prise de décisions de politiques publiques.

Fiche d'observation

A

Fiche d'observation

Heure d'arrivée du patient (hh:mm)

Informations recherchées

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

Nombres de clics de souris

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

☐

Heure de départ du patient (hh:mm)

Remarques du médecin (optionnel)

Questionnaire d'appréciation

B

Questionnaire

Diriez-vous que la prise en main de CoBoy est :

Très Facile	Plutôt Facile	Ni Facile, Ni Difficile	Plutôt Difficile	Très Difficile

Commentaires additionnels :

Diriez-vous que, en consultation, CoBoy est :

Vraiment Utile	Plutôt Utile	Ni Utile, Ni Inutile	Plutôt Inutile	Totalement Inutile

Commentaires additionnels :

Si les HCL intégraient CoBoy dans Easily, utiliseriez-vous CoBoy
durant vos consultations ?

Certainement	Plutôt Oui	Ne sais pas	Plutôt Non	Pas du tout

Commentaires additionnels :

Avez-vous des commentaires généraux à propos de CoBoy ?

.....

Avez-vous des suggestions d'amélioration ?

.....

Bibliographie

- ABDOLLAHI, Behnoush et Olfa NASRAOUI (2018). "Transparency in fair machine learning : The case of explainable recommender systems". In : *Human and Machine Learning*. Springer, p. 21–35 (cf. p. 97).
- AKKERMANS, Henk, Paul BOGERD et Jan VAN DOREMALEN (2004). "Travail, transparency and trust : A case study of computer-supported collaborative supply chain planning in high-tech electronics". In : *European Journal of Operational Research* 153.2, p. 445–456 (cf. p. 97).
- AMIRIBESHELI, Mohsen, Mahmood HOSSEINI et Hamid BOUCHACHIA (2016). "A principle-based transparency framework for intelligent environments". In : *Proceedings of the 30th International BCS Human Computer Interaction Conference : Fusion !* BCS Learning & Development Ltd., p. 49 (cf. p. 97).
- ANAND, Vibha, Paul G BIONDICH, Gilbert C LIU, Marc B ROSENMAN, Stephen M DOWNS *et al.* (2004). "Child health improvement through computer automation : the CHICA system." In : *Medinfo*, p. 187–191 (cf. p. 18).
- ANANNY, Mike et Kate CRAWFORD (2018). "Seeing without knowing : Limitations of the transparency ideal and its application to algorithmic accountability". In : *new media & society* 20.3, p. 973–989 (cf. p. 97).
- BARNETT, G Octo, James J CIMINO, Jon A HUPP et Edward P HOFFER (1987). "DXplain : an evolving diagnostic decision-support system". In : *Jama* 258.1, p. 67–74 (cf. p. 15, 18).
- BEAUCHAMP, Tom L, James F CHILDRESS *et al.* (2009). *Principles of biomedical ethics*. 7th. Oxford University Press, USA (cf. p. 171).
- BÉRANGER, Jérôme (2015). *Medical information systems ethics*. John Wiley & Sons (cf. p. 109).
- BERNASCONI, Andrea, Francois CRABBÉ, Ajibola Margret ADEDEJI *et al.* (2019). "Results from one-year use of an electronic Clinical Decision Support System in a post-conflict context : An implementation research". In : *PloS one* 14.12 (cf. p. 18, 24, 26).
- BERNER, Eta S (2016). *Clinical decision support systems*. 3rd. Springer (cf. p. 12).
- BERNER, Eta S et Mark L GRABER (2008). "Overconfidence as a cause of diagnostic error in medicine". In : *The American journal of medicine* 121.5, S2–S23 (cf. p. 16).
- BERTILLOT, Hugo (2016). "Comment l'évaluation de la qualité transforme l'hôpital. Les deux visages de la rationalisation par les indicateurs". In : *Cahiers internationaux de sociologie de la gestion* 15, p. 11–48 (cf. p. 31).

- BESSAT, Cécile, Noël Adannou ZONON et Valérie D'ACREMONT (2019). "Large-scale implementation of electronic Integrated Management of Childhood Illness (eIMCI) at the primary care level in Burkina Faso : a qualitative study on health worker perception of its medical content, usability and impact on antibiotic prescription and resistance". In : *BMC public health* 19.1, p. 449 (cf. p. 18, 24, 26).
- BEYNON-DAVIES, P et Michael LLOYD-WILLIAMS (1998). "Health information systems, 'safety' and organizational learning". In : *Health Informatics Journal* 4.3-4, p. 128–137 (cf. p. 25).
- BEZDEK, James C (1981). *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media (cf. p. 125).
- BOSE, RP Jagadeesh Chandra et Wil VAN DER AALST (2011). "Analysis of Patient Treatment Procedures." In : *Business Process Management Workshops (1)*. T. 99, p. 165–166 (cf. p. 68).
- BOUTELL, Matthew R, Jiebo LUO, Xipeng SHEN et Christopher M BROWN (2004). "Learning multi-label scene classification". In : *Pattern recognition* 37.9, p. 1757–1771 (cf. p. 132).
- BOUYSSOU, Denis et Thierry MARCHANT (2007). "An axiomatic approach to noncompensatory sorting methods in MCDM, II : More than two categories". In : *European Journal of Operational Research* 178.1, p. 246–276 (cf. p. 109).
- BOWMAN, Sue (2013). "Impact of electronic health record systems on information integrity : quality and safety implications". In : *Perspectives in health information management* 10.Fall (cf. p. 31, 33).
- BRIGGS, F., Y. HUANG, R. RAICH *et al.* (2013). "The 9th annual MLSP competition : New methods for acoustic classification of multiple simultaneous bird species in a noisy environment". In : *2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, p. 1–8 (cf. p. 132).
- BUCHANANN, B. et E. SHORTLIFFE (1984). *Rule-Based Expert Systems : The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, Reading, MA (cf. p. 15).
- CABANA, Michael D, Cynthia S RAND, Neil R POWE *et al.* (1999). "Why don't physicians follow clinical practice guidelines ? : A framework for improvement". In : *Jama* 282.15, p. 1458–1465 (cf. p. 29, 37).
- CABITZA, Federico, Raffaele RASOINI et Gian Franco GENSINI (2017). "Unintended consequences of machine learning in medicine". In : *Jama* 318.6, p. 517–518 (cf. p. 32).
- CARROLL, Aaron E, Paul BIONDICH, Vibha ANAND, Tamara M DUGAN et Stephen M DOWNS (2013). "A randomized controlled trial of screening for maternal depression with a clinical decision support system". In : *Journal of the American Medical Informatics Association* 20.2, p. 311–316 (cf. p. 18, 23).
- CARUANA, Rich, Yin LOU, Johannes GEHRKE *et al.* (2015). "Intelligible models for healthcare : Predicting pneumonia risk and hospital 30-day readmission". In : *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, p. 1721–1730 (cf. p. 97, 106).
- CASTELLUCCIA, Claude et Daniel LE MÉTAYER (2019). *Understanding algorithmic decision-making : Opportunities and challenges*. European Parliament (cf. p. 97).

- CASTILHO, Richard Eckart de, Giulia DORE, Thomas MARGONI, Penny LABROPOULOU et Iryna GUREVYCH (2018). “A legal perspective on training models for natural language processing”. In : *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (cf. p. 93).
- CHADWICK, DR, C HALL, C RAE *et al.* (2017). “A feasibility study for a clinical decision support system prompting HIV testing”. In : *HIV medicine* 18.6, p. 435–439 (cf. p. 25).
- CHALLEN, Robert, Joshua DENNY, Martin PITT *et al.* (2019). “Artificial intelligence, bias and clinical safety”. In : *BMJ Qual Saf*, bmjqs–2018 (cf. p. 32).
- CHANG, Phei Lang, Yu-chuan LI, Shih Tsong HUANG, Ta Min WANG et Ming Li HSIEN (1996). “Effects of a medical expert system on differential diagnosis of renal masses : A prospective study”. In : *Computerized medical imaging and graphics* 20.1, p. 43–48 (cf. p. 23).
- CHASE, Melissa et Sarah MEIKLEJOHN (2016). “Transparency overlays and applications”. In : *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, p. 168–179 (cf. p. 97).
- CHAUDHRY, Basit, Jerome WANG, Shinyi WU *et al.* (2006). “Systematic review : impact of health information technology on quality, efficiency, and costs of medical care”. In : *Annals of internal medicine* 144.10, p. 742–752 (cf. p. 21).
- CHOULAK, Mourad, Damien MARAGE, Manon GISBERT, Mélanie PARIS et Yves MEINARD (2019). “A meta-decision-analysis approach to structure operational and legitimate environmental policies–With an application to wetland prioritization”. In : *Science of the Total Environment* 655, p. 384–394 (cf. p. 42).
- CLANCY, Carolyn M (2009). “Ten years after to err is human”. In : *American Journal of Medical Quality* 24.6, p. 525–528 (cf. p. 22).
- CLÉMENT, Valérie, Christine LECLAINCHE et Daniel SERRA (2008). *Economie de la justice et de l'équité* (cf. p. 171).
- COHEN, William W. (1995). “Fast Effective Rule Induction”. In : *Twelfth International Conference on Machine Learning*. Morgan Kaufmann, p. 115–123 (cf. p. 95, 112, 119).
- CRAMER, Henriette, Vanessa EVERS, Satyan RAMLAL *et al.* (2008). “The effects of transparency on trust in and acceptance of a content-based art recommender”. In : *User Modeling and User-Adapted Interaction* 18.5, p. 455 (cf. p. 97).
- CROSKERRY, Pat (2009). “A universal model of diagnostic reasoning”. In : *Academic medicine* 84.8, p. 1022–1028 (cf. p. 81, 83).
- DALABA, Maxwell Ayindenaba, Patricia AKWEONGO, John WILLIAMS *et al.* (2014). “Costs associated with implementation of computer-assisted clinical decision support system for antenatal and delivery care : case study of Kassena-Nankana district of northern Ghana”. In : *PLoS One* 9.9 (cf. p. 24, 26).
- DATTA, Anupam, Shayak SEN et Yair ZICK (2016). “Algorithmic transparency via quantitative input influence : Theory and experiments with learning systems”. In : *2016 IEEE symposium on security and privacy (SP)*. IEEE, p. 598–617 (cf. p. 97).

- DE DOMBAL, FT (1987). "Ethical considerations concerning computers in medicine in the 1980s." In : *Journal of medical ethics* 13.4, p. 179–184 (cf. p. 33).
- DE FAUW, Jeffrey, Joseph R LEDSAM, Bernardino ROMERA-PAREDES *et al.* (2018). "Clinically applicable deep learning for diagnosis and referral in retinal disease". In : *Nature medicine* 24.9, p. 1342–1350 (cf. p. 20).
- DE MARCHI, Giada, Giulia LUCERTINI et Alexis TSOUKIÀS (2016). "From evidence-based policy making to policy analytics". In : *Annals of Operations Research* 236.1, p. 15–38 (cf. p. 41).
- DEIG, Christopher R, Aasheesh KANWAR et Reid F THOMPSON (2019). "Artificial Intelligence in Radiation Oncology". In : *Hematology/Oncology Clinics* 33.6, p. 1095–1104 (cf. p. 19, 46).
- DINKA, David, James M NYCE et Toomas TIMPKA (2006). "The need for transparency and rationale in automated systems". In : *Interacting with computers* 18.5, p. 1070–1083 (cf. p. 97).
- DIPLARIS, Sotiris, Grigorios TSOUMAKAS, Pericles A MITKAS et Ioannis VLAHAVAS (2005). "Protein classification with multiple algorithms". In : *Panhellenic Conference on Informatics*. Springer, p. 448–456 (cf. p. 132).
- DOHERTY, Joshua, Laura A NOIROT, Jennie MAYFIELD *et al.* (2006). "Implementing Germ-Watcher™, an Enterprise Infection Control Application". In : *AMIA Annual Symposium Proceedings*. T. 2006. American Medical Informatics Association, p. 209 (cf. p. 12).
- DONALDSON, Molla S, Janet M CORRIGAN, Linda T KOHN *et al.* (2000). *To err is human : building a safer health system*. T. 6. National Academies Press (cf. p. 15).
- DORAN, Derek, Sarah SCHULZ et Tarek R BESOLD (2017). "What does explainable AI really mean ? A new conceptualization of perspectives". In : *arXiv preprint arXiv :1710.00794* (cf. p. 33, 97, 100).
- DOŠILOVIĆ, Filip Karlo, Mario BRČIĆ et Nikica HLUPIĆ (2018). "Explainable artificial intelligence : A survey". In : *2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE, p. 0210–0215 (cf. p. 114).
- DUA, Sumeet, U Rajendra ACHARYA et Prerna DUA (2014). *Machine learning in healthcare informatics*. T. 56. Springer (cf. p. 19).
- DUCRAY, Victoria, Anna VLACHOMITROU, Maude BOUSCAMPBET-DUCHAMP *et al.* (2020). "Chest CT for rapid triage of patients in multiple emergency departments during COVID-19 epidemic : experience report from a large French university hospital". In : *European Radiology*, p. 1432–1084 (cf. p. 2).
- ECCLES, Martin, Elaine MCCOLL, Nick STEEN *et al.* (2002). "Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care : cluster randomised controlled trial". In : *Bmj* 325.7370, p. 941 (cf. p. 23).
- ELSNER, Peter, Andrea BAUER, Thomas Ludwig DIEPGEN *et al.* (2018). "Position paper : Telemedicine in occupational dermatology—current status and perspectives". In : *JDDG : Journal der Deutschen Dermatologischen Gesellschaft* 16.8, p. 969–974 (cf. p. 20).

- ELSTEIN, Arthur S, Charles P FRIEDMAN, Fredric M WOLF *et al.* (1996). "Effects of a decision support system on the diagnostic accuracy of users : a preliminary report". In : *Journal of the American Medical Informatics Association* 3.6, p. 422–428 (cf. p. 23).
- ESTEVA, Andre, Brett KUPREL, Roberto A NOVOA *et al.* (2017). "Dermatologist-level classification of skin cancer with deep neural networks". In : *Nature* 542.7639, p. 115–118 (cf. p. 20, 24, 34).
- ETTER, Dillon J, Allison MCCORD, Fangqian OUYANG *et al.* (2018). "Suicide screening in primary care : use of an electronic screener to assess suicidality and improve provider follow-up for adolescents". In : *Journal of Adolescent Health* 62.2, p. 191–197 (cf. p. 18).
- FERRETTI, Valentina, Irene PLUCHINOTTA et Alexis TSOUKIÀS (2019). "Studying the generation of alternatives in public policy making processes". In : *European Journal of Operational Research* 273.1, p. 353–363 (cf. p. 41).
- FIELD, Marilyn J, Kathleen N LOHR *et al.* (1990). *Clinical practice guidelines : directions for a new program*. National Academies Press (cf. p. 43).
- FIESCHI, M (1986). "Expert systems for medical consultation". In : *Health Policy* 6.2, p. 159–173 (cf. p. 14).
- FRANCO, John et John MARTIN (2009). "A History of Satisfiability." In : *Handbook of satisfiability* 185, p. 3–74 (cf. p. 93).
- GARCIA, Megan (2016). "Racist in the machine : The disturbing implications of algorithmic bias". In : *World Policy Journal* 33.4, p. 111–117 (cf. p. 34).
- GARG, Amit X, Neill KJ ADHIKARI, Heather McDONALD *et al.* (2005). "Effects of computerized clinical decision support systems on practitioner performance and patient outcomes : a systematic review". In : *Jama* 293.10, p. 1223–1238 (cf. p. 13, 22).
- GIBSON, Mark, K Neil JENKINGS, Rob WILSON et Ian PURVES (2006). "Verbal prescribing in general practice consultations". In : *Social Science & Medicine* 63.6, p. 1684–1698 (cf. p. 67, 82).
- GIORDANENGO, Alain, Eirik ÅRSAND, Ashenafi Zebene WOLDAREGAY *et al.* (2019). "Design and Prestudy Assessment of a Dashboard for Presenting Self-Collected Health Data of Patients With Diabetes to Clinicians : Iterative Approach and Qualitative Case Study". In : *JMIR diabetes* 4.3, e14002 (cf. p. 21, 159, 169).
- GOEBEL, Randy, Ajay CHANDER, Katharina HOLZINGER *et al.* (2018). "Explainable AI : the new 42 ?" In : *International Cross-Domain Conference for Machine Learning and Knowledge Extraction*. Springer, p. 295–303 (cf. p. 97).
- GONZALVO, Carlota Mestres, Hugo AJM de WIT, Brigit PC VAN OIJEN *et al.* (2017). "Supporting clinical rules engine in the adjustment of medication (SCREAM) : protocol of a multicentre, prospective, randomised study". In : *BMC geriatrics* 17.1, p. 35 (cf. p. 18).
- GOODFELLOW, Ian, Yoshua BENGIO et Aaron COURVILLE (2016). *Deep learning*. MIT press (cf. p. 95).
- GOODMAN, Kenneth W (2016). "Ethical and legal issues in decision support". In : *Clinical decision support systems*. Springer, p. 131–146 (cf. p. 33).

- GORDON, Erynn S, Deepti BABU et Dawn A LANEY (2018). "The future is now : Technology's impact on the practice of genetic counseling". In : *American Journal of Medical Genetics Part C : Seminars in Medical Genetics*. T. 178. Wiley Online Library, p. 15–23 (cf. p. 20).
- GÖRITZLEHNER, R, Clark BORST, Joost ELLERBROEK *et al.* (2014). "Effects of transparency on the acceptance of automated resolution advisories". In : *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, p. 2965–2970 (cf. p. 97).
- GRÄSSER, Felix, Stefanie BECKERT, Denise KÜSTER *et al.* (2017). "Therapy decision support based on recommender system methods". In : *Journal of healthcare engineering 2017* (cf. p. 21).
- GROTH, Paul (2013). "Transparency and reliability in the data supply chain". In : *IEEE Internet Computing 17.2*, p. 69–71 (cf. p. 97).
- GUAH, MW (1998). "Evaluation and analysis of multimedia information system design and implementation at the Coloscopy Unit of St. James University Hospital, Leeds, UK". Thèse de doct. UMIST School of Management (cf. p. 25).
- GUITOUNI, Adel et Jean-Marc MARTEL (1998). "Tentative guidelines to help choosing an appropriate MCDA method". In : *European Journal of Operational Research 109.2*, p. 501–521 (cf. p. 107, 109).
- GUNNING, David (2017). "Explainable artificial intelligence (xai)". In : *Defense Advanced Research Projects Agency (DARPA), nd Web 2* (cf. p. 33, 97, 101).
- GÜNTHER, Christian W et Wil VAN DER AALST (2007). "Fuzzy mining–adaptive process simplification based on multi-perspective metrics". In : *International conference on business process management*. Springer, p. 328–343 (cf. p. 76).
- HALL, Louise H, Judith JOHNSON, Ian WATT, Anastasia TSIPA et Daryl B O'CONNOR (2016). "Healthcare staff wellbeing, burnout, and patient safety : a systematic review". In : *PloS one 11.7*, e0159015 (cf. p. 32).
- HAND, David J (2007). "Principles of data mining". In : *Drug safety 30.7*, p. 621–622 (cf. p. 68).
- HECKERLING, PS, AS ELSTEIN, CG TERZIAN et MS KUSHNER (1991). "The effect of incomplete knowledge on the diagnoses of a computer consultant system". In : *Medical Informatics 16.4*, p. 363–370 (cf. p. 23).
- HEDBOM, Hans (2008). "A survey on transparency tools for enhancing privacy". In : *IFIP Summer School on the Future of Identity in the Information Society*. Springer, p. 67–82 (cf. p. 97, 105).
- HEEKS, Richard (2006). "Health information systems : Failure, success and improvisation". In : *International journal of medical informatics 75.2*, p. 125–137 (cf. p. 35, 37).
- HEEKS, Richard, David MUNDY et Angel SALAZAR (1999). "Why health care information systems succeed or fail". In : *Information Systems for Public Sector Management* (cf. p. 25, 35).

- HELBIG, Natalie, Evgeny STYRIN, Donna CANESTRARO et Theresa PARDO (2010). "Information and transparency : learning from recovery act reporting experiences". In : *Proceedings of the 11th Annual International Digital Government Research Conference on Public Administration Online : Challenges and Opportunities*. Digital Government Society of North America, p. 59–65 (cf. p. 97).
- HERLOCKER, Jonathan L, Joseph A KONSTAN et John RIEDL (2000). "Explaining collaborative filtering recommendations". In : *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. ACM, p. 241–250 (cf. p. 97).
- HOCHREITER, Sepp (1998). "The vanishing gradient problem during learning recurrent neural nets and problem solutions". In : *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 6.02, p. 107–116 (cf. p. 114).
- HOFFER, Edward P, Mitchell J FELDMAN, Richard J KIM, Kathleen T FAMIGLIETTI et G Octo BARNETT (2005). "DXplain : patterns of use of a mature expert system". In : *AMIA Annual Symposium Proceedings*. T. 2005. American Medical Informatics Association, p. 321 (cf. p. 15, 18).
- HOLLON, Todd C, Spencer LEWIS, Balaji PANDIAN *et al.* (2018). "Rapid intraoperative diagnosis of pediatric brain tumors using stimulated Raman histology". In : *Cancer research* 78.1, p. 278–289 (cf. p. 20).
- HOLZINGER, Andreas, Chris BIEMANN, Constantinos S PATTICHIS et Douglas B KELL (2017). "What do we need to build explainable AI systems for the medical domain?" In : *arXiv preprint arXiv :1712.09923* (cf. p. 93, 97).
- HOLZNER, Burkart et Leslie HOLZNER (2006). *Transparency in global change : The vanguard of the open society*. Pittsburgh. PA : University of Pittsburgh Press (cf. p. 97).
- HONAKER, Sarah Morsbach, Tamara DUGAN, Ameet DAFTARY *et al.* (2018). "Unexplained practice variation in primary care providers' concern for pediatric obstructive sleep apnea". In : *Academic pediatrics* 18.4, p. 418–424 (cf. p. 18).
- HORNIK, Kurt, Maxwell STINCHCOMBE, Halbert WHITE *et al.* (1989). "Multilayer feedforward networks are universal approximators." In : *Neural networks* 2.5, p. 359–366 (cf. p. 45).
- HUNT, Dereck L, R Brian HAYNES, Steven E HANNA et Kristina SMITH (1998). "Effects of computer-based clinical decision support systems on physician performance and patient outcomes : a systematic review". In : *Jama* 280.15, p. 1339–1346 (cf. p. 22).
- ITANI, Sarah, Fabian LECRON et Philippe FORTEMPS (2019). "Specifics of medical data mining for diagnosis aid : A survey". In : *Expert systems with applications* 118, p. 300–314 (cf. p. 33, 97).
- JANSSENSWILLEN, Gert (2020). *processmapR : Construct Process Maps Using Event Data*. [https ://www.bupar.net](https://www.bupar.net), [https ://github.com/bupaverse/processmapr](https://github.com/bupaverse/processmapr) (cf. p. 69).
- JAROSZEWSKI, Adam C, Robert R MORRIS et Matthew K NOCK (2019). "Randomized controlled trial of an online machine learning-driven risk assessment and intervention platform for increasing the use of crisis services." In : *Journal of consulting and clinical psychology* 87.4, p. 370 (cf. p. 20, 26).

- JEANMOUGIN, Martin, Camille DEHAIS et Yves MEINARD (2017). "Mismatch between habitat science and habitat directive : Lessons from the French (counter) example". In : *Conservation Letters* 10.5, p. 634–644 (cf. p. 41).
- JOHN, George H. et Pat LANGLEY (1995). "Estimating Continuous Distributions in Bayesian Classifiers". In : *Eleventh Conference on Uncertainty in Artificial Intelligence*. San Mateo : Morgan Kaufmann, p. 338–345 (cf. p. 112, 116).
- JOHNSTON, Mary E, Karl B LANGTON, R Brian HAYNES et Alix MATHIEU (1994). "Effects of computer-based clinical decision support systems on clinician performance and patient outcome : a critical appraisal of research". In : *Annals of internal medicine* 120.2, p. 135–142 (cf. p. 22).
- JOO, Segyeong, Yoon Seok YANG, Woo Kyung MOON et Hee Chan KIM (2004). "Computer-aided diagnosis of solid breast nodules : use of an artificial neural network based on multiple sonographic features". In : *IEEE transactions on medical imaging* 23.10, p. 1292–1300 (cf. p. 20).
- KARSENTY, Laurent et Valérie BOTHEREL (2005). "Transparency strategies to help users handle system errors". In : *Speech Communication* 45.3, p. 305–324 (cf. p. 97).
- KARUSH, William (2014). "Minima of functions of several variables with inequalities as side conditions". In : *Traces and Emergence of Nonlinear Programming*. Springer, p. 217–245 (cf. p. 121).
- KATARIA, Suchitra et Vinod RAVINDRAN (2018). "Digital health : a new dimension in rheumatology patient care". In : *Rheumatology international* 38.11, p. 1949–1957 (cf. p. 20).
- KAUSHAL, Rainu, Kaveh G SHOJANIA et David W BATES (2003). "Effects of computerized physician order entry and clinical decision support systems on medication safety : a systematic review". In : *Archives of internal medicine* 163.12, p. 1409–1416 (cf. p. 22).
- KEERTHI, S.S., S.K. SHEVADE, C. BHATTACHARYYA et K.R.K. MURTHY (2001). "Improvements to Platt's SMO Algorithm for SVM Classifier Design". In : *Neural Computation* 13.3, p. 637–649 (cf. p. 112, 120).
- KIM, Taemie et Pamela HINDS (2006). "Who should I blame ? Effects of autonomy and transparency on attributions in human-robot interaction". In : *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, p. 80–85 (cf. p. 97, 100).
- KINGSLAND, Lawrence C, Donald AB LINDBERG et Gordon C SHARP (1983). "AI/RHEUM". In : *Journal of medical systems* 7.3, p. 221–227 (cf. p. 15).
- KIRBY, Amanda M, Bradley KRUGER, Renuka JAIN, PO DANIEL et Bradi B GRANGER (2018). "Using clinical decision support to improve referral rates in severe symptomatic aortic stenosis : a quality improvement initiative". In : *CIN : Computers, Informatics, Nursing* 36.11, p. 525–529 (cf. p. 18, 23).
- KIRCHNER, Kathrin, Nico HERZBERG, Andreas ROGGE-SOLTI et Mathias WESKE (2012). "Embedding conformance checking in a process intelligence system in hospital environments". In : *Process Support and Knowledge Representation in Health Care*. Springer, p. 126–139 (cf. p. 68).

- KOHONEN, Teuvo (1988). "An introduction to neural computing". In : *Neural networks* 1.1, p. 3–16 (cf. p. 112).
- KOSTOPOULOU, Olga, Talya PORAT, Derek CORRIGAN, Samhar MAHMOUD et Brendan C DELANEY (2017). "Diagnostic accuracy of GPs when using an early-intervention decision support system : a high-fidelity simulation". In : *Br J Gen Pract* 67.656, e201–e208 (cf. p. 23).
- KULIKOWSKI, Casimir A (1988). "Artificial intelligence in medical consultation systems : A review". In : *IEEE Engineering in Medicine and Biology Magazine* 7.2, p. 34–39 (cf. p. 15).
- (2019). "Beginnings of Artificial Intelligence in Medicine (AIM) : Computational Artifice Assisting Scientific Inquiry and Clinical Art—with Reflections on Present AIM Challenges". In : *Yearbook of medical informatics* 28.01, p. 249–256 (cf. p. 19, 38).
- KUPERMAN, Gilad J et Richard F GIBSON (2003). "Computer physician order entry : benefits, costs, and issues". In : *Annals of internal medicine* 139.1, p. 31 (cf. p. 12).
- LEAPE, Lucian L (2000). "Institute of Medicine medical error figures are not exaggerated". In : *Jama* 284.1, p. 95–97 (cf. p. 16).
- LEAPE, Lucian L et Donald M BERWICK (2005). "Five years after To Err Is Human : what have we learned ?" In : *Jama* 293.19, p. 2384–2390 (cf. p. 22).
- LEAPER, DJ, PW GILL, JR STANILAND, Jane C HORROCKS et FT DE DOMBAL (1973). "Clinical diagnostic process : an analysis". In : *British medical journal* 3.5880, p. 569 (cf. p. 82, 85).
- LÓPEZ, Marta Manovel, Miguel Maldonado LÓPEZ, Isabel de la TORRE Díez, José Carlos Pastor JIMENO et Miguel LÓPEZ-CORONADO (2017). "MHealth app for iOS to help in diagnostic decision in ophthalmology to primary care physicians". In : *Journal of medical systems* 41.5, p. 81 (cf. p. 18).
- LYONS, Joseph B (2013). "Being transparent about transparency : A model for human-robot interaction". In : *2013 AAAI Spring Symposium Series* (cf. p. 97).
- MANS, Ronny S, MH SCHONENBERG, Minseok SONG, Wil VAN DER AALST et Piet JM BAKKER (2008). "Application of process mining in healthcare—a case study in a dutch hospital". In : *International joint conference on biomedical engineering systems and technologies*. Springer, p. 425–438 (cf. p. 68).
- MASUD, Rafia, Mona AL-REI et Cynthia LOKKER (2019). "Computer-Aided Detection for Breast Cancer Screening in Clinical Settings : Scoping Review". In : *JMIR medical informatics* 7.3, e12660 (cf. p. 26).
- MEINARD, Y et O CAILLOUX (2020). "On justifying the norms underlying decision support". In : *European Journal of Operational Research* (cf. p. 42).
- MEINARD, Yves et Gilles THÉBAUD (2019). "L'identification syntaxonomique dans les démarches de gestion et/ou de restauration d'espaces naturels en France : pour ou contre ?" In : *Naturae* 2019.6, p. 165–173 (cf. p. 171).
- MEINARD, Yves et Alexis TSOUKIÀS (2019). "On the rationality of decision aiding processes". In : *European Journal of Operational Research* 273.3, p. 1074–1084 (cf. p. 42, 44–46, 163).

- MICHENER, Greg et Katherine BERSCH (2013). "Identifying transparency". In : *Information Polity* 18.3, p. 233–242 (cf. p. 97).
- MILLER, RA et FE MASARIE JR (1990). "The demise of the "Greek Oracle" model for medical diagnostic systems". In : *Methods of information in medicine* 29.01, p. 1–2 (cf. p. 15).
- MILLER, Randolph A (2010). "A history of the INTERNIST-1 and Quick Medical Reference (QMR) computer-assisted diagnosis projects, with lessons learned". In : *Yearbook of medical informatics* 19.01, p. 121–136 (cf. p. 15, 18).
- (1994). "Medical diagnostic decision support systems—past, present, and future : a threaded bibliography and brief commentary". In : *Journal of the American Medical Informatics Association* 1.1, p. 8–27 (cf. p. 14, 15).
- MILLER, Randolph A, Melissa A MCNEIL, Sue M CHALLINOR, Fred E MASARIE JR et Jack D MYERS (1986). "The INTERNIST-1/quick medical REFERENCE project—Status report". In : *Western Journal of Medicine* 145.6, p. 816 (cf. p. 15, 18).
- MIOTTO, Riccardo, Fei WANG, Shuang WANG, Xiaoqian JIANG et Joel T DUDLEY (2017). "Deep learning for healthcare : review, opportunities and challenges". In : *Briefings in bioinformatics* 19.6, p. 1236–1246 (cf. p. 19).
- MITCHELL, Imogen, Anne SCHUSTER, Katherine SMITH, Peter PRONOVOST et Albert WU (2016). "Patient safety incident reporting : a qualitative study of thoughts and perceptions of experts 15 years after "To Err is Human"". In : *BMJ Qual Saf* 25.2, p. 92–99 (cf. p. 31).
- MOKDAD, Ali H, James S MARKS, Donna F STROUP et Julie L GERBERDING (2004). "Actual causes of death in the United States, 2000". In : *Jama* 291.10, p. 1238–1245 (cf. p. 16).
- MONTAVON, Grégoire, Wojciech SAMEK et Klaus-Robert MÜLLER (2018). "Methods for interpreting and understanding deep neural networks". In : *Digital Signal Processing* 73, p. 1–15 (cf. p. 97, 100).
- MORELLI, Ralph A, Joseph D BRONZINO et John W GOETHE (1987). "Expert systems in psychiatry". In : *Journal of medical systems* 11.2-3, p. 157–168 (cf. p. 15).
- MURPHY, Gwendolyn C, Charles P FRIEDMAN, Arthur S ELSTEIN *et al.* (1996). "The influence of a decision support system on the differential diagnosis of medical practitioners at three levels of training." In : *Proceedings of the AMIA Annual Fall Symposium*. American Medical Informatics Association, p. 219 (cf. p. 23).
- MUSEN, Mark A, Blackford MIDDLETON et Robert A GREENES (2014). "Clinical decision-support systems". In : *Biomedical informatics*. Springer, p. 643–674 (cf. p. 12).
- NADARZYNSKI, Tom, Jake BAYLEY, Carrie LLEWELLYN, Sally KIDSLEY et Cynthia Ann GRAHAM (2020). "Acceptability of artificial intelligence (AI)-enabled chatbots, video consultations and live webchats as online platforms for sexual health advice". In : *BMJ Sexual & Reproductive Health* (cf. p. 26).
- ONEGA, Tracy, Erin J Aiello BOWLES, Diana L MIGLIORETTI *et al.* (2010). "Radiologists' perceptions of computer aided detection versus double reading for mammography interpretation". In : *Academic radiology* 17.10, p. 1217–1226 (cf. p. 25).

- OSHEROFF, Jerome A, Jonathan M TEICH, Donald LEVICK *et al.* (2012). *Improving outcomes with clinical decision support : an implementer's guide*. HIMSS (cf. p. 42).
- OVERHAGE, J Marc, William M TIERNEY, Xiao-Hua ZHOU et Clement J McDONALD (1997). "A randomized trial of "corollary orders" to prevent errors of omission". In : *Journal of the American Medical Informatics Association* 4.5, p. 364–375 (cf. p. 25).
- OZAYDIN, Bunyamin, J Michael HARDIN et David C CHHIENG (2016). "Data mining and clinical decision support systems". In : *Clinical Decision Support Systems*. Springer, p. 45–68 (cf. p. 19).
- PASQUALI, Paola, Sidharth SONTALIA, David MORENO-RAMIREZ *et al.* (2020). "Tele dermatology and its current perspective". In : *Indian Dermatology Online Journal* 11.1, p. 12 (cf. p. 20).
- PATEL, Vimla L, Andre W KUSHNIRUK, Seungmi YANG et Jean-Francois YALE (2000). "Impact of a computer-based patient record system on data collection, knowledge organization, and reasoning". In : *Journal of the American Medical Informatics Association* 7.6, p. 569–585 (cf. p. 21).
- PEARCE, Christopher, Adam MCLEOD, Natalie RINEHART *et al.* (2019). "POLAR Diversion : using general practice data to calculate risk of emergency department presentation at the time of consultation". In : *Applied clinical informatics* 10.01, p. 151–157 (cf. p. 20).
- PÉGNY, Maël et Issam IBNOUHSEIN (2018). "Quelle transparence pour les algorithmes d'apprentissage machine ?" working paper or preprint (cf. p. 97, 103).
- PEIFFER-SMADJA, Nathan, Timothy Miles RAWSON, Raheelah AHMAD *et al.* (2019). "Machine learning for clinical decision support in infectious diseases : a narrative review of current applications". In : *Clinical Microbiology and Infection* (cf. p. 19, 24, 26, 37, 46).
- PELACCIA, Thierry, Jacques TARDIF, Emmanuel TRIBY et Bernard CHARLIN (2011). "An analysis of clinical reasoning through a recent and comprehensive approach : the dual-process theory". In : *Medical Education Online* 16.1, p. 5890 (cf. p. 81, 83).
- PESTIAN, John P, Christopher BREW, Paweł MATYKIEWICZ *et al.* (2007). "A shared task involving multi-label classification of clinical free text". In : *Proceedings of the Workshop on BioNLP 2007 : Biological, Translational, and Clinical Language Processing*. Association for Computational Linguistics, p. 97–104 (cf. p. 132).
- PIVOVAROV, Rimma et Noémie ELHADAD (2015). "Automated methods for the summarization of electronic health records". In : *Journal of the American Medical Informatics Association* 22.5, p. 938–947 (cf. p. 19, 88).
- PLATT, J. (1998). "Fast Training of Support Vector Machines using Sequential Minimal Optimization". In : *Advances in Kernel Methods - Support Vector Learning*. Sous la dir. de B. SCHOELKOPF, C. BURGESS et A. SMOLA. MIT Press (cf. p. 112, 120).
- PLUCHINOTTA, Irene, Alessandro PAGANO, Raffaele GIORDANO et Alexis TSOUKIÀS (2018). "A system dynamics model for supporting decision-makers in irrigation water management". In : *Journal of environmental management* 223, p. 815–824 (cf. p. 41).

- PLUCHINOTTA, Irene, Akin O KAZAKÇI, Raffaele GIORDANO et Alexis TSOUKIÀS (2019). "Design theory for generating alternatives in public decision making processes". In : *Group Decision and Negotiation* 28.2, p. 341–375 (cf. p. 42).
- POELS, Patrick JP, Tjard RJ SCHERMER, Daan PA SCHELLEKENS *et al.* (2008). "Impact of a spirometry expert system on general practitioners' decision making". In : *European Respiratory Journal* 31.1, p. 84–92 (cf. p. 23).
- PORAT, Talya, Brendan DELANEY et Olga KOSTOPOULOU (2017). "The impact of a diagnostic decision support system on the consultation : perceptions of GPs and patients". In : *BMC medical informatics and decision making* 17.1, p. 79 (cf. p. 26, 37).
- POVYAKALO, Andrey A, Eugenio ALBERDI, Lorenzo STRIGINI et Peter AYTON (2013). "How to discriminate between computer-aided and computer-hindered decisions : a case study in mammography". In : *Medical Decision Making* 33.1, p. 98–107 (cf. p. 31, 32, 34, 47).
- QU, Hongyang et Sandor M VERES (2016). "Verification of logical consistency in robotic reasoning". In : *Robotics and Autonomous Systems* 83, p. 44–56 (cf. p. 93).
- QUINLAN, J. Ross (1986). "Induction of decision trees". In : *Machine learning* 1.1, p. 81–106 (cf. p. 95).
- QUINLAN, Ross (1993). *C4.5 : Programs for Machine Learning*. San Mateo, CA : Morgan Kaufmann Publishers (cf. p. 95, 112, 118).
- REBUGE, Álvaro, Luís Velez LAPÃO, Alberto FREITAS et Ricardo CRUZ-CORREIA (2013). "A process mining analysis on a Virtual Electronic Patient Record system". In : *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*. IEEE, p. 554–555 (cf. p. 68, 82).
- REIDER, Jacob M (2016). "Impact of National Policies on the Use of Clinical Decision Support". In : *Clinical Decision Support Systems*. Springer, p. 111–130 (cf. p. 16).
- RICHARD, Antoine, Brice MAYAG, Yves MEINARD, François TALBOT et Alexis TSOUKIÀS (2018). "How AI could help physicians during their medical consultations : An analysis of physicians' decision process to develop efficient decision support systems for medical consultations". In : *PfIA 2018*. Nancy, France (cf. p. 6, 63, 81, 163).
- RICHARD, Antoine, Brice MAYAG, François TALBOT, Alexis TSOUKIÀS et Yves MEINARD (2020a). "Transparency of Classification Systems for Clinical Decision Support". In : *Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Sous la dir. de Marie-Jeanne LESOT, Susana VIEIRA, Marek Z. REFORMAT *et al.* Cham : Springer International Publishing, p. 99–113 (cf. p. 6, 94, 164).
- RICHARD, Antoine, Brice MAYAG, François TALBOT, Alexis TSOUKIÀS et Yves MEINARD (2020b). "What does it mean to provide decision support to a responsible and competent expert ?" In : *EURO journal on decision processes* (cf. p. 5, 162).
- ROCKAFELLAR, R Tyrrell (1993). "Lagrange multipliers and optimality". In : *SIAM review* 35.2, p. 183–238 (cf. p. 121).
- ROJAS, Eric, Jorge MUNOZ-GAMA, Marcos SEPÚLVEDA et Daniel CAPURRO (2016). "Process mining in healthcare : A literature review". In : *Journal of biomedical informatics* 61, p. 224–236 (cf. p. 68).

- ROY, Bernard et Denis BOUYSSOU (1993). *Aide multicritère à la décision : méthodes et cas*. Economica Paris (cf. p. 107, 109).
- RUDIN, Cynthia et Joanna RADIN (2019). “Why Are We Using Black Box Models in AI When We Don’t Need To? A Lesson From An Explainable AI Competition”. In : *Harvard Data Science Review* 1.2 (cf. p. 33).
- SAMEK, Wojciech, Thomas WIEGAND et Klaus-Robert MÜLLER (2017). “Explainable artificial intelligence : Understanding, visualizing and interpreting deep learning models”. In : *arXiv preprint arXiv :1708.08296* (cf. p. 114).
- SANDVIG, Christian, Kevin HAMILTON, Karrie KARAHALIOS et Cedric LANGBORT (2016). “When the algorithm itself is a racist : Diagnosing ethical harm in the basic components of software”. In : *International Journal of Communication* 10, p. 19 (cf. p. 34).
- SCHAPIRE, Robert E. et Yoram SINGER (2000). “BoosTexter : A boosting-based system for text categorization”. In : *Machine Learning* 39.2/3, p. 135–168 (cf. p. 133).
- SHANNON, Claude Elwood (1948). “A mathematical theory of communication”. In : *Bell system technical journal* 27.3, p. 379–423 (cf. p. 118).
- SHORTLIFFE, Edward (2012). *Computer-based medical consultations : MYCIN*. T. 2. Elsevier (cf. p. 15).
- SHORTLIFFE, Edward H et James J CIMINO (2014). *Biomedical informatics*. 4th. Springer (cf. p. 11, 25).
- SHORTLIFFE, Edward H, Stanton G AXLINE, Bruce G BUCHANAN, Thomas C MERIGAN et Stanley N COHEN (1973). “An artificial intelligence program to advise physicians regarding antimicrobial therapy”. In : *Computers and Biomedical Research* 6.6, p. 544–560 (cf. p. 15).
- SIMON, George, Courtney D DiNARDO, Koichi TAKAHASHI *et al.* (2019). “Applying artificial intelligence to address the knowledge gaps in cancer care”. In : *The oncologist* 24.6, p. 772–782 (cf. p. 21).
- SINHA, Rashmi et Kirsten SWEARINGEN (2002). “The role of transparency in recommender systems”. In : *CHI’02 extended abstracts on Human factors in computing systems*. ACM, p. 830–831 (cf. p. 93, 97, 100).
- SITTIG, Dean F, Michael A KRALL, Richard H DYKSTRA, Allen RUSSELL et Homer L CHIN (2006). “A survey of factors affecting clinician acceptance of clinical decision support”. In : *BMC medical informatics and decision making* 6.1, p. 6 (cf. p. 25).
- SLAIN, Tamara, Sherry RICKARD-AASEN, Janice L PRINGLE *et al.* (2014). “Incorporating screening, brief intervention, and referral to treatment into emergency nursing workflow using an existing computerized physician order entry/clinical decision support system”. In : *Journal of Emergency Nursing* 40.6, p. 568–574 (cf. p. 22).
- SPAGNOLLI, Anna, Lily E FRANK, Pim HASELAGER et David KIRSH (2017). “Transparency as an ethical safeguard”. In : *International Workshop on Symbiotic Interaction*. Springer, p. 1–6 (cf. p. 97, 101).
- STAUSBERG, Jürgen et Michael PERSON (1999). “A process model of diagnostic reasoning in medicine”. In : *International Journal of Medical Informatics* 54.1, p. 9–23 (cf. p. 81, 83).

- TAX, David MJ et Pavel LASKOV (2003). "Online SVM learning : from classification to data description and back". In : *2003 IEEE XIII Workshop on Neural Networks for Signal Processing (IEEE Cat. No. 03TH8718)*. IEEE, p. 499–508 (cf. p. 121).
- TAYLOR, Ben, Michael DINH, Raymond KWOK *et al.* (2008). "Electronic interface for emergency department management of asthma : a randomized control trial of clinician performance". In : *Emergency Medicine Australasia* 20.1, p. 38–44 (cf. p. 23).
- TAYLOR, TR, J AITCHISON et Edward M MCGIRR (1971). "Doctors as decision-makers : a computer-assisted study of diagnosis as a cognitive skill". In : *British Medical Journal* 3.5765, p. 35 (cf. p. 81, 82, 85).
- THELISSON, Eva (2017). "Towards Trust, Transparency and Liability in AI/AS systems." In : *IJCAI*, p. 5215–5216 (cf. p. 97).
- TIERNEY, William M, J Marc OVERHAGE, Michael D MURRAY *et al.* (2003). "Effects of computerized guidelines for managing heart disease in primary care". In : *Journal of General Internal Medicine* 18.12, p. 967–976 (cf. p. 25).
- TINTAREV, Nava et Judith MASTHOFF (2007). "Effective explanations of recommendations : user-centered design". In : *Proceedings of the 2007 ACM conference on Recommender systems*. ACM, p. 153–156 (cf. p. 97).
- TITANO, Joseph J, Marcus BADGELEY, Javin SCHEFFLEIN *et al.* (2018). "Automated deep-neural-network surveillance of cranial images for acute neurologic events". In : *Nature medicine* 24.9, p. 1337–1341 (cf. p. 20).
- TROHIDIS, Konstantinos, Grigorios TSOUMAKAS, George KALLIRIS et Ioannis P VLAHAVAS (2008). "Multi-label classification of music into emotions." In : *ISMIR*. T. 8, p. 325–330 (cf. p. 132).
- TSAI, Theodore L, Douglas B FRIDSMA et Guido GATTI (2003). "Computer decision support as a source of interpretation error : the case of electrocardiograms". In : *Journal of the American Medical Informatics Association* 10.5, p. 478–483 (cf. p. 30, 32, 34, 47).
- TSOUKIÀS, Alexis (2020). *Social Responsibility of Algorithms*. arXiv : 2012.03319 [cs.CY] (cf. p. 33, 97).
- TSOUKIÀS, Alexis, Gilberto MONTIBELLER, Giulia LUCERTINI et Valerie BELTON (2013). "Policy analytics : an agenda for research and practice". In : *EURO Journal on Decision Processes* 1.1-2, p. 115–134 (cf. p. 41, 42).
- TSOUMAKAS, Grigorios et Ioannis KATAKIS (2007). "Multi-label classification : An overview". In : *International Journal of Data Warehousing and Mining (IJDWM)* 3.3, p. 1–13 (cf. p. 88).
- TSOUMAKAS, Grigorios, Eleftherios SPYROMITROS-XIOUFIS, Jozef VILCEK et Ioannis VLAHAVAS (2011a). "Mulan : A Java Library for Multi-Label Learning". In : *Journal of Machine Learning Research* 12, p. 2411–2414 (cf. p. 132).
- TSOUMAKAS, Grigorios, Ioannis KATAKIS et Ioannis VLAHAVAS (2011b). "Random k-Labelsets for Multi-Label Classification". In : *IEEE Transactions on Knowledge and Data Engineering* 23.7, p. 1079–1089 (cf. p. 89, 116, 123, 132).

- TURNBULL, Douglas, Luke BARRINGTON, David TORRES et Gert LANCKRIET (2008). "Semantic annotation and retrieval of music and sound effects". In : *IEEE Transactions on Audio, Speech, and Language Processing* 16.2, p. 467–476 (cf. p. 132).
- VAN DER AALST, Wil (2011). *Process mining : discovery, conformance and enhancement of business processes*. T. 2. Springer (cf. p. 68, 72).
- VAN DER AALST, Wil, Hajo A REIJERS, Anton WEIJTERS *et al.* (2007). "Business process mining : An industrial application". In : *Information Systems* 32.5, p. 713–732 (cf. p. 68).
- VAN DER AALST, Wil, Arya ADRIANSYAH, Ana Karla Alves DE MEDEIROS *et al.* (2011). "Process mining manifesto". In : *International Conference on Business Process Management*. Springer, p. 169–194 (cf. p. 68).
- VAN DER AALST, Wil, Anton WEIJTERS et Laura MARUSTER (2004). "Workflow mining : Discovering process models from event logs". In : *IEEE Transactions on Knowledge and Data Engineering* 16.9, p. 1128–1142 (cf. p. 68).
- VAN DONGEN, Boudewijn F, Ana Karla A de MEDEIROS, HMW VERBEEK, Anton WEIJTERS et Wil VAN DER AALST (2005). "The ProM framework : A new era in process mining tool support". In : *International conference on application and theory of petri nets*. Springer, p. 444–454 (cf. p. 76).
- VERBEEK, HMW, JCAM BUIJS, BF VAN DONGEN et Wil VAN DER AALST (2010). "Prom 6 : The process mining toolkit". In : *Proc. of BPM Demonstration Track* 615, p. 34–39 (cf. p. 76).
- VERDOORN, Sanne, Henk-Frans KWINT, Petra HOOGLAND, Jacobijn GUSSEKLOO et Marcel L BOUVY (2018). "Drug-related problems identified during medication review before and after the introduction of a clinical decision support system". In : *Journal of clinical pharmacy and therapeutics* 43.2, p. 224–231 (cf. p. 23).
- VOIGT, Paul et Axel Von dem BUSSCHE (2017). *The EU general data protection regulation (GDPR)*. Springer (cf. p. 35).
- WACHTER, Robert M (2004). "The End Of The Beginning : Patient Safety Five Years After 'To Err Is Human' Amid signs of progress, there is still a long way to go." In : *Health affairs* 23.Suppl1, W4–534 (cf. p. 22).
- WARNER, Homer R, Charles M OLMSTED et Barry D RUTHERFORD (1972). "HELP—a program for medical decision-making." In : *Computers and biomedical research, an international journal* 5.1, p. 65–74 (cf. p. 12).
- WARNER, Homer R, Peter HAUG, Omar BOUHADDOU *et al.* (1988). "ILIAD as an expert consultant to teach differential diagnosis". In : *Proceedings of the Annual Symposium on Computer Application in Medical Care*. American Medical Informatics Association, p. 371 (cf. p. 23).
- WARNER JR, HR (1989). "Iliad : moving medical decision-making into new frontiers". In : *Methods of information in medicine* 28.04, p. 370–372 (cf. p. 23).
- WATROUS, Raymond L, W Reid THOMPSON et Stacey J ACKERMAN (2008). "The impact of computer-assisted auscultation on physician referrals of asymptomatic patients with heart murmurs". In : *Clinical Cardiology : An International Indexed and Peer-Reviewed Journal for Advances in the Treatment of Cardiovascular Disease* 31.2, p. 79–83 (cf. p. 23).

- WEIJTERS, Anton, Wil VAN DER AALST et AK Alves DE MEDEIROS (2006). "Process mining with the heuristics miner-algorithm". In : *Technische Universiteit Eindhoven, Tech. Rep. WP 166*, p. 1–34 (cf. p. 73).
- WEINGART, Saul N, Maria TOTH, Daniel Z SANDS *et al.* (2003). "Physicians' decisions to override computerized drug alerts in primary care". In : *Archives of internal medicine* 163.21, p. 2625–2631 (cf. p. 25).
- WEST, Colin P, Liselotte N DYRBYE et Tait D SHANAFELT (2018). "Physician burnout : contributors, consequences and solutions". In : *Journal of internal medicine* 283.6, p. 516–529 (cf. p. 32).
- WOOLF, Steven H (1993). "Practice guidelines : a new reality in medicine : III. Impact on patient care". In : *Archives of internal medicine* 153.23, p. 2646–2655 (cf. p. 44).
- YANASE, Juri et Evangelos TRIANTAPHYLLOU (2019). "A systematic survey of computer-aided diagnosis in medicine : Past and present developments". In : *Expert Systems with Applications*, p. 112821 (cf. p. 16, 19, 24).
- YANG, Min-Suk, Sang-Il CHOI, Woo-Jung SONG *et al.* (2018). "Electronic consultation support system for radiocontrast media hypersensitivity changes clinicians' behavior". In : *Allergy, asthma & immunology research* 10.2, p. 165–171 (cf. p. 18).
- YAO, Zheng, Peng LIU, Lei LEI et Junjie YIN (2005). "R-C4. 5 Decision tree model and its applications to health care dataset". In : *Proceedings of ICSSSM'05. 2005 International Conference on Services Systems and Services Management, 2005*. T. 2. IEEE, p. 1099–1103 (cf. p. 95).
- ZADEH, Lotfi A (1965). "Fuzzy sets". In : *Information and control* 8.3, p. 338–353 (cf. p. 125).
- ZADEH, Lotfi Asker, George J KLIR et Bo YUAN (1996). *Fuzzy sets, fuzzy logic, and fuzzy systems : selected papers*. T. 6. World Scientific (cf. p. 15).
- ZARSKY, Tal (2013). "Transparency in data mining : From theory to practice". In : *Discrimination and Privacy in the Information Society*. Springer, p. 301–324 (cf. p. 97, 106).
- ZHANG, Kai, Xiyang LIU, Fan LIU *et al.* (2018). "An interpretable and expandable deep learning diagnostic system for multiple ocular diseases : qualitative study". In : *Journal of medical Internet research* 20.11, e11144 (cf. p. 20, 46).
- ZHANG, Min-Ling et Zhi-Hua ZHOU (2007). "ML-KNN : A lazy learning approach to multi-label learning". In : *Pattern Recogn.* 40.7, p. 2038–2048 (cf. p. 89, 112, 115).
- (2006). "Multi-label neural networks with applications to functional genomics and text categorization". In : *IEEE Transactions on Knowledge and Data Engineering* 18, p. 1338–1351 (cf. p. 89, 112, 113).
- ZIER, JL, AB SPAULDING, M FINCH, T VERSCHAETSE et R TARRAGO (2017). "Improved time to notification of impending brain death and increased organ donation using an electronic clinical decision support system". In : *American Journal of Transplantation* 17.8, p. 2186–2191 (cf. p. 22).
- ZOU, James et Londa SCHIEBINGER (2018). *AI can be sexist and racist—it's time to make it fair* (cf. p. 34).

ZURADA, Jacek M (1992). *Introduction to artificial neural systems*. T. 8. West St. Paul (cf. p. 15, 112).

Colophon

This thesis was typeset with $\text{\LaTeX}$ 2 ϵ . It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc.

Download the *Clean Thesis* style at <http://cleanthesis.der-ric.de/>.

RÉSUMÉ

Réalisée en collaboration avec le GIE Hopsis et les employés des Hospices Civils de Lyon (HCL), cette thèse a pour objectif de proposer une réflexion sur les contraintes et les enjeux liés à l'introduction d'outils d'aide à la décision dans le processus de travail de médecins lors de consultations médicales. Nos travaux se sont organisés autour de trois axes principaux. Une étude des outils actuellement employés pour soutenir le personnel soignant dans leurs processus de décision, qui nous a permis de mettre en évidence les limites des approches actuelles pour l'aide aux décisions médicales. En second lieu, une analyse du processus de décision de médecins travaillant aux HCL, qui nous a permis de mettre en évidence le besoin en informations des médecins afin de prendre des décisions concernant leurs patients. Et enfin, la proposition d'un outil d'aide à la décision, qui vise à l'apprentissage et à l'anticipation des besoins en informations des médecins durant leurs consultations médicales coutumières.

MOTS CLÉS

Aide à la décision, Systèmes d'information hospitaliers, Analyse du processus de décision, Systèmes de classification multi-labels, Transparence des algorithmes

ABSTRACT

Conducted in partnership with the GIE Hopsis and the employees of the Hospitals of Lyon (HCL), this thesis proposes a reflection about the constraints and the challenges linked to the introduction of decision support systems in the workflow of physicians during medical consultations. Our work is organized into three main axes. Firstly, the study of current decision support systems used in healthcare contexts to support clinicians during their decision processes, which allowed us to highlight the limits of current approaches used to support decisions in customary clinical situations. Secondly, the analysis of HCL's physicians' decision process, which allowed us to highlight the physicians' need for patients' information to be able to take relevant decisions. And lastly, the proposal of a decision support system, which aims to learn and to anticipate the patients' information needed by physicians during their customary medical consultations.

KEYWORDS

Decision support, Health Information Systems, Decision Analysis, Multi-label Classification Systems, Algorithmic Transparency