



**HAL**  
open science

# Apports de la modélisation ontologique pour l'analyse des ruptures de parcours de soins dans la Sclérose Latérale Amyotrophique

Sonia Filomena Cardoso

► **To cite this version:**

Sonia Filomena Cardoso. Apports de la modélisation ontologique pour l'analyse des ruptures de parcours de soins dans la Sclérose Latérale Amyotrophique. Médecine humaine et pathologie. Sorbonne Université, 2019. Français. NNT: 2019SORUS587 . tel-02429414v2

**HAL Id: tel-02429414**

**<https://theses.hal.science/tel-02429414v2>**

Submitted on 14 Sep 2021

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

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

# THÈSE DE DOCTORAT DE SORBONNE UNIVERSITÉ

SPÉCIALITÉ : INFORMATIQUE MÉDICALE

ÉCOLE DOCTORALE 393  
PIERRE LOUIS DE SANTÉ PUBLIQUE À PARIS :  
ÉPIDÉMIOLOGIE ET SCIENCES DE L'INFORMATION BIOMÉDICALE

*Présentée par*

**Sonia Filomena Cardoso**

Pour obtenir le grade de

**DOCTEUR de l'UNIVERSITÉ SORBONNE  
UNIVERSITÉ**

**Apports de la modélisation ontologique pour  
l'analyse des ruptures de parcours de soins dans la  
Sclérose Latérale Amyotrophique**

Soutenue le 19 Novembre 2019 devant le jury composé de :

Directeur	Jean CHARLET	- AP-HP, INSERM/LIMICS
Directeur	Pr Vincent MEININGER	- Clinique des Peupliers/ICM
Rapporteurs	Pr Sandra BRINGAY	- LIRMM, Montpellier
	Pr Philippe CORCIA	- CHRU de Tours
Examineurs	Pr Corinne ISNARD-BAGNIS	- Sorbonne Université
	Pr Chantal REYNAUD	- LRI, Paris-Sud



« Je sais grâce à la théorie, mais je sens grâce à la pratique.  
La théorie enrichit l'intellect,  
alors que la pratique donne des couleurs aux sentiments,  
raffermit la volonté »  
Janusz Korczak *Théorie et pratique* (1925)



## Remerciements

Mes premiers remerciements s'adressent au Professeur Vincent Meininger et à Jean Charlet qui ont dirigé et soutenu ce travail de recherche, au Professeur David Grabli et à Xavier Aimé qui ont accompagné ce travail. Chacun d'entre eux s'est mobilisé à différents moments, pour faire que ce projet puisse se réaliser, je les remercie sincèrement de leur confiance. Ces accompagnements, ces regards croisés sur le travail, m'ont permis d'avancer et de mener le travail à son terme.

Je tiens à remercier l'Institut du Cerveau et de la moelle épinière (ICM) pour le financement de cette thèse. Le réseau SLA Île-de-France, qui accompagne chaque jour les patients et qui a fourni le matériel indispensable pour la réalisation de ce travail.

Je remercie le Professeur Sandra Bringay et le Professeur Philippe Corcia d'avoir accepté d'être rapporteurs de cette thèse. Je remercie également le Professeur Corinne Isnard Bagnis et le Professeur Chantal Reynaud d'avoir accepté de faire partie du jury de thèse.

Je remercie également Marie-Christine Jaulent et l'équipe du LIMICS pour l'accueil qui m'a été fait. Je les remercie pour ces rencontres et ces moments d'échanges, de discussions et de rires qui ont ponctué et animé ces trois années. Ce travail de thèse a surtout été un travail de collaboration avec de nombreuses personnes grâce à leur participation le travail a avancé. Merci au Pr Stephan Darmoni et Cyril Grouin pour le partage de leurs outils; à Pierre Meneton pour sa disponibilité, ces conseils, son expertise; à Kevin Bretonnel Cohen pour son aide. Merci à Felipe Melo, Aurélien Chick, Loane Danès et Gilles Guezennec avec qui j'ai eu plaisir à travailler, qui ont toujours pris du temps pour répondre aux nombreuses questions et qui ont développé les outils utilisés dans ce travail. Je remercie les experts et les étudiants qui sont intervenus dans l'évaluation et dans le travail de traduction. Ils ont pris de leur temps, permettant d'apporter un regard différent et important sur le travail. Merci à Catherine M, Cécile C, Dominique H, Isabelle S, Charlotte G, Ilan Z et Ilan B. Une pensée pour mes collègues doctorants avec qui nous avons partagé de nombreux moments Mary, Naiara, Eric, Troska, Marion.

Pour conclure, je tiens à remercier ma famille et en particulier mes parents, pour leur soutien, les valeurs transmises qui m'ont permis d'avancer "muito obrigado". Je ne peux oublier mes amis, pour tous cette question « quand est-ce que tu finis la thèse ? » le temps est enfin arrivé. Il me sera très difficile de remercier et citer tout le monde, c'est grâce à l'aide et au soutien de nombreuses personnes que j'ai pu mener cette thèse à son terme. Je les remercie tous pour leur écoute, leur patience, leurs remarques, leurs conseils, leurs bienveillances, ces moments de pause et de convivialité indispensables pour avancer et cheminer. Merci aux personnes qui ont pris du temps pour lire, corriger et reformuler certaines phrases trop longues. Une grande pensée pour Gwendoline partie trop tôt. Paul Eluard a écrit « *il n'y a pas de hasard il n'y a que des rendez-vous* », merci à tous pour ces rendez-vous qui m'ont accompagné et fait avancer.



# Table des matières

<b>Introduction</b>	<b>0</b>
Contexte de recherche et problématiques . . . . .	1
Organisation du manuscrit . . . . .	5
<b>I État de l’art</b>	<b>7</b>
<b>1 La SLA, parcours et coordination</b>	<b>9</b>
1.1 Introduction . . . . .	11
1.2 Épidémiologie de la SLA . . . . .	12
1.2.1 Prévalence et incidence . . . . .	12
1.2.2 Âge et sexe . . . . .	13
1.2.3 Délais diagnostic et délais de survie . . . . .	14
1.3 Caractéristiques cliniques . . . . .	15
1.3.1 Formes à début spinal . . . . .	15
1.3.2 Forme à début bulbaire . . . . .	16
1.3.3 Troubles cognitifs . . . . .	16
1.3.4 Symptômes . . . . .	17
1.4 Prise en charge sociale dans la SLA . . . . .	20
1.4.1 Personnes de moins de 60 ans . . . . .	21
1.4.2 Personnes de plus de 60 ans . . . . .	22
1.5 Aides techniques / Produits d’assistance . . . . .	23
1.6 Organisation de la prise en charge de la SLA . . . . .	25
1.6.1 La filière SLA . . . . .	25
1.7 Les parcours et la coordination . . . . .	26
1.7.1 Définition des parcours . . . . .	27
1.7.2 La coordination . . . . .	30
1.8 Le réseau SLA Île-de-France . . . . .	31
1.8.1 Mode de fonctionnement du Réseau SLA Île-de-France (Réseau SLA IDF) . . . . .	32

1.8.2	Impact du Réseau SLA IDF dans la prise en charge . . . . .	34
1.8.3	Base de données du Réseau SLA IDF . . . . .	35
1.9	Synthèse et hypothèse de travail . . . . .	39
<b>2</b>	<b>Les Systèmes d'Organisation des Connaissances</b>	<b>41</b>
2.1	Les systèmes d'organisation des connaissances (SOC) . . . . .	42
2.1.1	Définition des Système d'Organisation des Connaissances (SOC)	43
2.1.2	Typologie des SOC . . . . .	44
2.2	Les ontologies informatiques . . . . .	48
2.2.1	Types d'ontologies . . . . .	48
2.2.2	Les composants d'une ontologie . . . . .	50
2.3	Les RTO dans le domaine de la neurologie, de la coordination et du handicap . . . . .	53
2.3.1	Les ressources en neurologie . . . . .	53
2.3.2	Une ontologie de la coordination de soins infirmiers . . . . .	54
2.3.3	La Classification Internationale du Fonctionnement du Handicap et de la Santé (CIF) . . . . .	55
<b>3</b>	<b>Construction d'ontologies informatiques</b>	<b>57</b>
3.1	Les principes de construction d'une ontologie . . . . .	58
3.2	Méthode de construction d'ontologies . . . . .	59
3.2.1	Développement d'ontologies selon <b>Uschold &amp; King (1995)</b> . . . . .	59
3.2.2	Méthode Methontology proposé par <b>López (1999)</b> . . . . .	59
3.2.3	Construction d'ontologie à partir de textes . . . . .	59
3.2.4	Methode ARCHONTE . . . . .	60
3.3	Définition de la modularité d'une ontologie . . . . .	62
3.4	Objectif de la modularité ontologique . . . . .	62
3.5	Construire une ontologie modulaire . . . . .	63
3.6	Synthèse . . . . .	65
<b>II</b>	<b>Travaux réalisés / Contribution Pratique</b>	<b>67</b>
<b>4</b>	<b>Traitements Informatiques du corpus</b>	<b>69</b>
4.1	Aspect législatif du traitement des données . . . . .	70
4.2	Anonymisation et pseudonymisation . . . . .	71
4.3	Les méthodes d'anonymisation . . . . .	72
4.4	Traitement du corpus pour la pseudonymisation . . . . .	73
4.4.1	Préparation des données . . . . .	74
4.4.2	Utilisation de méthode à base d'apprentissage statistique . . . . .	76
4.4.3	Utilisation de règles . . . . .	78
4.4.4	Traitement des dates . . . . .	79
4.4.5	Résultat et discussion de la pseudonymisation . . . . .	79
4.5	Traitement Orthographique . . . . .	80

4.5.1	Types d’erreurs rencontrées	80
4.5.2	Pipeline de correction orthographique : Spell Checking	81
4.6	Synthèse	84
<b>5</b>	<b>Construction de l’ontologie ONTOPARON</b>	<b>87</b>
5.1	Principes de construction de ONTOPARON	88
5.2	Extraction des candidats termes avec BIOmedical Term EXtraction (Bio- Tex)	89
5.2.1	Résultats de l’extraction	90
5.2.2	Sélection des termes de BioTex	91
5.2.3	Sélection des termes et antériorité du travail	92
5.3	Construction de l’ontologie	93
5.4	Passage d’une ontologie monolithique à une ontologie modulaire	96
5.4.1	Les concepts définis	99
5.4.2	Le module noyau : ONTOPARONNOY	100
5.4.3	Le module médical : ONTOPARONMED	103
5.4.4	Le module socio-environnemental : ONTOPARONSOC	106
5.4.5	Le module de coordination : ONTOPARONCOORD	108
5.4.6	Le module de consolidation	110
5.5	Synthèse	110
<b>6</b>	<b>Annotation sémantique et OnBaSAM</b>	<b>111</b>
6.1	L’annotation sémantique	112
6.2	Le développement d’ONBASAM	113
6.2.1	Les pipelines et Processing Resource de ONBASAM	115
6.3	Négation et hypothèse	117
6.4	Export des annotations	118
6.5	ONTOPARON et ONBASAM	119
6.6	Synthèse	119
<b>7</b>	<b>Pronto outil d’évaluation</b>	<b>121</b>
7.1	Présentation et objectifs de l’outil d’évaluation PRONTO	122
7.1.1	Présentation des différentes parties	123
7.1.2	Modalité de validation et niveaux de validation	124
7.1.3	Ajout d’annotation manuelle	126
7.1.4	Métriques d’évaluation	127
7.2	Méthodologie d’évaluation du corpus Sclérose latérale Amyotrophique (SLA)	128
7.2.1	Protocole d’évaluation	128
7.2.2	Définition des critères pour la création du corpus d’évaluation	130
7.3	Résultats	131
7.3.1	Résultats qualitatifs : point de vue des experts	132
7.3.2	Résultats quantitatifs	133
7.3.3	Phase d’amélioration	136

7.3.4	Limites des résultats . . . . .	137
<b>III</b>	<b>Résultats</b>	<b>139</b>
<b>8</b>	<b>Données, analyses et résultats</b>	<b>141</b>
8.1	Description des données . . . . .	142
8.1.1	Présentation des différents types de données . . . . .	142
8.1.2	Exclusion de dossiers . . . . .	145
8.1.3	La temporalité . . . . .	146
8.1.4	Données socio-démographiques . . . . .	146
8.1.5	Les données issues des annotations des corpus . . . . .	147
8.1.6	Aligner les ressources aux besoins : la demande et la recherche de professionnel . . . . .	151
8.1.7	Synthèse des observations . . . . .	156
8.2	Sous populations spécifiques . . . . .	157
8.2.1	Patients vivant en institution . . . . .	157
8.2.2	Patients avec une trachéotomie . . . . .	159
8.2.3	Patients sans besoins et demandes, et sans actions de coordination . . . . .	160
8.2.4	Synthèse . . . . .	161
8.3	Analyses globales . . . . .	161
8.3.1	Liens entre les évènements et les actions de coordinations . . . . .	161
8.3.2	Analyse en correspondance multiple . . . . .	162
8.3.3	Le niveau de complexité . . . . .	163
8.3.4	Patients avec ruptures de parcours identifiés et des difficultés de prise en charge . . . . .	163
8.4	Synthèse des analyses . . . . .	165
	<b>Conclusion et perspective</b>	<b>167</b>
	Liste des acronymes	175
	Bibliographie	179
<b>IV</b>	<b>Annexes</b>	<b>191</b>
<b>A</b>	<b>Publications</b>	<b>193</b>
<b>B</b>	<b>Information globales</b>	<b>195</b>
B.1	Les aides techniques dans la SLA . . . . .	195
B.2	Captures d'écrans volets Réseau SLA IDF . . . . .	197

<b>C</b>	<b>Requêtes SPARQL utilisées dans CORESE</b>	<b>203</b>
<b>D</b>	<b>Classes, relations, concepts définis et notions thématiques</b>	<b>205</b>
D.1	Les classes et les relations formelles . . . . .	205
D.2	OWL-DL et les concepts définis . . . . .	205
D.2.1	Conditions nécessaires et suffisantes (classes primitives et définies)	206
D.2.2	Classes et raisonneurs . . . . .	206
D.3	Les concepts définis de ONTOPARON . . . . .	207
D.3.1	Exemple du concept définis : Domaine de l'épuisement . . . . .	207
<b>E</b>	<b>Exemple d'annotation des corpus du Réseau SLA IDF fait par d'autres systèmes d'annotation.</b>	<b>211</b>
E.1	Annotation réalisée par le SIFR Annotator du LIRMM . . . . .	211
E.2	Annotations réalisées par l'ECMTv3 . . . . .	213
<b>F</b>	<b>Données liées aux analyses</b>	<b>217</b>
F.1	Données des patients par année d'inclusion . . . . .	217
F.2	Carte de densité des professionnels en région Île-de-France . . . . .	219
F.3	Densité des kinésithérapeutes . . . . .	219
F.4	Densité des médecins . . . . .	220
F.5	Densité des orthophonistes . . . . .	220
F.6	Données socio-démographiques des patients sans besoins ni demandes .	221
F.7	Classement en tertile . . . . .	223



# Liste des figures

1	Présentation schématique du travail de thèse. . . . .	4
1.1	Maillage des centres SLA en France. . . . .	26
1.2	Représentation des parcours . . . . .	29
1.3	Professionnels et structures intervenants dans le parcours du patient ayant une SLA. . . . .	33
1.4	Capture d'écran de la fiche évènement de la base évènementielle. . . . .	38
2.1	Types d'ontologies, selon (Guarino, 1997) . . . . .	49
2.2	Types d'ontologies, selon (Roussey <i>et al.</i> , 2011) . . . . .	49
2.3	Triangle sémiotique . . . . .	51
3.1	Etapes de construction d'une ontologie selon Bachimont <i>et al.</i> (2002) . . . . .	60
3.2	Patron d'importation des modules ontologiques de Ben Abbès <i>et al.</i> (2012). . . . .	64
4.1	Processus de pseudonymisation réalisé dans le travail de thèse. . . . .	73
4.2	Illustration d'un évènement principal et des évènements de suivis . . . . .	75
4.3	Catégories définies pour la pseudonymisation. . . . .	77
4.4	Exemple de correction orthographique réalisée dans ONBASAM. . . . .	81
4.5	Processus mis en place pour la correction orthographique. . . . .	82
4.6	Exemple de repérage d'abréviation et de leur transformation. . . . .	84
5.1	Processus de construction de ONTOPARON. . . . .	89
5.2	Exemple des <code>skos:preflabel</code> , <code>skos:altlabel</code> du concept de fauteuil roulant manuel. . . . .	94
5.3	Schématisation des dimensions intervenants dans l'accompagnement de personne ayant une SLA. . . . .	97
5.4	Vue d'ensemble du schéma d'héritage des concepts d'ONTOPARON. . . . .	98
5.5	Illustration des caractéristiques et des restrictions pour le concept défini <i>DomaineEpuisement</i> . . . . .	100
5.6	Visualisation des concepts du module noyau. . . . .	101
5.7	Visualisation des relations, définies dans le module noyau. . . . .	102
5.8	Visualisation des concepts présents dans le module médical. . . . .	105

5.9	Visualisation des concepts présents dans le module médical . . . . .	106
5.10	Visualisation d’une partie des concepts présents dans le module ONTO- PARONSOC . . . . .	107
6.1	Illustration des API Java intégrées dans GATE. . . . .	114
6.2	Illustration des ressources dans un pipeline GATE. . . . .	115
6.3	Schéma présentant les éléments constitutifs des pipelines créés dans ON- BASAM. . . . .	116
6.4	Illustration d’un évènement annoté par la chaîne de traitement « seman- tic Tagging » dans ONBASAM. . . . .	117
6.5	Exemple de détection d’une classe de l’ontologie ONTOPARON niée dans le corpus annoté. . . . .	118
6.6	Tableau des annotations. . . . .	119
7.1	Interface PRONTO pour le gestionnaire de projet. . . . .	124
7.2	Interface PRONTO pour l’évaluateur. . . . .	125
7.3	Ajout d’une annotation manuelle à partir de l’outil PRONTO. . . . .	126
7.4	Extrait de la liste des annotations manuelles faites par les experts. . . .	127
7.5	Interface PRONTO des résultats quantitatifs obtenus pour un projet. . .	134
7.6	Processus réalisé par PRONTO. . . . .	136
8.1	Élément temporel dans le parcours des patients SLA. . . . .	146
8.2	Analyse en correspondance multiple montrant les associations entre les données issues de l’annotation et les caractéristiques des patients. . . .	162
8.3	Exemple d’évènement de patients suivis dans un réseau parkinson, an- notés avec ONBASAM et l’ontologie ONTOPARON. . . . .	170
8.4	Exemple d’évènement de patients suivis dans un réseau traumatisé crâ- nien, annotés avec ONBASAM et l’ontologie ONTOPARON. . . . .	171
B.1	Exemples d’aides techniques pouvant être proposées dans la SLA. . . .	196
B.2	Fiche information patient de la base du Réseau SLA IDF. . . . .	198
B.3	Fiche volet médical de la base du Réseau SLA IDF. . . . .	199
B.4	Fiche volet social de la base du Réseau SLA IDF. . . . .	200
B.5	Fiche intervenants de la base du Réseau SLA IDF. . . . .	201
D.1	Notion thématique présente dans ONTOPARONNOY. . . . .	207
D.2	Définition du concept défini : Domaine de l’épuisement. . . . .	208
D.3	Ensemble des classes inférées après utilisation du raisonneur HermiT, sous le concept définis du Domaine de l’épuisement. . . . .	208
E.1	Capture d’écran de l’interface du SIFR. . . . .	212
E.2	Résultats des annotations réalisées par le SIFR. . . . .	212
E.3	Capture d’écran de l’interface de l’ECMTv3. . . . .	213
F.1	Carte de densité des kinésithérapeutes en Île-de-France. . . . .	219
F.2	Densité des médecins en Île-de-France. . . . .	220

F.3 Densité des orthophonistes en Île-de-France. . . . .	220
--	-----



# Liste des tableaux

4.1	Exemple d'un évènement modifié par la pseudonymisation. . . . .	77
5.1	Illustration des listes de candidats termes obtenue par BioTex . . . . .	91
5.2	Métriques des différents modules de l'ontologie ONTOPARON . . . . .	102
5.3	Présentation des types d'activités de coordination inclus dans ONTOPARON et dans le Care Coordination Measures Atlas. Les activités de coordination sont regroupées par typologie. . . . .	108
7.1	Catégorisation des valeurs pour le calcul des métriques de rappel, précision et F-Mesure. . . . .	128
7.2	Caractéristiques des dossiers évalués sur Pronto. . . . .	131
7.3	Nombre de dossiers et d'évènements évalués par expert. . . . .	132
7.4	Résultats quantitatifs des évaluations des annotations sémantiques par les 5 experts. . . . .	134
8.1	Données socio-démographiques de la population étudiée. . . . .	144
8.2	Répartition du nombre d'évènements et de la moyenne de complexité pour l'ensemble de la population. . . . .	145
8.3	Données issues de l'annotation sémantique des concepts définis du module médical. . . . .	148
8.4	Données issues de l'annotation sémantique des concepts définis du domaine de coordination. . . . .	149
8.5	Demande et recherche de professionnel. . . . .	152
8.6	Données issues de l'annotation sémantique des concepts définis du module noyau. . . . .	153
8.7	Comparaison des moyennes d'annotations des concepts définis entre la population globale et les patients trachéotomisés. . . . .	159
F.1	Données de la population par année d'inclusion au Réseau SLA IDF. . . . .	218
F.2	Caractéristiques des patients sans demandes ni besoins dans le domaine social et médical. . . . .	222
F.3	Moyenne par classement en Tertile. . . . .	223



# Introduction

## Contexte de recherche et problématiques

### Les maladies neuro-dégénératives, la Sclérose Latérale Amyotrophique (SLA) et les parcours de santé.

Les maladies neuro-dégénératives, telles que la Sclérose Latérale Amyotrophique (SLA), la Maladie de Parkinson ou la Maladie d'Alzheimer sont pour les politiques de santé une priorité nationale qui a mené à la création d'un plan national, le « Plan Maladies neuro-dégénératives 2014-2019 ». L'augmentation de l'espérance de vie et l'amélioration des moyens thérapeutiques développés font que ces pathologies sont de plus en plus fréquentes dans la population, et qu'elles représentent un véritable enjeu de Santé publique. Les maladies neuro-dégénératives présentent des caractéristiques communes : ce sont des affections qui provoquent des incapacités et des situations de handicaps, qui n'ont pour la plupart, pas de traitement curatif, bien que des traitements médicamenteux et non médicamenteux puissent agir sur les symptômes. Les handicaps liés aux maladies neuro-dégénératives sont polymorphes (atteintes motrices, respiratoires, cognitives, etc.) nécessitant l'intervention de différents acteurs du domaine médical, social et médico-social tout au long du parcours du patient mais aussi auprès des aidants et des professionnels en termes d'orientation, de conseils de formation et d'actions à réaliser et mise en place de dispositifs. Ces pathologies entraînent une forte complexité de prise en charge, par l'intervention de nombreux acteurs, s'inscrivant dans le cadre de la définition des parcours complexes tels que défini dans la Loi de Modernisation du système de santé du 26 janvier 2016. La [Haute Autorité de Santé \(HAS\)](#) a tenté d'optimiser la prise en charge des maladies chroniques par le parcours de soins coordonné par le médecin traitant. L'objectif des parcours proposé par l'[HAS](#) est de permettre une fluidité des parcours de santé, afin de limiter et éviter les événements de « rupture ». Toutefois, la fréquence des ruptures de parcours de soins dans ces pathologies reste élevée sans qu'il soit possible d'identifier la nature et les causes de ces ruptures. Le virage ambulatoire souhaité par les pouvoirs publics et l'amélioration des

parcours de santé, en limitant les ruptures, est devenue un enjeu central de la politique de Santé avec des bénéfices convergents pour les patients (amélioration de la prise en charge) et le système de santé (limitation des hospitalisations et actes médicaux inutiles). Or, notre connaissance des causes médicales et sociales de ruptures est partielle car l'information concernant le parcours de soin est fragmentée entre différents acteurs et le plus souvent non traçable. Pour limiter ces ruptures et améliorer le parcours de santé il semble important de comprendre leur origine et voir s'il existe une typologie de patients à risque.

La SLA est une maladie neuro-dégénérative, qui touche les motoneurones. Elle présente comme caractéristique une durée moyenne de survie de 36 mois après l'annonce diagnostic (Couratier *et al.*, 2016). L'âge moyen d'apparition est de 63 ans. À Paris, à l'hôpital de la Pitié-Salpêtrière, existe un centre expert SLA pour lequel est associé un réseau de coordination ville-hôpital. Le rôle du Réseau SLA IDF est d'assurer la coordination et de faire le relai entre les structures hospitalières et les intervenants au domicile. Pour assurer la liaison entre les différents intervenants, ce réseau a développé une base de données textuelles que l'on nomme « base événementielle », regroupant les textes écrits par les coordinateurs et détaillant au quotidien et de manière exhaustive les demandes et les besoins des patients et des professionnels de premier recours.

L'hypothèse de départ de notre recherche, est que les ruptures doivent avoir lieu essentiellement au domicile. S'il est possible de tracer les actes de soin dont bénéficient les patients à l'hôpital ou à domicile à travers le PMSI<sup>1</sup> ou actes de soin remboursés par la sécurité sociale par le SNIIRAM<sup>2</sup> il n'existe pas d'outils permettant de connaître les besoins et les demandes formulés par les patients, aidants ou professionnels au domicile. Si pour les parcours de soins, il est possible de tracer les actes effectués par le PMSI, cela est plus complexe dans le parcours de santé. Le parcours de santé inclut le parcours de soins et positionne la personne dans son environnement. La coordination proposée au Réseau SLA IDF se situe dans le cadre du parcours de santé, c'est-à-dire à la fois au domicile mais aussi lors des phases d'hospitalisations. Selon notre hypothèse, en analysant et en exploitant les données de la base événementielle des patients inclus au Réseau SLA IDF il serait possible de comprendre et de déterminer s'il existe des facteurs de risque de ruptures de parcours, et de comprendre les besoins des patients, en ayant une approche centrée sur le patient et non sur les dispositifs existants.

## Les outils informatiques au service de l'exploitation de corpus textuels

Pour exploiter ces corpus, l'utilisation d'outils informatiques est indispensable, car le volume de données est trop important pour envisager un travail manuel. Dans ce travail de thèse, nous avons fait le choix d'utiliser comme système d'organisation des connaissances une ontologie. Une ontologie permet de modéliser et conceptualiser les

---

1. Programme de Médicalisation des Systèmes d'Information

2. Système National d'Information Inter-Régimes de l'Assurance Maladie

connaissances d'un domaine, elle est selon la définition de [Gruber \(1995\)](#) une formalisation d'une conceptualisation partagée. Les parcours de santé se situent au carrefour de plusieurs domaines : *a)* le domaine médical par les aspects liés à la maladie et aux conséquences qu'elle provoque ; *b)* le domaine socio-environnemental car le patient est avant tout une personne, qui a une histoire, un projet de vie, un logement et qui exprime des choix ; *c)* le dernier domaine est celui de la coordination de parcours qui est proposé aux patients suivis au centre SLA et inclus au Réseau SLA IDF. Pour prendre en compte ces dimensions, nous nous sommes orientés vers une ontologie modulaire. Une ontologie modulaire est composée par un ensemble de module ontologique qui sont autonomes mais présente des associations avec d'autres modules d'ontologie ([Pathak et al., 2009](#)). Une ontologie informatique permet notamment, par la création de relation sémantique des raisonnements formels. Il est possible de créer dans une ontologie des concepts définis par des conditions suffisantes et nécessaires. Ceci permet d'inférer sous un concept définis l'ensemble des classes qui partagent une même relation, présents dans les différents modules ontologiques. La modélisation et la structuration des connaissances sont importantes, cependant la construction d'une ontologie doit répondre à un objectif. Dans notre travail l'ontologie est une ressource utilisée comme ressource afin de réaliser l'annotation sémantique des évènements des parcours patients. Pour réaliser, le processus d'annotation, nous avons développé au sein du laboratoire LIMICS : un outil d'annotation ONBASAM (Ontology Based Semantic Annotation Module). Pour son développement nous avons utilisé les ressources disponibles sur General Architecture for Text Engineering (GATE). Les ressources disponibles ont été adaptées aux spécificités de notre corpus, permettant le développement de différentes chaînes de traitement.

La figure 1 montre l'ensemble des processus réalisés au cours du travail de thèse, permettant la transformation de données brutes nominatives et non structurées présentes dans la base de données du portail SLA, en données annotées et structurées. Ces données sont utilisées secondairement pour réaliser des analyses statistiques dans l'objectif de comprendre les parcours de santé des patients inclus au sein du Réseau SLA IDF. Parmi ces processus, nous retrouverons dans les différents chapitres de ce manuscrit : le processus de pseudonymisation, la construction de l'ontologie à partir de l'extraction des candidats termes, la correction orthographique, la phase de validation des annotations par les experts du domaine et la phase d'analyse des résultats.

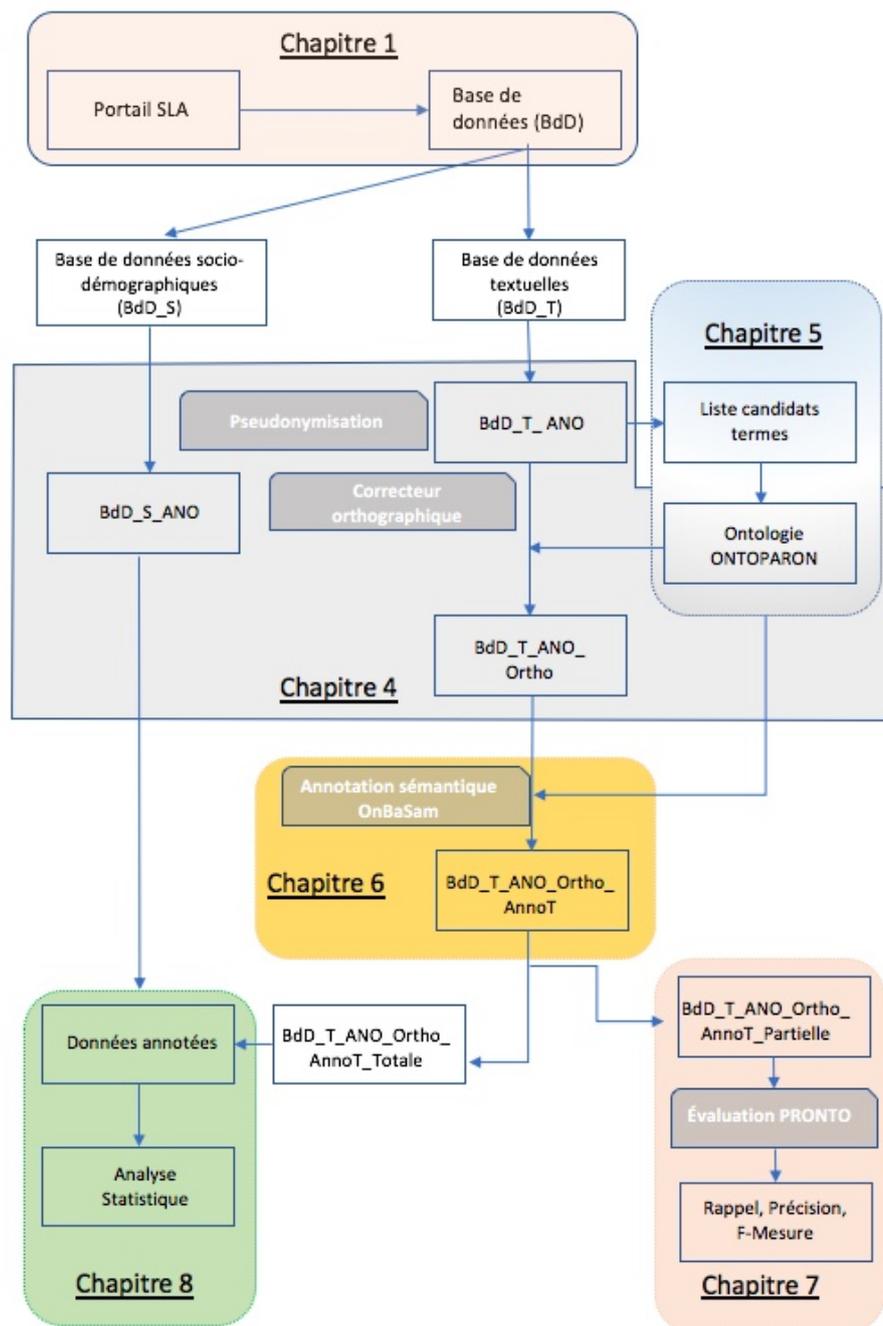


FIGURE 1 – Présentation schématique du travail réalisé au cours de cette thèse, correspondant aux différents chapitres du manuscrit. (BdD : Base de données extraite du portail SLA au format excel ; BdD-T : Base de données textuelles, extraction des données textuelles des évènements et transformation en fichier texte ; BdD-T-ANO : Base de données textuelle pseudonymisée, ; BdD-T-Ortho : Base de données textuelles pseudonymisées et corrigées orthographiquement ; BdD-S-ANO, Base de données socio-démographiques pseudonymisée ; BdD-T-ANO-Ortho-AnnoT : Base de données textuelles pseudonymisée, corrigées orthographiquement et annotées ; BdD-T-ANO-Ortho-AnnoT<sub>4</sub> : Corpus pseudonymisé, corrigé et annoté).

## Organisation du manuscrit

La **première partie** de ce manuscrit fait état des recherches dans les domaines connexes à notre étude.

**Le Chapitre 1** expose les caractéristiques de la [SLA](#). Ce premier chapitre permet de comprendre et expliciter dans un premier temps les données physiopathologiques et épidémiologiques de la maladie et son processus d'évolution. Une section est dédiée à l'organisation et la structuration de la prise en charge des patients, ainsi que les spécificités du système français, prenant comme critère l'âge de la personne pour l'accès aux aides médico-sociales. Une troisième section présente les notions de parcours de soins, parcours de santé et parcours de vie, ainsi que le concept de la coordination. La dernière partie de ce chapitre présente le [Réseau SLA IDF](#) et sa base de données, matériel primordial dans la réalisation de ce travail. Ce premier chapitre nous semble important pour que le lecteur puisse avoir une représentation de la pathologie, de la diversité des situations qui peuvent exister et être rencontrées dans le cadre de la coordination des parcours au [Réseau SLA IDF](#).

**Le Chapitre 2** présente et permet de définir les différents systèmes d'organisation des connaissances existants comme les classifications, les thésaurus etc. Une section de ce chapitre explicite plus spécifiquement les ontologies informatiques et les types d'ontologies existantes comme les top-ontologies, les ontologies de domaine ou encore les ontologies d'applications. La dernière section présente certaines des ressources terminologiques existantes dans les thématiques de notre travail, que sont la coordination, les pathologies neuro-dégénératives et le handicap.

**Le Chapitre 3** présente, différentes méthodes de construction d'ontologie et notamment la méthode ARCHONTE. C'est cette méthode que nous avons choisi d'utiliser lors de la construction de l'ontologie ONTOPARON. Ce chapitre permet de présenter les définitions et les spécificités des ontologies modulaires et des concepts définis.

La **deuxième partie** de ce manuscrit présente les travaux réalisés au cours de cette thèse. Les travaux menés et les outils développés ont été rendus possible par la collaboration avec de nombreux intervenants ayant les compétences informatiques pour les développer.

**Le Chapitre 4** présente les travaux réalisés sur deux dimensions que sont la pseudonymisation et le traitement des erreurs orthographiques du corpus. Le traitement de données personnelles, nécessite de prendre des mesures de sécurité afin de respecter la réglementation française et européenne. Cette phase préalable à toute exploitation des données est nécessaire et fondamentale. Pour garder un niveau d'information sur les acteurs intervenants, nous avons fait le choix de réaliser une pseudonymisation transformant les noms propres des personnes en étiquette indiquant leurs fonctions (assistante

sociale, entourage, patient, coordinateur sla etc.). Les corpus présentent des spécificités, comme l'utilisation d'abréviation, la présence d'erreurs de saisies, et l'utilisation parfois de style syntaxique télégraphique. Pour optimiser le processus d'annotation, nous avons décidé de créer une chaîne de traitement qui puisse traiter ces spécificités.

**Le Chapitre 5** expose la méthodologie utilisée dans le cadre de la construction de l'ontologie ONTOPARON, ainsi que la présentation des différents modules la constituant. Nous présentons et explicitons notre choix de modélisation et la création des concepts définis.

**Le Chapitre 6** présente l'outil ONBASAM (Ontology Based Semantic Annotation Module) développé à partir de la plate-forme open source General Architecture for Text Engineering (GATE). Différents pipelines furent développés, chacun répondant à des objectifs spécifiques. Spell checking pour permettre la correction orthographique des corpus présenté au chapitre 4. Semantic Tagging pour permettre l'annotation sémantique, et Frequency extractor pour réaliser l'export des annotations des dossiers patients.

**Le Chapitre 7** décrit l'outil d'évaluation PRONTO, qui a pour objectif de permettre aux experts du domaine d'évaluer les annotations faites sur le corpus dans l'outil ONBASAM. Les résultats obtenus seront présentés.

**Enfin, le Chapitre 8** expose les résultats obtenus à partir de l'annotation sémantique des corpus constitués d'un ensemble de 931 patients inclus entre 2013 et 2017, représentant 31 260 évènements annotés. Nous tenterons, dans cette partie, de répondre aux hypothèses émises lors de l'initiation de ce travail de thèse. La première section présente de manière descriptive les données socio-démographiques issues de la base SLA et les données obtenues par l'annotation. La seconde section présente les analyses réalisées.

Première partie

État de l'art



# La sclérose latérale amyotrophique, parcours et coordination

## Sommaire

---

<b>1.1</b>	<b>Introduction</b>	<b>11</b>
<b>1.2</b>	<b>Épidémiologie de la SLA</b>	<b>12</b>
1.2.1	Prévalence et incidence	12
1.2.2	Âge et sexe	13
1.2.3	Délais diagnostic et délais de survie	14
<b>1.3</b>	<b>Caractéristiques cliniques</b>	<b>15</b>
1.3.1	Formes à début spinal	15
1.3.2	Forme à début bulbaire	16
1.3.3	Troubles cognitifs	16
1.3.4	Symptômes	17
<b>1.4</b>	<b>Prise en charge sociale dans la SLA</b>	<b>20</b>
1.4.1	Personnes de moins de 60 ans	21
1.4.2	Personnes de plus de 60 ans	22
<b>1.5</b>	<b>Aides techniques / Produits d'assistance</b>	<b>23</b>
<b>1.6</b>	<b>Organisation de la prise en charge de la SLA</b>	<b>25</b>
1.6.1	La filière SLA	25
<b>1.7</b>	<b>Les parcours et la coordination</b>	<b>26</b>
1.7.1	Définition des parcours	27

1.7.2	La coordination . . . . .	30
<b>1.8</b>	<b>Le réseau SLA Île-de-France . . . . .</b>	<b>31</b>
1.8.1	Mode de fonctionnement du Réseau SLA IDF . . . . .	32
1.8.2	Impact du Réseau SLA IDF dans la prise en charge . . . . .	34
1.8.3	Base de données du Réseau SLA IDF . . . . .	35
<b>1.9</b>	<b>Synthèse et hypothèse de travail . . . . .</b>	<b>39</b>

---

*La Sclérose Latérale Amyotrophique (SLA) est la plus fréquente des maladies rares, l'apparition des symptômes et l'évolution de la pathologie peuvent être hétérogènes d'une personne à une autre. Il s'agit d'une pathologie d'évolution constante dont la durée moyenne de survie après le diagnostic est de 36 mois. C'est une pathologie qui va engendrer de nombreuses incapacités, plaçant la personne en situation de handicap lors de la réalisation des activités de vie quotidienne (mobilité, dextérité, communication, déplacement etc.). On estime que la SLA est responsable, en France, de l'apparition de 4 nouveaux cas par jour et de 2 décès. En 2005 en France, suite au livre noir de la SLA, il fut décidé de structurer les soins et la prise en charge de façon spécifique avec la création de centres experts référents. L'accompagnement nécessite l'intervention de nombreux professionnels des secteurs sanitaire, médico-social et social tout au long du parcours de santé du patient. En Île-de-France les patients peuvent bénéficier d'un accompagnement dans la coordination de leurs parcours de santé par un réseau de santé, le Réseau SLA IDF. Pour tracer les demandes et les besoins émis par les patients, l'entourage et les professionnels, le Réseau SLA IDF a créé une base de données textuelles. Elle permet de regrouper les demandes, les difficultés, les actions mises en place au cours du parcours de santé des patients inclus dans le Réseau SLA IDF. La première section de ce chapitre est consacrée aux données épidémiologiques de la pathologie, la seconde à l'aspect clinique, la troisième porte sur les traitements, la quatrième section expose les conséquences fonctionnelles de la pathologie. La dernière partie de ce chapitre décrit la coordination et les parcours ainsi que la structuration de la prise en charge de la SLA en France. Une section aborde particulièrement le Réseau SLA IDF et la base de données, dite base événementielle, qui est l'élément central de ce travail de thèse. La description de la pathologie, de ses formes et de son expression clinique, par les répercussions qu'elle engendre, nous semble importante à décrire, car ce sont ces éléments que l'on retrouve dans les corpus du Réseau SLA IDF. La pathologie n'a pas uniquement une dimension médicale, elle a des conséquences dans la vie quotidienne de la personne et de son entourage, qui nécessite des démarches sociales qui seront différentes en fonction de l'âge de la personne.*

## 1.1 Introduction

La SLA encore appelée « Maladie de Charcot » ou maladie de « Lou Gehrig » fut décrite la première fois par le Professeur Jean Martin Charcot au XIX<sup>e</sup> siècle. Cette pathologie est caractérisée par la dégénérescence des motoneurones, ou neurones moteurs, qui assurent la commande de la contraction musculaire. L'atteinte de ces motoneurones va conduire progressivement mais de façon constante à une perte des fonctions de la motricité affectant le contrôle volontaire de la musculature squelettique et éventuellement une perte de fonction des muscles de la ventilation, entraînant la mort entre 24 et 36 mois (Corcia *et al.*, 2008).

La SLA est la plus fréquente des maladies du motoneurone avec un incidence de 2/100 000 personnes-années (Couratier *et al.*, 2014). En France, la HAS estime que la SLA concerne entre 4 500 à 6 000 personnes. Du fait de sa prévalence elle fait partie

des maladies rares. L'atteinte des neurones centraux et périphériques vont provoquer des paralysies progressives et extensives, plaçant la personne en situation de handicap progressif et chronique. L'accompagnement des patients et de leur entourage va nécessiter une prise en charge pluri-professionnelle au niveau médical, social et médico-social. L'apparition progressive des déficiences s'accompagne de mise en place d'aides qui peuvent être de différentes natures, comme des aides humaines et des [Aides Techniques \(AT\)](#). Un accompagnement social pour la mise en place de financement des aides sera nécessaire ([Soriani & Desnuelle, 2017](#)). 90% des personnes ayant une SLA présentent une forme sporadique c'est-à-dire sans causes connues et 10% présentent une forme familiale ([Riancho et al., 2019](#)). L'âge moyen d'apparition est de 63 ans dans les formes sporadiques, il se situe autour de 50 ans dans les formes familiales.

À ce jour, il n'existe pas de traitement curatif, seul le riluzole est préconisé dans le traitement de la [SLA](#). Il permettrait de retarder l'évolution de la pathologie ([Bensimon et al., 1994](#)).

En France, la prise en charge des patients s'appuie sur des recommandations de la [HAS](#) ainsi que sur des structures (centre expert, centre de compétence, réseau de santé) regroupées au sein de la [Filière de santé maladies rares Sclérose Latérale Amyotrophique et maladies du neurone moteur \(FILSLAN\)](#). En Île-de-France, les patients peuvent bénéficier d'un accompagnement médico-social par une structure dédiée à la coordination du parcours de santé, le [Réseau SLA IDF](#). Ce réseau créé en 2005 possède une base de données textuelles qui permet de tracer l'ensemble des événements intervenant dans le parcours du patient lorsque ce dernier est inclus dans le réseau. L'hypothèse de notre travail est qu'en exploitant ces bases de données il est possible de comprendre les causes de rupture de parcours de santé, ou d'identifier les patients à risque de rupture.

## 1.2 Épidémiologie et diagnostic de la [SLA](#)

La [SLA](#) présente deux étiologies, une étiologie sporadique regroupant 90% des cas et une forme familiale regroupant 10 % des cas ([Kiernan et al., 2011](#)). Les formes familiales sont caractérisées par un début plus précoce de 10 ans en moyenne (extrêmes de 15 ans et 85 ans). Actuellement vingt gènes identifiées permettraient d'expliquer 60 à 70% des formes familiales ([Couratier et al., 2014](#)). Bien que l'origine de la SLA ne soit pas totalement déterminée, les mutations génétiques et les facteurs environnementaux sont deux pistes explorées depuis plusieurs années, mais à ce jour les principaux facteurs sont inconnus. La mutation génétique dans le gène de la SOD1, dans les formes familiales et le gène C9ORF72 sont retrouvés dans 40% des cas, et dans 6% des formes sporadiques ([Pradat et al., 2016](#)).

### 1.2.1 Prévalence et incidence

Plusieurs études ont démontré une variation de l'incidence, de la mortalité et de la prévalence de la SLA selon les zones géographiques et les populations ([Luna et al.,](#)

2017).

### 1.2.1.1 Prévalence

Les données sur la prévalence de la SLA varient selon les études. Elle serait comprise entre 3,3 et 7,9 pour 100 000 habitants (Couratier *et al.*, 2014). La SLA fait partie des maladies rares. La HAS définit, dans le document « Avis sur la prise en charge des maladies rares dans le cadre du dispositif ALD »<sup>1</sup>, les maladies rares comme les pathologies dont la prévalence est inférieure à 5/10 000 personnes, c'est à dire moins de 31 000 personnes concernées par la même pathologie en France. La SLA est une pathologie neurologique rare qui représente 2500 nouveaux cas par an en France<sup>2</sup>. En 2012, 1 729 décès furent dénombrés en France (923 hommes et 806 femmes)<sup>3</sup>.

### 1.2.1.2 Incidence

Chiò *et al.* (2013) montrent la variabilité du taux d'incidence : il est de 0,5/100 000 à Belgrade (Yougoslavie) et de 3,6/100 000 dans l'île Féroé. En Europe le taux se situe autour de 2,1/100 000 (Logroscino *et al.*, 2010). L'incidence est variable en fonction de l'âge, avec une faible incidence avant l'âge de 40 ans, puis une augmentation progressive dans la tranche d'âge des 70 à 80 ans, suivie d'une forte baisse entre 80 et 100 ans (Marin *et al.*, 2018).

## 1.2.2 Âge et sexe

De manière globale, la SLA se déclare entre 40 et 70 ans, avec une moyenne d'âge se situant à 60 ans pour les formes sporadiques et à 50 ans dans le cas des formes familiales (Soriani & Desnuelle, 2009). La pathologie débiterait plus tardivement chez les femmes, avec une prédominance à début bulbaire pour un âge moyen de 68,4 ans, contre un âge moyen de 61,5 ans pour les hommes qui présentent plus un début à forme spinale (Pradat *et al.*, 2016).

Le sex-ratio est de 1,5, cependant des variations du sex-ratio existent en fonction de l'âge : il serait de 2,5 pour les moins de 40 ans et de 0,75 après 80 ans (Pradat *et al.*, 2016).

---

1. [https://www.has-sante.fr/portail/upload/docs/application/pdf/avis\\_ald\\_rapport.pdf](https://www.has-sante.fr/portail/upload/docs/application/pdf/avis_ald_rapport.pdf)

2. [https://www.has-sante.fr/portail/upload/docs/application/pdf/2015-11/pnds-sclerose\\_laterale\\_amyotrophique\\_sla.pdf](https://www.has-sante.fr/portail/upload/docs/application/pdf/2015-11/pnds-sclerose_laterale_amyotrophique_sla.pdf)

3. <http://invs.santepubliquefrance.fr/fr../layout/set/print/Dossiers-thematiques/Maladies-chroniques-et-traumatismes/Maladies-neurodegeneratives/La-sclerose-laterale-amyotrophique>

### 1.2.3 Délais diagnostic et délais de survie

L'annonce du diagnostic est un moment très important et sensible, à la fois pour le patient et son entourage (Kiernan *et al.*, 2011), il inscrit le patient et son entourage dans le cadre d'une pathologie progressive dont il n'existe pas à ce jour de traitement curatif. « Un malade atteint de sla impose des charges lourdes aux familles, et non par une simple présence active à ses côtés, en effet les familles sont confrontées à une multitude de problèmes tout à la fois sociaux, financiers, médicaux et affectifs. Cela va de rebutantes démarches aux appels sans réponse » (Fussellier, 2006).

Le diagnostic de la SLA peut être complexe à poser essentiellement en début de maladie, ainsi que dans les formes atypiques. Le diagnostic doit être évoqué lors de la combinaison d'une atteinte des **Neurones Moteurs Centraux (NMC)** et des **Neurones Moteurs Périphériques (NMP)** associé à une diffusion progressive des symptômes. Cette progression de la symptomatologie au cours du temps est un critère indispensable. Le diagnostic de SLA va reposer sur des critères cliniques lors de l'évaluation neurologique ainsi que sur les résultats de l'**Electro-neuro-myogramme (ENMG)**. De part sa diversité d'expression clinique et la difficulté de diagnostic, il existe un délai entre l'apparition des premiers symptômes et l'annonce diagnostic, qui se situerait entre 7 et 12 mois (Couratier *et al.*, 2014, Logroscino *et al.*, 2010). Ce délai peut être expliqué en partie par le parcours réalisé par le patient, en effet le premier intervenant de proximité auquel va se référer le patient sera le médecin traitant. Une étude de 2006 définit un délai moyen de 4,7 mois après l'apparition des symptômes pour la première consultation avec le médecin traitant, et l'expertise par un neurologue serait réalisée 7 mois après cette consultation (Torny *et al.*, 2006). Une fois le diagnostic posé, le patient va bénéficier d'un suivi trimestriel auprès d'un centre référent suivant les recommandations du **Protocole National de Diagnostic et de Soins (PNDS)**<sup>4</sup> de 2015 de l'HAS. L'évolution clinique et l'apparition des symptômes seront différentes selon les patients, cependant la seule variable connue sera une aggravation et une majoration progressive du handicap conduisant au décès du patient.

La survie est variable, de quelque mois à plusieurs années, 50% des patients décèdent dans les 30 mois suivant l'apparition des premiers symptômes, 15 à 20% sont en vie à 5 ans et un faible pourcentage survit au-delà de 10 ans (Talbot, 2009). L'espérance de vie dans la SLA se situe entre 3 et 5 ans après l'apparition des symptômes, la principale cause de décès est l'insuffisance respiratoire terminale dans 77% des cas, suivies par les pneumopathies qui peuvent être dues à la présence de troubles de déglutition. Des facteurs pronostiques ont été identifiés comme l'âge de la personne, le siège initial de l'atteinte, le délai diagnostic, la présence d'une insuffisance respiratoire restrictive et la dégradation de l'état nutritionnel (Couratier *et al.*, 2016). Des facteurs psychosociaux ont aussi montré un impact sur la durée de survie : c'est le cas de la présence d'un état psychologique altéré (stress, dépression, colère, manque d'espoir) induisant une durée de survie plus courte (Couratier *et al.*, 2014). Quels que soient la forme et le mode de début, une aggravation progressive des atteintes et une majoration des difficultés vont

---

4. [https://www.has-sante.fr/portail/upload/docs/application/pdf/2015-11/pnds-sclerose\\_laterale\\_amyotrophique\\_sla.pdf](https://www.has-sante.fr/portail/upload/docs/application/pdf/2015-11/pnds-sclerose_laterale_amyotrophique_sla.pdf)

apparaître au cours du temps. Parmi celles-ci, on peut mentionner une dépendance accrue pour les activités de vie quotidienne comme l'alimentation, la mobilité et la communication, ainsi que des difficultés respiratoires. Des suppléances pourront être proposées comme la **Ventilation Non Invasive (VNI)** ou la gastrostomie. La **VNI** va améliorer la qualité de vie, et le gain de survie peut dépasser 12 mois. La présence d'une perte de poids de 5% au moment du diagnostic multiplie par 2 le risque de décès, alors que le gain de 2,5 kg sera associé à une réduction du risque de 10%, la question de la nutrition entérale par gastrostomie est donc fondamentale (**Danel-Brunaud et al., 2017**).

La question des **Directives Anticipées (DA)** sur le choix et le positionnement du patient concernant les suppléances, les soins palliatifs et les situations de fins de vie devront être abordées avec le patient et son entourage avant la perte de communication ou bien l'apparition de troubles cognitifs. Les **DA**<sup>5</sup> consistent en une déclaration écrite par le patient pour exprimer ses souhaits concernant sa fin de vie. Les **DA** permettent au médecin de prendre les décisions sur les soins à donner au moment venu, si la personne ne peut plus exprimer ses volontés; néanmoins seuls 20% des patients rédigent des **DA** (**Danel-Brunaud et al., 2017**). Pour certains auteurs, la prise en charge palliative commence au moment du diagnostic de SLA, par les caractéristiques d'évolutivité et d'engagement du pronostic vital spécifique de cette pathologie (**Brocq & Bioy, 2006, Campana-Salort, 2006**).

## 1.3 Caractéristiques cliniques

Selon l'atteinte du motoneurone central ou périphérique touché et les symptômes associés, on distingue différentes formes cliniques de la SLA. Si l'atteinte concerne le premier motoneurone, l'atteinte sera à prédominance « centrale », si elle concerne le second motoneurone, l'atteinte sera à prédominance « périphérique ».

La **SLA** se présente sous deux formes cliniques principales : la forme « spinale » qui débute par l'atteinte d'un membre, et la forme « bulbaire » qui débute par une faiblesse des muscles bulbaires, une dysarthrie, une dysphagie et des fasciculations de la langue (**Hamidou, 2015**). La forme spinale concerne 2/3 des patients et principalement les hommes, la forme bulbaire concerne 1/3 des personnes et touche préférentiellement les femmes (**Hermet, 2018**).

### 1.3.1 Formes à début spinal

Les formes à début spinal sont caractérisées par l'apparition d'un déficit moteur au niveau d'un membre. Le déficit aura un début distal et unilatéral qui peut débiter au niveau brachial ou crural. Si le déficit est bilatéral, il sera alors asymétrique. Les formes à début spinal représente 70% des cas de **SLA** (**Marin, 2011**). Les formes spinales vont avoir pour expressions cliniques : une faiblesse musculaire, des crampes, des

---

5. <https://www.service-public.fr/particuliers/vosdroits/F32010>

fasciculations.

**L'atteinte du membre supérieur** est la plus fréquente et représente 40% des cas. Cette atteinte à début brachial sera distale, se caractérisant par une faiblesse de la pince pouce-index ainsi que par une amyotrophie de l'éminence thénar, hypothénar et des interosseux provoquant un déficit des gestes fins ([Hamidou, 2015](#)).

**L'atteinte du membre inférieur** est caractérisée par une atteinte distale unilatérale se situant au niveau du pied, provoquant un pied tombant ou un steppage pouvant être responsable d'une gêne à la marche et/ou de troubles de l'équilibre. 29% des patients présentent une forme spinale de SLA débutant au membre inférieur ([Hermet, 2018](#)).

### 1.3.2 Forme à début bulbaire

La forme bulbaire est caractérisée par un tableau de paralysie labio-glosso-pharyngolaryngée. Il est observé des troubles de la phonation, de l'élocution (voix nasonnée), des troubles de la déglutition ainsi qu'une dysarthrie et des fasciculations de la langue. Le début bulbaire de la maladie est associé avec un pronostic péjoratif par rapport à un début spinal ([Couratier et al., 2014](#)). Cette forme va concerner essentiellement les femmes (2 à 3 femmes pour un homme) de 60-65 ans et son évolution est plus rapide, la médiane de survie étant généralement inférieure à 30 mois ([Hamidou, 2015](#), [Hermet, 2018](#)).

### 1.3.3 Troubles cognitifs

L'association entre SLA et troubles cognitifs est significative dans plusieurs études, allant de 30% à 50% des cas de SLA ([Benbrika et al., 2019](#), [Prete, 2016](#)). L'altération des fonctions exécutives, et la présence d'une atteinte frontale se traduisent par une diminution de la fluence verbale, une apathie et des troubles de la cognition sociale ([Viader et al., 2014](#)). Dans 5% des cas, les patients développent une **Démence Fronto-Temporale (DFT)**, associant des troubles du comportement à une atteinte du langage et des fonctions exécutives ([Pradat et al., 2016](#)). L'atteinte cognitive porte avant tout sur les fonctions exécutives, le langage et la cognition sociale, les troubles psycho-comportementaux incluent l'apathie, la désinhibition, un manque d'empathie, de l'impulsivité. Les troubles cognitifs et psycho-comportementaux aggravent le pronostic de la SLA ([Viader et al., 2014](#)).

La détection des troubles cognitifs est un élément important dans la prise en charge des patients ayant une SLA, car l'altération des fonctions cognitives et la présence de troubles du comportement peuvent avoir des conséquences sur les aidants, mais aussi sur la formulation et la prise de décision concernant la formulation des DA. La détection de ces troubles peut permettre d'expliquer aux aidants que l'attitude de la personne est altérée par la maladie et qu'elle n'est pas délibérée et totalement involontaire. « Ils aggravent le fardeau de l'aidant, qui voit ses efforts ignorés ou vécus par le patient comme une source de désagrément ou de contrainte » ([Viader et al., 2014](#)). Toutefois,

les formes et les évolutions de la SLA sont très variées, faisant de chaque malade un cas unique. Il est impossible de prédire la vitesse et les modalités d'évolution de la maladie.

La symptomatologie rencontrée dans les différentes formes de la maladie est variée, mais peut avoir de nombreux points communs comme l'atteinte motrice, l'atteinte des fonctions respiratoires ou encore la présence de troubles de déglutition et de nutrition. Dans tous les cas, la prise en charge des symptômes va être importante.

#### 1.3.4 Symptômes

A ce jour il n'existe pas de traitement curatif de la SLA. Le « Riluzole » (Rilutek) ayant seul prouvé son efficacité dans le ralentissement de l'évolution de la pathologie (Bensimon *et al.*, 1994), il bénéficie d'une Autorisation de Mise sur le Marché (AMM) pour le traitement de la SLA. Au-delà de ce traitement à visée étiologique, il est proposé différentes thérapeutiques (médicamenteuses et non médicamenteuses) afin d'agir sur les symptômes pour améliorer la qualité de vie et éventuellement la survie des patients (Cintas, 2006). Le traitement des symptômes est un axe important de la prise en charge de la SLA, celle-ci ayant une dimension pluridisciplinaire par la mise en place de traitements médicamenteux (contre la spasticité, la xérostomie, les douleurs, etc.), mais aussi par un accompagnement paramédical (infirmier, orthophonique, kinésithérapique, ergothérapique, etc.) et un accompagnement social. Cette prise en charge va se réaliser tout au long du parcours de santé du patient et nécessite une réévaluation fréquente (Kiernan *et al.*, 2011).

**Les douleurs** peuvent être de deux types dans la SLA : les crampes fréquentes aux stades initiaux et intermédiaires et les douleurs secondaires c'est-à-dire survenant à la suite de rétractions musculo-squelettiques, ou les douleurs liées aux points de pressions dus à l'immobilité. Les douleurs par excès de nociception sont les plus fréquentes. Elles sont le plus souvent secondaires à des rétractions myotendineuses, à des rétractions articulaires ou à la présence de points de compressions liés à l'immobilisation (Cintas, 2006). Le traitement des douleurs va avoir deux orientations, d'une part une prise en charge médicamenteuse du ressort du médecin, la seconde sur les douleurs liées aux rétractions musculo-squelettiques ou secondaire aux points de pression liée à l'immobilité relevant des soins de kinésithérapie, de soins de nursing et d'ergothérapie afin de trouver un positionnement optimal et antalgique.

**La spasticité** est à l'origine de douleurs, elle aggrave le handicap moteur en augmentant les difficultés ou l'impossibilité de la déambulation et des difficultés de positionnement. Il n'existe pas de traitement spécifique ayant une AMM pour le traitement de la spasticité, cependant Il peut être proposé des injections de toxines botuliques.

**Les troubles moteurs** vont avoir pour conséquence directe une perte progressive des capacités motrices, plaçant la personne en situation de handicap. Ces difficultés motrices peuvent engendrer des chutes, une tête tombante, une perte de la dex-

térité et de la marche, limitant la réalisation des [Activité de Vie Quotidienne \(AVQ\)](#). Un accompagnement en kinésithérapie et en ergothérapie est nécessaire. Il pourra être préconisé la mise en place de matériel orthétique (collier cervical, releveur de pieds, etc.) ou la mise en place et l'utilisation d'aides techniques (rollator, fauteuil roulant, lit médicalisé, etc.). La mise en place de ces aides techniques interviendra tout au long du parcours, nécessitant des évaluations fréquentes afin de proposer les aides les plus adaptées aux besoins du patient. Le financement de ces produits d'assistance, ou l'adaptation de l'environnement, nécessiteront la création de dossiers de financement auprès de divers organismes.

**Les troubles trophiques** : si les troubles cutanés de type escarres sont rares, dans la [SLA](#), il n'en est pas de même pour les troubles vasomoteurs fréquents, de type œdèmes des membres inférieurs, cyanose ([Clavelou & Guy, 2006](#)).

**La fatigue** est un élément très présent dans le contexte de la [SLA](#) ([McElhiney et al., 2009](#)). Son origine est méconnue mais probablement multifactorielle. Son retentissement délétère sur la qualité de vie est indéniable.

**Les troubles de déglutition** peuvent engendrer des conséquences importantes comme une perte de poids ([Bouteloup, 2006](#)) ou encore une pneumopathie d'inhalation. L'accompagnement par une orthophoniste et une diététicienne est important, pour prévenir les fausses routes et limiter le risque de pneumopathie d'inhalation, mais aussi pour prévenir une dénutrition.

**Les troubles nutritionnels** peuvent être d'origines diverses et multifactorielles. Comme par exemple, la dysphagie, les troubles de la sécrétion salivaire et l'anorexie secondaire à la détresse psychologique. Au cours de l'évolution de la [SLA](#), 16 à 50 % des patients présentent un déclin de l'état nutritionnel. La prévalence de la dénutrition est évaluée comme touchant entre 9 à 55% des patients dès le début de la maladie ou au cours de son évolution. Un suivi régulier sur l'état nutritionnel du patient est important, car la perte de poids est un facteur de mauvais pronostic. La prise en charge diététique et nutritionnelle des patients atteints de [SLA](#) est essentielle dès le diagnostic et durant toute l'évolution de la maladie ([Lehéricy et al., 2012](#)). La prise en charge des troubles nutritionnels est encore une fois pluridisciplinaire, et graduée dans le temps en fonction des thérapeutiques proposées par le médecin et acceptées ou non par le patient et selon l'évolution de la [SLA](#). L'orthophoniste interviendra à la fois sur un plan de rééducation pour limiter les fausses routes mais aussi par la proposition d'adaptation des textures, l'ergothérapeute pour la proposition d'aides techniques, de positionnement et d'aménagement de l'environnement favorisant la prise des repas, l'assistante sociale pour permettre la mise en place d'aides humaines ou de dispositifs « sociaux » tels que le portage des repas ou la mise en place d'aides humaines et la diététicienne assurant un suivi nutritionnel. Des moyens de suppléance peuvent être proposés aux patients comme la nutrition entérale et la gastrostomie, ces traitements nécessitent une adhésion du patient mais aussi de son entourage qui devront être informés et conseillés sur les bénéfices et les risques de ces traitements ainsi que leur incidence dans la vie quotidienne à court et à long terme.

**Altération du poids** : il est fréquent de rencontrer des altérations de la masse corporelle chez les patients ayant une SLA, qui peuvent être dues à un amaigrissement par la baisse des apports nutritionnels lors de la présence de trouble de déglutition, ou un hyper-métabolisme dont le mécanisme n'est pas connu (Pradat *et al.*, 2016).

**Les troubles salivaires** résultent de l'interaction de plusieurs facteurs comme : la modification de la consistance salivaire, la présence de trouble de la posture et de troubles de déglutition, ainsi que d'un défaut d'occlusion des lèvres provoquant une sécheresse buccale (xérostomie) ou d'une stase salivaire. « Les troubles salivaires affectent plus de 20% des patients atteints de SLA et sont à l'origine d'inconfort, d'importantes répercussions sociales et exposent le patient à un risque d'inhalation » (Cintas, 2006). Les traitements peuvent être médicamenteux comme l'injection de toxine botulique, la radiothérapie parotidienne, la pose de traitements transdermiques ou le recours à des mesures physiques telles qu'une installation favorisant la déglutition, les soins de bouche et les aspirations mécaniques.

**Les troubles digestifs** sont très fréquents et quasi constants (Clavelou & Guy, 2006). La prise en charge repose avant tout sur l'application des mesures hygiéno-diététiques, des massages et une verticalisation.

**La labilité émotionnelle (rire et pleurs spasmodique)** est un phénomène fréquent dans le cadre du syndrome pseudo bulbaire. Ce symptôme est présent chez 50% des patients. Le mécanisme de ces manifestations est inconnu. Son incidence sur la vie sociale de la personne et de sa famille est importante, et il doit être différencié des états dépressifs.

**Dépression et anxiété** : les troubles de l'humeur sont fréquents dans la SLA et présents à tous les stades de la maladie. De façon spécifique, l'anxiété est un phénomène commun lors de pathologies engageant le pronostic vital. « Bien qu'il s'agisse d'une réponse naturelle face au décès, elle peut résulter de symptômes insuffisamment traités comme la douleur et surtout la détresse respiratoire » (Cintas, 2006). La SLA va générer une angoisse importante pour le patient car il va de façon progressive devenir dépendant du milieu dans lequel il vit, à la fois en matière d'outils et de personnes (Brocq & Bioy, 2006). Le traitement repose sur la proposition d'un suivi et d'un accompagnement psychologique, ainsi que sur l'utilisation de traitements médicamenteux.

**Les troubles du sommeil** peuvent avoir des origines multiples : les difficultés de mobilisation dans le lit, le syndrome des jambes sans repos, la présence de douleurs, de troubles sphinctériens, des difficultés pour éliminer les sécrétions, l'hypersialorhée, les troubles de déglutition, l'anxiété, la dépression, l'insuffisance respiratoire mais aussi des épisodes de désaturation nocturne (Boentert, 2019) Le traitement des troubles du sommeil repose donc sur l'identification de la cause. L'objectif du traitement est de diminuer son incidence afin de favoriser la qualité de vie du patient et de son entourage.

**Les troubles respiratoires** peuvent être majorés par la présence d'autres symp-

tômes, comme les troubles de déglutition, une insuffisance de toux responsable d'encombrement et de complications broncho-pulmonaires infectieuses (Danel-Brunaud *et al.*, 2008), mais aussi par la présence de troubles posturaux. Les symptômes cliniques des troubles respiratoires peuvent être divers : insuffisance de la toux, dyspnée, désaturation, décompensation respiratoire, ils peuvent aussi s'exprimer par la présence de troubles du sommeil et des céphalées matinales. Les différents examens existants pour objectiver ces troubles sont la spirométrie, la gazométrie artérielle, l'oxymétrie nocturne et l'exploration spécifique des muscles respiratoires (fonction diaphragmatique) (Danel-Brunaud *et al.*, 2008). Les traitements utilisés pour les troubles respiratoires peuvent être différents selon les symptômes et l'évolution de la pathologie et sa forme initiale, une prise en charge kinésithérapique peut être proposée notamment pour réaliser des exercices d'accélération du flux expiratoire, tout comme l'utilisation d'instrument d'aide à la toux ou de désencombrement (Cough assist, percussionnaire ou technique d'hyper-inflation). A des stades plus avancés il peut être proposé d'autres thérapeutiques comme la ventilation non invasive (VNI) et la trachéotomie ayant chacune des répercussions différentes. Il est important de pouvoir aborder la question des suppléances respiratoires avec le patient précocement et de connaître ces choix afin de ne pas aller à l'encontre de ses désirs et choix de vie. Une enquête, faite auprès de 50 malades sur les derniers mois de vie, a montré que 62% se plaignent de difficultés pour communiquer mais surtout que 56 et 42% se plaignent de dyspnée et de troubles du sommeil (Gonzalez-Bermejo *et al.*, 2010).

## 1.4 Prise en charge sociale dans la SLA

La SLA fait partie de la liste des Affections Longue Durée (ALD), elle est inscrite au titre de l'ALD 9, c'est-à-dire les « formes graves des affections neurologiques et musculaires » permettant une exonération du ticket modérateur. Comme énoncé dans la description clinique de la pathologie, la SLA va avoir pour conséquences l'apparition progressive de différentes incapacités. Les sphères de la motricité, de la communication, de l'alimentation vont être touchées, induisant une limitation progressive dans la réalisation des AVQ et dans le cas d'atteintes cognitives associées, les Activités Instrumentales de la Vie Quotidienne (AIVQ) seront aussi touchées. Ces limitations et situations de handicap vont nécessiter la mise en place d'aides humaines pour la réalisation des AVQ mais aussi la mise en place d'AT afin de compenser les incapacités des personnes. La « loi du 11 février 2005 pour l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées », définit dans l'article 14 le terme de handicap : *Art. L. 114. - Constitue un handicap, au sens de la présente loi, toute limitation d'activité ou restriction de participation à la vie en société subie dans son environnement par une personne en raison d'une altération substantielle, durable ou définitive d'une ou plusieurs fonctions physiques, sensorielles, mentales, cognitives*

ou psychiques, d'un polyhandicap ou d'un trouble de santé invalidant.<sup>6</sup>

Si le financement des soins médicaux est du ressort de l'assurance-maladie, le financement des aides et l'octroi d'allocations de prestations sociales sont sous le régime de structures sociales et médico-sociales. En France, la prise en compte de l'âge de la personne au moment du diagnostic est un critère important qui va définir l'orientation de la personne et le type d'allocation auxquelles elle peut prétendre au regard de sa situation de handicap. Lorsque la pathologie se déclare avant 60 ans, la personne sera orientée vers la [Maison Départementale des Personnes Handicapées \(MDPH\)](#) et peut percevoir la [Prestation de Compensation du Handicap \(PCH\)](#) ; si elle a plus de 60 ans, elle sera orientée vers d'autres structures : Conseil général et [Méthodes d'Accompagnement à l'Autonomie \(MAIA\)](#), elle peut percevoir l'[Allocation Personnalisée d'Autonomie \(APA\)](#). Les montants alloués au titre de la [PCH](#) ou de l'[APA](#) sont différents bien que les personnes puissent présenter les mêmes besoins. Les prestations sociales sont importantes pour financer tout au long du parcours les aides. Ces aides peuvent être de plusieurs natures : elles peuvent être humaines (aides ménagères, aides à domicile pour la toilette, l'habillement, la préparation des repas, etc.), techniques (fauteuil roulant, synthèse vocale, contrôle d'environnement, etc.) ou financières pour financer l'aménagement du domicile, d'un véhicule, ou encore d'un poste de travail lorsque la personne est en activité. L'une des spécificités de la [SLA](#) est son caractère évolutif, provoquant parfois une aggravation rapide et importante. Les aides doivent donc être réévaluées très fréquemment pour s'adapter à l'état du patient. Si la pathologie progresse, la mise en place des aides nécessite des démarches administratives pour réévaluer les plans d'aides et les plans personnalisés de compensation. Cependant les délais administratifs ne sont pas les mêmes que les temps d'évolution de la pathologie. Le délai moyen de traitement des demandes à la [MDPH](#) est de 4 mois et demi<sup>7</sup>. Cependant dans le cadre de la [PCH](#) il est possible dans certaines situations de faire une demande de [PCH](#) en urgence ([CNSA, 2017](#)).

Selon l'étude « Aide sociale » menée par la [Direction de la recherche, des études, de l'évaluation et des statistiques \(DRESS\)](#) fin 2017, les départements ont délivré 2 millions de prestations d'aide sociale à des personnes âgées ou handicapées sur le territoire français, représentant une augmentation de 0,5% en un an ([Abdouni, 2018](#)).

### 1.4.1 Personnes de moins de 60 ans

La « Loi du 11 février 2005 pour l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées » a permis la création de structures départementales qui sont le guichet unique pour les démarches administratives concernant l'attribution d'allocations et d'aides pour les personnes en situation de handicap de moins de 60 ans, les [MDPH](#). La loi pose plusieurs principes importants dont celui de la compensation, selon lequel toute personne a le droit à une compensation du handicap,

---

6. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT00000809647&categorieLien=id>

7. réf de la cnsa mettre lien

quelles que soient la cause et la nature de la déficience. Elle s'inscrit dans une démarche d'évaluation individualisée des besoins et du projet de vie de la personne. Les aides et prestations du ressort de la MDPH sont diverses comme la PCH, l'Allocation Adulte Handicapé (AAH), la Reconnaissance Qualité de Travailleur Handicapé (RQTH) et ou bien encore la délivrance de Carte Mobilité Inclusion (CMI). Suite à une évaluation en utilisant le GEVA (Guide d'évaluation des besoins de compensation des personnes handicapées) et de la formulation du projet de vie de la personne, la MDPH va proposer un Plan Personnalisé de Compensation (PPC). Ce sont les Commissions Départementales des Personnes Handicapées (CDAPH) qui sont les instances décisionnelles qui décident de l'attribution des prestations et des orientations vers les établissements et les services médico-sociaux comme l'orientation vers un Foyer d'Accueil Médicalisé (FAM), ou vers un Services d'Accompagnement Médico-Social pour Adultes Handicapés (SAMSAH). Les MDPH sont donc des structures indispensables dans le parcours du patient. L'une des principales allocations intervenant dans le parcours du patient ayant une SLA sera la PCH. Cette allocation va permettre le financement de différents types d'aides : les aides humaines, les aides techniques, les aménagements de logement ou de véhicule et des aides spécifiques ou aides animalières. En 2017, il fut octroyé 521 500 aides sociales aux personnes en situation de handicap. La PCH représente 56% des aides, c'est-à-dire 290 400 bénéficiaires (Abdouni, 2018) sur le territoire français. Le montant de la PCH varie en fonction du niveau de ressources de la personne et de la nature de l'aide à financer<sup>8</sup>. Pour la PCH aides humaines et selon le mode de conventionnement choisi (emploi direct, service mandataire, dédommagement de l'aidant familial etc.), le tarif horaire peut aller de 3,80 €/h à 13,61 €/h. Dans le cas de la PCH aides techniques, le montant maximal attribué pour les aides est 3960 € par période de 3 ans. Elle sera pour les aménagements de logement, d'un montant maximal de 10 000 € (pour une période de 10 ans) et de 5 000 € pour l'aménagement du véhicule.

#### 1.4.2 Personnes de plus de 60 ans

L'APA fut instaurée par la Loi du 20 juillet 2001, permettant à « toute personne âgée résidant en France qui se trouve dans l'incapacité d'assumer les conséquences du manque ou de la perte d'autonomie liées à son état physique ou mental, a droit à une allocation personnalisée d'autonomie permettant une prise en charge adaptée de ses besoins »<sup>9</sup>. La « Loi relative à l'adaptation de la société au vieillissement du 28 décembre 2015 »<sup>10</sup> a permis l'intégration dans les plans d'aide l'accès aux aides techniques et l'aménagement du domicile ainsi que le soutien aux proches par le droit au répit. L'APA est à destination des personnes d'au moins 60 ans et en perte d'autonomie, elle peut être de deux types : « l'APA domicile » permettant de payer en partie ou en totalité les

---

8. [http://vos-droits.apf.asso.fr/files/Baremes/tableaux\\_tarifs\\_pch\\_2018\\_v2.pdf](http://vos-droits.apf.asso.fr/files/Baremes/tableaux_tarifs_pch_2018_v2.pdf)

9. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000406361&categorieLien=cid>

10. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000031700731&categorieLien=id>

dépenses nécessaires pour permettre le maintien à domicile et « l'APA établissement » permettant de financer une partie du tarif dépendance de l'établissement médico-social dans lequel vit la personne. Elle est versée par les conseils départementaux. Fin 2017, 1,4 million de prestations d'aide sociale départementale furent versées aux personnes âgées, dont 785 200 (55%) en aides à domiciles (Abdouni, 2018). Pour établir le degré de perte d'autonomie, une évaluation est réalisée par une équipe médico-sociale qui utilise un outil d'évaluation : la grille AGGIR. Cette grille permet de classer les degrés de perte d'autonomie des personnes en 6 groupes dits « iso-ressources » (Gir)<sup>11</sup>. Les degrés de perte d'autonomie vont du Gir 1 (perte d'autonomie la plus importante) au Gir 6 (perte d'autonomie la plus faible), seules les personnes se situant entre les Gir 1 et Gir 4 peuvent percevoir l'APA. A chaque groupe correspond un montant mensuel maximum de l'APA. Les montants alloués sont pour le Gir 1 : 1737,14 € ; pour le Gir 2 : 1394,86 € ; pour le Gir 3 : 1007,83 € ; pour le Gir 4 : 672,26 €. Si cette allocation est en un droit en France, une étude de la DRESS en 2016 montre que certaines personnes qui y ont droit, ne recourent pas à cette allocation. Les raisons de non-recours sont : une méconnaissance du dispositif, un découragement des personnes face aux démarches administratives, le montant du reste à charge, mais aussi un choix personnel (Ramos-Gorand, 2016). Un rapport récent de l'IDRES<sup>12</sup> montre que le reste à charge est toujours supérieur pour les personnes de plus de 60 ans en comparaison des personnes de moins de 60 ans (Penneau *et al.*, 2019).

Qu'il s'agisse de l'attribution de l'APA ou de la PCH, il existe des disparités territoriales dans les montants alloués comme l'indique la Cour des Comptes dans son rapport annuel de février 2018 (des comptes, 2018). Le montant moyen de l'APA est de 4 374 € par bénéficiaire en 2014, mais des écarts importants sont présents, le montant le plus faible étant celui de la Lozère à 3 461 € et le plus élevé celui de la Guyane à 8 606 €. Pour la PCH, le montant moyen par bénéficiaire au niveau national se situe à 5 555 €, avec des écarts conséquent entre les départements allant de 2 812 € à 9 407 €. Si des disparités existent dans les montants, il existe aussi des disparités sur les délais de traitement des dossiers. Au-delà de la gestion de la pathologie et de ses symptômes, il est nécessaire que le patient et/ou son entourage réalisent des démarches administratives auprès de différentes structures (MDPH, mutuelle, mairie, sécurité sociale, etc.). La connaissance des différents dispositifs existants est parfois complexe à appréhender pour le patient et son entourage.

## 1.5 Aides techniques / Produits d'assistance

Selon la norme NF ISO 9999, « est considéré comme produit d'assistance (ou aide technique) tout produit (y compris tout dispositif, équipement, instrument et logiciel) fabriqué spécialement ou généralement sur le marché, utilisé par ou pour les personnes

---

11. <https://www.pour-les-personnes-agees.gouv.fr/beneficier-daides/lallocation-personnalisee-dautonomie-apa/comment-le-gir-est-il-determine>

12. Institut de recherche et documentation en économie de la santé

en situation de handicap, pour favoriser la participation ; protéger, soutenir, entraîner, mesurer ou remplacer les fonctions organiques, les structures anatomiques et les activités, ou prévenir les déficiences, les limitations d'activité et les restrictions de la participation<sup>13</sup> ». Les aides techniques pouvant utilisées dans le parcours du patient ayant une SLA sont nombreuses, elles peuvent servir pour les activités de toilette, de déplacement ou bien pour la communication (comme illustré dans la figure B.1). Selon le type d'aide technique nécessaire, le processus d'accès ne sera pas le même et les coûts seront différents. Pour l'obtention d'un fauteuil roulant manuel, seule une prescription médicale est nécessaire. Pour l'obtention d'un fauteuil roulant électrique il est nécessaire de recourir à une évaluation des capacités du patient. Cette évaluation doit être réalisée par une équipe pluridisciplinaire constituée au minimum d'un médecin de médecine physique et de réadaptation et d'un professionnel paramédical (kinésithérapeute ou ergothérapeute). L'accès à ces consultations peut être long. Si des aides techniques, comme les fauteuils roulants, ont un remboursement (total ou partiel) par l'assurance maladie car elles font partie de la liste des produits et prestations remboursables (LPP), d'autres aides techniques ne bénéficient d'aucun remboursement, c'est le cas du système de commande oculaire « Eye Mobile Mini<sup>14</sup> » dont le prix de base est de 3550 €. Bien que des aides financières existent, parfois le reste à charge est trop important et l'accès à ces aides techniques est rendu très difficile.

Pour pallier ces difficultés, des associations de patients peuvent intervenir, en réalisant un prêt d'aides techniques. Une étude réalisée en 1998 (Corr *et al.*, 1998) montre que les aides techniques les plus utilisées et achetées par l'association SLA/MND en Irlande sont : les fauteuils, les fauteuils électriques, les lits électriques, les matelas prévention d'escarre, des coussins prévention d'escarre et des lève-personnes. Ces aides sont mises à disposition gratuitement et rapidement par l'association, favorisant ainsi leur utilisation par les patients et permettant de limiter les temps d'attente. L'un des acteurs majeur intervenant dans ce processus d'évaluation et de réévaluation sera l'ergothérapeute (Corr *et al.*, 1998). En France l'une des association qui intervient très fréquemment dans le prêt d'aides techniques auprès des patients en Île-de-France est l'Association pour la Recherche sur la Sclérose Latérale Amyotrophique et autres maladies du Motoneurone (ARSLA).

La pathologie provoque des handicaps qui se succèdent, nécessitant une adaptation constante dans la mise en place d'aides (humaines, techniques), de dispositifs médicaux (traitements, VNI, gastrostomie...), d'aménagement de l'environnement (Moutet, 2006). A ces modifications constantes s'associent des démarches administratives multiples et complexes. Le patient et sa famille sont constamment sollicités, créant parfois un épuisement de l'aidant (Fussellier, 2006). Face au constat des difficultés de prise en charge à divers niveaux (médical, social et médico-social) il a été organisé une filière permettant la mise en place de processus dans la prise en charge et dans l'accompagnement des patients atteint de SLA en France.

---

13. <https://www.cnsa.fr/compensation-de-la-perde-dautonomie/aides-techniques/les-aides-techniques>

14. <http://www.handicat.com/at-num-27893.html>

## 1.6 Organisation de la prise en charge de la SLA

### 1.6.1 La filière SLA

La prise en charge de la SLA a évolué au cours du temps, l'un des principaux moments fut la création des centres experts en 2002. Aujourd'hui la filière SLA, **FILSLAN**, est intégrée dans le plan national maladie rare. A la suite du Livre Noir de la SLA élaboré par l'**ARSLA** (anciennement ARS), et des Etats généraux de la SLA en octobre 2000 et 2001 faisant le constat des grandes difficultés auxquelles étaient confrontées les familles et les patients, les pouvoirs publics ont décidé de créer des structures médicales spécifiques adaptées à la prise en charge des patients (**Pouget, 2013**). En 2002 la **Direction de l'Hospitalisation et de l'Organisation des Soins (DHOS)** élabore une circulaire<sup>15</sup> définissant l'organisation de la prise en soin des patients atteints de SLA. Son but est d'homogénéiser les soins, d'augmenter la qualité des soins et donc d'améliorer la qualité de vie des personnes. Dans le même temps est créé par le ministère de la santé, les centres de compétence SLA et les centres experts SLA. La circulaire précise parallèlement les missions des centres de référence : « En plus du rôle de centre de compétence, un centre de référence, de vocation inter-régionale est relié par convention avec les centres de compétence de sa zone d'attraction, assume des fonctions de centralisation des données et de coordination de la recherche par l'élaboration ou la participation à des protocoles de recherche médicale et en soins infirmiers ou paramédicaux ».

Suite au Plan Maladie Rare (PNMR) 2011-2014 est créée la filière nationale de santé maladie rares SLA et Maladies du Neurone Moteur (**FILSLAN**)<sup>16</sup>, destinée à la prise en charge de personnes atteintes de maladies neuro-dégénératives ou d'origine génétique affectant la commande des muscles squelettiques. L'objectif de la filière **FILSLAN** est de favoriser l'organisation de la prise en charge des maladies rares, de coordonner l'expertise, de coordonner les actions entre les structures et les acteurs des secteurs sanitaires, médico-sociaux, sociaux et de l'accompagnement, et de faciliter l'orientation des patients mais aussi des professionnels. L'arrêté du 25 novembre 2017 fait état de 19 centres SLA labellisés en France sur l'ensemble du territoire comme indiqué à la figure 1.1. Plusieurs réseaux dédiés à la SLA sont présents sur le territoire français, le **Réseau SLA IDF** en Île-de-France, un réseau SLA en PACA (Marseille, Nice), réseau Neuro Centre (Tours), réseau SEP/SLA Auvergne (Clermont-Ferrand), cependant d'autres réseaux interviennent aussi dans la prise en charge des patients SLA (**Desmuelle, 2016**).

Au delà de la structuration de la filière, des actions ont été réalisées dans le cadre de la prise en charge, notamment lors de la conférence de consensus en 2005, avec le partenariat de la **HAS** qui a permis la rédaction des recommandations professionnelles sur la prise en charge des patients SLA en France, afin d'homogénéiser la prise en charge<sup>17</sup>. En 2007 est établi le Protocole National de Diagnostic et de Soins **PNDS**

---

15. DHOS/0/DGS/SD 5D/DGAS n° 2002-229 du 17 avril 2002, <https://solidarites-sante.gouv.fr/fichiers/bo/2002/02-38/a0383084.htm>

16. <https://portail-sla.fr>

17. [https://www.has-sante.fr/portail/upload/docs/application/pdf/Sclerose\\_laterale\\_](https://www.has-sante.fr/portail/upload/docs/application/pdf/Sclerose_laterale_)

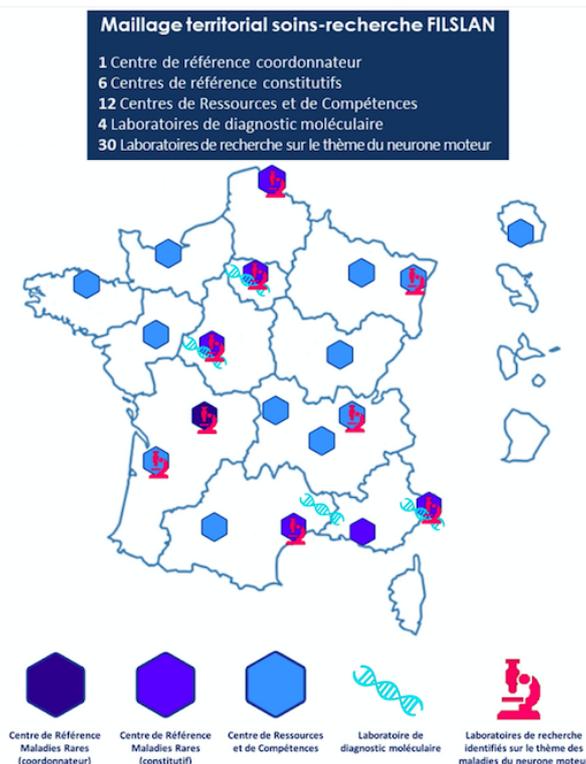


FIGURE 1.1 – Maillage des centres SLA en France.

pour la SLA <sup>18</sup>, actualisé tous les 4 ans, la dernière version date de novembre 2015. L'objectif du PNDS est de permettre aux professionnels de santé de connaître les modalités de prise en charge optimale pour les patients ayant une SLA. La prise en charge dans le PNDS prend en compte la dimension médicale, par le diagnostic, la prise en charge thérapeutique, mais aussi la dimension sociale et médico-sociale en prenant en considération les structures pouvant intervenir au domicile ou dans le financement des aides dans les parcours de santé.

## 1.7 Les parcours et la coordination

Comme nous l'avons décrit précédemment, les personnes souffrant de SLA nécessitent l'intervention de nombreux professionnels tout au long de leur parcours. La notion de parcours est liée à une façon de penser l'intervention des professionnels, auprès des personnes en difficulté du fait de la maladie, du handicap, de l'âge, des difficultés sociales et économiques (Bouquet *et al.*, 2017). Les parcours s'inscrivent dans des actions

amyotrophique\_long.pdf

18. [https://www.has-sante.fr/portail/upload/docs/application/pdf/2015-11/pnds-sclerose\\_laterale\\_amyotrophique\\_sla.pdf](https://www.has-sante.fr/portail/upload/docs/application/pdf/2015-11/pnds-sclerose_laterale_amyotrophique_sla.pdf)

nationale, dont par exemple le dispositif Paerpa, « Parcours des personnes âgées »<sup>19</sup>, à destination des personnes âgées de plus de 75 ans pour permettre aux personnes de recevoir les bons soins par les bons professionnels, dans les bonnes structures au bon moment, le tout au meilleur coût<sup>20</sup>. En France les termes « parcours de soins et parcours de santé » sont des termes apparus dans les textes de loi et devenus des paradigmes, visant à placer le patient au centre de l'organisation des soins. Au-delà de ces deux parcours, une troisième notion intervient, celle du « parcours de vie » (Piveteau *et al.*, 2014). Elle englobe les deux autres et prend en compte des dimensions plus larges que le soin ou la santé. Ces différents parcours nécessitent l'intervention de professionnels issues de trois secteurs : le secteur sanitaire, le secteur social et le secteur médico-social (qui est une spécificité française). A ces notions de parcours est associée la notion de « rupture du parcours ». Bloch & Hénaut (2014) montrent que les ruptures du parcours présentent plusieurs dimensions :

- un moment difficile à vivre pour la personne et son entourage ;
- un problème à résoudre pour les professionnels de première ligne et les pouvoirs publics ;
- un problème résultant du cloisonnement des différents secteurs d'intervention.

Ces notions de parcours, posent d'une part la question de la coordination des interventions des différents professionnels et d'autre part les questions de coûts et d'efficacité de l'accompagnement proposé aux personnes en situation de maladie chronique ou de handicap. « Ainsi, chercher à ce que toutes les compétences se coordonnent toujours mieux pour réduire les occasions de ruptures dans les parcours de soins et de santé vécues par les personnes malades ou en perte d'autonomie, c'est d'abord mieux soigner, c'est tout autant mieux dépenser. » Cordier (2014). À ces notions génériques viennent s'ajouter les spécificités des territoires, en termes d'offre de soins, de structures disponibles dans le territoire, d'Hospitalisation à Domicile (HAD), d'Hospitalisation à domicile de réadaptation (HAD-R), de Services de Soins Infirmiers A Domicile (SSIAD), de MDPH, de Centre Local d'Information et de Coordination (CLIC), équipe mobile, réseau de coordination, etc.. Au regard de ces nombreuses abréviations on peut légitimement se demander comment les patients et leurs entourages se retrouvent-ils dans la multitude de structures, de financements et des démarches administratives ? L'objectif de cette partie est de définir les concepts de parcours et de coordination, afin de comprendre les champs d'intervention du Réseau SLA IDF.

### 1.7.1 Définition des parcours

Pour illustrer les différents types de parcours : parcours de soins, parcours de santé, parcours de vie et leurs champs d'application, il est fréquent d'utiliser le schéma 1.2<sup>21</sup>

19. Personnes Agées En Risque de Perte d'Autonomie

20. <https://solidarites-sante.gouv.fr/systeme-de-sante-et-medico-social/parcours-des-patients-et-des-usagers/le-parcours-sante-des-aines-paerpa/article/le-dispositif-paerpa>

21. <https://www.nouvelle-aquitaine.ars.sante.fr/quelle-strategie-regionale-pour-reduire-les-risq>

proposé par l'Agence Régionale de Santé (ARS).

### 1.7.1.1 Le parcours de soins

Le « parcours de soins coordonnés », fut institué par la loi du 13 août 2004. Il nécessite la désignation d'un médecin traitant qui doit être consulté préférentiellement avant toute consultation d'un autre professionnel. L'objectif du parcours de soins coordonné est de « faire bénéficier chaque usager d'un suivi médical coordonné, d'une gestion rigoureuse du dossier médical et d'une prévention personnalisée. »<sup>22</sup> Le parcours de soins est axé sur la dimension médicale, avec l'accès aux consultations de premier recours et, quand cela est nécessaire, aux autres lieux de soins : hospitalisation programmée ou non (urgences), HAD, Soins de Suite et de Réadaptation (SSR), Unité de soins de longue durée (USLD) et Etablissement d'Hébergement pour Personnes Âgées Dépendantes (EHPAD). Les parcours de soins répondent aux besoins sanitaires. Ils intègrent les soins ambulatoires et hospitaliers (soins de premiers recours et hospitalisation, hospitalisation à domicile, soins de suite et de réadaptation, unités de soins de longue durée...). Dans la logique de parcours, le médecin traitant tient un rôle central dans la gestion, l'organisation et la collaboration des différents acteurs, il a une mission d'évaluation des besoins médicaux et d'orientation des patients vers les acteurs adéquats (Talon-Chrétien *et al.*, 2016).

### 1.7.1.2 La parcours de santé

L'article 1<sup>er</sup> de la Loi de modernisation de notre système de santé modifiant article L. 1411-1 (Modifié par Ordonnance n°2017-31 du 12 janvier 2017 - art. 5) du code de la santé publique<sup>23</sup>, met au point 5 de la politique de santé : « 5° L'organisation des parcours de santé. Ces parcours visent, par la coordination des acteurs sanitaires, sociaux et médico-sociaux, en lien avec les usagers et les collectivités territoriales, à garantir la continuité, l'accessibilité, la qualité, la sécurité et l'efficacité de la prise en charge de la population, en tenant compte des spécificités géographiques, démographiques et saisonnières de chaque territoire, afin de concourir à l'équité territoriale. » Les parcours de santé prennent en compte les soins, l'accompagnement médico-social et social, éléments nécessaires pour le retour et maintien à domicile .

### 1.7.1.3 Le parcours de vie

Dans son rapport d'activité de 2012 (CNSA, 2012), la Caisse Nationale de Solidarité pour l'Autonomie (CNSA) donne la définition suivante du parcours de vie : « La no-

---

22. <https://solidarites-sante.gouv.fr/systeme-de-sante-et-medico-social/parcours-de-sante-vos-droits/liberte-de-choix-et-acces-aux-soins/article/qu-est-ce-que-le-parcours-de-soins>

23. <https://www.legifrance.gouv.fr/affichCodeArticle.do?idArticle=LEGIARTI000033865687&cidTexte=LEGITEXT000006072665&dateTexte=20170114>



FIGURE 1.2 – Représentation des parcours

tion de parcours de vie recouvre le cheminement individuel de chaque personne dans les différentes dimensions de sa vie : personnelle et relationnelle, professionnelle et sociale, familiale et citoyenne. [...] la notion de parcours de vie désigne l'ensemble des événements intervenant dans la vie d'une personne et les différentes "périodes" et transitions qu'elle connaît. Ces derniers affectent son bien-être physique, mental et social, sa capacité à prendre des décisions ou à maîtriser ses conditions de vie, ses interactions avec son entourage, sa participation à la vie sociale. » Le parcours de vie prend en compte la personne dans son environnement : au niveau familial, au niveau professionnel et dans son lieu de vie (logement, ville), etc.

Au vue de ces définitions, il est possible de dire que les coordinateurs du Réseau SLA IDF vont intervenir sur les trois dimensions des parcours. Ils interviennent lors du parcours de soins, dans le parcours de santé lors de la recherche et la mise en place de ressources (professionnel, structure médico-sociale, financement, etc.) et dans le parcours de vie sur l'accompagnement dans les projets de vie des personnes. Ces accompagnements par les coordinateurs du Réseau SLA IDF se font en lien avec l'ensemble des intervenants (ceux de proximité, ceux du centre expert, etc.).

#### 1.7.1.4 Les ruptures de parcours

Si les concepts de parcours de vie, parcours de soins et parcours de santé sont définis de façon spécifique, il n'existe pas de définition unanime de la « rupture de parcours ». Le parcours renvoie à une succession d'évènements au cours desquels, peuvent apparaître des périodes de transition qui peuvent entraîner des impacts négatifs et provoquer des ruptures (ANESM, 2018). La loi n° 2016-41 du 26 janvier 2016<sup>24</sup> de modernisation de notre système de santé consacre son titre II aux parcours de santé, la rupture de parcours se retrouve dans l'article 74 par la création des plateformes territoriales d'appui (PTA) dont l'objectif est de contribuer à prévenir les hospitalisations inutiles ou

24. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000031912641&categorieLien=id>

évitables ainsi que les ruptures de parcours de santé complexe.

Les ruptures de parcours de soins sont associées à la non continuité des soins, comme par exemple l'absence ou l'arrêt d'une prise en charge (par l'absence d'un professionnel : médecin traitant, kinésithérapeute, etc.). Les hospitalisations évitables ou inutiles sont aussi des ruptures de parcours, cependant leurs identifications sont complexes car des critères médicaux doivent être clairement définis. Seul un médecin peut définir si une hospitalisation est évitable ou nécessaire. Il n'existe pas de définition de « rupture de parcours de santé », dans notre travail nous considérons que la présence d'un épuisement peut être considéré comme une rupture de parcours, car elle nécessite la mise en place de séjours de répit. De même l'arrêt d'une prise en charge médico-sociale par une structure, un professionnel, ou un manque d'aide humaine

### 1.7.2 La coordination

La coordination est un élément important, car elle concerne de nombreux intervenants qui habituellement travaillent isolément et en silo. Dans l'accompagnement des personnes il est nécessaire de prendre en compte les dimensions sociales et environnementales du patient et de sa famille (Brun *et al.*, 2016, Hénaut & Bloch, 2016). L'objectif de la coordination est de favoriser la coopération entre les professionnels et d'éviter la fragmentation des éléments entre le monde médical et social (Hénaut & Bloch, 2016). La HAS s'appuie sur une définition nord américaine de (McDonald *et al.*, 2007) pour définir la coordination : « la coordination est l'organisation délibérée des activités de soins à un patient entre plusieurs acteurs du système de santé. Elle implique la mobilisation de personnels et d'autres ressources nécessaires à la prise en charge du patient et est souvent gérée par l'échange d'informations entre les participants responsables de différents aspects des soins. ». Une étude de (Fishman-Bosc *et al.*, 2016) met en avant les 6 attributs de la coordination des soins : *a*) la planification proactive (qui est l'adaptation des plans de soins en fonctions des ressources et des besoins de la personne et la gestion de conflits) ; *b*) centrée sur la personne, en prenant en compte la dimension holistique de la personne : à savoir la réalisation d'une évaluation des besoins, le prise en compte des choix et préférence de la personne et de son entourage ; *c*) la corrélation qui est utilisée pour faire référence à la dimension de communication et de la collaboration, entre les professionnels mais aussi avec le patient ; *d*) le partage : partage de l'information, le partage des connaissances, le partage de pouvoir entre les professionnels et le patient ; *e*) la transversalité qui prends en compte les ressources de proximité (soins à domicile) et les ressources plus spécialisées (soins hospitaliers) ; *f*) la facilitation, lors des périodes de transfert ou de transition, elle optimise l'orientation, l'un des acteurs de cette facilitation est le coordinateur.

Bloch & Hénaut (2014) définissent la coordination comme « le besoin d'assurer la cohérence des différentes tâches dans un souci d'efficacité ». Les auteurs font état des difficultés que rencontrent les personnes et les acteurs du système de santé :

- un manque de visibilité de l'offre pour les personnes, dû à « la fragmentation de l'offre en une myriade d'administrations, d'institutions et de prestations relevant de secteur et de statut variés » ;

- la difficulté d'accès à l'information et aux services répondant aux besoins de leurs proches, élément qui augmente parfois leur épuisement, car cela s'ajoute à l'aide quotidienne qu'ils doivent prodiguer à leurs proches ;
- le manque de transmission d'informations, d'harmonisation des pratiques, la redondance des évaluations qui peuvent parfois être délétères pour les personnes si les évaluations sont contradictoires ;
- le cloisonnement administratif, qui provoque des délais d'attente parfois longs pour l'obtention d'aides financières et techniques ;
- pour les pouvoirs publics, la fragmentation et le cloisonnement du système de santé et d'aide ont pour conséquences :
  1. une mauvaise maîtrise des dépenses publiques (due à la redondance des évaluations) et des hospitalisations évitables (dues aux manques de coordination) ; cette difficulté de maîtrise des coûts vient aussi du manque « de fongibilité entre crédits sanitaires et médico-sociaux. »
  2. des difficultés de planification de l'offre résultant du manque de données sur les besoins des populations vers les décideurs, du manque d'interopérabilité entre les différents secteurs dans les systèmes d'informations, mais aussi des contraintes liées au respect de la confidentialité et à la sécurité des données personnelles.

Pour [Desnuelle \(2016\)](#) « La coordination est le maître-mot pour pallier les difficultés d'adaptation des soins de proximité vis-à-vis d'une pathologie souvent méconnue, lourde en soins et en charge émotionnelle ; elle est nécessaire au maintien à domicile afin d'éviter les hospitalisations inutiles, longues ou inadaptées. » C'est pour intervenir entre la ville et l'hôpital et faire le lien entre les différents professionnels intervenant que les réseaux SLA ont été créés. La section suivante va présenter de façon plus spécifique le Réseau SLA IDF.

## 1.8 Le réseau SLA Île-de-France

Le Réseau SLA IDF créé en décembre 2005, est défini comme un réseau ressource se situant au carrefour entre la ville et l'hôpital, dont la mission est d'accompagner les patients suivis au centre SLA et inclus au réseau. Il fut créé dans les suites de la circulaire<sup>25</sup> relative à l'organisation des soins pour la prise en charge des patients souffrant de sclérose latérale amyotrophique. L'objectif du réseau est de répondre aux exigences formulées dans le chapitre 3 des missions des centres de prise en charge, définissant les missions comme « assurer le transfert de connaissances et de savoir-faire auprès des acteurs de proximité et la mise en oeuvre des moyens nécessaires sur le lieu de vie » ([Corcia & Bourdillon, 2008](#)). Dans cette circulaire, il est indiqué que le suivi des patients se situe essentiellement à domicile, et qu'il est important, compte tenu de l'évolution et de la polymorphie des conséquences de la SLA, d'assurer un suivi sur le

---

25. DHOS/0/DGS/SD 5D/DGAS n° 2002-229 du 17 avril 2002

lieu de vie par des professionnels expérimentés, en concertation avec le médecin et les acteurs de proximité. « Les soins étant réalisés à domicile, la prise en charge doit intégrer l’objectif du maintien à domicile et de facilitation de la vie quotidienne, dès le début du suivi. Le suivi sur le lieu de vie par des professionnels expérimentés deviendra en outre d’autant plus indispensable qu’avec l’évolution de la maladie et de la dépendance fonctionnelle ou sanitaire (assistance respiratoire), les problèmes sont de plus en plus complexes, et que les déplacements au centre de consultation hospitalier deviennent plus difficilement réalisables et moins bien acceptés ». Les missions du réseau sont en direction des patients, des familles mais aussi vers les professionnels intervenant dans l’accompagnement des patients.

Les missions du Réseau SLA IDF sont [Abdelnour & Cordesse \(2008\)](#) :

- accompagner les patients, leur entourage et les intervenants de proximité tout au long de la maladie ;
- répondre aux diverses questions des patients ;
- améliorer la coordination entre les intervenants de proximité et hospitaliers ;
- diffuser l’information portant sur les avancées médicales concernant la maladie ;
- former les différents intervenants de proximité.

Le [Réseau SLA IDF](#) se traduit physiquement par des locaux se situant dans l’enceinte de l’hôpital de la Pitié-Salpêtrière à Paris. Cette localisation géographique permet aux coordinateurs une proximité avec le centre de référence SLA, favorisant les échanges d’information et le suivi des patients. Les coordinateurs peuvent participer aux consultations multidisciplinaires réalisées aux centres tous les 3 mois. Dans la mesure où la prise en charge des patients atteints de SLA nécessite des soins permanents, il est indispensable d’assurer une interface entre la prise en charge hospitalière multidisciplinaire trimestrielle et la prise en charge quotidienne de proximité. Cette communication entre ces deux modes d’exercice légitime l’organisation de réseaux de soins afin de promouvoir le maintien des patients au domicile et d’apporter soutien et formation aux professionnels de santé de proximité ([Corcia & Bourdillon, 2008](#)). De nombreux intervenants sont nécessaires lors du parcours de soins du patient, comme indiqué dans la figure 1.3 ; une coordination entre ces professionnels est nécessaire et peut être assurée par le Réseau SLA IDF.

### 1.8.1 Mode de fonctionnement du Réseau SLA IDF

Les critères d’inclusion des patients au Réseau SLA IDF sont :

- les patients doivent vivre en région Île-de-France ;
- l’inclusion doit être demandée par un neurologue (du centre expert ou non) ;
- le diagnostic de SLA doit avoir été posé ;
- il n’existe pas de limite d’âge ;
- l’inclusion est effective lors de la signature par le patient et de l’envoi des documents d’adhésion au Réseau SLA IDF.

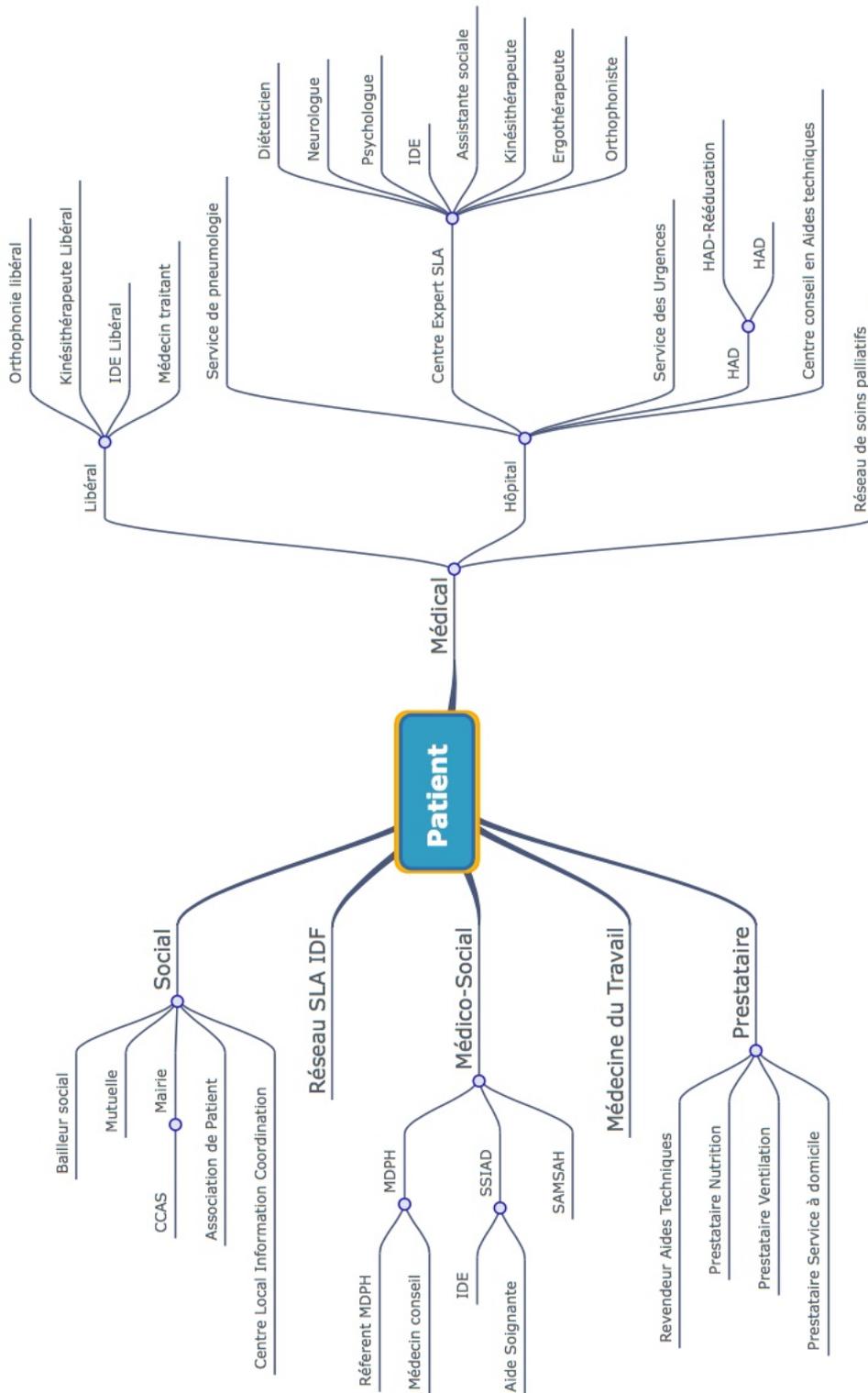


FIGURE 1.3 – Liste (non exhaustive) des professionnels et des structures intervenants dans le parcours du patient atteint de SLA. Cette représentation permet d’observer le nombre conséquent de professionnel pouvant intervenir auprès des patients.

Le neurologue informe le patient de l'existence du Réseau SLA IDF, lors de la consultation d'annonce et de la consultation du retour d'annonce, et des possibilités d'adhésion. Il est proposé au patient et à sa famille une consultation dénommée « retour d'annonce » dans les 15 jours qui suivent l'annonce diagnostique. Il est fréquent que les coordinateurs rencontrent les patients et leurs familles lors de cette consultation (Corcia & Bourdillon, 2008). Lors de la première rencontre, le coordinateur va exposer les objectifs et les modalités d'intervention du Réseau SLA IDF. Cette rencontre permet un premier recueil de données sur la situation médico-sociale de la personne. Au terme de cette rencontre il est proposé au patient d'adhérer au Réseau SLA IDF. Le patient est libre d'y adhérer ou non. La troisième étape est l'inclusion au Réseau SLA IDF. Cette dernière est actée lorsque le réseau reçoit le formulaire d'adhésion signé par le patient. Cette étape est un préalable important et nécessaire à toute action de coordination, sans autorisation du patient les coordinateurs ne réalisent aucune démarche. Les coordinateurs vont agir principalement sur sollicitation du patient, des professionnels ou bien de la famille. La sortie définitive du patient du Réseau SLA IDF peut résulter :

- de la demande du patient qui souhaite quitter la structure ;
- du déménagement du patient vers une autre région ;
- du décès du patient.

### 1.8.2 Impact du Réseau SLA IDF dans la prise en charge

Les patients vivent la majeure partie de leur temps à domicile, les modifications dues à l'évolution de la pathologie (perte des capacités motrices, apparition des troubles de déglutition, modification du caractère, apparition de trouble de l'humeur, etc.) vont apparaître à domicile. Dans ces périodes, le Réseau SLA IDF devient donc un interlocuteur privilégié pour les patients, leurs familles et les intervenants afin de trouver une aide ou une réponse rapide pour la résolution d'une problématique. Dans deux rapports successifs, l'un en mai 2008, « Evaluation du centre de référence national de la prise en charge de la sclérose latérale amyotrophique de Paris » (Corcia & Bourdillon, 2008), et le second en février 2009, « Evaluation des dispositifs mucoviscidose et sclérose latérale amyotrophique », le rôle et la valeur ajoutée du Réseau SLA IDF sont mis en avant. Le premier rapport met en avant l'efficacité de la complémentarité des actions entre le centre et le Réseau SLA IDF favorisant un accompagnement optimal des patients ; le second indique le rôle d'interface favorisant les actions entre le centre SLA et les acteurs locaux, permettant une optimisation de la coordination des actions de chacun (Alcimed *et al.*, 2009). Le positionnement du Réseau SLA IDF en réseau ressource se situant au carrefour entre le centre expert et le domicile permet d'assurer la coordination, favorisant la continuité des soins, mais aussi d'apporter une réponse rapide aux patients.

Les coordinateurs du réseau sont des professionnels paramédicaux : infirmières et ergothérapeute et un psychologue, qui ont tous une expérience clinique. Ils ont des compétences multiples, à la fois dans le domaine médical sur les troubles et la symp-

tomatologie, mais aussi des connaissances dans le domaine médico-social et social. L'objectif des coordinateurs n'est pas de remplacer les acteurs-clés intervenant auprès des patients (neurologue, assistante sociale, orthophoniste, etc.) mais d'être un premier niveau de réponse et d'orientation. Les coordinateurs ont une vision globale (vision holistique) de la situation de la personne, du contexte de vie, de l'évaluation clinique et de la situation sociale. Pour tracer les actes réalisés une base événementielle a été créée, décrite dans la partie 1.8.3 de ce manuscrit. Cette base de données permet d'inscrire l'ensemble des informations sur le patient et sa situation, et les solutions et actions de coordinations proposées aux personnes au cours de leurs parcours. Lorsque le réseau est sollicité (quel que soit le demandeur) différents cas peuvent se présenter :

- le coordinateur peut apporter une réponse directe à la demande ;
- le coordinateur ne peut apporter une réponse immédiate, il doit solliciter un expert en fonction de la problématique (neurologue, assistante sociale, ergothérapeute, etc.). Il pourra par la suite proposer une réponse au demandeur.

Les réponses et les actions de coordination peuvent être pluri-thématiques (médicale, sociale) et de plusieurs niveaux, comme par exemple : une transmission d'information ; une explication concernant des gestes techniques à faire en lien avec une suppléance ; une évaluation de la situation ; une préconisation ou une écoute.

En 2015, une étude a été réalisée, dans l'objectif de comprendre l'impact des actions de coordinations, sur les hospitalisations et le pronostic dans le cadre de l'accompagnement des patients ayant une SLA (Cordesse *et al.*, 2015). Les données utilisées, lors de cette étude (celles du [Système National d'Information Interrégimes de l'Assurance Maladie \(SNIIRAM\)](#)) permettent de mettre en évidence qu'en 2012, 623 patients prennent du riluzole en Île-de-France. Au cours de cette période le réseau a accompagné 570 patients, donc 91,5% des patients ayant une SLA en Île-de-France sont inclus au Réseau SLA IDF. Cette étude a mis en évidence, par la comparaison des deux groupes (l'un avec coordination et l'autre sans), l'impact de la coordination des soins sur la diminution du nombre d'hospitalisations et de l'altération fonctionnelle, ainsi qu'une amélioration de la durée de survie.

### 1.8.3 Base de données du Réseau SLA IDF

Le [Réseau SLA IDF](#) a mis en place un dossier de coordination informatisé, permettant un recueil de données. Chaque patient va posséder un dossier individuel de coordination. Le dossier est composé de différentes parties. Certaines sont structurées et permettent un recueil des données socio-démographiques et médico-sociales du patient (volet), une autre partie est un espace de saisie de texte libre pour la transcription des demandes et des actions de coordination qui sont mises en place dans le cadre de l'accompagnement des patients.

Cette base est enregistrée auprès de la Commission Nationale Informatique et des Libertés (CNIL). Elle est accessible à partir du portail sla<sup>26</sup>, dans l'onglet réseau sla<sup>27</sup>.

26. <https://portail-sla.fr>

27. <http://reseau-idf.portail-sla.fr/login.0>

L'accès à cette base nécessite un identifiant et un mot de passe individuel. Afin de comprendre le contenu de cette base et l'exploitation que nous en avons fait, nous allons en décrire ci-dessous les différentes parties. Le dossier de coordination doit permettre de regrouper l'ensemble des informations nécessaires à la coordination du parcours des patients mais il doit aussi regrouper l'ensemble des propositions faites par le Réseau SLA IDF, afin que chaque coordinateur intervenant sur le dossier patient puisse apporter des réponses aux patients. En effet, il n'existe pas de sectorisation des coordinateurs, tous peuvent intervenir sur les différentes situations.

Le dossier de coordination est structuré sous la forme de huit volets thématiques :

**Un volet fiche (en annexe B.2) :** cette partie regroupe les données socio-démographiques de la personnes (nom, prénom, adresse, nom de la personne de confiance, etc.).

**Un volet médical (en annexe B.3) :** permet de renseigner les données sur l'histoire de la maladie (la date de diagnostic, la date des premiers symptômes, la forme initiale, etc.) mais aussi le positionnement du patient (acceptation, refus) concernant les suppléances que sont la [VNI](#), la gastrostomie et la trachéotomie.

**Un volet social (en annexe B.4) :** permet de renseigner la situation sociale et familiale du patient. Il permet aussi d'identifier les prestations ou allocations sociales perçues par la personne [Allocation Compensatrice Tierce Personne \(ACTP\)](#), [PCH](#), [APA](#). Ces données ne sont cependant, pas toujours renseignées.

**Un volet ergothérapie et un volet psychologie :** ils sont peu, voire pas utilisés par les coordinateurs.

**Un volet intervenant (en annexe B.5) :** regroupe l'ensemble des coordonnées des professionnels de proximité qui interviennent dans la situation ; il peut s'agir de personnes physiques (infirmiers libéraux, kinésithérapeute ou orthophoniste, etc.) mais il peut aussi s'agir de structure de types [SSIAD](#) ou [SAMSAH](#).

**Un volet évènement (présenté à la figure 1.4).** La partie évènement est différente des autres volets, elle permet la saisie de texte libre. Dans cette partie, les coordinateurs vont saisir des résumés des demandes, des rencontres faites avec les patients, leurs familles ou avec les professionnels. L'ensemble des informations saisies permettent de retracer l'ensemble du parcours de la personne de son inclusion à sa sortie du Réseau SLA IDF. Il existe une grande hétérogénéité : sur le nombre d'évènements saisis par patient mais aussi sur les demandes et les besoins exprimés.

Au 26 août 2019, il est indiqué que la base du Réseau SLA IDF est constituée de 2 684 dossiers de patients, dont 528 actifs (c'est-à-dire le nombre de patients en vie), le nombre d'évènements est de 42 935 évènements (principaux) et la base contient les coordonnées de 12 258 intervenants. Il est possible, à partir du portail de la base du Réseau SLA IDF, de réaliser un export des informations contenues dans la base vers un fichier excel. Dans le cadre de notre travail, nous avons utilisé cette option pour extraire l'ensemble des évènements des patients. Nous avons choisi de sélectionner pour notre étude les patients inclus entre le 1 janvier 2013 et le 31 décembre 2017. Le choix de ces années résulte de l'initiation de ce travail qui a débuté de façon formelle en 2015, et dont la première phase fut l'extraction de candidats termes à partir des corpus, comme

indiqué dans le chapitre 6. Pour avoir un corpus conséquent pour réaliser les analyses statistiques, mais aussi pour avoir un nombre de parcours « complet », nous avons fait le choix de prendre l'ensemble des patients inclus entre 2013 et 2017.

Portail de la coordination des centres et des réseaux de prise en charge de la Sclérose Latérale Amyotrophique

**SLA PRO** Portail SLA Pro - Accès Membres

- User Dernière connexion le 10/01/2019 Déconnexion

Nouveau patient Patients actifs Archives Alertes Bilans & Exports

Patient : 473 Sélectionner

Fiche Médical Social Ergo. Psy. Intervenants Situ. Évènements

Ajouter un évènement Filtre Pages : << < 1 2 3 4 5 6 7 8 > >>

**• Suivi fermé**

Date [redacted]  
 Intervenant [redacted]  
 Évènement **appel reçu**  
 Nature **médical**  
 Gravité "3-4"  
 Interlocuteur -

Appel de [redacted] concernant un mail envoyé à [redacted] -  
 [redacted] A besoin d'un certificat médical du Docteur  
 [redacted] pour constituer son dossier MPDH.  
 Docteur [redacted] est en congés jusqu' [redacted] et [redacted]  
 jusqu'au [redacted]

ai envoyé un mail à [redacted] lui expliquant la demande de  
 [redacted] et lui demandant de la recontacter à son retour de  
 congés.;  
 Lui ai précisé la demande de RQTH et l'ai interrogé sur la PCH  
 AT.

Date [redacted]  
 Intervenant [redacted]

CR de CMD du [redacted]  
 Parésie prédominant franchement sur l'hémicorps gauche avec  
 composante spastique et instabilité posturale marquée source de  
 plusieurs chutes dans les escaliers.  
 poids stabilisé et BR normal  
 majoration le Riluzole à ½ cp matin et 1 cp soir sous réserve de  
 la tolérance clinique et biologique (transaminases à surveiller une  
 fois par mois)  
 autonome dans les AVQ atteinte mbre sup G, risque de chute:  
 proposition canne pliante  
 texture nmale  
 vit avec encore 2 de ses fils + mari, accessibilité du logement  
 avec marche  
 travaille à tps complet et voudrait teletravail  
 social: a le 100% et a demandé à [redacted] de lancer le dossier  
 MDPH CI et CES

**Supprimer l'évènement**

*Pas de suivi nécessaire*

Date [redacted]  
 Intervenant [redacted]  
 Évènement **autre**  
 Nature **médical**  
 Complexité 1  
 Gravité "1-2"  
 Interlocuteur [redacted]

Mail reçu de [redacted] demandant au [redacted] qu'il lui  
 prescrive, un arrêt de travail à compter du [redacted]

Mail transférer au [redacted] fait l'arrêt de travail et l'envoi à la  
 patiente. Arrêt à compter du [redacted]

**Supprimer l'évènement**

*Pas de suivi nécessaire*

FIGURE 1.4 – Capture d'écran de l'interface de la base évènementielle du Réseau SLA IDF.

## 1.9 Synthèse et hypothèse de travail

La description de la SLA a permis de voir la diversité des symptômes, le caractère évolutif de la pathologie ainsi que la diversité des professionnels qui doivent intervenir dans les dimensions médicales, sociales et médico-sociales. L'interaction de ces différentes dimensions, qui dépendent de systèmes et d'organisations spécifiques (MDPH, conseil général, centre expert, structure de soins de ville, etc.), la rapidité et la diversité d'apparition des situations de handicap (moteurs, cognitifs, respiratoires, etc.) rendent la prise en charge complexe. Le « parcours de santé complexe » est défini comme la nécessité d'intervention de plusieurs catégories de professionnels de santé, sociaux ou médico-sociaux, en raison de l'état de santé ou de la situation sociale du patient<sup>28</sup>. La SLA, va positionner le patient dans le cadre d'un parcours de santé complexe, de plus elle nécessite la mise en place de ressources de façon rapide et adaptée (Cordesse *et al.*, 2013) pour répondre aux besoins. Des ruptures de parcours peuvent apparaître, comme l'absence ou l'arrêt de prise en charge (par une structure ou un professionnel), la présence d'hospitalisations évitables, ou bien encore la présence de l'épuisement des aidants.

L'hypothèse de notre travail est qu'en analysant les événements saisis par les coordinateurs du Réseau SLA IDF, il est possible de comprendre le parcours des patients. Des données normalisées, indiquant les médicaments prescrits ou les actes de soins réalisés existent, comme le Programme de Médicalisation du Système d'Informations (PMSI) ou le SNIIRAM. Cependant, ces données n'indiquent pas les besoins et les demandes faites par le patient ou son entourage, par exemple le fait qu'ils n'aient plus de médecin traitant, de kinésithérapeute, ou bien que les patients aient besoin de plus d'aides humaines, ou que l'entourage soit en épuisement, etc. Les événements, sont sources d'informations sur le parcours des patients. Ils sont rédigés en textes libres sans structuration spécifique, pour annoter ces événements par un système informatique il est nécessaire de recourir à un système d'organisation des connaissances. Le chapitre suivant va présenter les systèmes d'organisation des connaissances existant et en particulier les ontologies. Nous avons choisi dans notre travail, de construire une ontologie, pour exploiter les informations présentes dans les événements pour tenter de comprendre les parcours des patients ayant une SLA.

---

28. <https://www.legifrance.gouv.fr/affichLoiPreparation.do?idDocument=JORFDOLE000029589477&type=contenu&id=2&typeLoi=proj&legislature=14>



# Chapitre 2

## Les Systèmes d'Organisation des Connaissances

### Sommaire

---

<b>2.1</b>	<b>Les systèmes d'organisation des connaissances (SOC)</b>	<b>42</b>
2.1.1	Définition des SOC	43
2.1.2	Typologie des SOC	44
<b>2.2</b>	<b>Les ontologies informatiques</b>	<b>48</b>
2.2.1	Types d'ontologies	48
2.2.2	Les composants d'une ontologie	50
<b>2.3</b>	<b>Les RTO dans le domaine de la neurologie, de la coordination et du handicap</b>	<b>53</b>
2.3.1	Les ressources en neurologie	53
2.3.2	Une ontologie de la coordination de soins infirmiers	54
2.3.3	La Classification Internationale du Fonctionnement du Handicap et de la Santé (CIF)	55

---

*Comme nous l'avons vu dans le chapitre précédent, nous souhaitons identifier les causes de rupture de parcours des patients ayant une SLA accompagné dans le cadre du Réseau SLA IDF. Pour identifier les éléments intervenant dans ces ruptures, la matière dont nous disposons est l'ensemble des données textuelles contenues dans les évènements, relatant de façon chronologique les demandes et les actions réalisées au cours du parcours. Notre hypothèse est que ces transcriptions contiennent un ensemble d'informations permettant d'émettre des hypothèses sur les causes de rupture. Comme nous l'avons indiqué dans le chapitre 1 le parcours de santé a une dimension multidisciplinaire, à la fois sur un accompagnement médical, médico-social et social en lien avec le l'environnement de la personne. Le texte est un élément important dans la médecine, il est le vecteur d'information auprès des patients, des professionnels, il peut être normalisé ou non. Le traitement de l'information médicale peut être utilisé dans un objectif différent, pour la recherche clinique ou bien pour le recueil sur l'activité et la gestion d'un établissement de santé, cependant elle nécessite un recueil de données normalisées.*

*Si l'on souhaite exploiter et extraire de la connaissance sur les parcours des patients, il est nécessaire d'annoter les évènements. Pour cela il est nécessaire de structurer et organiser les connaissances présentes dans les corpus. La première section de ce chapitre présente les différents systèmes d'organisation des connaissances (SOC) existant comme les thésaurus, les classifications et les ontologies formelles. Dans le travail mené, la construction et l'utilisation d'une ontologie nous a semblé être le système d'organisation des connaissances le plus adapté à notre recherche. Nous présentons dans la seconde section, les différentes ontologies selon leurs niveaux de modélisation, ainsi que les composants d'une ontologie.*

## **2.1 Les systèmes d'organisation des connaissances (SOC)**

Les évènements textuels de la base de données du Réseau SLA IDF contiennent différentes connaissances : des connaissances contextuelles sur la situation du patient à un moment donné qui sont spécifiques de son histoire, de son environnement, mais aussi des connaissances métier (professionnelles), qui au regard de la présence de certains symptômes, problématiques ou situations nécessitent la mise en place d'actions spécifiques par les coordinateurs. La connaissance est définie dans le dictionnaire Larousse comme : « Action, fait de comprendre, de connaître les propriétés, les caractéristiques, les traits spécifiques de quelque chose. Opération par laquelle l'esprit humain procède à l'analyse d'un objet, d'une réalité et en définit la nature ». L'organisation des connaissances est une activité ancienne et naturelle de l'être humain (Ermine, 2000, Hudon & Hadi, 2010), qui lui permet d'apprendre, d'agir sur son environnement et de résoudre des problématiques. La connaissance s'inscrit dans un contexte, elle possède un aspect pragmatique lié à la situation dans laquelle elle s'inscrit. Une connaissance est de l'information qui prend une certaine signification dans un contexte donné : « une connaissance n'existe pas seulement parce qu'elle a une forme et une signification don-

née, mais aussi parce que cette forme et cette signification sont données dans un milieu qui lui donne sa richesse et sa pertinence » (Ermine, 2000). Pour structurer et organiser cette connaissance, différents SOC existent comme les thésaurus, les glossaires, les nomenclatures et les ontologies.

### 2.1.1 Définition des SOC

Les SOC sont créés pour organiser des documents, des travaux et des concepts (Hjørland, 2008), leur utilisation est importante pour partager une connaissance, les SOC servent de référentiel dans un domaine. Ils vont être construits et choisis en fonction de leur objectif de leur finalité. Les SOC peuvent être de natures différentes et présenter une structuration spécifique, comme le glossaire, une classification, un dictionnaire ou une ontologie formelle. Si leur nature et leur structuration peuvent être différentes, leur utilisation peut être variée (Grosjean, 2014, Vandenbussche, 2011) comme :

- représenter les connaissances d'un domaine spécifique, ce qui permet la normalisation d'un vocabulaire commun pour faciliter l'échange d'information ;
- rechercher l'information et effectuer des statistiques, comme le codage de dossiers patients ;
- inférer des faits, des règles pour réaliser des tâches de raisonnement.

Afin d'avoir une vision globale des SOC nous avons choisi de présenter plusieurs définitions de ce que peut être un SOC.

**Définition de Hodge (2000) :** « The term knowledge organization systems is intended to encompass all types of schemes for organizing information and promoting knowledge managements. Knowledge organization systems include classification schemes that organize materials at a general level (such as books on a shelf), subject headings that provide more detailed access, and authority files that control variant versions of key information (such as geographic names and personal names). They also include less-traditional schemes, such as semantic networks and ontologies. »<sup>1</sup> A cette définition, Hodge (2000) ajoute les caractéristiques communes aux SOC :

1. un SOC impose une vision particulière du monde à une collection et à ses éléments ;
2. la même entité peut être caractérisée de différentes manières, en fonction du SOC utilisé ;

---

1. Le terme « systèmes d'organisation de la connaissance » vise à englober tous les dispositifs pour organiser l'information et promouvoir la gestion des connaissances. Les systèmes d'organisation de la connaissance incluent des dispositifs de classification qui organisent des documents à un niveau général (tels que des livres sur une étagère), des rubriques qui fournissent un accès plus détaillé et des fichiers d'autorité qui contrôlent les versions des informations clés (telles que les noms géographiques et les noms de personne). Ils incluent également des dispositifs moins traditionnels, tels que les réseaux sémantiques et les ontologies.

3. il doit exister suffisamment de points communs entre le concept exprimé dans un SOC et l'objet du monde réel auquel ce concept fait référence, à savoir qu'une personne bien informée pourrait appliquer le système avec une fiabilité raisonnable. De même, une personne cherchant du matériel pertinent en utilisant un SOC doit pouvoir associer son concept à sa représentation dans le système.

Selon Grosjean (2014), les SOC sont des vocabulaires plus ou moins structurés qui sont utilisés dans les systèmes d'information, principalement en science ; ils sont des ressources permettant de stocker et d'exploiter la connaissance.

Pour les auteurs Zacklad & Giboin (2010) « Les SOC visent à définir des principes de description d'un domaine pour faciliter les opérations de classement et de recherche d'items plus ou moins abstraits : documents, personnes, lieux, produits, opinions ou activités. »

Pour Vandenbussche (2011), un SOC est un « ensemble de connaissances en interaction, représentées et regroupées au sein d'une structure dans le but de répondre à des besoins et d'atteindre des objectifs déterminés ». Leur construction répond à un besoin d'usage.

L'ensemble de ces définitions permettent de mettre en avant des points communs des SOC :

1. un processus de structuration des connaissances qui peut être plus ou moins complexe,
2. les SOC sont une conception en vue d'un objectif d'exploitation.

## 2.1.2 Typologie des SOC

Nous allons présenter dans cette partie certains SOC qui sont les plus couramment utilisés, afin de comprendre leurs spécificités et de pouvoir positionner notre objet d'étude, les ontologies.

### 2.1.2.1 Glossaire

Hodge (2000) définit ce type de SOC de la manière suivante « Un glossaire est une liste de termes, généralement avec des définitions. Les termes peuvent provenir d'un domaine spécifique ou d'une œuvre particulière. Les termes sont définis dans un environnement spécifique et incluent rarement des significations différentes ». Nous pouvons citer un exemple de glossaire regroupant l'ensemble des sigles utilisés dans le domaine de la Sécurité Sociale<sup>2</sup>. Le sigle « CCAS » fait référence à « Centre Communal d'Action Sociale ».

### 2.1.2.2 Nomenclature

Une nomenclature est une liste méthodique, systématique des objets, des éléments d'un ensemble. Il s'agit d'une des formes les plus simples d'organisation de la connaissance (Vandenbussche, 2011). La structuration d'une nomenclature a pour objectif

---

2. <http://www.annuaire-secu.com/glossaire.html>

l'exhaustivité dans un domaine donné. Nous pouvons citer deux exemples des nomenclatures utilisées dans le domaine de la médecine :

- La [Nomenclature Générale des Actes Professionnels \(NGAP\)](#)<sup>3</sup>, regroupe les dispositions générales et la liste des actes pris en charge par l'assurance maladie. Cette nomenclature permet de connaître, par exemple, les actes spécifiques que réalisent les masseurs-kinésithérapeutes dans le cadre de la rééducation et de la réadaptation fonctionnelle.
- La [Liste des Produits et Prestations \(LPP\)](#)<sup>4</sup> regroupe l'ensemble des dispositifs médicaux pour traitements et matériels d'aide à la vie, les aliments diététiques et articles pour pansement, les orthèses et prothèses externes, les dispositifs médicaux implantables et les véhicules pour handicapés physiques, bénéficiant d'un remboursement par l'Assurance Maladie. Cette nomenclature permet par exemple de connaître les montants de remboursement des aides techniques telles que les fauteuils roulants manuels.

### 2.1.2.3 Thésaurus

Le thésaurus est défini par [Vandenbussche \(2011\)](#) comme « un ensemble structuré de termes normalisés organisés au sein d'une hiérarchie de concepts liés par des relations sémantiques ».

[Charlet et al. \(2004\)](#) propose la définition suivante, « Un thésaurus est un ensemble de termes normalisés fondé sur une structuration hiérarchisée. Les termes y sont organisés de manière alphabétique et conceptuelle et reliés entre eux par des relations sémantiques. Organisés alphabétiquement, ils forment un répertoire alphabétique de termes normalisés pour l'analyse de contenu, le classement et donc l'indexation de documents d'information. »

Pour [Hodge \(2000\)](#) « Les thésaurus sont basés sur des concepts et montrent les relations entre les termes ». Les relations peuvent être hiérarchiques, d'équivalence, d'association ou de parenté.

Pour illustrer le thésaurus dans le domaine médical, nous pouvons citer le [Medical Subject Headings \(MeSH\)](#)<sup>5</sup> développé par la NLM (National Library of Medicine), utilisé pour indexer et interroger la base de donnée medline/PubMed. Il est traduit en français par l'INSERM et sert de thesaurus au CiSMéF.

### 2.1.2.4 Terminologie

Une terminologie est définie par [Charlet et al. \(2004\)](#), comme un « Ensemble des termes particuliers à une science, à un art, à un domaine. Les termes y sont également définis par un texte en langue naturelle et caractérisés par différentes propriétés

---

3. <https://www.ameli.fr/medecin/exercice-liberal/remuneration/nomenclatures-codage/ngap>

4. <https://www.ameli.fr/medecin/exercice-liberal/remuneration/nomenclatures-codage/liste-produits-prestations-lpp>

5. <http://mesh.inserm.fr/FrenchMesh/search/index.jsp>

linguistiques ou grammaticales suivant l'usage prévu de cette terminologie. Avec leur mise sur support informatique, les terminologies ont beaucoup évolué et sont parfois enrichies de relations entre termes, formant ainsi un réseau terminologique. »

Pour [Zweigenbaum \(1999\)](#) une terminologie s'intéresse à un domaine technique et a pour objectif l'échange d'information ; elle permet de recenser les concepts d'un domaine et les termes qui les désignent. Pour cela, un processus de normalisation des termes désignant les concepts est nécessaire, ainsi qu'un agencement spécifique, les concepts sont reliés entre eux par des relations.

On peut citer comme exemple de terminologie utilisée en médecine la SNOMED V3.5, qui contient 156 000 concepts.

### 2.1.2.5 Classification

La classification est l'une des méthodes d'organisation des connaissances. Elle est dépendante du processus de définition, permettant de déterminer ce qu'est ou n'est pas une entité ceci afin de permettre un partage. Une classification est définie par le Centre Nationale de Ressources Textuelles et Lexicales comme « une répartition systématique en classes, en catégories, d'être, de choses ou de notions ayant des caractères communs notamment afin d'en faciliter l'étude ; résultat de cette opération »<sup>6</sup>. Le processus de classification a donc pour objectif d'organiser des entités, il est basé sur deux principes importants que sont le regroupement (grouping) et le classement (ordering) ([Hudon & Hadi, 2010](#)). Le regroupement permet de rapprocher les entités qui ont des caractéristiques communes et le classement permet de les ordonner les unes par rapport aux autres. Nous pouvons donner un exemple des classifications utilisées dans le domaine de la médecine, en citant la [Classification Internationale des Maladies \(CIM\)-10](#), utilisée dans le monde pour le codage des maladies, traumatismes et l'ensemble des motifs de recours aux services de santé et causes de décès. Elle est publiée et mise à jour par l'[Organisation Mondiale de la Santé \(OMS\)](#).<sup>7</sup>

### 2.1.2.6 Les ontologies

Le terme « Ontologie » est originellement lié à un domaine de la philosophie initié par Aristote au XVII<sup>e</sup> siècle, qui a pour objectif de s'intéresser à la nature et à l'organisation de la réalité ([Guarino & Giaretta, 1995](#)), dans cette vision le Centre National de Ressource Textuelles et Lexicales (CNRTL) propose la définition suivante :« Partie de la philosophie qui a pour objet l'étude des propriétés les plus générales de l'être, telles que l'existence, la possibilité, la durée, le devenir. »<sup>8</sup>

Dans le cadre de l'Ingénierie des Connaissances et de l'Intelligence Artificielle de nombreuses définitions ont été proposées pour expliciter ce qu'est une ontologie infor-

---

6. <http://www.cnrtl.fr/definition/classification>

7. Une nouvelle version de la classification existe il s'agit de la CIM-11 <https://www.who.int/classifications/icd/en/>, publiée en juin 2018. Cette nouvelle version contient près de 55 000 codes, et devrait entrer en vigueur au 1<sup>er</sup> janvier 2022, après adoption par les états membres.

8. <http://www.cnrtl.fr/definition/ontologie>

matique. Nous avons choisi de présenter différentes définitions des ontologies et d'identifier les notions qu'elles partagent.

- La première définition est celle de **Gruber (1995)** qui propose comme définition : « an ontology is an explicit specification of a conceptualization. The term is borrowed from philosophy, where an ontology is a systematic account of Existence. For knowledge-based systems, what «exists» is exactly that which can be represented. »<sup>9</sup>
- Une ontologie est définie par **Blanc et al. (2010)** comme « un système formel dont l'objectif est de représenter les connaissances d'un domaine spécifique au moyen d'éléments de base, les concepts, définis et organisés les uns par rapport aux autres sous forme d'une taxonomie enrichie de relation associative. »
- Pour **Bachimont (2000)** « Définir une ontologie pour la représentation des connaissances, c'est définir, pour un domaine et un problème donné, la signature fonctionnelle et relationnelle d'un langage formel de représentation et la sémantique associée. »
- **Charlet (2002)** définit une ontologie comme « Ensemble des objets reconnus comme existant dans le domaine. Construire une ontologie c'est aussi décider de la manière d'être et d'exister des objets. »
- **Roche (2005)** propose la formulation suivante « Définie pour un objectif donné et un domaine particulier, une ontologie est pour l'ingénierie des connaissances une représentation d'une modélisation d'un domaine partagé par une communauté d'acteurs. Objet informatique défini à l'aide d'un formalisme de représentation, elle se compose principalement d'un ensemble de concepts définis en compréhension, de relations et de propriétés logiques ».
- **Vandenbussche (2011)** explique qu'une ontologie présente une vue d'un domaine particulier avec un but donné, c'est un artefact qui doit être opérationnel et compréhensible par une machine.

Ces différentes définitions mettent en avant les caractéristiques fondamentales des ontologies, qui doivent être explicites pour un humain et utilisable par une machine, en rendant compte des connaissances d'un domaine. En premier lieu, une ontologie est une conceptualisation, c'est-à-dire que l'on va y définir des concepts. Etant le reflet d'un point de vue (partiel) d'un domaine spécifique, les concepts doivent être partagés et faire l'objet d'un consensus entre les "experts du domaine". En outre, c'est une spécification formelle, devenant un artefact informatique. Elles doivent à ce titre posséder une sémantique formelle, une théorie logique pour laquelle le vocabulaire manipulé sera défini (**Aussenac-Gilles, 2005, Charlet, 2002**).

---

9. « Est une spécification explicite d'une conceptualisation. Le terme est emprunté à la philosophie, où une ontologie est une prise en compte systématique de l'Existence. Pour les systèmes à base de connaissances, ce qui « existe » est exactement ce qui peut être représenté ».

## 2.2 Les ontologies informatiques

Après avoir présenté les différentes définitions d'une ontologie informatique (dite ontologie formelle), nous allons présenter dans cette section les différents éléments qui la composent. Une ontologie sera constituée par des concepts ayant des propriétés, qui sont structurés et organisés par des relations, elle contiendra des instances et des axiomes (Gómez-Pérez, 1999). Il existe différents types d'ontologies, l'une des premières classifications prend en compte le niveau de formalisation.

### 2.2.1 Types d'ontologies

Les ontologies peuvent être classées différemment selon des critères comme leurs niveaux de modélisation ou bien leurs niveaux de granularité.

#### 2.2.1.1 Typologie d'ontologies selon le niveau de modélisation conceptuelle

Les ontologies peuvent être répertoriées en fonction de leur niveau de conceptualisation. Nous allons présenter certaines de ces ontologies. Gómez-Pérez (1999), Guarino (1997) et Roussey *et al.* (2011) proposent une classification des ontologies ayant des points de convergence et de divergence. La figure 2.1 de Guarino (1997) propose quatre types d'ontologies alors que Roussey *et al.* (2011) en figure 2.2 en distinguent cinq.

1. **Les ontologies de haut niveau, top-level ou upper-level ontologies**, décrivent des concepts qui vont être très généraux et à un haut niveau d'abstraction. Ces ontologies vont contenir les éléments relatifs au temps, à l'espace, à la matière, à l'objet (Zacklad, 2010). Elles devraient pouvoir être utilisées dans tous les domaines et être reliées au sommet des ontologies noyaux ou des ontologies du domaine. On peut citer l'ontologie Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE), qui décrit des éléments pouvant être un objet physique (endurant), des événements (perdurant), une qualité.
2. **Les ontologies génériques ou core-ontologies**, présentes dans le schéma proposé par Roussey *et al.* (2011) mais absentes dans le schéma proposé par Guarino (1997). Tout comme les ontologies de haut niveau, elles décrivent des concepts généraux indépendants d'un domaine, mais à un niveau d'abstraction moindre. L'objectif d'une ontologie noyau est d'être réutilisable et de permettre la construction d'un domaine à un niveau générique (Després & Szulman, 2009).
3. **Les ontologies de domaine** vont permettre de décrire les connaissances d'un domaine spécifique. Elles rendent compte du vocabulaire et du point de vue spécifique d'un domaine. Nous pouvons citer en exemple l'ontologie ONTOLURGENCES<sup>10</sup> pour la modélisation des urgences médicales, ou Bilingual Ontology of Alzheimer's Disease and Related Diseases<sup>11</sup> (Drame, 2014), spécifique à la maladie d'Alzheimer.

---

10. <https://bioportal.bioontology.org/ontologies/ONTOLURGENCES>

11. <https://bioportal.bioontology.org/ontologies/ONTOAD>

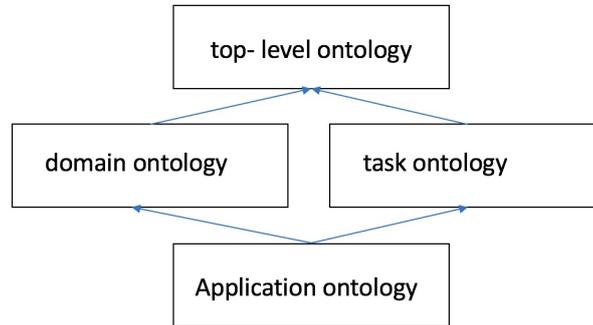


FIGURE 2.1 – Types d’ontologies, selon (Guarino, 1997)

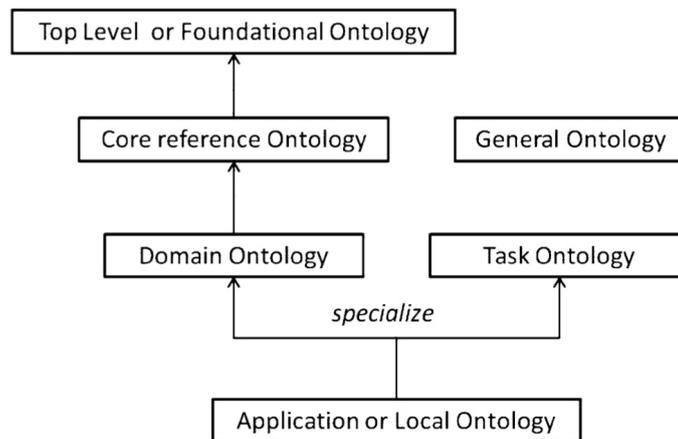


FIGURE 2.2 – Types d’ontologies, selon (Roussey et al., 2011)

4. **Les ontologies de tâche** contiennent des connaissances permettant d’accomplir une tâche.
5. **Les ontologies d’application** sont les plus spécifiques, elles décrivent à la fois les concepts qui dépendent d’un domaine et d’une tâche particulière pour une application donnée.

### 2.2.1.2 Typologie d’ontologie selon le niveau de granularité

Le second critère pouvant être utilisé pour caractériser une ontologie est son niveau de granularité qui peut être fin ou large.

**Ontologie de granularité fine :** correspond à une ontologie très détaillée, possédant un vocabulaire très riche capable de permettre une description détaillée des concepts pertinents d’un domaine ou d’une tâche.

**Ontologie de granularité large** : correspond à une ontologie ayant un vocabulaire moins détaillé. Les ontologies de haut niveaux ont une granularité large, car les notions qu'elles contiennent peuvent être précisées par des notions plus spécifiques.

Au vu de ces différentes descriptions de typologies d'ontologies, le travail mené dans cette thèse se situe dans le champ des ontologies noyau, de domaine et d'application. En effet ONTOPARON est : une ontologie d'application car elle est utilisée pour réaliser l'annotation sémantique des corpus, une ontologie de domaine car nous avons créé trois modules spécifiques de domaine particulier (médical, socio-environnemental et de coordination). Pour lier les différents modules ontologiques de domaines, nous avons utilisé une ontologie noyau OntoMénélas.

## 2.2.2 Les composants d'une ontologie

### 2.2.2.1 Concept, terme et instance

Le concept est défini par le dictionnaire Larousse comme une « idée générale et abstraite que se fait l'esprit humain d'un objet de pensée concret ou abstrait, et qui lui permet de rattacher à ce même objet les diverses perceptions qu'il en a, et d'en organiser les connaissances ». [Gómez-Pérez \(1999\)](#) propose comme définition « Concepts are used in a broad sense. They can be abstract or concret, elementary (electron) or composite (atom), real or fictitious. In short, a concept can be anything about which something is said and, therefore, could also be the description of a task, function, action, strategy, reasoning process etc. »<sup>12</sup> Ces deux définitions mettent en avant la notion que les concepts peuvent être abstraits (par exemple l'autonomie) ou bien des concepts concrets (par exemple un médicament).

[Ermine \(2000\)](#) propose une définition permettant de lier les notions de concepts, aux objets et aux attributs. Pour lui « un objet décrit une entité du monde réel. De même que cette entité est unique, un objet est unique : il possède donc une identité propre. Un concept désigne une catégorie d'objets ayant des propriétés communes. Un objet appartient forcément à un concept : on dit qu'il est une instance de ce concept. Les propriétés communes aux objets d'un concept sont représentées par des attributs. Les attributs permettent par exemple de modéliser que toute instance du concept voiture a une couleur, roule à une certaine vitesse, et possède 4 roues. Le concept voiture à comme attributs : couleur, vitesse... qui sont eux même des concepts ». Pour illustrer ces différentes notions, nous pouvons reprendre l'exemple du concept « voiture » utilisé par [Ermine \(2000\)](#). Lors de l'évocation du concept « voiture », il est possible d'avoir une représentation mentale, associant voiture à un tout constitué d'une carrosserie, de vitres, de roues et ayant la capacité de rouler. Pour évoquer et exprimer ce concept auprès d'un tiers, nous pouvons utiliser différentes formes linguistiques (qui sont les termes) comme « caisse », « véhicule », « bagnole », etc.

---

12. Les concepts sont utilisés dans un sens large. Ils peuvent être abstraits ou concrets, élémentaires (électron) ou composites (atome), réels ou fictifs. En bref, un concept peut être n'importe quoi sur lequel quelque chose est dit, par conséquent il pourrait également être la description d'une tâche, d'une fonction, d'une action, d'une stratégie, d'un processus de raisonnement, etc.

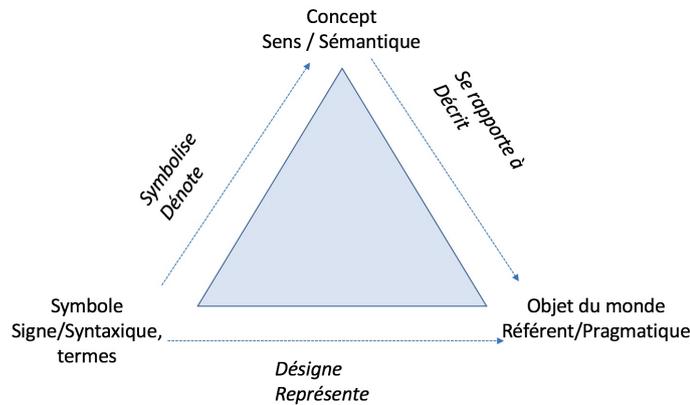


FIGURE 2.3 – Triangle sémiotique

Pour illustrer ces notions de concepts et de termes, il est possible d'utiliser le triangle sémiotique d'après (Ogden & Richards, 1923) comme indiqué à la figure 2.3.

#### 2.2.2.2 Les relations

Les relations sont importantes dans une ontologie, elles permettent de donner un sens aux concepts de l'ontologie. Plusieurs types de relations peuvent être utilisés dans une ontologie, des relations hiérarchiques et partitives ou bien des relations associatives ou transversales (Drame, 2014).

**La relation de subsomption** (ou d'hyponymie) *is-a; est un*, définit un lien de généralisation, elle est utilisée pour structurer l'ontologie. Cette relation permet l'héritage de propriétés, car elle va lier un concept supérieur (le concept le plus générique) à un concept inférieur (concept spécifique); elle doit être complétée par d'autres relations pour exprimer la sémantique du domaine (Charlet *et al.*, 2004). « La relation de subsomption organise les concepts par abstraction de caractères communs pour aboutir à une hiérarchie de concepts correspondant à une organisation taxonomique des objets. » (Roche, 2005). Par exemple, dans l'ontologie ONTOPARON, Le concept *Agent* subsume le concept *Agent médical* qui subsume le concept *Médecin*. Il est possible de dire que le concept *Médecin* spécialise le concept *Agent médical*, et que ce dernier est différent du concept *Agent De Coordination*.

**Les relations sémantiques** elles permettent d'énoncer des relations entre classes ou de leur attribuer des propriétés. Ainsi la relation **a\_pour\_couleur** pourra relier le concept de voiture à la couleur rouge.

**La relation de méronymie** *est une partie de*, est une relation sémantique importante utilisée dans les ontologies, et notamment en médecine concernant les parties du corps humain, ainsi il est possible d'exprimer que *Main* est une partie de

*Membre supérieur.* Il s'agit d'une relation qui lie un couple de concepts dont l'un fait partie de l'autre.

### 2.2.2.3 Les labels

Habituellement les ontologies utilisent un langage en OWL <sup>13</sup>, mais il est possible de compléter les propriétés de OWL, en utilisant en complément un langage comme SKOS <sup>14</sup>. Le SKOS est une recommandation du W3C pour représenter des thésaurus, classifications ou d'autres types de vocabulaire contrôlés comme les ontologies. Il s'appuie sur le modèle de données RDF, son principal objectif est de permettre la publication facile de vocabulaire structurés pour leur utilisation dans le cadre du Web sémantique. Les représentations conceptuelles réalisées à l'aide de SKOS peuvent être utilisées par des systèmes d'information fermés, mais aussi, dans la perspective du web sémantique, être publiées sur le web et alignées avec d'autres systèmes d'organisation de connaissances, de façon à rendre ces différents langages plus interopérables. Dans notre cas d'usage, nous avons intégré et défini certaines propriétés en SKOS afin de définir des termes préférés : `prefLabel` en français (fr) et en anglais (en) qui seront uniques et les autres termes (synonymes, abréviations, acronymes etc.) en `altLabel` comme illustré par l'exemple médecin traitant :

- `skos:prefLabel` "médecin traitant" fr
- `skos:prefLabel` "doctor" en
- `skos:altLabel` "med tt" fr
- `skos:altLabel` "médecin de famille" fr
- `skos:altLabel` "MDT" fr
- `skos:prefLabel` "general practitioner" en

### 2.2.2.4 Le langage des ontologies

Les ontologies sont généralement décrites en OWL2, standardisé sous l'égide du W3C <sup>15</sup>. A OWL2, correspond plusieurs sous-dialectes que nous ne développerons pas ici <sup>16</sup>. Quand les chercheurs veulent exprimer une ontologie sur laquelle on puisse faire des raisonnements efficaces et simples, ils utilisent en première intention le sous-dialecte OWL-DL. C'est lui qui permet de faire fonctionner les « raisonneurs » utiles à la validation formelle des ontologies et à la classification automatique des concepts primitifs sous les concepts définis (Cf. annexe D).

---

13. Web Ontology Language [https://fr.wikipedia.org/wiki/Web\\_Ontology\\_Language](https://fr.wikipedia.org/wiki/Web_Ontology_Language)

14. Simple Knowledge Organization System [https://fr.wikipedia.org/wiki/Simple\\_Knowledge\\_Organization\\_System](https://fr.wikipedia.org/wiki/Simple_Knowledge_Organization_System)

15. <https://www.w3.org/TR/2012/REC-owl2-overview-20121211/>

16. [https://fr.wikipedia.org/wiki/Web\\_Ontology\\_Language](https://fr.wikipedia.org/wiki/Web_Ontology_Language), (Richard, 2017)

## 2.3 Les ressources termino-ontologiques (RTO) dans le domaine de la neurologie, de la coordination et du handicap

Le domaine médical possède un grand nombre de ressources terminologiques et ontologiques, construites pour répondre à des objectifs et à des besoins différents. Nos travaux portant sur la SLA, sur la coordination de soin (prenant en compte la dimension sociale) et le handicap, nous avons souhaité voir quelles étaient les ressources existantes dans ces domaines.

### 2.3.1 Les ressources en neurologie

Des ontologies ont été développées dans le domaine de la pathologie neurologique comme l'ontologie de la maladie d'Alzheimer ONTOAD ([Dramé et al., 2014](#)), la neurological disease ontology (ND) ([Jensen et al., 2013](#)), ou bien encore une ontologie pour la maladie de Parkinson (PDON) ([Younesi et al., 2015](#)).

#### 2.3.1.1 L'ontologie pour la maladie d'Alzheimer

ONTOAD<sup>17</sup> est une ontologie bilingue (anglais ; français) constitué de 5765 concepts. Elle concerne la maladie d'Alzheimer et définit des termes très spécifiques du domaine biomédical, comme la « Photoradiation », les « Facteurs de risques » mais aussi des éléments du domaine clinique comme le « comportement d'une personne ». Si les concepts faisant référence à des processus physiopathologique sont absents dans ONTOPARON, il existe cependant des points communs entre les deux ontologies, comme nous le verrons dans la partie 5.

#### 2.3.1.2 L'ontologie de la maladie de Parkinson

L'ontologie pour la maladie de Parkinson<sup>18</sup> (PDON) a pour objectif de modéliser les caractéristiques moléculaires et cliniques de la maladie de Parkinson (MP). Elle compte 632 concepts. Les auteurs et les experts du domaine ont fait le choix de structurer l'ontologie en 9 "super-classes" :

- Les aspects cliniques, qui se composent des concepts de « Diagnostic » (regroupant les éléments pouvant intervenir dans l'élaboration du diagnostic de la pathologie comme les biomarqueurs, et les éléments de diagnostic différentiel), de « Symptomatologie » (présentant les caractéristiques motrices et non motrices spécifiques de la maladie) et de « Traitement » (contenant les aspects liés aux différents traitements pouvant être proposés aux patients).
- L'étiologie de la maladie de Parkinson, décrivant les facteurs génétiques et environnementaux connus de la maladie.

---

17. <https://bioportal.bioontology.org/ontologies/ONTOAD>

18. <https://bioportal.bioontology.org/ontologies/PDON>

- Modèle de la maladie de Parkinson qui contient les divers modèles de maladies in vivo et in vitro utilisés dans la recherche sur la maladie de Parkinson.
- Neuropathologie qui décrit les mécanismes de la pathologie.
- Maladie neuro-dégénérative familiale] qui contient les informations spécifique aux troubles héréditaires cliniquement associés à la MP, tels que les maladies de Huntington et de Wilson.

Cette ontologie est utilisée principalement pour la recherche et l'annotation concernant l'aspect génétique. Les dimensions de coordination de parcours et de dimensions sociale prenant en compte les structures et les allocations spécifiques en France sont absentes.

### 2.3.1.3 L'ontologie des maladies neurologiques

La Neurological Disease Ontology (ND) est constituée de 450 termes, mais elle importe de nombreux termes (700) issus d'autres ontologies. Elle a pour objectif de représenter « avec exactitude pour chaque maladie ses origines moléculaire, génétique et environnementale, les processus impliqués dans son étiologie et son évolution, ainsi que sa présentation clinique et ses phénotypes, y compris les signes associés, les symptômes, les syndromes, les critères de diagnostic, le traitement et les méthodes d'essai ». Son développement a pour objectif l'annotation et l'analyse de données présentes dans les dossiers pour des patients présentant une maladie d'Alzheimer, une sclérose en plaque ou un accident vasculaire cérébrale. La ND utilise comme top ontologie la Basic Formal Ontology (BFO).

### 2.3.2 Une ontologie de la coordination de soins infirmiers

La Nurse Care Coordination Ontology (NCCO)<sup>19</sup> est une ontologie de coordination de soins, constituée de 394 concepts. Elle permet de quantifier les actes et les types de coordination délivrés aux patients âgés (Popejoy *et al.*, 2015), dans le cadre d'une étude menée aux Etats Unis. L'ontologie est structurée en 5 classes principales qui sont :

- les activités de coordination, regroupant notamment la communication et l'évaluation, etc. ;
- les objets de coordination, c'est-à-dire les thématiques sur lesquelles interviennent les infirmières, comme le soin, le suivi, etc. ;
- les personnes, qui interviennent dans le soin des patients, comme les intervenants de premiers recours, les paramédicaux, etc. ;
- les problèmes, qui permettent d'identifier les thématiques des problèmes rencontrés comme les problèmes moteurs, ou les problèmes cognitifs ,etc. ;
- les lieux, qui correspondent aux lieux où les infirmières peuvent intervenir.

Cette ontologie présente des points communs avec notre sujet d'étude, elle nous a aidé à structurer une partie des concepts des activités de coordination dans la module de coordination de ONTOPARON, comme nous le verrons dans la partie 5.

---

19. <https://bioportal.bioontology.org/ontologies/NCCO>

### 2.3.3 La Classification Internationale du Fonctionnement du Handicap et de la Santé (CIF)

La [Classification Internationale du Fonctionnement, du Handicap et de la Santé \(CIF\)](#)<sup>20</sup> a été élaborée par l'OMS pour fournir un langage uniformisé et un cadre pour la description et l'organisation des informations relatives au fonctionnement et au handicap. Le principe de la [CIF](#) est de permettre d'identifier les interactions entre les différents composants pouvant créer des situations de handicap :

- les fonctions organiques et les structures anatomiques des individus ;
- les activités et les domaines de la vie dans lesquels participent les individus ;
- les facteurs environnementaux qui peuvent influencer la participation sociale des personnes ;
- les facteurs personnels qui sont les facteurs propres de la personnes.

L'utilisation de la [CIF](#) nécessite le recours à des codes qualitatifs qui permettent de rendre compte des déficiences, des limitations de participation ou bien encore d'identifier les facteurs facilitateurs ou bien les obstacles et la situation de handicap qui en résulte. Nous nous sommes inspirés de la [CIF](#) pour organiser certaines classes de ONTOPARON, notamment dans le domaine des activités de vie quotidienne et des aides techniques.

Comme présenté dans ce chapitre, il existe différents systèmes d'organisation des connaissances, chacun répondant à des définitions et des objectifs d'utilisation différents. Nous avons fait le choix, pour exploiter les données textuelles présentes dans les corpus du Réseau SLA IDF, de prendre comme système de représentation des connaissances une ontologie formelle. Le chapitre suivant va aborder les méthodes de construction d'ontologie.

---

20. <https://www.who.int/classifications/icf/en/>



# Chapitre 3

## Construction d'ontologies informatiques

### Sommaire

---

<b>3.1</b>	<b>Les principes de construction d'une ontologie</b>	<b>58</b>
<b>3.2</b>	<b>Méthode de construction d'ontologies</b>	<b>59</b>
3.2.1	Développement d'ontologies selon <b>Uschold &amp; King (1995)</b>	59
3.2.2	Méthode Methontology proposé par <b>López (1999)</b>	59
3.2.3	Construction d'ontologie à partir de textes	59
3.2.4	Methode ARCHONTE	60
<b>3.3</b>	<b>Définition de la modularité d'une ontologie</b>	<b>62</b>
<b>3.4</b>	<b>Objectif de la modularité ontologique</b>	<b>62</b>
<b>3.5</b>	<b>Construire une ontologie modulaire</b>	<b>63</b>
<b>3.6</b>	<b>Synthèse</b>	<b>65</b>

---

Dans la partie précédente nous avons présenté, certains *SOC*, et avons décrit les spécificités des ontologies. Pour rappel une ontologie est un objet informatique résultant d'une modélisation des connaissances dans un domaine spécifique, qui est créé pour répondre à un besoin et utilisée dans un objectif. Dans cette partie nous allons exposer les méthodologies existantes pour la construction d'ontologies.

La construction d'une ontologie résulte d'un processus de trois étapes principales que sont : l'acquisition de connaissances, la modélisation et la représentation des connaissances en un système formel. Différentes méthodes de construction existent, nous avons choisi d'utiliser la méthode de construction à partir de texte. La construction d'une ontologie doit répondre à des critères et des principes de construction que nous présenterons. L'ontologie créée peut être structurée de deux façons, en un seul bloc elle sera alors monolithique ou bien constituée de modules c'est une ontologie modulaire. Si initialement nous étions parti sur la construction d'une ontologie monolithique, rapidement après la modélisation nous nous sommes orientés vers une structuration modulaire. La dernière section de ce chapitre présente, les définitions et objectifs des ontologies modulaires et les modalités d'importation permettant d'obtenir une ontologie.

### 3.1 Les principes de construction d'une ontologie

Charlet (2002) et Gómez-Pérez (1999) rappellent quelques principes de construction d'une ontologie basés sur différents travaux, dont ceux de Gruber.

**La clarté, l'objectivité et la complétude :** l'ontologie doit fournir la signification de termes définis en donnant des définitions objectives ainsi qu'une documentation en langage naturel. Des définitions complètes exprimées par une condition nécessaire et suffisante doivent être préférées à des définitions partielles.

**La cohérence :** les axiomes de l'ontologie doivent être consistants. La cohérence des définitions en langage naturel doit être vérifiée autant que possible.

**L'extensibilité :** il doit être possible d'ajouter de nouveaux termes généraux ou spécialisés sans que cela remette en cause ce qui a déjà été fait.

**Un biais d'encodage minimal :** l'ontologie doit être conceptualisée indépendamment de tout langage d'implémentation, afin de permettre un partage de connaissances entre différentes applications.

**Un engagement ontologique minimal :** l'ontologie construite doit faire a minima des hypothèses sur le monde modélisé. Cependant ce principe peut être remis en perspective lors de la réalisation d'ontologies de domaine, car elle est et doit être spécifique à un domaine qui nécessite de faire des hypothèses.

**Principe de distinction ontologique :** qui signifie que les classes d'une ontologie doivent être disjointes.

**Minimiser la distance sémantique entre les concepts de la fratrie :** les concepts similaires sont regroupés et représentés sous forme de sous-classe et doivent être définis à l'aide des mêmes primitives, ceux qui sont plus distants doivent être représentés dans une autre part de la hiérarchie. Ce principe est en lien avec les principes différentiels proposés par Bachimont (2000).

## 3.2 Méthode de construction d'ontologies

### 3.2.1 Développement d'ontologies selon **Uschold & King (1995)**

Pour les auteurs, le développement d'une ontologie doit suivre quatre étapes principales :

**Identifier les objectifs**, c'est à-dire déterminer l'objectif pour lequel l'ontologie est créée, et par quels utilisateurs elle sera utilisée.

**Construire l'ontologie** : pour cela, il est nécessaire de capturer les connaissances du domaine en identifiant les concepts clés et leur relations, coder l'ontologie (c'est à-dire créer une représentation explicite de la conceptualisation réalisée à l'étape précédente) et intégrer des ontologies existantes.

**Evaluer** les ontologies dans leurs différentes dimensions.

**Documenter**, ce qui permet d'explicitier le but de l'ontologie et son contexte de création.

### 3.2.2 Méthode Methontology proposé par **López (1999)**

Methontology proposée par **López (1999)**, est développée dans le Laboratoire d'Intelligence Artificielle de l'Université Polytechnique de Madrid. La méthode proposée identifie trois types d'activités principales afin de guider l'utilisateur dans le développement et la manutention de l'ontologie. Les trois activités principales sont ainsi décrites, « Les activités de gestion de projet comprennent : la planification, le contrôle et l'assurance qualité ; Les activités axées sur le développement comprennent : la spécification, la conceptualisation, la formalisation, la mise en œuvre et la maintenance ; Les activités de support incluent : acquisition des connaissances, intégration, évaluation, gestion de la documentation et des configurations. Le cycle de vie de l'ontologie identifie l'ensemble des étapes de son déplacement au cours de sa vie, décrit les activités à exécuter à chaque étape et leur relation » **López et al. (2000)**.

### 3.2.3 Construction d'ontologie à partir de textes

Le paradigme des méthodologies de construction d'ontologie à partir de textes, se base sur le postulat que la connaissance d'un domaine va se situer dans le corpus utilisé et/ou rédigé par les experts du domaine. « Les textes fournissent des éléments stables, consensuels et partagés d'un domaine, comme des descriptions d'objets et de concepts tels qu'on peut en avoir besoin pour former un modèle du domaine ou une ontologie. » **Aussenac-Gilles (2005)**. Les connaissances s'expriment en langage naturel, pour **Zweigenbaum (1999)** la mise au point d'ontologies doit s'appuyer sur la linguistique. Dans le domaine médical, il est nécessaire d'utiliser les ressources tel que les comptes-rendus médicaux ou les articles scientifiques, qui vont rendre compte des notions manipulées par les acteurs de la médecine. Ces méthodes basées sur l'acquisition de termes à partir de ressources textuelles nécessitent l'utilisation d'outils du **Traitement Automatique de la langue Naturelle (TALN)**.

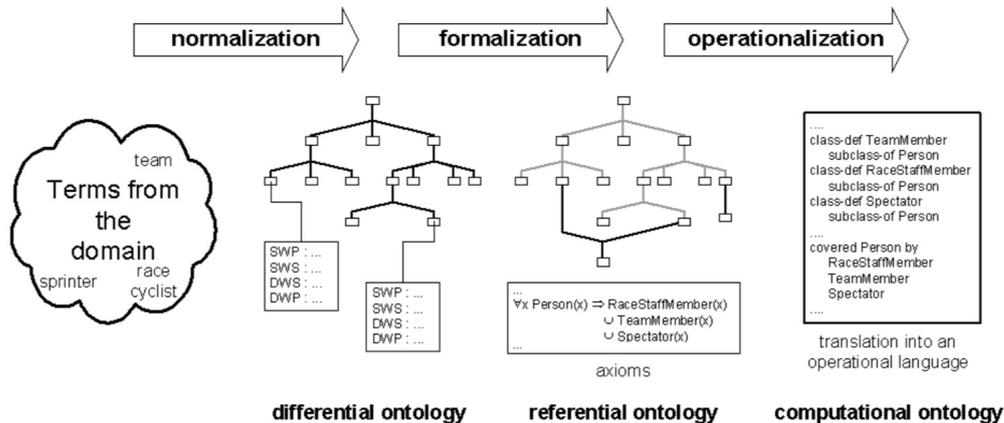


FIGURE 3.1 – Étapes de construction d'une ontologie selon Bachimont *et al.* (2002)

### 3.2.4 Methode ARCHONTE

La méthode ARCHONTE (ARCHitecture for ONTological Elaborating) fut décrite initialement par Bachimont (2000), Bachimont *et al.* (2002) comme une séquence de trois étapes indiquées à la figure 3.1. La première étape consiste en une phase de normalisation devant aboutir à une « ontologie différentielle », la seconde phase est la formalisation permettant d'obtenir une « ontologie référentielle », et la phase d'opérationnalisation permet d'obtenir une « ontologie computationnelle ». L'objectif d'une ontologie est de formaliser et conceptualiser les connaissances d'un domaine. Pour cela il est nécessaire « de passer de l'expression linguistique des connaissances telles que nous autres, humains, pouvons la considérer à une représentation formelle et calculable des connaissances propre à une exploitation informatique. » (Bachimont, 2000). Charlet *et al.* (2006) ont ultérieurement enrichi le modèle et le décrivent en quatre étapes.

**Constitution d'un corpus :** comme l'indique Bachimont (2000) et Bourigault & Aussenac-Gilles (2003), la constitution d'un corpus est une tâche délicate et primordiale qui conditionne fortement le type et la nature des traitements que l'on peut effectuer. Le choix d'un corpus va induire des biais qu'il est parfois difficile d'identifier. Aimé (2015) ajoute que le « corpus doit être suffisamment large pour couvrir tout le domaine, et consensuel pour répondre à l'objectif d'une ontologie qui est – par définition – la formalisation d'une conceptualisation consensuelle ». Une fois le corpus constitué, des outils du TALN tel que BioTex (Lossio-Ventura *et al.*, 2014b) sont utilisés pour extraire les candidats termes. Les termes (syntagmes nominaux pouvant être composés d'une à quatre unités) vont être classés en fonction de leur fréquence d'apparition dans le corpus. Il est alors nécessaire de sélectionner les termes pertinents. Cette sélection peut se faire sur un critère quantitatif en prenant uniquement en compte la fréquence, ou bien par l'expertise de l'ontologue et/ou de l'expert, en décidant de prendre des termes ayant une fréquence faible mais étant essentiels au domaine (Aimé, 2015).

**La normalisation sémantique**, ou engagement sémantique a pour objectif de fixer le sens du concept dans l'ontologie. L'étape précédente fournit des listes de candidats termes qui vont avoir un sens pour l'expert du domaine, cependant ces libellés peuvent présenter des ambiguïtés. La première phase sera de regrouper les candidats termes dénotant un même concept. Nous pouvons citer comme exemple dans notre corpus « l'IDE, l'infirmière, l'infirmière libérale, l'IDEC ». L'ontologie doit modéliser des connaissances ne présentant pas d'ambiguïté. Il est donc nécessaire de les formaliser, afin que les concepts choisis n'aient qu'une seule signification possible. Pour illustrer une des ambiguïtés rencontrée dans le corpus, nous utilisons le terme « kiné » : selon le contexte, ce terme peut prendre deux acceptations et faire référence : *a*) à l'activité d'un professionnel, « la kinésithérapie » ou *b*) désigner le professionnel, « le kinésithérapeute ». La compréhension et la définition de ces deux termes ne sont fondamentalement pas les mêmes. **Bachimont (2000)** rappelle qu'il est nécessaire de contraindre l'interprétation des libellés afin que tous les spécialistes associent la même signification au libellé ; si tel n'est pas le cas, le libellé ne peut être considéré comme un concept et ne correspond pas aux primitives recherchées.

La normalisation sémantique nécessite de spécifier les différences et caractéristiques communes, pour cela des principes différentiels peuvent être utilisés comme proposé par **Bachimont (2000)** :

1. Le principe de communauté avec le père : toute unité se détermine par l'identité qu'elle possède avec l'unité parente.
2. Le principe de différence avec le père : toute unité enfant se distingue de l'unité parente, sinon il n'y aurait pas lieu de la définir.
3. Le principe de différence avec les frères : toute unité se distingue de ses frères, sinon il n'y aurait pas lieu de la définir. Il faut donc expliciter la différence de l'unité avec chacune de ses unités soeurs.

A partir de cette phase, nous obtenons un arbre ontologique, contenant les primitives conceptuelles valables dans le domaine modélisé, dans une ontologie différentielle.

**L'engagement ontologique** est la formalisation des connaissances, c'est-à-dire le passage à une sémantique formelle. L'une des tâches sera de définir les relations entre les concepts en décidant des domaines (sujets) et co-domaines (objets) des relations. Dans cette formalisation, il est possible de construire des représentations conceptuelles plus complexes que sont les concepts définis (condition suffisante et nécessaire), en combinant des concepts primitifs (**Charlet, 2002**).

**L'opérationnalisation** consiste en la représentation de l'ontologie dans un langage de représentation des connaissances permettant le calcul et les inférences.

Cette méthode a été utilisée dans de nombreux cas d'usage aboutissant au développement de différentes ontologies comme ONTOPNEUMO dans le cadre de la pneumologie (**Baneyx, 2007**), ONTOLURGENCES dans le domaine des services d'urgences, ONTODPN (ontologie du diagnostic prénatal) dans le domaine de la médecine préna-

tale (Dhombres *et al.*, 2010) ou encore ONTOPSYCHIA dans le cadre de la psychiatrie (Richard, 2017).

### 3.3 Définition de la modularité d'une ontologie

Si des ontologies peuvent être construites en un seul composant, il est également possible de construire une ontologie à partir de modules ontologiques. La modularité peut être choisie pour différentes raisons comme la facilitation de gestion (les modules ayant des tailles plus petites), la spécification du domaine (chaque module étant spécifique d'un domaine).

Pathak *et al.* (2009) définissent une ontologie modulaire comme un ensemble de modules qui sont des « composants réutilisables d'une ontologie plus grande ou plus complexe, qui est autonome mais qui présente une association définie avec d'autres modules d'ontologie, y compris l'ontologie originale ».

Pour Grau *et al.* (2006), « les ontologies modulaires (1) facilitent la réutilisation de connaissances sur de multiples applications, (2) sont faciles à construire, à maintenir et à modifier, (3) permettent une ingénierie distribuée des modules sur différents champs d'expertise et (4) permettent une gestion et une navigation efficace dans les modules ». Atrash *et al.* (2014) ajoutent que la modularité facilite la gestion de l'ontologie et en diminue la complexité, chaque module d'ontologie étant un composant autonome, et l'ontologie qui intègre les modules doit posséder les liens pour exploiter les modules. Selon Bao et Honavar (2006), « comme les ontologies sont conçues pour des domaines spécifiques, des applications, ou des utilisateurs, elles nécessitent souvent des adaptations importantes avant de pouvoir être déployées avec succès dans des environnements connexes et proches. Il y a donc un besoin urgent de favoriser une réutilisation sélective et indépendante de modules au sein de ces ontologies. »

Les critères de taille, de facilité d'utilisation, de modification, de gestion et de spécification des champs d'expertise par module, ont orienté notre choix vers la construction d'une ontologie modulaire.

### 3.4 Objectif de la modularité ontologique

Menken *et al.* (2005) définissent les objectifs de la modularité ontologique de la manière suivante :

1. **L'évolutivité** est liée à la performance d'utilisation de l'ontologie ; un module de taille plus petit doit faciliter et améliorer l'utilisation de raisonneurs et ainsi améliorer la recherche d'information lors de requêtes.
2. **La gestion de la complexité** : par la décomposition en module, il doit être plus aisé à l'être humain de maîtriser la construction et l'exactitude des connaissances de l'ontologie, car il peut en avoir une vision totale.
3. **La compréhension** est un critère important, lors de l'utilisation d'une ontologie, un être humain doit pouvoir comprendre et appréhender les connaissances de

l'ontologie. Si le contenu d'une ontologie est de taille réduite, la compréhension par un tiers sera plus efficace et rapide.

4. **La personnalisation** : la modularité permet d'utiliser des modules existants et de les spécifier à l'objectif et à l'environnement choisis.
5. **La réutilisation** est un en lien avec les autres critères cités ci-dessus ; ainsi, plus un module sera compréhensible, personnalisable, plus il sera susceptible d'être utilisé dans d'autres ontologies ou applications.

Les objectifs de modularité sont en phase et complémentaires aux principes de construction d'une ontologie énoncés plus haut dans ce chapitre. Si la modularité ontologique peut avoir de nombreux avantages, elle nécessite cependant de suivre des processus spécifiques lors de l'importation des modules pour la création de l'ontologie finale.

### 3.5 Construire une ontologie modulaire

En pratique, quand on veut construire une ontologie modulaire, on se trouve devant trois cas :

1. On part de zéro (pas d'ontologie monolithique ou modulaire déjà construite) et on décide de construire une ontologie modulaire.
2. On a déjà une ontologie monolithique et on va décomposer cette ontologie en différents modules.
3. On est déjà dans un contexte d'ontologie modulaire. En conséquence, le module à construire doit s'inscrire parfaitement dans ce contexte. Et l'ontologue se doit de maîtriser ce contexte, à savoir *a*) connaître les autres modules, *i.e.* les autres ontologies, *b*) y compris les ontologies importées par les autres modules.

Dans notre travail, nous étions dans le contexte où nous avons élaboré une ontologie monolithique et nous souhaitions passer à une ontologie modulaire. Aux modules du domaine créés, il est nécessaire d'ajouter d'autres modules pour obtenir l'ontologie finale.

La construction modulaire s'opère sur trois types de modules : (1) un module top-core, (2) des modules spécifiques et (3) un module de consolidation. Comme suggéré dans [Aimé et al. \(2015\)](#), chaque module répond à un certain nombre de contraintes parmi lesquelles nous pouvons citer :

- tout module de consolidation possède son propre espace de nommage ;
- tout module de consolidation répond à une problématique (*i.e.* un besoin de modélisation d'un point de vue donné) ;
- tout module de consolidation ne peut hériter que d'un ou de plusieurs modules spécifiques, ou (non exclusif) d'un ou de plusieurs modules de consolidation ;
- toute relation d'objet d'un module de consolidation a pour domaine et co-domaine des concepts dans des modules différents (sinon elles doivent être créées au sein du module concerné).

Chaque module spécifique hérite directement de tous les concepts et relations de l'ontologie noyau, et indirectement de la top-ontologie si elle est définie. Chaque module spécifique exprime un point de vue ou un sous-domaine.

Le module de consolidation permet la spécialisation de concepts en provenance de modules spécifiques par l'ajout de labels, de relations avec des concepts ou d'attributs. Le module de consolidation permet de former des ensembles logiques et cohérents de modules spécifiques ou de consolidation en fonction de leur finalité d'usage. Ce jeu de modules consolidés peut ensuite être assemblé, permettant d'avoir une grande souplesse dans la diffusion de la base de connaissances. D'autre part, chaque modification dans un module spécifique (ou même de consolidation pour peu qu'il soit importé dans un autre module de consolidation) est automatiquement reportée dans l'ensemble des modules héritant directement ou indirectement (par transitivité).

**Ben Abbès *et al.* (2012)** ont étudié les caractéristiques de la modularité et ont proposé de classer la modularisation en quatre types de patrons :

1. 1 module important n modules,
2. n modules important 1 module,
3. n modules important n-1 modules,
4. un mixte de tous ces patrons.

La figure 3.2 précise les deux premiers patrons pour les importations. Ils sont les plus importants pour la construction modulaire puisqu'ils représentent, respectivement la modularité de l'ontologie via l'agrégation de plusieurs modules et l'héritage d'un module correspondant la plupart du temps à une ontologie noyau.

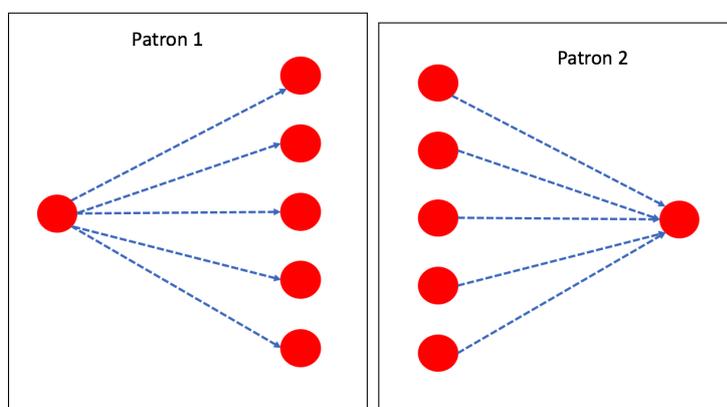


FIGURE 3.2 – Premier et second patrons de type d'importation de modules. Le premier patron correspond à une action d'agrégation où l'on construirait une ontologie des animaux en agrégeant des ontologies des mammifères, des oiseaux, etc. Le second schéma correspond à une action d'héritage où n modules se partagent les concepts d'un module plus général. C'est typique des actions de partage d'ontologie noyau. Schémas tirés de **Ben Abbès *et al.* (2012)**.

Dans le travail que nous avons mené nous avons choisi d'importer dans chaque module uniquement des modules sémantiquement plus généraux, représentant le patron 2 de la figure 3.2. Même si, techniquement, un éditeur comme Protégé autorise n'importe

quel import – quel que soit le niveau de granularité sémantique. Nous estimons en effet que l’importation d’un module se calque sur la notion d’héritage, telle que développée dans la programmation orientée objet par exemple. Ainsi une commande import dans une ontologie se comporte comme l’instruction « extends » dans une classe en Java. Le module hérite de l’ensemble des concepts et relations qu’il va spécifier (et non l’inverse).

## 3.6 Synthèse

Ce chapitre a abordé les différentes méthodes de construction d’ontologie et en particulier la méthode à partir de textes, c’est cette méthode que nous avons utilisé pour le développement de ONTOPARON. Nous avons présenté les ontologies modulaires et leurs nombreux avantages, notamment du fait de la taille des modules, qui permettent une gestion et un partage plus aisé. Cependant, la modularité ontologique impose aussi des contraintes, comme la présence d’un module noyau et d’un module de consolidation, nous aborderons ces aspects dans le chapitre 5 lors de la présentation de la construction de l’ontologie ONTOPARON. La prochaine partie de ce manuscrit va aborder les différents travaux réalisés au cours du travail de thèse, le pré-traitement des corpus (pseudonymisation, correction orthographique), la construction de l’ontologie ainsi que la présentation des outils développés, dont un pour l’annotation sémantique des corpus (ONBASAM) et l’autre pour l’évaluation des annotations (PRONTO).



## Deuxième partie

### Travaux réalisés / Contribution Pratique



Chapitre **4**

# Traitements Informatiques du corpus

## Sommaire

---

<b>4.1</b>	<b>Aspect législatif du traitement des données . . . . .</b>	<b>70</b>
<b>4.2</b>	<b>Anonymisation et pseudonymisation . . . . .</b>	<b>71</b>
<b>4.3</b>	<b>Les méthodes d’anonymisation . . . . .</b>	<b>72</b>
<b>4.4</b>	<b>Traitement du corpus pour la pseudonymisation . . . . .</b>	<b>73</b>
4.4.1	Préparation des données . . . . .	74
4.4.2	Utilisation de méthode à base d’apprentissage statistique . . .	76
4.4.3	Utilisation de règles . . . . .	78
4.4.4	Traitement des dates . . . . .	79
4.4.5	Résultat et discussion de la pseudonymisation . . . . .	79
<b>4.5</b>	<b>Traitement Orthographique . . . . .</b>	<b>80</b>
4.5.1	Types d’erreurs rencontrées . . . . .	80
4.5.2	Pipeline de correction orthographique : Spell Checking . . . .	81
<b>4.6</b>	<b>Synthèse . . . . .</b>	<b>84</b>

---

Nous avons décrit dans la partie 1.8.3 (p.35) la base de données événementielle du Réseau SLA IDF, les données extraites ne peuvent être utilisées à l'état brut. Les événements sont nominatifs, ils contiennent de nombreuses données personnelles (nom, prénom, lieux, date, etc.) ainsi que des données de santé (diagnostic, signes cliniques, etc.). Il est donc nécessaire de réaliser un traitement informatique préalable de pseudonymisation, pour retirer les informations permettant d'identifier les personnes. Pour cela nous avons utilisé des méthodes symboliques à base de lexiques et de règles et des méthodes à base d'apprentissage statistiques. Pour pseudonymiser nous avons fait le choix de transformer les données nominatives en données sémantiques indiquant les fonctions des personnes, « Mr Jean » est transformé en « Patient » et « Mme Jean » en « Entourage ».

Pour extraire de l'information à partir des corpus et comprendre les éléments intervenants au domicile des patients et dans les parcours patients, nous avons fait le choix d'annoter les corpus avec une ontologie du domaine ONTOPARON. Cependant si les données saisies en texte libre dans les corpus, présentent des erreurs de ponctuation, d'orthographe, l'annotation sémantique sera moins efficace générant une perte d'information. Pour limiter, l'incidence des erreurs d'orthographe lors de la phase d'annotation sémantique, nous avons mis en place une chaîne de traitement permettant de corriger les erreurs d'orthographe tout en prenant en compte les spécificités du corpus du Réseau SLA IDF. En effet les coordinateurs utilisent de nombreuses d'abréviations, ou termes spécifiques (marque et modèle d'aides techniques) qu'il est nécessaire de prendre en compte. Nous présentons dans une première section le travail réalisé pour garantir la pseudonymisation du corpus. La seconde section présente les outils et la méthodologie utilisés pour le traitement des fautes orthographiques présentes dans les corpus et les résultats obtenus.

### 4.1 Cadre législatif du traitement des données personnelles et données de santé pour l'anonymisation

Le 25 mai 2018 est entré en application le règlement n°2016/679 du parlement européen et du conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données<sup>1</sup>. Ce règlement de l'Union européenne, dénommé [Règlement Général sur la Protection des Données \(RGPD\)](#) (ou en anglais GDPR pour General Data Protection Regulation), est le texte de référence en matière de protection des données à caractère personnel. Il s'inscrit dans la continuité de la Loi Informatique et Libertés de 1978<sup>2</sup> consolidé par la loi n° 2018-493 du 20 juin 2018 relative à la protection des données personnelles<sup>3</sup>.

---

1. abrogeant la directive 95/46/CE (règlement général sur la protection des données)

2. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT00000886460>

3. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000037085952>

Une donnée à caractère personnel correspond à toute information se rapportant à une personne physique identifiée ou identifiable ; est réputée être « une personne physique identifiable » une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale.<sup>4</sup> La [Commission Nationale de l'Informatique et des Libertés \(CNIL\)](#) définit les données personnelles comme : « Toute information identifiant directement ou indirectement une personne physique (nom, numéro d'immatriculation, numéro de téléphone, date de naissance...). »

La [CNIL](#) et le [RGPD](#)<sup>5</sup> définissent les données de santé comme : « Les données à caractère personnel concernant la santé sont les données relatives à la santé physique ou mentale, passée, présente ou future, d'une personne physique (y compris la prestation de services de soins de santé) qui révèlent des informations sur l'état de santé de cette personne. » Cette définition des données de santé inclus différents types d'informations :

1. « Les informations relatives à une personne physique collectées lors de son inscription en vue de bénéficier de services de soins de santé ou lors de prestation de ces services. »
2. « Les informations obtenues lors du test ou de l'examen d'une partie du corps. »
3. « Les informations concernant une maladie, un handicap, les antécédents médicaux. »

Au regard de ces définitions la base de données du Réseau SLA IDF contient des données personnelles et des données de santé. Il est nécessaire de réaliser des traitements informatiques afin de transformer les données et ne pas pouvoir identifier les personnes comme indiqué dans l'article 26 du [RGPD](#), « il y a lieu d'appliquer les principes relatifs à la protection des données à toute information concernant une personne physique identifiée ou identifiable. »

## 4.2 Anonymisation et pseudonymisation

Il existe deux concepts fondamentaux lors du traitement des données à caractère personnel, l'anonymisation et la pseudonymisation. Dans le cadre de notre travail après traitement des données nous obtenons des données pseudonymisées.

**L'anonymisation** des données est définie par la [CNIL](#) comme le traitement permettant le retrait de l'ensemble des informations permettant l'identification d'une personne. « Une anonymisation irréversible consiste à supprimer tout caractère identifiant à un ensemble de données. Concrètement, cela signifie que toutes les informations directement ou indirectement identifiantes sont supprimées ou modifiées, rendant impossible

---

4. [https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=4941B11FB7C531A046EC7E6199A84A60.tplgfr24s\\_1?cidTexte=JORFTEXT000037187498&dateTexte=&oldAction=rechJO&categorieLien=id&idJO=JORFCONT000037186583](https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=4941B11FB7C531A046EC7E6199A84A60.tplgfr24s_1?cidTexte=JORFTEXT000037187498&dateTexte=&oldAction=rechJO&categorieLien=id&idJO=JORFCONT000037186583)

5. <https://www.cnil.fr/fr/quest-ce-que-une-donnee-de-sante>

toute ré-identification des personnes<sup>6</sup> ». Les données anonymes ne rentrent pas dans le champ d'application du RGPD.

La **pseudonymisation** est l'autre concept du traitement des données à caractère personnel. Il permet après un traitement des données, de ré-identifier la personne. La CNIL définit la pseudonymisation comme « une technique qui consiste à remplacer un identifiant (ou plus généralement des données à caractère personnel) par un pseudonyme. Cette technique permet la ré-identification ou l'étude de corrélations en cas de besoin particulier<sup>7</sup> ». Le travail réalisé dans cette thèse se situe dans le cadre de la pseudonymisation. Pour permettre une analyse des données, nous devons relier les données d'annotations sémantiques aux données socio-démographiques des patients (âge, sexe, situation sociale, etc.). Nous avons donc créé une table reliant les données nominatives des patients à une table de correspondance. Cette table de correspondance n'est accessible qu'à une seule personne, et nécessite l'utilisation d'un mot de passe.

### 4.3 Les méthodes d'anonymisation

Les méthodes d'anonymisation automatique reposent principalement sur deux méthodologies (Meystre *et al.*, 2010) : une approche basée sur des règles et une approche basée sur l'apprentissage statistique, cependant certains outils vont utiliser une combinaison des deux méthodes. Chacune de ces méthodes possède des avantages et des inconvénients.

Les méthodes à base de règles et de dictionnaires peuvent être modifiées rapidement permettant d'augmenter les performances du système en ajoutant des règles, des termes ou des expressions régulières. Néanmoins il nécessite une maintenance et la mise en place de nombreux dictionnaires mais aussi une bonne connaissance des différents modèles de données à anonymiser. Ces méthodes à base de dictionnaires ont de meilleurs résultats avec des termes rarement mentionnés dans les corpus.

Les méthodes à base d'apprentissage, nécessite un grand corpus annoté pour créer les algorithmes d'apprentissage automatique, ce sont des méthodes généralisables à d'autres corpus. Une étude de Uzun *et al.* (2007) explicite que les systèmes les plus performants sur leurs corpus sont les systèmes d'apprentissage statistiques basés sur des modèles de règles et secondairement les systèmes hybrides.

Pour rappel les événements de la base du Réseau SLA IDF sont nombreux, un travail d'anonymisation manuelle n'étant pas envisageable, un traitement automatique du corpus était nécessaire, pour cela nous avons collaboré avec un autre laboratoire le LIMSI<sup>8</sup> qui a développé une méthodologie et des outils pour l'anonymisation. L'approche choisie fut l'utilisation de systèmes en cascade (figure 4.1), c'est-à-dire l'application successive de deux systèmes, le second prenant en entrée la sortie du premier

---

6. [https://www.cnil.fr/sites/default/files/typo/document/FICHE10\\_PackConf\\_LOGEMENT\\_SOCIAL\\_web.pdf](https://www.cnil.fr/sites/default/files/typo/document/FICHE10_PackConf_LOGEMENT_SOCIAL_web.pdf)

7. [https://www.cnil.fr/sites/default/files/typo/document/FICHE10\\_PackConf\\_LOGEMENT\\_SOCIAL\\_web.pdf](https://www.cnil.fr/sites/default/files/typo/document/FICHE10_PackConf_LOGEMENT_SOCIAL_web.pdf)

8. Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur <https://www.limsi.fr/fr>

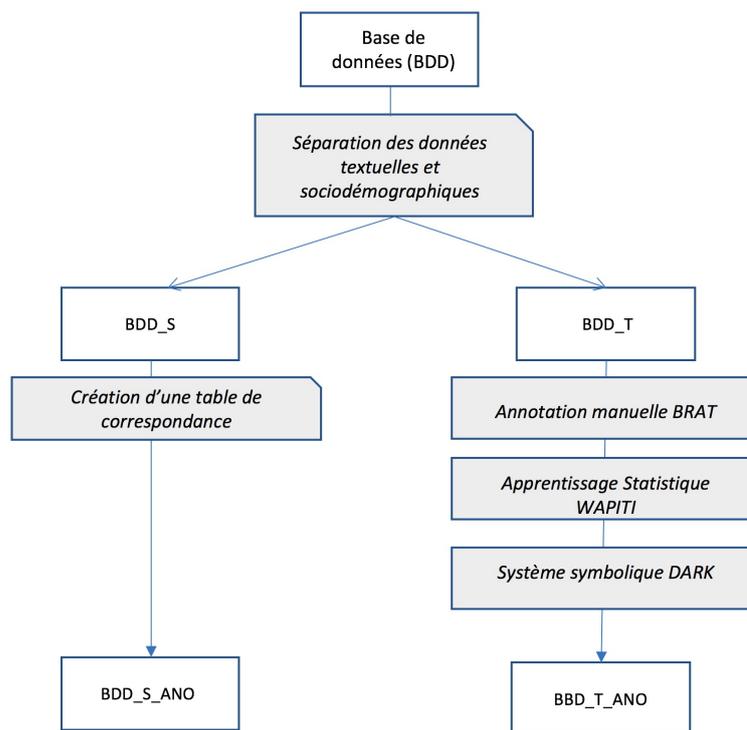


FIGURE 4.1 – Processus de pseudonymisation réalisé dans le travail de thèse, sur la base de donnée extraite du portail SLA. BDD-T : base de données textuelle; BDD-S : base de données socio-démographique; BDD-T-ANO : base de données textuelle anonymisée par la méthode symbolique; BDD-S-ANO : base de données socio-démographique anonymisée.

système. Le premier système utilise l'apprentissage automatique basé sur le formalisme des champs aléatoires conditionnels (CRF) utilisé dans WAPITI (Lavergne *et al.*, 2010). Le second système utilise un système à base de règles et de lexique [Data Annotation using Rules and Knowledge \(DARK\)](#).

Grouin (2013) explicite ce type d'hybridation « le premier système agit sur les informations qu'il est en mesure de traiter, tandis que le deuxième complète les traitements en agissant sur les informations qui n'auront pas été traitées par le premier. Dans ce schéma, chaque système agit sur son domaine de spécialité, dans le sens d'une spécialisation des systèmes en fonction des types d'information pour lesquels chaque système a été conçu. »

## 4.4 Traitement du corpus pour la pseudonymisation

La première partie du travail a consisté en l'extraction de l'ensemble des données présente dans la base événementielle du Réseau SLA IDF à partir de l'application

internet vers un fichier excel. Afin de garantir une sécurité des données, seule une personne a eu accès aux données nominatives des patients à partir de l'export des événements de la base du Réseau SLA IDF. L'ensemble des données extraites ont été placées sur un seul poste informatique, l'accès au poste nécessitant un identifiant et mot de passe, il fut ajouté un niveau de sécurité supplémentaire en verrouillant les dossiers contenant les données personnelles par un mot de passe. Lors de l'adhésion du patient au Réseau SLA IDF, le patient signe un document stipulant qu'il autorise le Réseau SLA IDF à utiliser les données à des fins de recherche. La base de donnée du Réseau SLA IDF fut inscrite à la [CNIL](#) lors de la création du réseau en 2005.

#### 4.4.1 Préparation des données

L'extraction des données à partir de la base SLA permet d'obtenir un fichier excel de dix-huit colonnes réparties ainsi :

1. N° Inclusion **001/98**
2. Nom patient **Cerveau**
3. Prénom Patient **Marc**
4. Patient Sortie **oui**
5. Date Sortie **01/02/01**
6. Cause Sortie **Décès**
7. Date évènement **23/12/00**
8. Intervenant réseau **Jeanne IDE**
9. Interlocuteur **Mme Cerveau épouse**
10. Type **Appel reçu**
11. Nature **Social**
12. Complexité **1**
13. Gravité **1-2**
14. Type suivi **Oui**
15. Etat du suivi **suivi fermé**
16. Question / Commentaire **appel de Mme Cerveau qui nous sollicite pour obtenir un certificat médical du neurologue pour le dossier Mdp. Demande si Marie Ergo peut la rappeler pour un problème de fauteuil**
17. Réponse / Action **Je transmets la demande à Marie Ergo et au neurologue**
18. Suivi **Réception ce jour du certificat du neurologue. Envoie au patient**

Les données textuelles se trouvent dans les colonnes 16 à 18 (Question / Commentaire; Réponse / Action et Suivi). Pour exploiter les corpus sur les différentes applications créées, il est nécessaire d'organiser et de transformer le fichier excel en une séquence de fichier texte, contenant un identifiant par évènement. Les événements saisis par les coordinateurs peuvent être de deux natures comme indiqué à la figure [4.2](#).

Question / Commentaire	Réponse/Action	Suivi
patient rencontré en pneumologie sur une hospit, à la demande du NEUROLOGUE	<p>Suivi depuis DATE, le patient presente une tétraparesie, ventilé 24/24 ( en journée masque narinaire), etait accompagné dans la chambre de son epouse + 2 filles mineures; le patient maîtrise mal le français et assez incomprehensible avec la vNI, l'epouse qu'il a voulu designer comme personne de confiance ne parle pas du tout français ( coupe XXXX)</p> <p>La jeune fille mineure accepte de faire la traduction auquotidien, mais aura des limites suivant le type d'echanges et en previent</p> <p>le patient vit dans une maison inadaptée t dont le bailleur refuse de faire des travaux.il voudrait un relogement dans le parc social. Son FRE n'est pas adapté car les dimensions ne permettent pas de circuler dans le logement ( couloir et WC) et la salle de B est petite</p> <p>il voudrait un futeuil confort pour la journée son lit medicalisé est trop étroit pour sa corpulence, en voudrait un plus large</p> <p>Il n'y a pas d'aide humaines en place, seulement un medecin traitant, et pas de kiné non plus qui se deplace au domicile</p> <p>ils m'ont exprimé :</p> <p>-que ASSISTANTE-SOCIALE faisait appui pour l'obtention d'un logement plus adapté HLM</p> <p>-Mr était OK aujourd'hui pour une aide les matins à la toilette : je voulais contacter le ssiad local</p> <p>-aide à la recherche de kiné</p>	<p>COORDINATEUR SLA (DATE):après consultation de la base portail, et identification des kiné possibles:</p> <p>COORDINATEUR SLA (DATE):après consultation de la base portail, et identification des kiné possibles:5 coordonnées dans 4 centres de santé: tous appelés, aucune VAD possible ai demandé également des pistes au ssiad local: aucune ressource</p> <p>COORDINATEUR SLA (DATE) :singalement fait pa messagerie au ssiad de VILLE pour PEC toilette matinale</p> <p>COORDINATEUR SLA (DATE):tentative de rappel au medecin traitant echoué: rappeler un autre jour, le matin</p> <p>COORDINATEUR SLA (DATE):info du ssiad de VILLE, PROFESSIONNEL: a débuté la PEC ce jour MAIS</p> <p>"le logement n'est pas du tout adapté et que la famille est dans l'attente de la mise en place du matériel. Nous ne pourrons pas lui faire la douche car la salle de bain n'est pas du tout adaptée.</p> <p>Son épouse s'occupe très bien de son mari mais prend beaucoup de risque notamment pour les transferts car c'est elle qui l'emmène aux toilettes à chaque fois sachant qu'il a une défaillance au niveau des membres inférieurs, qu'il pèse 103kg et qu'il existe un risque de chute non négligeable.</p> <p>Pouvez vous tenir au courant de l'avancé de la mise en place du matériel."</p> <p>Info donnée à ASSISTANTE-SOCIALE + NEUROLOGUE + PROFESSIONNEL CENTRE EXPERT, avec une demande à PROFESSIONNEL CENTRE EXPERT de rapeller le ssiad qui veut savoir où en sont les mises en place de materiel</p>

FIGURE 4.2 – Illustration d'un évènement principal et des évènements de suivis

Un **évènement principal** contenu dans la colonne 16 du fichier. Il s'agit de la demande initiale ou premier contact. La réponse ou l'action mise en place pour répondre à la demande est reportée à la colonne 17. Lorsque de nombreuses actions sont nécessaires pour apporter des réponses, un ou des évènements de suivi vont être créés.

Les **évènements de suivi** se trouvent dans la colonne 18. Ils permettent de suivre les différentes actions réalisées successivement pour répondre à la problématique initiale.

Afin de faciliter le traitement des corpus et permettre de lier les évènements de suivis à l'évènement initial, nous avons fait le choix de re-structurer le fichier excel. Pour cela un algorithme a été créé afin de :

1. Fusionner les question et les actions, c'est-à-dire mettre dans un même document l'évènement principal et la réponse apportée (fusion des colonnes « Question / Commentaire » (colonne 16) et les « Réponse / Action » (colonne 17)).
2. Séquencer les suivis en  $n$  documents.
3. Attribuer un identifiant commun pour relier les évènements de suivis à l'évènement initial.

L'identifiant crée permet de voir s'il s'agit d'un évènement principal ou d'un évènement de suivi comme illustré ci dessous :

- 17-002—2027-01-06—00—APPEL-REÇU—D-R : évènement principal
- 17-002—2027-01-06—00—APPEL-REÇU—D-R—2027-06-06—00—SUI : évènement de suivi n°1

- 17-002—2027-01-06—00—APPEL-REÇU—D-R—2027-06-15—00—SUI : évènement de suivi n°2
- 17-002—2027-01-06—00—APPEL-REÇU—D-R—2027-07-10—00—SUI : évènement de suivi n°3
- 17-002—2027-01-06—00—APPEL-REÇU—D-R—2027-11-07—00—SUI : évènement de suivi n°4

Dans l'exemple présenté à la figure 4.2 seul 5 des suivis sont présentés parmi les 9 initialement saisis dans la base du Réseau SLA IDF.

## 4.4.2 Utilisation de méthode à base d'apprentissage statistique

Dans cette section, nous présentons la méthode à base d'apprentissage statistique utilisée pour la dé-identification. Les événements contiennent de nombreuses données personnelles comme des données nominatives (nom et prénom de patient, de leur famille, de professionnel de santé et de professionnel social, etc.), des données numériques (date, numéro de dossier, numéro de téléphone, etc.) ainsi que des données de localisation (ville, nom de structure d'accueil, nom d'hôpital, etc.). Afin de répondre aux normes de protection des données personnelles, il est nécessaire de réaliser un traitement préalable de dé-identification pour secondairement utiliser les corpus à des fins de recherche.

L'utilisation de méthode à base d'apprentissage permet de traiter rapidement de grands volumes de données, pour que cette méthode soit optimale, il est nécessaire de posséder un corpus annoté représentatif. Ce corpus d'apprentissage doit couvrir chaque catégorie d'information et prendre en compte l'ensemble des variantes afin de permettre la construction d'un modèle robuste (Grouin, 2013).

### 4.4.2.1 Définition des entités nommées

Les événements contiennent de nombreux types de données à la fois sur le patient, l'entourage mais aussi sur les structures et les personnes intervenant à domicile. La première étape a consisté en la création de catégorie sémantiques à anonymiser. Nous avons défini une liste de vingt deux catégories faisant référence au type sémantique des données à modifier dans le corpus regroupé en cinq catégories comme indiqué à la figure 4.3.

Les annotations manuelles ont été réalisées à l'aide de l'outil d'annotation BRAT<sup>9</sup>. Lors du travail de pseudonymisation nous avons fait l'hypothèse que l'identification des « agents » par leurs fonctions (patients, entourage, neurologue, assistante-sociale) aurait un impact, qualitatif et quantitatif, sur le repérage des entités nommées lors de l'annotation, en les remplaçant directement par les catégories les subsumant. La transformation d'une donnée nominale *Jean Dupont* (4.1), non définie dans l'ontologie, en donnée identifiant un concept « agent » (Patient) défini dans l'ontologie permet

---

9. Brat Rapid Annotation Tool, <http://brat.nlplab.org/>.

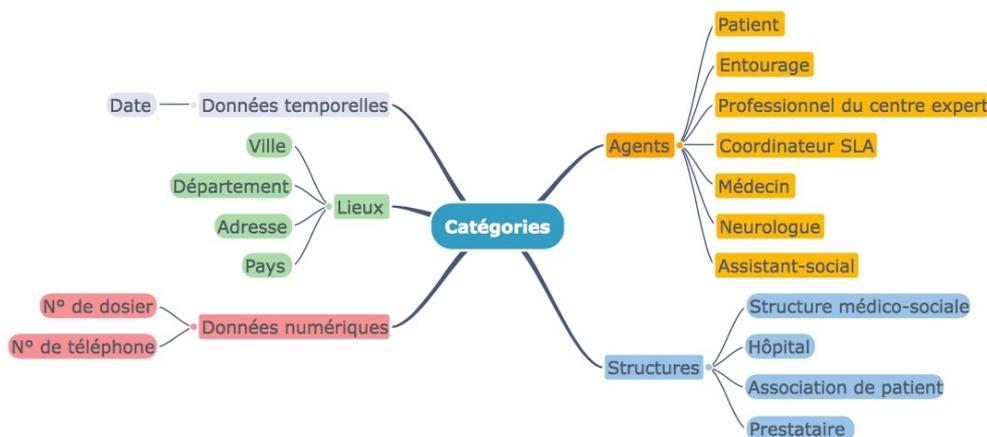


FIGURE 4.3 – Liste des catégories sémantiques à anonymiser dans les corpus.

TABLEAU 4.1 – Exemple d’un événement modifié par la pseudonymisation.

Corpus original	appel de <u>Jean Dupont</u> qui a sollicité <u>Claire Marche</u> , pour obtenir un certificat médical du <u>Dr Cerveau</u> pour son dossier Mdp. Demande si <u>Marie Ergo</u> peut le rappeler pour son problème de FRE
Corpus anonymisé	appel de <b>PATIENT</b> qui a sollicité <b>COORDINATEUR SLA</b> , pour obtenir un certificat médical du <b>NEUROLOGUE</b> pour son dossier Mdp. Demande si <b>COORDINATEUR SLA</b> peut le rappeler pour son problème de FRE

d’un point de vue conceptuel de repérer les interactions entre les agents et les actions. D’un point de vue qualitatif, l’anonymisation permet de mieux repérer les agents acteurs dans le parcours de santé. La coordination de parcours repose sur l’interaction d’agents réalisant des actions à destination d’autres agents. L’anonymisation peut aider à comprendre les interactions et les interdépendances des agents et des structures, mais aussi de la prépondérance de la présence versus absence de certains. Le type de structures sollicitées et le type d’agents impliqués dans la coordination sont probablement un facteur important de compréhension des parcours, connaître qui sont les demandeurs, quelles personnes donnent l’alerte sur la dégradation d’une situation clinique ou familiale telle que l’épuisement.

#### 4.4.2.2 Création du corpus de référence

La méthode à base d’apprentissage statistique reposant sur le repérage des entités à partir d’un corpus annoté, nécessite l’utilisation d’un corpus de référence. Pour cela nous avons extrait de façon aléatoire, de la base Réseau SLA IDF, cinquante quatre dossiers patients contenant 2 311 événements. Nous avons réalisé une annotation manuelle sur tous ces événements, nous avons vérifié au cours de l’annotation que chacune des catégories définie était bien présente et annotée afin de créer un modèle robuste contenant les différents contextes d’apparition des informations à anonymiser. Ce corpus a

servi de référentiel pour la création du modèle d'apprentissage utilisé dans WAPITI <sup>10</sup>.

Une fois le modèle d'apprentissage créé nous avons traité un nouveau corpus, rapidement des erreurs d'annotation et des oublis sont apparues. Les événements ne suivent pas une structuration homogène dans leurs contenus et dans leurs formes syntaxiques, rendant l'apprentissage plus complexe. Les erreurs rencontrées étaient « la confusion » dans l'annotation des personnes physiques ex : *Patient* annoté comme *Professionnel* ou bien des oublis d'annotation sur des dates ou des lieux (nom de ville, nom d'hôpitaux, etc.) et des oublis pour le nom de professionnel exerçant au sein du centre expert SLA. Pour pallier à ces erreurs il fut décidé d'utiliser une méthode à base de règles et de lexiques.

### 4.4.3 Utilisation de règles

« Les méthodes symboliques reposent sur l'utilisation de dictionnaires (dictionnaire de langue générale ou de termes métiers spécifiques), de listes (listes d'entités nommées : noms, prénoms, villes, etc.), de déclencheurs (des indices permettant de détecter des informations à anonymiser dans leur voisinage : “M., Mme, Dr” pour les noms de personnes, “CHU, clinique” pour les noms d'hôpitaux) et la définition de patrons syntaxiques » (Grouin, 2013). La méthode à base de lexiques et de règle offre des résultats de qualité et elle est optimale pour des documents ayant la même typologie (document issu d'un même service et du même type), cependant elle nécessite une maintenance dans la mise à jour des listes (Grouin, 2013).

La syntaxe utilisée par les coordinateurs est très hétérogène d'un coordinateur à l'autre. Pour dénommer une même personne, les coordinateurs peuvent utiliser une grande variabilité de forme syntaxique : *nom, prénom; initiales du nom et prénom; titre du professionnel et prénom; nom et initiale du prénom; prénom*, etc. Pour que le système d'apprentissage puisse être performant, il est nécessaire que ces différentes formes soient annotées et repérées en grande quantité. La variabilité des dénominations et la diversité des professionnels intervenants ont montré les limites du système d'apprentissage, qui ne reconnaissant pas toutes les formes. Après identification des erreurs sur les corpus de sortie de WAPITI nous avons fait le choix d'utiliser en complément une méthode à base de règles. Pour améliorer la détection des noms (Foufi *et al.*, 2017), nous avons identifié la liste des déclencheurs <sup>11</sup> que sont les titres comme par exemple « *Monsieur, M. ou Mr; Docteur, Dr, Professeur, Pr.* », de plus nous avons créé neuf listes de termes afin d'améliorer le repérage des termes à anonymiser :

- liste des noms des patients (contenant le nom et le prénom),
- liste des noms de professionnels du centre expert SLA (contenant le nom et prénom ainsi que les initiales et les différentes formes utilisées),
- liste des neurologues du centre expert SLA,
- liste des assistants sociaux du centre expert SLA,
- liste des noms des coordinateurs du Réseau SLA IDF,

---

10. <https://wapiti.limsi.fr>

11. Un déclencheur est un élément du texte servant d'indice pour détecter une entité nommée

- liste des villes de la région Île-de-France,
- liste des pays,
- listes des noms d'hôpitaux (sous forme d'acronyme et dans leur nom complet, par exemple HEGP : pour l'hôpital européen Georges-Pompidou),
- liste des déclencheurs.

À ces listes de termes, nous avons implémenté des expressions régulières, essentiellement pour les données numériques qui suivent une structuration homogène, comme les dates ayant un format JJ/MM/AA , JJ/MM ou des numéros de téléphone construit d'une séquence de 5 nombres (01 01 01 01 01).

### 4.4.4 Traitement des dates

Les corpus contiennent de nombreuses données temporelles, date de diagnostic, date d'intervention etc., et présentent différents formats : 19/11/2019 ; 19/11/19 ; 19 novembre 2019 ou encore le 19/11. Certaines méthodes permettent d'appliquer un décalage temporelle aux dates présentes dans les corpus textuels. Nous avons fait le choix de masquer les données temporelles présentes dans les corpus textuels en plaçant une étiquette « Date ». La temporalité des faits est importante dans la notion de parcours, afin d'avoir une vision chronologique des évènements intervenants pour les patients, nous avons créé un identifiant spécifique pour chaque évènement, dans lequel apparait la date. Ces informations contenues dans les noms de fichier peuvent nous permettre d'ordonner et de traiter chronologiquement les évènements principaux et les évènements de suivis qui sont liés.

### 4.4.5 Résultat et discussion de la pseudonymisation

Après avoir testé et apporté des modifications aux deux systèmes nous avons réalisé une évaluation. Un nouveau corpus constitué de 15 dossiers, soit 431 évènements, tiré de façon aléatoire sur la base, fut testé pour mesurer la performance des deux systèmes en utilisant successivement le système d'apprentissage et le système à base de règles. Les résultats obtenus donnent un Rappel de 0,81 ; une précision de 0,90 et une F-Mesure de 0,85 (Cardoso *et al.*, 2017).

Les erreurs résiduelles sont des heures annotées comme des dates du fait de leur similarité de format, par exemple : 12/24 (pour indiquer 12 H sur 24 H) et une date 12/12. Dans les corpus les heures correspondent au temps de ventilation ou bien aux temps d'heures d'aides humaines accordées aux patients dans le cadre du financement des aides.

Initialement nous avons créé, dans le corpus de test, une distinction entre les *Patients* et l'*Entourage*, car d'un point de vue clinique et dans l'objectif de comprendre le parcours des patients, il nous semblait important de comprendre quelles personnes sollicitaient le Réseau SLA IDF. Néanmoins, les outils utilisés et notamment l'emploi de lexique et en particulier les listes des noms des patients dans le système à base de règles ne permettent pas de faire la distinction entre *Mr Cerveau* et *Mme Cerveau*, dans

les deux cas l'étiquette mise en place sera *patient*. Si la mise en place d'étiquettes entre le patient et l'entourage ne fonctionne pas de façon optimale, l'utilisation de liste nominative permet une distinction pertinente entre les différents professionnels du centre expert. La distinction entre le neurologue et les assistantes-sociale, permet de mieux identifier les interactions entre le Réseau SLA IDF et les différents professionnels du centre expert. Si le premier traitement des données fut la pseudonymisation, le second traitement mis en place fut le traitement orthographique. Les événements présentent de nombreuses erreurs, dû au contexte de saisie des événements. La présence de ces fautes entraîne une diminution de la reconnaissance des concepts lors de la phase d'annotation sémantique dans ONBASAM. Pour augmenter les performances du système d'annotation nous avons souhaité mettre en place un module de correction orthographique, ceci afin d'avoir une meilleure compréhension du contexte de l'évènement ainsi qu'une augmentation du repérage des classes lors de l'annotation. La section suivante va présenter la méthodologie et les outils mis en place.

## 4.5 Traitement Orthographique

La lecture des événements lors de la réalisation du processus de pseudonymisation a mis en évidence la présence de nombreuses abréviations non homogènes entre les coordinateurs, de fautes lexicales, grammaticales et typographiques. Ces erreurs s'expliquent par le contexte de saisie des événements par les coordinateurs qui peuvent être en double tâche lors de l'écriture des événements, de plus, ils utilisent une formulation qui est à destination d'autres coordinateurs ne nécessitant pas un formalisme strict préférant parfois l'utilisation d'un style télégraphique.

Nous avons émis l'hypothèse que la présence de ces erreurs pouvait diminuer les capacités du système d'annotation sémantique ONBASAM à reconnaître des concepts clés. Par exemple la présence du terme « aggravation » s'il est mal orthographié ne sera pas annoté. L'objectif de la chaîne de traitement de correction orthographique est de corriger les erreurs. Pour limiter ces erreurs nous avons utilisé les outils du TALN, afin d'apporter des corrections aux corpus sans en changer la qualité clinique. La correction produite ne doit pas sur-corriger le texte, ce qui engendrerait la modification de terme pouvant modifier le sens du contenu. Il est préférable d'utiliser un outil qui sous corrige plutôt qu'un qui le sur-corrige (Baranes, 2012).

### 4.5.1 Types d'erreurs rencontrées

Nous avons constaté dans le corpus différents types d'erreurs :

- Ajout d'un caractère : ppour → pour
- Suppression d'un caractère : domicile → domicille
- Remplacement d'un caractère : gasionnaire → gestionnaire
- Oubli des accents : epoux → époux
- Permutation de caractère : patietn → patient.

```

file:/Users/sc/Desktop/Brat/brat-v1.3_Crunchy_Frog/data/sortie/059_14_ANOWAP/14-059_2016-01-04_00_APPEL-REÇU_D-
R_2016-04-06_00_SUI.anoto
reussi --> réussi
gestionnaire --> gestionnaire
domicile --> domicile
epouse --> épouse
epuisée --> épuisée
hypersollicitée --> hyperbolicité
epoux --> époux
absenteisme --> absentéisme
mere --> mère
epouse --> épouse
egalement --> également
gerer --> gérer
etait --> était
separer --> séparer
egalement --> également
pieces --> pièces
deployé --> déployé
gastionnaire --> gestionnaire
ppour --> pour
degager --> dégager
epouse --> épouse
reseau --> réseau
miseen --> mise en
domicile --> domicile
medical --> médical

```

FIGURE 4.4 – Exemple de correction orthographique réalisée dans ONBASAM. Cet exemple illustre les différents types d’erreurs qu’il est possible de rencontrer dans les corpus de coordination de parcours.

La figure 4.4 illustre les erreurs rencontrées dans les corpus. Dans cet exemple l’évènement est composé de 266 mots, le système identifie 25 erreurs dont 18 fois l’oubli d’accent, l’oubli d’un caractère apparaît 2 fois, le remplacement d’un caractère 1 fois, l’ajout d’un caractère 1 fois et la suppression d’un caractère (dont l’espace) 1 fois.

Nous avons souhaité mettre en place un pipeline permettant la correction automatique des erreurs pour l’ensemble du corpus. Pour cela nous avons construit un pipeline dénomé Spell Cheking dans ONBASAM qui s’est structuré progressivement. Si initialement nous souhaitions utiliser les outils génériques (dictionnaire Hunspell, tree tagger en français) nous avons dû intégrer des listes spécifiques au corpus et au contexte du Réseau SLA IDF, pour améliorer les performances du système.

## 4.5.2 Pipeline de correction orthographique : Spell Checking

Les outils développés dans le cadre de cette thèse dans le champs du TALN ont pris pour source les ressources existantes de General Architecture for Text Engineering (GATE) <sup>12</sup>, avec des modifications pour que les outils soient adaptés à nos corpus. Les données textuelles sont spécifiques du domaine et du contexte dans lesquelles elles sont écrites. Dans ce travail, les données sont saisies par des professionnels paramédicaux, qui utilisent quotidiennement des abréviations, des acronymes et un champ lexical spécifique (lié à la pathologie, aux aspects sociaux, etc.) et parfois une structuration de type télégraphique sans avoir de construction grammaticale. Pour réaliser les corrections nous avons créé un pipeline dans ONBASAM, qui va utiliser successivement un dictionnaire de la langue française (hunspell) un dictionnaire du domaine (vocabulaire UMLS en français), le lexique de notre ontologie, ainsi qu’une liste de terme que nous

12. <https://gate.ac.uk>

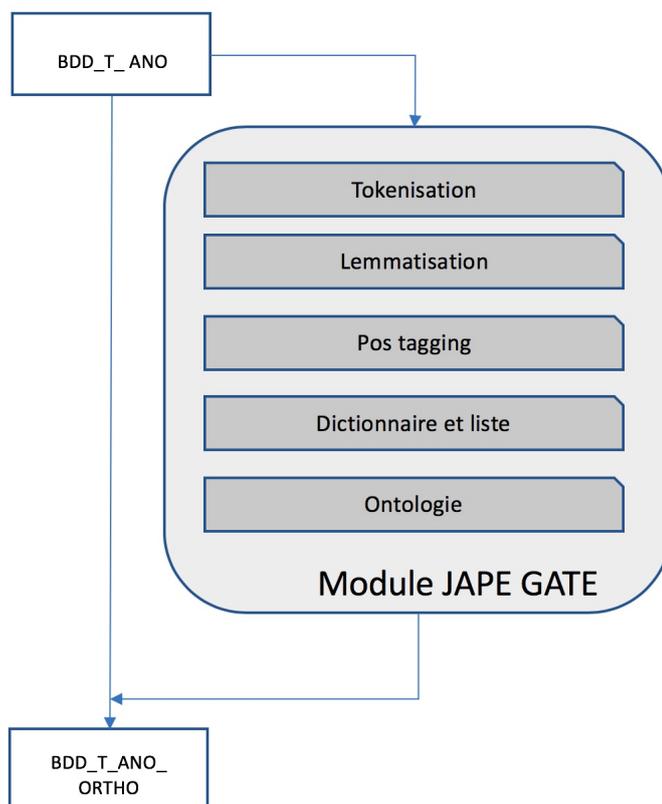


FIGURE 4.5 – Processus mis en place permettant de passer à une base de données textuelle anonymisée et corrigée. Pour réaliser les corrections nous avons créé une chaîne de traitement spécifique dans ONBASAM.

avons créée afin de limiter la sur-correction (exemple de liste : liste de marque de fauteuil et liste d’abréviations d’aides techniques). Pour cela nous utilisons [Java Annotation Patterns Engine \(JAPE\)](#) l’un des composants de GATE qui intervient sur les expressions régulières comme indiqué dans la figure 4.5.

Une fois le mot mal orthographié repéré, un ensemble de suggestions est proposé par le système en utilisant les termes dénotant les concepts de l’ontologie du domaine et le dictionnaire Hunspell. Nous appliquons pour ce faire, la distance d’édition Damerau-Levenshtein<sup>13</sup> à chaque mot mal orthographié et ses suggestions, pour déterminer le meilleur choix. Il convient de noter que la distance d’édition ne peut pas être plus grande qu’un seuil spécifié, dans ce cas, la correction ne sera pas faite et le mot restera mal orthographié (Cardoso *et al.*, 2017).

Les suggestions de corrections sont basées sur le concept de distance d’édition, qui

13. il s’agit d’une distance entre deux chaînes de caractères. Elle permet de calculer le nombre minimum d’opérations nécessaires pour transformer une chaîne de caractères en une autre, il existe quatre types d’opérations l’insertion, la suppression ou la substitution d’un simple caractère ou une transposition de deux caractères. Ces quatre opérations correspondent à plus de 80% des fautes d’orthographe humaines <http://fracademic.com/dic.nsf/frwiki/520710>

calcule le nombre de manipulations nécessaires pour retrouver, à partir d'un input erroné, une chaîne de caractères dans le dictionnaire qu'il s'agit de proposer à l'utilisateur (Fontenelle, 2006).

Les différentes étapes réalisées dans spell cheking sont :

1. normalisation correction des problèmes d'encodage ;
2. tokenisation : séparation des mots dans le texte en prenant en compte les signes (ponctuations, espaces) ;
3. découpage en phrases ;
4. lemmatisation pour réduire la dispersion. La lemmatisation des corpus consiste à remplacer chaque mot par sa forme canonique, l'un des avantages est la réduction du nombre de formes et l'augmentation des occurrences de chaque forme dans le corpus (Lemaire, 2008) ;
5. PoS tagging pour reconnaître la catégorie grammaticale ;
6. utilisation des dictionnaires et listes ;
7. utilisation de l'ontologie ONTOPARON ;
8. utilisation de Tree Tagger en version française pour le Part of Speech (PoS) tagging et la lemmatisation. TreeTagger est un étiqueteur morphosyntaxique basé sur un lexique propre de formes fléchies du français. Il permet de segmenter un texte en formes et d'attribuer à chaque forme une catégorie grammaticale et un lemme (Wolfarth *et al.*, 2017).

Si des erreurs dues aux inversions, remplacements ou ajouts de caractères sont présents, l'autre spécificité est la présence de nombreuses abréviations dans les corpus. Si ces abréviations sont reconnues par le système et transformées, le sens de l'évènement peut-être perdu. La figure 4.6 illustre le repérage de ces abréviations comme des erreurs par le système et leurs transformations. Ainsi les termes :

- *appartement* pour lequel les coordinateurs utilisent le terme *appt* est transformé en  $\rightarrow$  *appât*,
- le terme *Service* pour lequel les coordinateurs utilisent l'abréviation *sce* sera transformé en  $\rightarrow$  *Sec*,
- l'abréviation *FdV* correspondant au concept *Fin de vie* devient  $\rightarrow$  *VDF*, n'ayant aucune signification.

Ces exemples illustrent parfaitement la modification du sens premier des abréviations, nécessitant la mise en place d'outil pour limiter cette sur-correction. Pour pallier ces erreurs nous avons décidé d'implémenter des listes de termes et d'abréviations spécifiques rencontrées dans le corpus.

Nous avons créé deux listes de termes que nous avons implémentés dans le module du correcteur orthographique :

- une liste de terme indiquant les marques de certaines aides techniques, en effet parfois les aides techniques sont identifiées à partir de leurs marques et non à partir de leurs catégories (exemple : *spinx* pour un fauteuil roulant manuel actif, *netti* encore écrit *netty*, *salsa*, *tobbii*, *Ipad*, *roho*...);

```
file:/Users/sc/Desktop/anonymisationExport/FICHIERSLATRAITEMNT/FichierANONYMISES/2013/005_13%20AN0/005-13_13-06-2013_HDJ D-R_3545a67eeec4eccf90f43e178947eb8a.txt
appt --> appât
ème --> sème
sem --> stem
sem --> stem
Ø --> E
jrs --> jars
durogesic -->
oxynorme --> oxymore
dantrium -->
qttes --> gattes
ttes --> têtes
microlax --> microlux
FdV --> VDF
ajd --> ad
Sce --> Sec
cde --> cade
nyvelles --> nivelles
fte --> fête
ergos --> ergots
check --> cheik
list --> lits
```

FIGURE 4.6 – Exemple de repérage d’abréviation et de leur transformation par le système. Ce repérage et cette transformation sont extrait d’un exemple pour lequel nous n’avions pas encore intégré les listes indiquant les abréviations utilisées par les coordinateurs.

- une liste de terme qui reprend certains termes anglais utilisé dans les corpus (switch, check list, cought assist) ou week-end écrit de plusieurs manières week-end, seuls les termes français sont utilisés lors de l’annotation.

Nous avons appliqué ce pré-traitement de correction orthographique à l’ensemble du corpus. Une fois l’ensemble des évènements corrigés nous avons pu les utiliser dans la chaîne de traitement créé dans ONBASAM pour réaliser la phase d’annotation sémantique. La création des listes spécifiques et leurs implémentations ainsi que l’utilisation de l’ontologie dans le module ont permis une limitation des erreurs d’orthographe et de la sur-correction et donc une amélioration de l’annotation sémantique par les autres pipelines de ONBASAM.

## 4.6 Synthèse

Les processus de pseudonymisation et de correction orthographique des corpus, nous ont permis d’améliorer la qualité du corpus pour la phase d’annotation sémantique, en permettant un meilleur repérage des termes. Le traitement de pseudonymisation est un processus incontournable pour le traitement de corpus contenant des données personnelles. La transformation de données nominales en étiquettes sémantiques qui identifient le rôle de l’agent intervenant, apporte une source d’information dans le traitement et l’analyse des corpus et des concepts annotés. L’utilisation de listes à la fois pour la pseudonymisation et pour la correction orthographique a été rendue possible par l’accès à l’ensemble des données. Cependant l’utilisation de listes nécessite la mise à jour régulière des données, mais leur apport est important dans ces deux

processus. La transposition de ces systèmes à d'autres corpus est possible mais elle nécessite des adaptations afin de créer des listes (de termes, d'abréviations, etc.) qui soient adaptées au corpus à traiter.

Le chapitre suivant va présenter les différentes étapes pour la construction de l'ontologie ONTOPARON.



# Chapitre 5

## Construction de l'ontologie ONTOPARON

### Sommaire

---

<b>5.1</b>	<b>Principes de construction de ONTOPARON</b>	<b>88</b>
<b>5.2</b>	<b>Extraction des candidats termes avec BioTex</b>	<b>89</b>
5.2.1	Résultats de l'extraction	90
5.2.2	Sélection des termes de BioTex	91
5.2.3	Sélection des termes et antériorité du travail	92
<b>5.3</b>	<b>Construction de l'ontologie</b>	<b>93</b>
<b>5.4</b>	<b>Passage d'une ontologie monolithique à une ontologie modulaire</b>	<b>96</b>
5.4.1	Les concepts définis	99
5.4.2	Le module noyau : ONTOPARONNOY	100
5.4.3	Le module médical : ONTOPARONMED	103
5.4.4	Le module socio-environnemental : ONTOPARONSOC	106
5.4.5	Le module de coordination : ONTOPARONCOORD	108
5.4.6	Le module de consolidation	110
<b>5.5</b>	<b>Synthèse</b>	<b>110</b>

---

Comme nous l'avons indiqué dans le chapitre 3 une ontologie est une spécification explicite d'une conceptualisation. Quand la connaissance du domaine est représentée dans un formalisme déclaratif, l'ensemble des objets pouvant être représentés est appelé l'univers du discours. Les ontologies sont des outils permettant de partager et d'échanger de la connaissance entre des experts d'un domaine mais aussi de la transcrire dans un formalisme afin d'être utilisée par une machine. L'une des phases importantes dans la construction de l'ontologie est la phase de conceptualisation, ou il est nécessaire de passer du terme au concept. Le terme est une unité linguistique qui peut être polysémique ; cependant le concept ontologique doit, lui, répondre à un sens univoque et sera décrit dans l'ontologie de façon formelle. Dans ce chapitre nous allons décrire les différentes phases de conception de l'ontologie ONTOPARON. Dans le chapitre 2, nous avons défini les différents niveaux de parcours (parcours de vie, parcours de soins, parcours de santé) sur lesquels les coordinateurs vont d'une part saisir des informations mais aussi réaliser un accompagnement auprès des personnes ayant une SLA. Le suivi et l'accompagnement proposés sont holistiques, c'est-à-dire prenant en compte l'ensemble des dimensions intervenant dans le parcours du patient, qu'il s'agisse de la dimension médicale liée à la pathologie, de la dimension socio-environnementale prenant en compte le patient dans son milieu de vie, l'ensemble s'articulant sur des actions et des préconisations mises en place par les coordinateurs. Pour comprendre les éléments intervenant dans les parcours, il est nécessaire de modéliser ces trois dimensions dans l'ontologie. Si initialement nous avons construit l'ontologie sous un modèle monolithique, rapidement nous avons vu les limites de ce modèle et avons fait le choix de développer une ontologie modulaire, dans laquelle chaque dimension est représentée par un module ontologique différent.

### 5.1 Principes de construction de ONTOPARON

Pour concevoir l'ontologie, nous avons suivi la méthodologie générique illustrée par (Charlet *et al.*, 2006) qui combine une approche ascendante, avec la recherche de candidats termes à partir de corpus, et une approche descendante, par l'utilisation d'une top-ontologie. La figure 5.1 illustre les différentes étapes réalisées lors de la construction de ONTOPARON.

Comme nous l'avons vu au chapitre 1, les coordinateurs interviennent auprès des patients tout au long de leur parcours à la fois au domicile mais aussi sur les différents temps d'hospitalisations. Les champs d'intervention étant nombreux et transversaux, l'ontologie doit rendre compte de ces différents domaines.

L'ontologie permet d'élaborer, de normaliser et de standardiser la sémantique associée à l'information tout en permettant à l'utilisateur d'employer une terminologie qui lui est familière pour décrire cette information. On utilise pour cela des concepts et des relations.

Pour construire l'ontologie nous avons extrait les candidats termes à partir d'un extracteur de terme (BioTex), fait un travail de normalisation qui a abouti à une première ontologie monolithique que nous avons ensuite décomposée en modules.

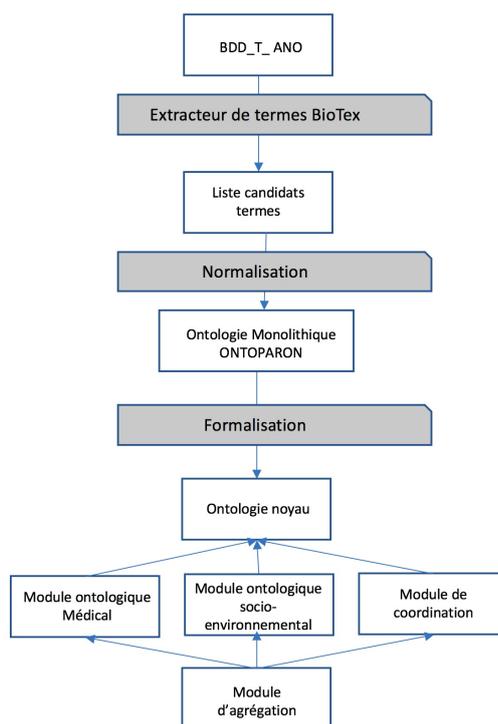


FIGURE 5.1 – Processus de construction de ONTOPARON.

## 5.2 Extraction des candidats termes avec *BioTex*

L'extraction de termes à partir d'un corpus spécifique est une étape importante pour l'acquisition des connaissances du domaine. Certaines méthodes permettent d'extraire automatiquement les termes pertinents d'un corpus (Lossio-Ventura *et al.*, 2014b). L'utilisation d'un outil de TALN est une aide à l'interprétation humaine, qui s'appuie sur la forme, sur des critères numériques et sur les distributions pour faire les propositions. Cependant, l'approche humaine est nécessaire pour passer à un niveau sémantique, c'est-à-dire donner du sens au terme (Aussenac-Gilles, 2005).

Pour réaliser l'extraction de termes des corpus du Réseau SLA IDF, nous avons utilisé le logiciel *BioTex*. Cet outil recourt à des méthodes linguistiques basées sur des patrons syntaxiques et des mesures statistiques. L'extraction de termes à partir d'un corpus peut être décomposée en différentes étapes (Lossio-Ventura *et al.*, 2014a) comprenant une annotation morpho-syntaxique (Part-of-speech tagging), l'extraction des candidats termes et leur classement. Les termes (syntagmes nominaux pouvant être composés d'une à quatre unités) vont être classés en fonction de leur fréquence d'apparition dans le corpus. Il est alors nécessaire de sélectionner les termes pertinents. Cette sélection peut se faire sur un critère quantitatif, en prenant en compte uniquement la fréquence, ou bien par l'expertise de l'ontologue et ou de l'expert, en décidant de prendre des termes ayant éventuellement une fréquence faible mais étant essentiels au domaine (Aimé, 2015).

Pour pouvoir utiliser le logiciel et obtenir des résultats optimaux, il est nécessaire de fournir au système un corpus. [Aimé \(2015\)](#) et [Bourigault & Aussenac-Gilles \(2003\)](#) proposent de suivre quelques recommandations lors de la création du corpus :

**Disponibilité des ressources** : dans le cadre de notre travail, nous avons accès à tout moment à la base de données du Réseau SLA IDF, ce qui nous a permis de réaliser autant d'exports que nécessaire. Plusieurs exports des données ont été réalisés à des moments différents en 2016, 2017 et 2018. [Aimé \(2015\)](#) propose de décliner cette notion de disponibilité en trois sous-critères :

1. **Un critère de connaissance**, qui nécessite la mise en place d'un travail en partenariat avec les experts du domaine. Cette collaboration fut mise en place, dans notre travail, à différents moments et spécifiquement lors de la phase d'évaluation des annotations sémantiques, comme indiqué au chapitre 7.
2. **Un critère légal**, qui prend en compte la pseudonymisation des données et l'enregistrement de la base auprès de la [CNIL](#). La base du Réseau SLA IDF répond aux critères administratifs puisqu'elle a été enregistrée auprès de la [CNIL](#). Le travail de pseudonymisation a été réalisé comme indiqué au chapitre 4.
3. **Un critère social**, qui nécessite d'avoir une connaissance de l'écosystème.

**Un corpus suffisamment consensuel** : dans notre cas d'usage, le corpus utilisé vient uniquement de la base du Réseau SLA IDF. Il est adapté à l'objectif spécifique de comprendre les causes de rupture de parcours de soins de patients atteints de [SLA](#) suivis par le réseau Réseau SLA IDF. La consensualité de notre corpus ne s'applique qu'au cadre de la coordination au sein de cette structure. Il pourrait être remis en question par d'autres experts réalisant d'autres formes de coordination.

**Une taille convenable** : elle est requise afin que le corpus puisse couvrir largement le domaine étudié. Pour répondre à ce critère, nous avons extrait l'ensemble des événements rédigés entre 2005 et 2015, soit 31 000 événements, dans l'objectif d'obtenir un maximum de représentativité des termes et syntagmes nominaux présents dans le corpus.

### 5.2.1 Résultats de l'extraction

[BioTex](#) a produit quatre fichiers contenant les candidats termes ([Bourigault et al., 2004](#)), classés par fréquence d'apparition. Le premier fichier contient les termes unigrammes (composé d'une seule unité), le second les termes bigrammes (composé de deux unités), le troisième les 3-grammes (composé de quatre unités) et enfin le dernier fichier contient tous les termes de plus de 4-grammes. Le tableau 5.1 illustre par des exemples (pris de manière aléatoire) les termes obtenus dans chacun des fichiers, ainsi que le nombre total de candidats termes obtenus par fichier. Nous pouvons y observer que certains symboles ou abréviations peuvent être utilisés dans les corpus :  $\emptyset$  ; *frm* (*pour fauteuil roulant manuel*) ou bien par exemple *pec kiné* (*pour prise en charge en*

TABLEAU 5.1 – Illustration des listes de candidats termes obtenue par *BioTex* .

unigramme	bigramme	3-grammes	4-grammes
<b>appel</b>	médecin traitant	prise en charge	filles de la patiente
patiente	fausses routes	mise en place	aide à la toilette
<b>demande</b>	assistante sociale	auxiliaire de vie	troubles de la déglutition
vie	aide humaine	séjour de répit	<b>appel de la fille</b>
frm	compléments alimentaires	perte de poids	échange sur la situation
kiné	pec kiné	sejour de repit	demande de séjour de répit
cmd	certificat médical	appel de mme	fausses routes aux liquides
oxymétrie	med ttt	gaz du sang	øtroubles du sommeil
17 177	19 782	24 399	19 304

*kinésithérapie*). Au-delà de l’aspect quantitatif indiquant la fréquence d’apparition, les listes de termes nous permettent d’étudier et d’identifier les différents sens que peut avoir un terme dans le domaine étudié, et identifier les relations entre eux ([Aussenac-Gilles, 2012](#)). Dans la première liste, nous pouvons voir les termes *appel* et *demande* qui tous deux ont une fréquence importante. La liste de 4-grammes permet d’identifier le contexte d’apparition et la relation entre les termes. Ainsi le terme *appel* fait référence à une action qui est liée à une personne, comme dans l’exemple *appel de la fille*.

La première liste permet de repérer les termes les plus représentatifs dans les corpus, cependant, le sens peut être large. Le terme *vie* est présent dans la première liste, et la liste à 3-grammes permet d’identifier et de relier *vie* au terme *auxiliaires de vie*, mais d’autres situations peuvent utiliser le terme *vie*, par exemple le terme *fin de vie*. L’utilisation des quatre listes nous permet d’obtenir de nombreuses informations pour contextualiser l’utilisation du terme et ainsi initier un travail de catégorisation sémantique.

### 5.2.2 Sélection des termes de *BioTex*

Pour construire la hiérarchie de l’ontologie, il est nécessaire de sélectionner à partir des candidats termes, les éléments de connaissances spécifiques au domaine de la coordination de parcours de patients ayant une *SLA*. Le domaine traité est vaste (coordination de soins, pathologie neuro-dégénérative, aspects médicaux-sociaux et aspects environnementaux, etc.). Une première observation des listes de *BioTex* permet de voir que certains termes font référence à des personnes : *patiente*, *kiné*, *auxiliaire de vie*, *assistante sociale*. D’autres vont correspondre à des documents ou des objets : *certificat médical*, *compléments alimentaires*, *fauteuil roulant*; et d’autres à des actions ou des processus *mise en place*, *perte de poids*, *demande de séjour de répit*, *aide à la toilette*. Ces premières observations ont permis de réaliser un premier niveau de catégorisation comprenant :

1. **Les agents**, pour les termes faisant référence aux personnes quels que soient leurs rôles et leurs fonctions.
2. **Les actions**, pour les termes en lien avec la réalisation d’une action quel que soit l’agent effecteur.

3. **Les états**, pour les termes faisant référence à un état comme les *signes* et *symp-tômes* des pathologies, ou bien encore *l'état civil* de la personne ou son *état de santé*.
4. **Les lieux**, pour les termes faisant référence à des lieux qu'ils soient géographiques : comme les *villes*, un *pays*, l'adresse du *domicile*, institutionnel comme un *hôpital*, un *service de soins*, etc.
5. **Les objets**, pour les objets physiques énumérés dans les corpus comme les *aides techniques*, les *médicaments* etc.

La coordination de parcours de santé consiste en l'interaction d'*agents* (patients, entourage, neurologues, coordinateur SLA etc.) réalisant des *actions* (communication, sollicitation, mise en place, évaluation etc.) à destination d'autres *agents* afin d'intervenir sur des *états* ou des *objets* dans des *lieux* spécifiés. Pour sélectionner les candidats termes, nous avons choisi les plus représentatifs, c'est-à-dire ayant une fréquence d'apparition importante dans le corpus. Nous avons décidé de retirer les termes ayant une fréquence faible et ceux pour lesquels la compréhension n'était pas possible sans le contexte dans lequel ils avaient été écrits comme par exemple « x8 ; vii ; plae, etc. » Dans un deuxième temps un travail a été réalisé avec les neurologues pour étudier les candidats termes ayant une fréquence plus faible mais qui peuvent être importants d'un point de vue clinique et qu'il est nécessaire de prendre en compte. Comme le rappelle [Aussenac-Gilles \(2005\)](#), les concepts vont rendre compte du domaine et de l'usage dans lesquelles ils sont utilisés.

### 5.2.3 Sélection des termes et antériorité du travail

Le travail mené dans le cadre de cette thèse s'inscrit dans la continuité d'un travail antérieur mené en Master 2. Lors de cette première phase, une analyse manuelle des corpus, nous avait conduit à structurer les connaissances sous forme de huit catégories principales (médicale, sociale, aide technique, transmission etc.), dans un **SOC** de type glossaire, permettant d'identifier dans un premier temps la thématique de l'évènement saisi par les coordinateurs. Chacune de ces huit catégories était subdivisée en sous-catégories. Nous souhaitons illustrer les éléments faisant partie de la catégorie médicale, car cette classification a été utilisée en partie lors du travail de normalisation sémantique et la mise en place des principes différentiels proposé dans la méthode ARCHONTE. Les 14 sous-catégories faisaient référence à la typologie des symptômes :

**Respiratoire** pour l'ensemble des éléments concernant l'appareil respiratoire liés à la SLA.

**Nutrition** pour l'ensemble des informations liées à l'apport des nutriments.

**Douleur** regroupant l'ensemble des sensations pénibles, désagréables, ressenties dans une partie du corps et exprimées par le patient.

**Troubles psychologiques et cognitifs**, pour les signes et symptômes liés à une modification ou altération de l'état psychologique ou cognitif de la personne.

**Troubles du sommeil** , pour l'ensemble des termes faisant référence à la modification ou à l'altération du sommeil nocturne.

**Urinaire** concernant les fonctions d'élimination de l'urine.

**Digestif** pour les modifications faisant référence aux processus d'assimilation et de dégradation des aliments.

**Moteur** pour les signes et symptômes concernant les capacités et incapacités motrices de la personne.

**Fatigue** pour l'ensemble des termes se rapportant à une faiblesse générale.

**Cutané vaso-moteur** , pour l'ensemble des termes spécifiant une atteinte de l'intégrité de la peau.

**Général** pour l'ensemble des symptômes et maladies intercurrentes.

**Fin de vie** pour spécifiée les situations de traitements et soins palliatifs.

**Bilans et constantes biologiques** pour l'ensemble des examens et résultats d'examen renseignant l'état de santé du patient.

**Traitement à visée étiologique** concernant les informations faisant référence de spécifiquement au traitement de la SLA que sont l'utilisation du Rilutek et Toco.

Cette première classification fut une aide, dans la phase de regroupement des termes issus des listes obtenues par [BioTex](#). Cependant, cela a nécessité de restructurer les connaissances en les modélisant de façon plus élaborée en changeant de paradigme, passant d'un modèle de communication à un modèle de structuration des connaissances plus formalisé, ne laissant pas de possibilité à l'ambiguïté. Dans le travail initial, la lecture et le traitement et l'annotation des événements étaient fait de façon manuelle. La lecture par un être humain permet de comprendre les éléments implicites, ce qui n'est pas permis lors du traitement des corpus par une machine. Si la première classification réalisée a été une aide, il a cependant fallu la restructurer pour qu'elle puisse répondre aux critères des ontologies.

## 5.3 Construction de l'ontologie

Dans le chapitre 3, nous avons présenté les méthodologies de construction d'ontologies. Au cours de ce travail, nous nous sommes inspirés de la méthode ARCHONTE, dont nous avons suivi les principales étapes.

L'objectif de la normalisation sémantique est de parvenir à un consensus afin de définir les étiquettes utilisées pour nommer les concepts ([Charlet et al., 2006](#)). Pour cela, il est important de définir et d'identifier les similitudes et les différences de chaque terme. Pour chaque terme obtenu à partir de la phase d'extraction des candidats termes, nous avons réalisé un travail de spécification et de définition, qui a suivi plusieurs étapes :

**Regroupement des termes pour un même concept.** Nous avons tout d'abord recherché dans les corpus l'ensemble des synonymes, acronymes, abréviations désignant un même concept. Au regard des informations contenues dans le tableau 5.1, il est possible de voir que certains concept vont avoir des formes orthographiques différentes. Ainsi *med ttt* est une forme d'abréviation pour le

FIGURE 5.2 – Exemple des `skos:preflabel`, `skos:altlabel` du concept de fauteuil roulant manuel.

concept de *médecin traitant*. Les évènements saisis par les coordinateurs du Réseau SLA IDF ont pour spécificités, qu'un même concept peut être exprimé par différents termes mais aussi par différentes abréviations ou symbole, sans qu'il y ait réellement de consensus sur leurs utilisation entre les coordinateurs. Ainsi le concept *Médecin Traitant* peut par exemple être écrit de huit manières différentes : *méd traitant*, *médecin de famille*, *généraliste*, *MT*, *MDT*, *med tt*, *méd traitant* ou encore *med ttt*.

Si l'utilisation d'abréviations et d'acronymes est très fréquente dans le corpus du Réseau SLA IDF, une spécificité est apparue dans le cadre des aides techniques et de certaines suppléments nutritionnelles. En effet pour désigner certaines aides techniques comme le « fauteuil roulant manuel », les coordinateurs peuvent utiliser ; la forme complète *fauteuil roulant manuel*, une abréviation *FRM*, *Fr manuel*, ou bien utiliser la désignation commerciale du produit *Action 3*. Des exemples similaires apparaissent fréquemment : *mytoobi* pour désigner une *synthèse vocale*, *Aquatec* pour un fauteuil de douche ou encore *Fresubin* pour désigner les suppléments oraux. Nous avons fait le choix, au regard de ces éléments, de créer des `skos:preflabel` pour les concepts et de placer en `skos:altlabel` l'ensemble des synonymes, abréviations et désignations commerciales comme indiqué dans la figure 5.2.

**Définir et structurer les concepts entre eux.** Le regroupement des termes de l'étape précédente a permis la réalisation d'un premier niveau de structuration sous forme d'une taxinomie contenant : les agents, les processus, les objets etc. Pour réaliser cette structuration et la formalisation des connaissances, nous nous sommes

inspirés des classifications et ressources termino-ontologiques existantes et disponibles que sont la CIF<sup>1</sup> pour organiser les activités de vie quotidienne, la Nurse Care Coordination Ontology (NCCO)<sup>2</sup> pour les activités de coordination, ONTOLURGENCES<sup>3</sup> pour structurer les éléments médicaux comme les signes et symptômes et ONTOPSYCHIA<sup>4</sup> (Richard *et al.*, 2015) pour certains composants environnementaux. Pour développer cette taxinomie, nous avons utilisé l'éditeur d'ontologie Protégé<sup>5</sup> (Musen, 2015).

**Alignement des concepts.** Pour réaliser l'alignement avec les terminologies existantes nous avons utilisé la plateforme Health Terminology Ontology Portal (HeTOP)<sup>6</sup> (Grosjean *et al.*, 2014) de l'équipe du CHU de Rouen, qui contient 56 terminologies et ontologies du domaine médical, représentant 1 951 834 concepts et 6 636 654 termes. Cette plateforme contient de nombreuses ressources, mais elle ne recouvre pas toutes les ontologies présentes sur Bioportal<sup>7</sup>, qui rassemble 783 ontologies et 9 434 599 classes. Nous avons utilisé ces deux outils de façon complémentaire : HeTOP nous a permis d'utiliser des définitions et d'attribuer des codes (UMLS, MesH etc.) existants pour certains concepts, et de proposer un terme préféré en anglais ; Bioportal nous a permis d'accéder à des ressources, de voir la structuration et l'organisation des classes de ressources non disponibles sur HeTOP.

Pour chaque concept de ONTOPARON, nous avons recherché dans la plateforme HeTOP les identifiants Unified Medical Language System (UMLS) ou bien l'identifiant de la classe à laquelle le concept défini comme identique dans notre domaine faisait référence. Pour cela nous avons créé une annotation `rdfs:seeAlso` pour les différentes classifications que nous avons utilisées :

1. *CIF-Id* pour les concepts présents dans la Classification Internationale du Fonctionnement, du handicap et de la santé<sup>8</sup> ;
2. *MeSH-Id* pour les termes présents dans le Medical Subject Headings en version française<sup>9</sup> ;
3. *Cladimed-Id* pour les codes issus de la Classification des dispositifs médicaux<sup>10</sup> ;
4. *NCIt-Id* pour le National Cancer Institute Thesaurus<sup>11</sup> ;
5. *ICNP-Id* pour International Classification for Nursing Practice<sup>12</sup> ;

- 
1. <https://bioportal.bioontology.org/ontologies/ICF>
  2. <https://bioportal.bioontology.org/ontologies/NCCO>
  3. <https://bioportal.bioontology.org/ontologies/ONTOLURGENCES>
  4. <https://bioportal.bioontology.org/ontologies/ONTOPSYCHIA>
  5. <https://protege.stanford.edu>
  6. <https://www.hetop.eu/hetop/>
  7. <https://bioportal.bioontology.org>
  8. [https://www.hetop.eu/hetop/?rr=CIF\\_CO\\_ARBO&tab=1](https://www.hetop.eu/hetop/?rr=CIF_CO_ARBO&tab=1)
  9. [https://www.hetop.eu/hetop/?rr=MSH\\_D\\_ARBO&tab=1](https://www.hetop.eu/hetop/?rr=MSH_D_ARBO&tab=1)
  10. [https://www.hetop.eu/hetop/?rr=CLA\\_FA\\_ARBO&tab=1](https://www.hetop.eu/hetop/?rr=CLA_FA_ARBO&tab=1)
  11. [https://www.hetop.eu/hetop/?rr=NCI\\_CO\\_V-NCI&tab=1](https://www.hetop.eu/hetop/?rr=NCI_CO_V-NCI&tab=1)
  12. [https://www.hetop.eu/hetop/?rr=ICN\\_F\\_10014477&tab=1](https://www.hetop.eu/hetop/?rr=ICN_F_10014477&tab=1)

## 6. UMLS-Id<sup>13</sup>.

A la fin de cette première étape de normalisation sémantique, nous avons une structuration des concepts dans une seule ontologie. L'ontologie obtenue était conséquente en nombre de classes (plus de 2000). Lors du travail de collaboration avec les experts, il était difficile d'avoir une vision globale et pragmatique des concepts présents dans l'ontologie. Cette difficulté de partage, et de taille, nous ont orienté vers la modularité ontologique. Les critères de taille de l'ontologie et de manutention sont des critères favorisant la modularité ontologique, comme le rappelle Parent & Spaccapietra (2009).

## 5.4 Passage d'une ontologie monolithique à une ontologie modulaire

La définition des concepts et l'étape de formalisation des connaissances en sémantique formelle nous ont fait changer de point de vue de modélisation pour structurer l'ontologie sous forme modulaire (Cardoso *et al.*, 2018). Pour conceptualiser les domaines des modules que nous devons modéliser, nous avons pris comme point de départ et point central la personne qui est au coeur de la situation. Il nous a semblé que l'accompagnement et la prise en charge de la personne étaient réalisées à deux niveaux : le premier dans la dimension médicale, car la personne est atteinte d'une maladie neuro-dégénérative, elle y a un statut de « patient » ; le second niveau dans la dimension sociale car le patient accompagné est une personne qui a une culture, des habitudes de vie, un logement, et pour laquelle la pathologie va créer des incapacités et des situations de handicap etc. Nous avons souhaité illustrer ces dimensions dans la figure 5.3. La coordination vient s'inscrire dans cette schématisation par l'accord de la personne à être incluse dans le Réseau SLA IDF, pour bénéficier d'un accompagnement dans les aspects médicaux mais aussi sociaux. Il n'existe pas de dichotomie dans l'accompagnement de la personne cependant, en France, la structuration de la prise en charge se fait sur trois plans, d'un côté le sanitaire, le social et de l'autre le médico-social.

A partir de cette schématisation en trois domaines, nous avons décomposé l'ontologie monolithique initiale en trois modules ontologiques.

1. Un **module médical** faisant référence aux concepts liés aux aspects médicaux de la pathologie et de la prise en charge médicale, dénommé ONTOPARONMED.
2. Un **module socio-environnemental** concernant les éléments liés à l'environnement personnel du patient (domicile, travail, projet de vie, famille, etc.), les institutions et professionnels des domaines sociaux et médico-sociaux intervenants dans le parcours de santé des patients, ainsi que les éléments liés à la prise en charge sociale, dénommé ONTOPARONSOC.
3. Un **module de coordination** liés aux activités et missions de coordination réalisées par les coordinateurs du Réseau SLA IDF dénommé ONTOPARONCOORD.

---

13. [https://www.hetop.eu/hetop/?rr=UML\\_SG\\_ARBO&tab=1](https://www.hetop.eu/hetop/?rr=UML_SG_ARBO&tab=1)

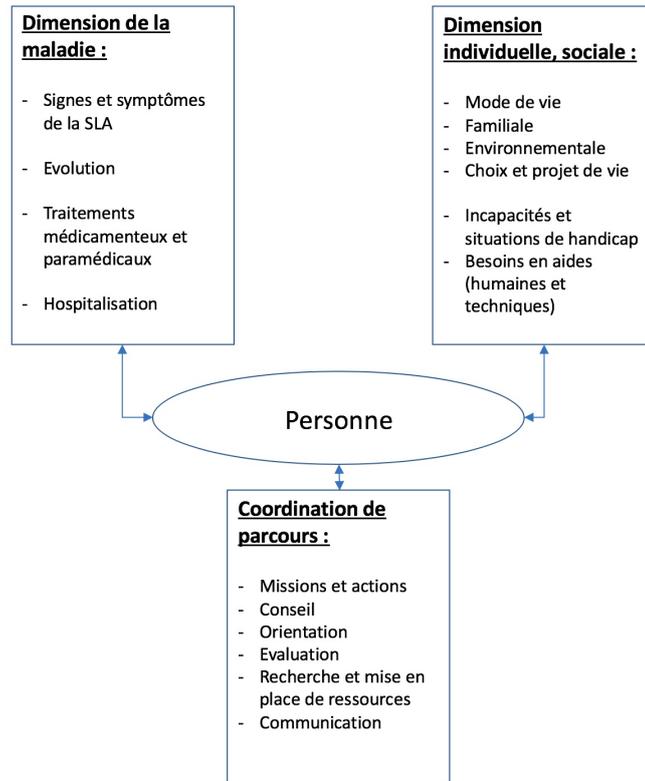


FIGURE 5.3 – Schématisation des dimensions intervenant dans l'accompagnement de personne ayant une SLA.

Le passage d'une ontologie monolithique à une ontologie modulaire nécessite l'utilisation d'une ontologie noyau. Ainsi trouver des concepts qui subsument des actes de différents professionnels médicaux ou de coordination requiert la notion d'acte non spécifique. Et ce type de constat se répétant  $n$  fois, l'ontologie noyau devient indispensable. Au laboratoire LIMICS nous avons à disposition une ontologie noyau subsumée par une top-ontologie qui est utilisée habituellement, TOP-MENELAS<sup>14</sup> et c'est celle-ci qui a été retenue pour ONTOPARON avec quelques aménagements. Le module noyau est dénommé dans ONTOPARON : ONTOPARONNOY. Au-delà de la nécessité d'utiliser une ontologie noyau qui corresponde aux concepts communs aux trois modules (et qui est importée par chacun d'eux), il est nécessaire d'avoir un module d'agrégation qui rassemble tous les modules. Le module de consolidation ne modélise aucun concept, il ne sert qu'à agréger. La figure 5.4 schématise les liens d'importation de nos principaux modules. Au total nous avons cinq modules ontologiques qui constituent l'ontologie finale ONTOPARON.

Aimé *et al.* (2014) propose de respecter des règles lors de la création d'ontologie modulaire :

14. <http://bioportal.bioontology.org/ontologies/TOP-MENELAS>

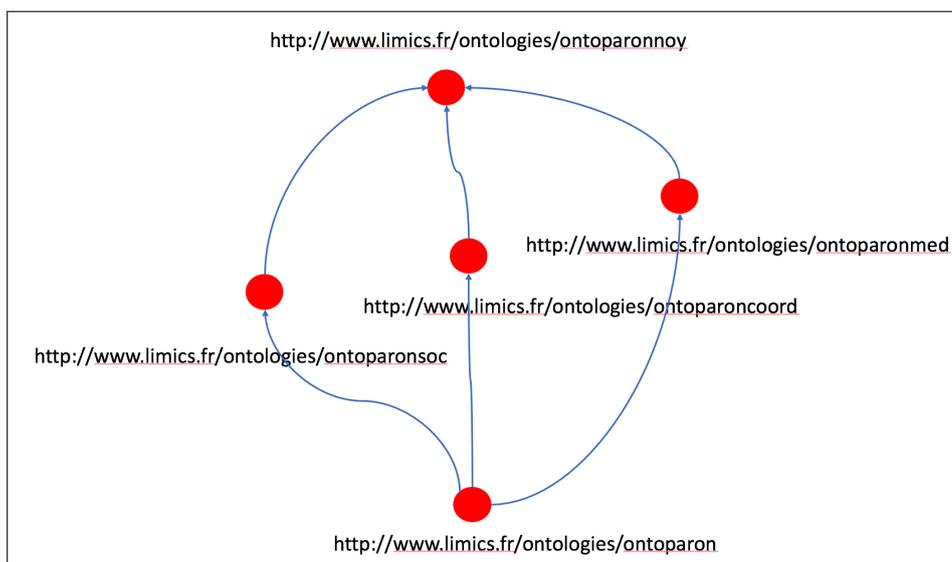


FIGURE 5.4 – Vue d’ensemble du schéma d’héritage des concepts d’ONTOPARON et les URI correspondantes. Les ontologies correspondant aux annotations de l’ontologie et des concepts ne sont pas représentées (*e.g.* SKOS ou *Dublin Core*). Les processus d’importation se font du bas vers le haut, la flèche reliant l’ontologie important à l’ontologie importée.

- tout module spécifique est autonome dans sa conceptualisation, dans le sens où les concepts de ce type de module n’ont de relations de (subsumption ou autre) avec aucun concept appartenant à un autre module (hormis *core-ontology*) ;
- tout concept d’un module spécifique doit pouvoir être défini en langage naturel et autant que possible en langage formel ;
- tout concept d’un module spécifique est spécifique au sous-domaine modélisé dans ce module ;
- tout concept d’un module spécifique possède au moins un label en français ;
- au sein d’un module spécifique, tout label dans une langue donnée ne dénote qu’un seul concept ;
- toute relation d’un module spécifique est sous-relation des relations terminales de la *core-ontology* ;
- toute relation d’objet d’un module spécifique a pour domaine et co-domaine un concept du même module spécifique.

Pour valider en partie ces règles nous avons utilisé des requêtes SPARQL (une partie des requêtes sont présentées en annexe C) et l’outil CORESE<sup>15</sup>. Nous avons utilisé ces règles pour chaque module. En présence d’erreurs, nous avons apporté les corrections nécessaires. Comme nous l’avons indiqué précédemment, nous avons ré-utilisé des concepts présents dans d’autres ontologies. Afin de respecter ces imports de

15. COnceptual REsource Search Engine

classes nous avons gardé l'URI<sup>16</sup> de l'ontologie initiale à partir de laquelle le concept était extrait. Nous obtenons donc 8 URI différentes dans l'ontologie en plus de celle de ONTOPARON :

- [limics.fr/ontologies/menelastop](http://limics.fr/ontologies/menelastop), lorsque les concepts sont importés de l'ontologie MÉNÉLAS ;
- [limics.fr/ontologies/ontopsychia](http://limics.fr/ontologies/ontopsychia), lorsque les concepts viennent de ONTOPSYCHIA ;
- [limics.fr/ontologies/ontolurgences](http://limics.fr/ontologies/ontolurgences), pour les concepts provenant de ONTO-LURGENCES.

### 5.4.1 Les concepts définis

L'organisation des connaissances sous forme d'ontologie, nous permet de créer des concepts définis (*equivalent class*). L'utilisation du raisonneur<sup>17</sup> dans Protégé, permet d'inférer sous le concept définis, l'ensemble des concepts liés par une même relation. ONTOPARON est une ontologie qui contient au total 2446 classes. L'étude des annotations de chacune de ces classes, ne nous semblait pas pertinente à mettre en place, dans le cadre de la compréhension de parcours de santé. Des classes peuvent être présentes dans différents modules, mais font partie de la même thématique lorsque l'on analyse la situation. Ainsi par exemple, *l'épuisement de l'aidant* est décrit dans le module socio-environnementale, *la recherche de structure de répit* est une des activités des coordinateurs du Réseau SLA IDF, cependant la rédaction du document pour demander l'hospitalisation est du ressort du médecin. D'un point de vue théorique, il est aisé de comprendre la présence de ces classes dans les différents modules, cependant dans une vision clinique et pragmatique ces 3 classes se rejoignent et font parties de la même thématique, celle du *domaine de l'épuisement*.

L'objectif de notre travail étant de comprendre les parcours de santé des patients, nous souhaitons avoir la possibilité d'avoir, à la fois une vision globale pour tenter de répondre à certaines questions par exemple existe t-il de nombreuses situations d'épuisement dans le parcours des patients ? ; tout en ayant la possibilité d'avoir une vision plus fine sur certaines thématiques comme par exemple, la demande et la recherche de kinésithérapeute est elle une activité récurrente pour les coordinateurs du Réseau SLA IDF ? Pour cela nous avons créés différents concepts définis. Des concepts définis de haut niveaux, qui sont communs aux différentes ontologies du domaine, présents dans l'ontologie noyau, par exemple *Domaine de l'épuisement*. Et des concepts définis spécifiques à chaque domaines par exemple *Actions de coordination* dans ONTOPARONCOORD ; *État nutritionnel* dans ONTOPARONMED, ou bien encore *Processus social* dans ONTOPARONSOC. L'ensemble des concepts définis créés sont indiqués à l'annexe D.3.

Nous avons créé dans les *ObjectProperty* la relation **a\_pour\_thématique** qui a pour domaine Thing et pour Range (co-domaine) : Notion Thématique. Nous avons

---

16. Uniform Resource Identifier

17. Le raisonneur que nous avons utilisé est Hermit 1.3.8.413.

```

<!-- http://www.limics.fr/ontologies/ontoparonnoy#DomaineEpuisement -->
<owl:Class rdf:about="http://www.limics.fr/ontologies/ontoparonnoy#DomaineEpuisement">
  <owl:equivalentClass>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
        <rdf:Description rdf:about="http://www.limics.fr/ontologies/
          menelastop#ObjetAbstrait"/>
        <owl:Restriction>
          <owl:onProperty rdf:resource="http://www.limics.fr/ontologies/
            ontoparonnoy#aPourThematique"/>
          <owl:someValuesFrom rdf:resource="http://www.limics.fr/ontologies/
            ontoparonnoy#ThematiqueEpuisement"/>
        </owl:Restriction>
      </owl:intersectionOf>
    </owl:Class>
  </owl:equivalentClass>
  <rdfs:comment xml:lang="fr">concerne de manière spécifique les éléments faisant
    référence à l'&apos;épuisement des aidants.</rdfs:comment>
  <skos:prefLabel xml:lang="fr">domaine de l'&apos;épuisement</skos:prefLabel>
  <skos:prefLabel xml:lang="en">field of exhaustion</skos:prefLabel>
</owl:Class>

```

FIGURE 5.5 – Illustration des caractéristiques et des restrictions pour le concept défini *DomaineEpuisement*

créé 10 notions thématiques comme indiqué en annexe D.1. Tous les concepts définis sont construits de la même manière, avec des restrictions de propriétés, **a\_pour\_thematique someValuesFrom ThematiqueXX**. La figure 5.5 illustre ces caractéristiques pour le concept défini *DomaineEpuisement*.

### 5.4.2 Le module noyau : ONTOPARONNOY

L'objectif de l'ontologie noyau, comme indiqué dans la partie 2.2, est de décrire des concepts qui soient génériques et communs aux différents modules ontologies. Ainsi le concept *Agent* du module noyau, permet dans les autres modules de spécifier le domaine : *Agent médical* pour l'ensemble des intervenants dans le domaine médical ; *Agent Social* pour les personnes du domaine social (épouse, enfants, assistant social etc.) ; *Agent de coordination* pour spécifier les coordinateurs du Réseau SLA IDF par exemple. Ce modèle s'est répété pour l'ensemble des concepts des différents modules. Nous avons choisi de définir les concepts définis de haut niveau dans l'ontologie noyau afin que l'ensemble des modules puissent en hériter lors de l'importation. Nous avons fait le choix d'utiliser comme ontologie noyau, MÉNÉLAS<sup>18</sup> qui fut développée au sein du laboratoire. Cependant des adaptations furent réalisées sur cette ontologie : en effet au cours de la formalisation nous nous sommes rendu compte que de nombreux concepts et relations n'étaient pas utiles dans notre travail, de plus le temps d'appropriation de l'ensemble des concepts était complexe, engendrant une difficulté dans l'utilisation de certains concepts. La figure 5.6, illustre les principaux concepts présents dans ONTOPARONNOY.

Le tableau 5.2 présente les métriques des différents modules de ONTOPARON. On peut observer de grandes disparités entre le nombre de classes par module ontolo-

18. <https://bioportal.bioontology.org/ontologies/TOP-MENELAS>

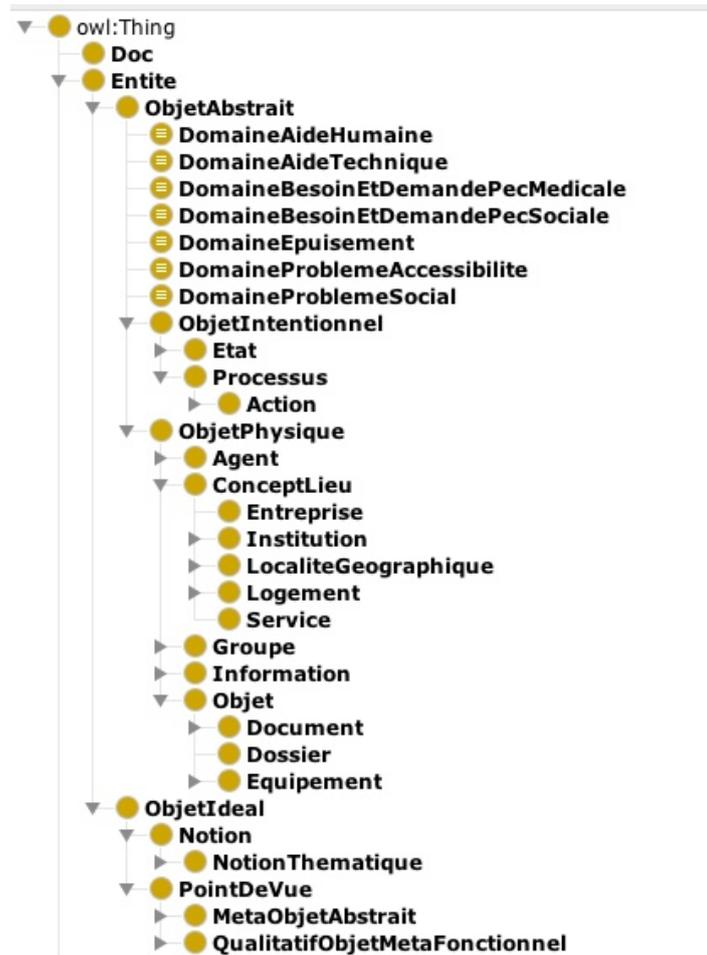


FIGURE 5.6 – Visualisation des concepts du module noyau. Les concepts définis sont identifiés par le signe ≡ dans la visualisation de l'ontologie.

gique. Ces différences s'expliquent en grande partie par la quantité de connaissance à formaliser dans ces modules. Ainsi le module médical est le plus important car les connaissances doivent porter sur la pathologie, les traitements, les processus, les agents médicaux, les états de santé et les différentes structures pouvant intervenir etc.. Le second module en terme de quantité de classes est le module social, là aussi les structures, les processus et les agents sont nombreux. Chacun de ces modules sera présenté, dans la partie suivante.

De même que l'ontologie noyau possède les concepts communs aux différentes ontologies, elle possède des relations, des propriétés d'objets qui permettent de lier des concepts entre eux afin de formaliser la connaissance. Le module noyau comporte l'ensemble des propriétés d'objets qui sont utilisées dans les différents modules de l'ontologie. Les propriétés d'objets sont au nombre de 32. La figure 5.7 illustre les différentes propriétés d'objets contenue dans l'ontologie. Pour illustrer les relations nous pouvons citer deux exemples :

1. **a\_pour\_projet\_de\_vie** : cette relation relie le concept *patient* au concept de

TABLEAU 5.2 – Métriques des différents modules de l'ontologie ONTOPARON

	Noyau	Médical	Social	Coordination	ONTOPARON
Nombre de classes	388	1052	751	321	2446
Nombre de relations	32	0	0	0	32
Concept défini	7	17	10	9	43



FIGURE 5.7 – Visualisation des relations, définies dans le module noyau.

*projet de vie*. Elle permet, d'exprimer la relation *un patient a pour projet de vie un déménagement*.

2. **est\_realisé\_par** : permet de relier le concept *action* et *agent*, ainsi une action est réalisée par un agent. Dans l'ontologie noyau les relations sont très génériques, cependant dans les modules elles sont utilisées plus spécifiquement. Ainsi les *actions de coordinations* **est\_realisé\_par** *les coordinateurs du Réseau SLA IDF* et les *actions médicales* **est\_realisé\_par** *les agents médicaux*.
3. **a\_pour\_thématique** lie le concept à l'une des notions thématiques définies, permettant leurs inférences lors du raisonnement. Par exemple le concept *épuisement de l'aidant à pour thématique thématique épuisement*.

#### 5.4.2.1 Concepts définis présents dans ONTOPARONNOY

Les concepts définis créés, permettent de regrouper l'ensemble des concepts, quels que soient leurs natures (processus, objet, état), liés à leur thématique. Nous avons défini 7 thématiques principales : l'aide humaine, les aides techniques, les besoins et demandes liées à la prise en charge sociale, les besoins et demandes liées à la prise

en charge médicale, le domaine de l'épuisement, les problèmes d'accessibilité et les problèmes sociaux.

**Domaine des aides humaines :** permet de regrouper l'ensemble des concepts, liés aux aides humaines comme par exemple : Besoin en aide humaine, recherche d'auxiliaire de vie, financement des aides humaines, mise en place des aides humaines, etc.

**Domaine des aides techniques :** ce concept définis lie, l'ensemble des classes qui concernent les aides techniques comme par exemple : Les demandes d'aides techniques, les problèmes de confort, la mise en place des aides techniques, etc.

**Domaine des besoins et demandes de prise en charge médicale :** concerne les demandes d'hospitalisations, les demandes de prise en charge médicale, comme les demandes de consultation, ou demandes de certificat médicaux, les échanges sur les suppléances ou les directives anticipées.

**Domaine des besoins et demandes de prise en charge sociale :** permet de regrouper les concepts qui vont indiquer un besoin d'aide dans les démarches administratives mais aussi les demandes de documents sociaux, les demandes d'entrée en structure médico-sociale comme une Maison d'accueil spécialisé (MAS), une demande d'allocation (APA, PCH), etc.

**Domaine de l'épuisement :** lie les concepts qui vont indiquer directement ou indirectement la présence d'un épuisement qu'il soit de l'aidant familial ou des aidants professionnels. Nous y trouverons donc les concepts de demande de séjour de répit, l'épuisement de la prise en charge, les propositions de soutiens, etc..

**Domaine Problème accessibilité :** permet de lier les concepts qui vont indiquer la nécessité d'agir sur l'environnement, afin de l'adapter pour qu'il devienne accessible à la situation de handicap de la personne. Nous y trouverons les concepts de : logement inaccessible, aménagement de l'environnement (voiture, salle de bain, poste de travail, etc.).

**Domaine Problème social :** vont regrouper les concepts qui font état d'un problème de type social, comme l'isolement de la personne, la présence d'une maltraitance, un statut social qui peut être complexe (Sans domicile Fixe ; en situation irrégulière, etc.), l'ensemble des processus liés à la protection de la personne comme la sauvegarde de justice.

### 5.4.3 Le module médical : ONTOPARONMED

Ce module a été le premier à avoir été construit. Le travail de mené en Master 2, ainsi que la description dans la littérature (décrits dans le chapitre 1 de ce manuscrit) des symptômes dans la *SLA*, ont permis de d'organiser les concepts. Le module médical, tout comme les autres, importe ONTOPARONNOY sur lequel nous venons agréger et spécifier les concepts du domaine médical.

La figure 5.8 illustre une partie des concepts présents dans ONTOPARONMED. Il contient les connaissances sur :

**Les agents médicaux.** Qui sont l'ensemble des professionnels médicaux et paramédicaux, comme : les médecins ; les médecins spécialisés (pneumologue, neurologue, chirurgien etc.) ; les kinésithérapeutes, infirmiers, ergothérapeutes, diététiciens etc..

**Les objets du domaine médical.** Dans lesquels nous trouvons les dispositifs médicaux comme : les suppléances respiratoires, les suppléances nutritionnelles ; l'ensemble des compléments alimentaires qui peuvent être prescrits. Mais aussi les médicaments, les structures anatomiques qui permettent de situer et spécifier les atteintes des patients (Fracture du poignet, atteinte du membre inférieur droit etc.).

**Les processus médicaux.** Qui concernent, par exemple, la mise en place de suppléance (mise en route de la VNI, mise en place de la gastrostomie), la réalisation de soins (injection de toxine botulique, mise en place de sédation profonde et continue etc.).

**L'état de santé.** C'est-à-dire l'ensemble des signes, des symptômes, des troubles que l'on peut rencontrer dans la SLA, ainsi que les pathologies dont peut souffrir le patient (autres que la SLA) comme du diabète, un cancer etc..

Comme nous l'avons indiqué dans la partie 5.4.1, chaque module ontologique, contient des concepts définis. La figure 5.8, illustre certains de ces concepts définis du module médical. Nous avons créé des concepts définis sur les « états », afin de relier l'ensemble des signes, des troubles et des symptômes qui indiquent une altération d'une fonction. Ces classes sont proches de la première classification et de la description des symptômes que l'on trouve dans la littérature sur la SLA.

#### 5.4.3.1 Les concepts définis du module médical

**État Médical.** Regroupe l'ensemble des classes liées à l'état médical du patient que ces signes soient positifs (fièvre) ou négatifs (apyrétique). L'« EtatMedical » regroupe l'ensemble des classes spécifiées ci-dessous.

**État respiratoire.** Regroupe les signes et symptômes et pathologies intervenant dans le domaine de la fonction respiratoire par exemple : une plainte respiratoire, une pneumopathie et les signes cliniques respiratoires.

**État nutritionnel.** Regroupe l'ensemble des classes indiquant une altération de la fonction nutritionnelle comme la perte de poids, la dénutrition, etc..

**État moteur.** Permet de voir l'ensemble des signes qui expriment une altération de la fonction motrice de la personne, comme des troubles de préhension, de coordination et les troubles de l'équilibre etc.

**État cutané.** Indique la présence d'altération au niveau de la peau, comme l'escarre, des plaies etc.

**État psychique.** Permet de regrouper les classes indiquant la présence de signes comme l'anhédonie, l'apathie les rires et pleurs spasmodiques, etc.



FIGURE 5.8 – Visualisation des concepts présents dans le module médical.

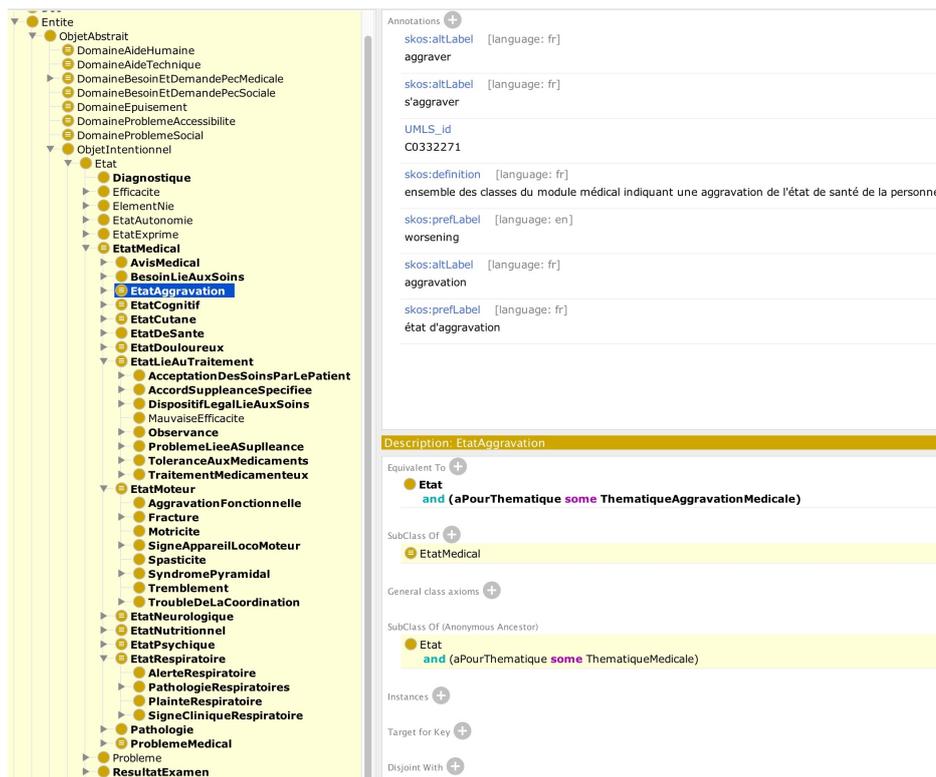


FIGURE 5.9 – Visualisation des concepts présents dans le module médical

**État cognitif.** Regrouper les classes indiquant une altération des fonctions cognitives de la personne, comme un trouble du comportement ou par exemple la présence d’une Démence Fronto-Temporales (DFT).

**État d’aggravation.** Ce concept a été créé car les corpus contiennent de nombreuses fois, le fait que le patient s’aggrave sans toujours le relier à une fonction spécifique. Nous pouvons ainsi rencontrer les termes « aggravation récente, aggravation progressive, aggravation importante ». Nous avons souhaité créer de façon spécifique cet état d’aggravation pour voir s’il est en lien avec d’autres situations.

**État lié au traitement.** Ce concept regroupe l’ensemble des classes qui font référence à la thématique des traitements, qu’il s’agisse des difficultés liées aux suppléances, ou le refus des soins par le patient.

#### 5.4.4 Le module socio-environnemental : ONTOPARONSOC

Le module socio-environnemental rassemble l’ensemble des concepts liés à la vie de la personne dans son environnement familial et social. Nous y trouvons par exemple :

**Les agents sociaux,** se sont les personnes de l’entourage du patient (épouse, conjoint, enfant, parent, etc.), les professionnels du domaine social et médico-social (les assistantes sociales, les auxiliaires de vie (avs), les aides ménagères, etc.).

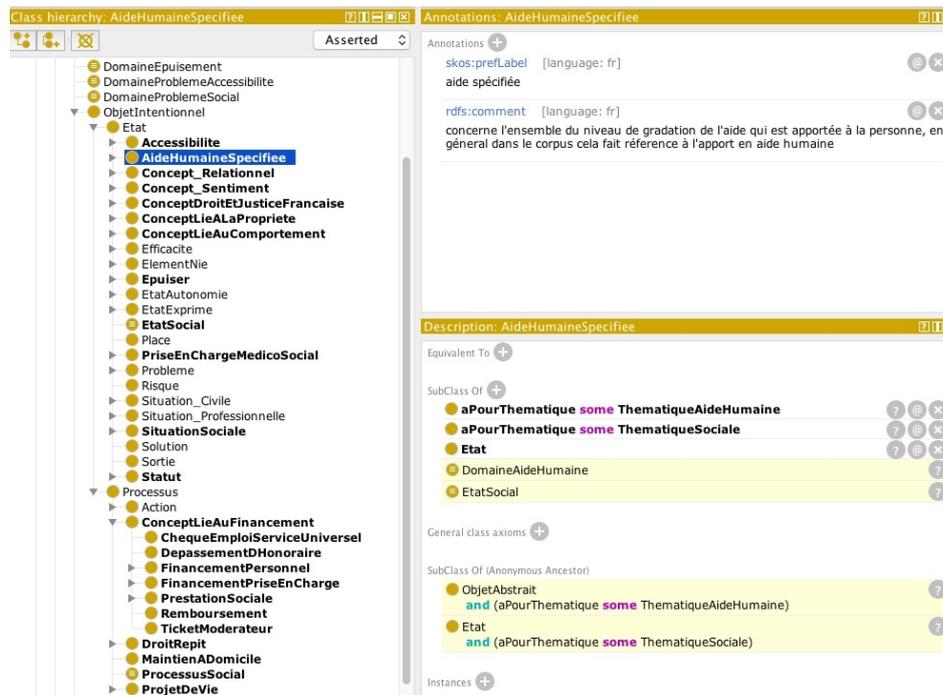


FIGURE 5.10 – Visualisation d'une partie des concepts présents dans le module ONTOPARONSOC

**Les lieux et structures du domaine social**, sont être le logement, les structures qui financent les allocations auxquelles la personne peut prétendre (MDPH, Conseil Général, mutuelle, etc.); les structures qui viennent en soutien ou qui peuvent intervenir à domicile (association d'auxiliaire de vie, SAMSAH, association de patient, etc.), et les structures et organismes de conseil en aides techniques (la Plate-Forme Nouvelles Technologies "PFNT", le Centre de Ressources et d'Innovation Mobilité Handicap "CEREMH", etc.).

**Les processus sociaux**, regroupe l'ensemble des processus liés aux prestations sociales (APA ; PCH, aide médicale d'état etc.), les demandes de carte d'invalidité, demande de logement, et les processus liés à l'aménagement du logement etc.

La figure 5.10 illustre une partie des concepts présent dans le module ONTOPARONSOC.

#### 5.4.4.1 Les concepts définis du module socio-environnemental

**État social** : lie l'ensemble des concepts qui indiquent un état social, comme le niveau d'autonomie, les concepts liés à la propriété (locataire, propriétaire), ceux liés aux problèmes financier, à la situation civile ou la situation sociale.

**Processus social** : regroupe les concepts qui indiquent un processus du domaine social, l'ensemble des actions sociales, les prestations sociale, les projets de vie.

**Action sociale** : regroupe l'ensemble des demandes liées à la thématique médico-sociale, comme les demandes d'évaluation, les demandes de documents, les dé-

TABLEAU 5.3 – Présentation des types d'activités de coordination inclus dans ONTOPARON et dans le Care Coordination Measures Atlas. Les activités de coordination sont regroupées par typologie.

ONTOPARON	Care Coordination Measures Atlas <sup>20</sup>
	Establish Accountability or Negotiate Responsibility
Action de communication	Communicate Facilitate Transitions
Évaluation de coordination Demande reçue en coordination FormationCoordination	Assess Needs and Goals, Create a Proactive Plan of Care, Monitor, Follow Up, and Respond to Change Support Self-Management Goals
Aligner les ressources aux besoins	Link to Community Resources Align Resources with Patient and Population Needs

marches sociales, les évaluations sociales faites par des équipes de l'APA ou de la MDPH.

#### 5.4.5 Le module de coordination : ONTOPARONCOORD

Le module de coordination modélise l'ensemble des connaissances liées aux activités de coordination réalisées par les coordinateurs du Réseau SLA IDF. Pour modéliser ces actions, nous nous sommes inspirés de la NCCO et du document Care Coordination Measures Atlas McDonald *et al.* (2014)<sup>19</sup> proposé par l'« Agency for Healthcare Research and quality ». Dans ce dernier il est proposé 9 types d'activités de coordination. Le tableau 5.3 positionne les actions de coordination de ONTOPARON et celles proposées par le Care Coordination Measures Atlas. Les thématiques entre les deux modélisations sont proches, cependant une différence s'observe, dans ONTOPARON la notion d'établir ou négocier la responsabilité de la défaillance des soins est une thématique absente. Cependant nous avons dans le module socio-environnemental créé la classe concernant les difficultés et l'absence de prise en charge, ce qui peut permettre d'identifier certaines défaillances. L'objectif de ces deux ontologies sont différents, dans ONTOPARON l'objectif est de comprendre les parcours de santé des patients, ayant une SLA inclus dans un réseau régionale, en prenant en compte les actions de coordination qui sont mises en place. Dans l'autre système, l'objectif est d'étudier spécifiquement la coordination des soins, et non la notion de parcours et cela dans le système de santé américain.

---

19. [https://www.ahrq.gov/sites/default/files/publications/files/ccm\\_atlas.pdf](https://www.ahrq.gov/sites/default/files/publications/files/ccm_atlas.pdf)

#### 5.4.5.1 Les concepts définis du module de coordination

ONTOPARONCOORD est le plus petit des modules de ONTOPARON, car il ne prend en compte que les concepts liés à la coordination de parcours. Les activités de coordination sont en lien avec les fonctions et objectifs du Réseau SLA IDF, elles peuvent être de différentes natures comme la communication, l'évaluation, la mise en place de ressources, faire du lien et faciliter les échanges entre les professionnels intervenants et le patient Mboko (2009), Meininger *et al.* (2014). Pour structurer les *actions de coordination* nous avons définis 7 types d'activités qui correspondent aux concepts définis de ce module :

**Les actions de communication.** La communication est un élément important dans la coordination, qu'il s'agisse de recevoir une information, ou de la transmettre à un tiers (patient, entourage, structure, neurologue etc.). Transmettre l'information à un kinésithérapeute, ou un SSIAD que le patient est hospitalisé suite à une pneumopathie d'inhalation, par exemple. Le Réseau SLA IDF intervient dans les situations lorsqu'une sollicitation est faite par les patients et ou les intervenants. Pour comprendre quels étaient les personnes qui sollicitaient le Réseau SLA IDF, et quels étaient les interlocuteurs que le Réseau SLA IDF contactait, nous avons défini deux niveaux de communication :

1. les actions de communications entrantes, qui vont identifier l'émetteur de la sollicitation : *Appel reçu d'un agent médical ; Appel reçu de l'entourage ou du patient ; Appel reçu d'une structure médicale etc.*, et
2. les actions de communication sortantes, qui permettent d'identifier le récepteur des informations que transmettent les coordinateurs du Réseau SLA IDF : *Appel à l'entourage et patient ; Appel à une structure sociale ; Appel à un agent social etc.*

**Les demandes reçues.** Permettent de regrouper l'ensemble des demandes qui sont faites par un *agent* aux coordinateurs du Réseau SLA IDF. Cela peut être une *demande de prise en charge* par une structure (*Demande de SSIAD* ou *demande de séjour de répit*) ou par un professionnel (*demande de recherche d'un kinésithérapeute*). Elle peut aussi concerner une *demande de formation* par un établissement afin de former l'équipe à la prise en charge d'une personne ayant une SLA.

**Les actions d'évaluations de coordination.** Le rôle des coordinateurs n'est pas uniquement de répondre à une demande. Lors de l'appel d'un patient ou de l'entourage, le coordinateur va : évaluer et analyser la situation et il peut faire des préconisations, recommandations qui soient en lien ou non avec la sollicitation initiale. Par exemple, l'appel initial concerne la posologie d'un traitement, en questionnant la personne le coordinateur détecte une situation d'épuisement, il peut alors proposer différentes solutions, comme la mise en place d'aide humaine ou bien la mise en place d'un séjour de répit. Les évaluations et les actions de coordination ne se limitent pas à un domaine spécifique (médical, social ou médico-social), mais prennent en compte la globalité de la personne et de sa situation.

**Les actions d'aligner les ressources aux besoins.** Lors de l'analyse de la situation par le coordinateur, une ou des ressources peuvent être mises en place pour répondre aux besoins identifiés. La détection d'un manque ou d'une absence d'un professionnel dans la prise en charge (per exemple un manque de kinésithérapeute), nécessite la recherche et la mise en place. Cette classe permet de regrouper l'ensemble des actions de coordinations qui vont dans ce sens, qu'il s'agisse de rechercher une structure ou un professionnel, de mettre en place des aides nécessaires (humaines, techniques etc.), ou de réaliser un argumentaire pour justifier de la situation de la personne auprès des structures de financement.

**La formation par les coordinateurs.** Il peut être demandé ou proposé, que les coordinateurs interviennent en structure afin de faire une formation pour informer les professionnels des spécificités de prise en charge des patients ayant une SLA.

**Les alertes en coordination.** Ceci permet d'identifier l'ensemble des alertes reçues sur la situation, souvent complexe du patient.

### 5.4.6 Le module de consolidation

Le module de consolidation ne modélise aucun concept, il importe les différents modules. Une fois les différents modules importés, le raisonneur HermiT (1.3.8.413) de Protégé est utilisé afin d'inférer les différentes classes sous les concepts définis. Une version raisonnée de l'ontologie avec l'ensemble des concepts inférés sous les concepts définis, est exportée pour être utilisée dans les deux outils que nous avons développés : ONBASAM pour l'annotateur sémantique présenté au chapitre 6, et pour PRONTO, l'outil d'évaluation des annotations présenté au chapitre 7.

## 5.5 Synthèse

La construction de l'ontologie ONTOPARON nous a confronté aux questions de modélisation, nous faisant passer d'une ontologie monolithique à une ontologie modulaire. Le choix de la modularité, au-delà de ces aspects de fonctionnalité, de taille et de manipulation nous a confronté à la question de l'objectif de l'ontologie : permettre l'annotation de corpus afin de comprendre les ruptures de parcours de santé des patients. Pour regrouper des concepts présents dans différents modules et pour avoir une approche qui puisse être à la fois globale et spécifique, la conception de concepts définis nous a semblé être l'approche la plus pertinente. Si l'extraction des candidats termes s'est réalisée à partir des corpus du Réseau SLA IDF, il nous semble que la modularité et les thématiques abordées dans les différents modules de ONTOPARON, peuvent permettre la ré-utilisation des modules pour l'annotation d'autres corpus de coordination de parcours d'autres pathologies neuro-dégénératives. L'objectif du travail mené, est l'annotation des corpus du Réseau SLA IDF, il est nécessaire de développer un système permettant l'annotation sémantique en prenant comme ressource l'ontologie ONTOPARON. Le chapitre suivant va présenter l'outil développé.

# Chapitre 6

## Annotation sémantique avec l'Ontology Based Semantic Annotation Module (OnBaSAM)

### Sommaire

---

<b>6.1</b>	<b>L'annotation sémantique</b> . . . . .	<b>112</b>
<b>6.2</b>	<b>Le développement d'ONBASAM</b> . . . . .	<b>113</b>
6.2.1	Les pipelines et Processing Resource de ONBASAM . . . . .	115
<b>6.3</b>	<b>Négation et hypothèse</b> . . . . .	<b>117</b>
<b>6.4</b>	<b>Export des annotations</b> . . . . .	<b>118</b>
<b>6.5</b>	<b>ONTOPARON et ONBASAM</b> . . . . .	<b>119</b>
<b>6.6</b>	<b>Synthèse</b> . . . . .	<b>119</b>

---

L'ontologie permet une modélisation des connaissances dans un domaine spécifique. L'annotation des textes de manière manuelle est une tâche couteuse en temps et difficile lors de l'annotation de grandes quantités de données. L'utilisation d'outils de Traitement Automatique de la Langue Naturelle (TALN) permettent d'automatiser ces tâches. C'est dans cet objectif que nous avons développé l'outil ONBASAM (Ontology Based Semantic Annotation Module). Le principe d'ONBASAM est de prendre comme base de connaissances l'ontologie développée (ONTOPARON) et de réaliser l'annotation sémantique des corpus pour en extraire de la connaissance. Pour optimiser la qualité des annotations, nous avons dû réaliser des prétraitements sur les corpus du Réseau SLA IDF. En effet les corpus présentaient certaines spécificités : documents non structurés, présence de fautes d'orthographe, utilisation de style télégraphique, utilisation d'abréviations. La détection de ces erreurs et leurs corrections, permettent d'améliorer la reconnaissance des concepts à partir de l'ontologie. Les annotations sont importantes, elles peuvent nous permettre de rechercher la présence de « problématique ou dysfonctionnement » dans le parcours des patients. Mais pour cela les données issues de l'annotation doivent être reliées aux données socio-démographiques des patients. L'analyse de l'ensemble de ces données nous permet de faire des hypothèses pour tenter de comprendre les causes de rupture. Ce sont ces annotations et leurs fréquences que nous avons utilisées dans l'analyse des résultats, présenté au chapitre 8.

Ce chapitre présentera en première partie l'objectif et l'intérêt de l'annotation sémantique. La seconde partie présentera l'outil ONBASAM et les chaînes de traitements qui ont été développées dans le cadre de notre projet. Le développement et la mise en place de ONBASAM ont été réalisés par des ingénieurs informatiques. Notre rôle dans le développement de cet outil a été de travailler en étroite collaboration avec les ingénieurs, afin d'identifier les possibles erreurs du système pour tenter de les corriger. Il ne sera pas explicité les mécanismes spécifiques, ni les algorithmes utilisés, car dans le cadre de cette thèse, ONBASAM a servi d'outil et non de domaine de recherche.

## 6.1 L'annotation sémantique

Le Centre de Ressources Textuelles et Lexicales propose comme définition du terme annotation : « remarques manuscrites notées en marge d'un texte »<sup>1</sup>, l'annotation consiste donc à créer une nouvelle donnée à partir d'un document. Pour Ma *et al.* (2009) l'objectif de l'annotation sémantique est de formaliser l'interprétation qui peut être faite des textes sous la forme de métadonnées attachées aux textes ou à certains de leurs segments. Les annotations vont permettre de décrire le contenu du document annoté, à partir de concepts et/ou de relations représentées dans une ontologie. L'ontologie va permettre de fournir le vocabulaire et la connaissance du domaine annoté (Tissaoui *et al.*, 2013). Pour Jonquet *et al.* (2016) l'utilisation d'ontologie pour indexer et intégrer les données, permet de valoriser la connaissance d'un domaine et facilite la recherche et la fouille de données. L'annotation sémantique permet d'améliorer et d'extraire de la connaissance contenue dans des données (Abacha & Zweigenbaum, 2010), cependant

---

1. <https://www.cnrtl.fr/definition/annotation>

cette tâche si elle est faite manuellement peut-être couteuse en temps et fastidieuse lorsque le volume est conséquent (Jonquet *et al.*, 2016).

Des outils d'annotation sémantique existent, nous avons choisi d'en présenter deux. L'annexe E présente des captures d'écran des interfaces et les résultats d'annotation obtenus par les deux systèmes.

**Le SIFR BioPortal Annotator<sup>2</sup> du LIRMM<sup>3</sup>**, utilise les ontologies présentes sur le site pour réaliser l'annotation<sup>4</sup>. BioPortal Annotator propose de réaliser l'annotation des textes, en définissant certains paramètres : *a*) le choix de l'ontologie à utiliser, *b*) la définition des caractéristiques de reconnaissance des mots (long, partiel, exclure les numéros etc.), *c*) l'activation ou non, de la détection de négation, de la temporalité et de « l'experiencer » (personne sur qui porte les informations).

**L'Extracteur de Concept Mutli-Terminologique<sup>5</sup> (ECMT) du CHU de Rouen.**

Pour extraire les concepts, l'ECMT utilise les terminologies incluses dans HeTOP (Cabot, 2017). La réponse fournie est un fichier XML qui indique pour chaque concept identifié son identifiant et son type sémantique s'il est inclus dans le Mé-tathesaurus UMLS. Il est possible comme pour le SIFR de paramétrer certaines options (si *c* = Vrai, les spécialités médicales ainsi que le type sémantique du concept sont inclus dans la réponse.)

Ces deux plateformes, permettraient des traitements de nos corpus, pour peu qu'elles soient installées en local. Cherchant beaucoup de souplesse et d'autonomie pour des recherches spécifiques, nous avons fait le choix de développer un annotateur pour le projet qui a moyen terme puisse servir pour d'autres projets au sein du laboratoire. L'objet ici n'est pas de comparer les performances des différents annotateurs. Comme l'indique Aussenac-Gilles (2012) les résultats de l'annotation dépendent en grande partie de l'ontologie utilisée. Nous n'avons pas fait l'expérience de soumettre notre ontologie aux deux systèmes pour réaliser les annotations et comparer les résultats.

## 6.2 Le développement d'ONBASAM

L'objectif de notre travail est de tenter de comprendre quels peuvent être les indicateurs pouvant expliquer les causes de rupture de parcours de santé des patients inclus au Réseau SLA IDF. Nous avons pour cela développé une chaîne de traitement qui permet l'annotation sémantique de l'ensemble des corpus en prenant comme référentiel ONTOPARON. Pour cela nous avons utilisé les ressources disponibles sur General Architecture for Text Engineering (GATE)<sup>6</sup>. C'est une plateforme d'analyse de texte, mettant à disposition des ressources en open source, développée par l'université de

---

3. Laboratoire d'Informatique de Robotique et de Microélectronique de Montpellier

4. 30 ontologies françaises sont disponibles, dont la Classification Internationale des Soins Primaires deuxième édition (CISP-2), La classification Internationale des Maladies 10ème révision (CIM-10), Le Medical Subject Headings, version française (MSHFRE), la Liste des Produits et Prestations (LPP) par exemple.

6. <https://gate.ac.uk>

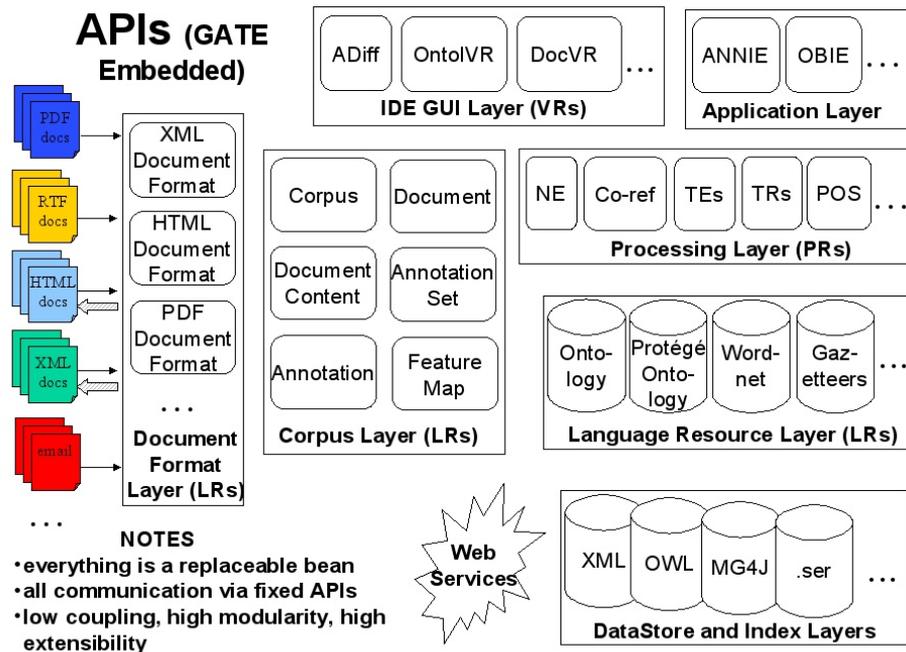


FIGURE 6.1 – Cette figure illustre les différentes API Java intégrées dans GATE Embedded. Figure extraite du site de GATE.

Sheffield. Nombreux sont les projets de recherche dans le domaine biomédical, utilisant GATE. Nous pouvons citer l'exemple du South London and Maudsley Hospital en Angleterre, qui utilise les outils de la plateforme pour rechercher des informations spécifiques sur les dossiers patients en psychiatrie (Cunningham *et al.*, 2013). GATE constitue une boîte à outil permettant essentiellement la construction de chaînes d'annotation de documents textuels. La figure 6.1 montre les différentes API présentes. A la fin d'une chaîne d'annotation les documents textuels d'un corpus sont enrichis par des meta-données d'annotation généralement représentées par des balises XML incluses dans le document annoté. Les outils disponibles sous la forme de ressources qui peuvent être de trois types.

**Les Processing resources (PR).** Un PR est un programme Java prenant en entrée un document (annoté ou non) et produit en sortie un document annoté, son comportement est adaptable en lui fournissant des paramètres. Le plus souvent, ces paramètres sont constitués par des LR ou par de simple indication de comportement (tenir compte ou non des majuscules par exemple, définir un nom d'annotation, etc.). Un PR peut être soit disponible en standard avec la distribution de GATE, soit développé spécifiquement pour les besoins d'une chaîne particulière. Les PRs disponibles sont regroupés avec leurs LR associées dans des plugins.

**Les Language resources (LR).** Un LR peut être constitué par exemple par des listes de termes, une ontologie, une grammaire définissant des patrons à recon-

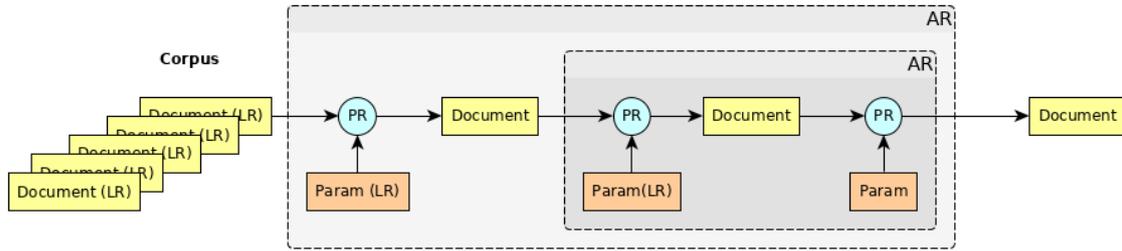


FIGURE 6.2 – Illustration de Language Resource (LR), Processing Resource (PR) et Application Resource (AR) dans un pipeline GATE.

naître, etc. Ainsi un document ou un ensemble de documents (corpus) sont eux mêmes des LRs.

**Les Application resources (AR).** Une AR définit un enchaînement (séquentiel) de PRs traitant des LRs. Chaque PR prend en entrée le document annoté produit par le PR précédent d'où leur nom de Pipeline. Une AR peut être utilisée comme un PR et donc être imbriquée dans une autre AR et ce de manière récursive.

La figure 6.2 illustre les éléments d'un pipeline. Par exemple, un gazetteer (PR) prendra pour paramètre une liste de localités (LR), définira un type d'annotation « Location » et ajoutera ce type d'annotation dans le document d'entrée à chaque fois qu'un mot du texte correspondra à un terme présent dans la liste. Développer une chaîne d'annotation avec GATE revient donc à définir un enchaînement de traitements à appliquer aux documents et à les paramétrer. L'outil Gate Developer permet de mettre au point et de tester ces enchaînements par le biais d'une interface graphique. Il permet de sauvegarder le résultat sous la forme d'une « application » qui pourra alors être rejouée sur un corpus en mode batch (sans interface). Selon les besoins, on peut utiliser les ressources standard ou en développer de nouvelles. Cette possibilité rend l'outil flexible. Celui-ci est par nature très sensible à l'ordre d'exécution des traitements.

### 6.2.1 Les pipelines et Processing Resource de ONBASAM

Pour réaliser les différents traitements dont nous avons besoin : correction de l'orthographe, annotation et export des annotations, différents pipelines ont été créés :

**Lemma Finder**, est le premier et sert de pipeline de base à l'ensemble des autres pipelines.

**Semantic Tagging**, englobe Lemma Finder et d'autres plugins. Il permet de réaliser l'annotation sur les corpus en prenant comme référentiel de connaissance ONTOPARON. La figure 6.3 présente les pipelines : Lemma et Semantic Tagging. La capture d'écran présente à la figure 6.4 illustre l'interface de cette chaîne de traitement.

**Spell Cheking**, permet de réaliser la correction orthographique des corpus. Cette chaîne de traitement a été présentée à la section 4.5 du chapitre 4 de ce manuscrit.

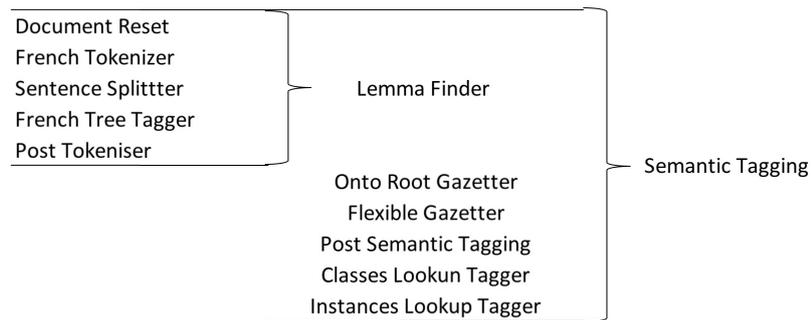


FIGURE 6.3 – Schéma présentant les éléments constitutifs des pipelines créés dans ONBASAM.

Il utilise le pipeline Semantic Tagging et Hunspell Checker.

**Frequency Extractor.** Est un Processing Resource, situé en fin de chaîne. Il permet de réaliser un export des annotations, vers un fichier au format « csv », en regroupant les annotations de tous les évènements des patients. Chaque classe de l'ontologie y est représentée en colonne et chaque ligne correspond à un dossier patient annoté (dossier constitué de l'ensemble des évènements).

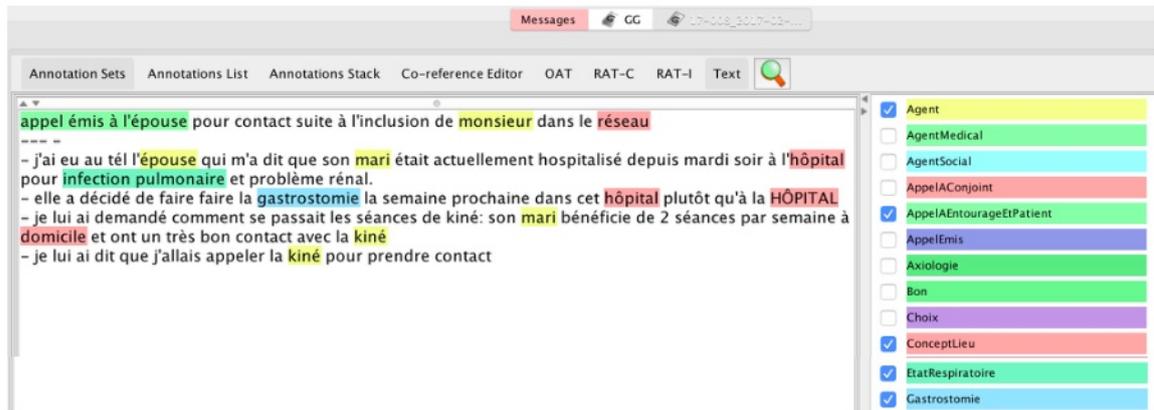


FIGURE 6.4 – Illustration d'un évènement annoté par la chaîne de traitement « semantic Tagging » dans ONBASAM. La partie centrale contient l'évènement annoté. La partie droite contient l'ensemble des classes présentes dans ONTOPARON et annotées dans le texte. Dans l'exemple proposé nous avons sélectionné certaines classes (elles apparaissent surlignées dans le texte) : a) en jaune les « Agents », b) en rose les « Lieux », c) en vert clair « l'action de communication sortante » du Réseau SLA IDF, d) en vert foncé la classe indiquant la présence d'un signe de l'« état respiratoire » du patient résultant d'une infection pulmonaire.

## 6.3 Négation et hypothèse

La détection de la négation et de l'hypothèse sont des points importants à prendre en compte lors de l'annotation des corpus. Sans cela, l'analyse peut être faussée. Il y a, en effet, une grande différence entre « effet secondaire » et « absence d'effet secondaire ». De la même façon, il nous faut trouver des solutions pour l'analyse des conditionnels (donc l'expression des hypothèses) dans les textes afin de ne pas fausser les résultats. La prise en compte de ces deux modalités, apporte des informations différentes, par exemple « le patient n'a pas de médecin traitant, ou bien il est probable qu'il soit hospitalisé » apportent des informations qui vont générer différentes actions de coordination. Au cours du travail, nous avons pris conscience rapidement de l'importance de ces deux modalités, cependant l'intégration de la négation dans la chaîne d'annotation des corpus n'a pu être faite que tardivement. L'intégration de l'algorithme « ConText » [Harkema \*et al.\* \(2009\)](#) dans la chaîne a permis la détection et le repérage des classes de ONTOPARON qui sont niées dans les corpus. L'algorithme ConText, est un algorithme fondé sur des règles qui recherche des indices lexicaux afin de déterminer la négation, l'expérience, l'état temporel ou même hypothétique des conditions cliniques en anglais. Nous avons adapté la liste de termes déclencheurs et des règles de repérage, initialement en anglais à la langue française. La figure 6.5 illustre l'utilisation de ConText dans ONBASAM. Le syntagme « Pas de » est repéré et la classe « déplacement » est niée.

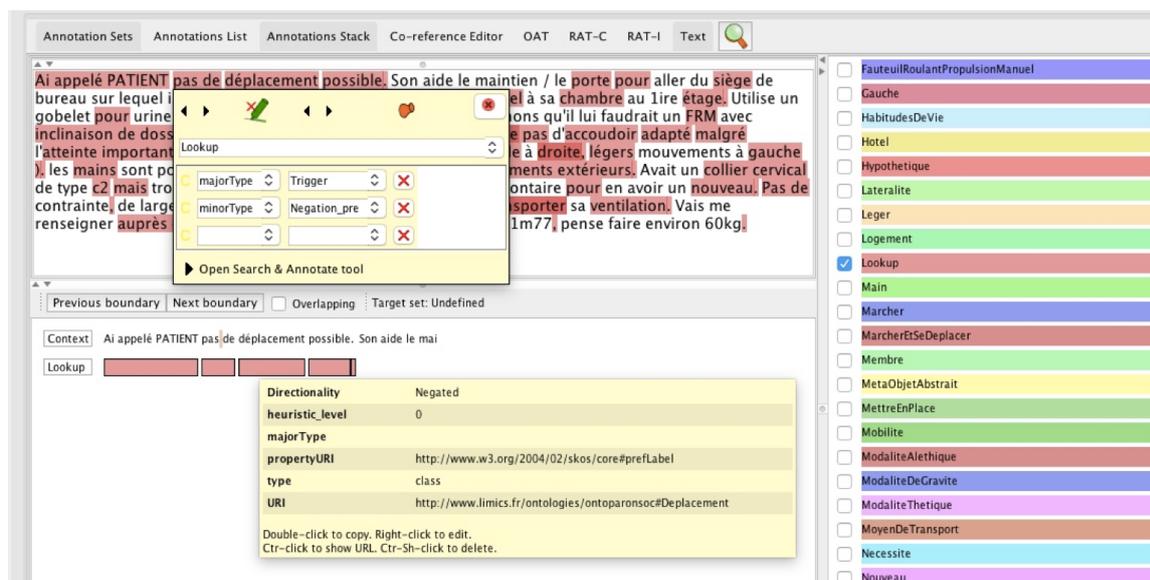


FIGURE 6.5 – Exemple de détection d’une classe de l’ontologie ONTOPARON niée dans le corpus annoté.

## 6.4 Export des annotations

Les annotations permettent d’obtenir de la connaissance sur les documents traités. La figure 6.6 illustre une petite partie du tableau obtenu lors de l’export des annotations. Chaque classe de l’ontologie est représentée par une colonne qui indique le nombre d’annotation de la classe dans le corpus.

Les annotations seules donnent des informations, mais lorsqu’elles sont reliées à la situation du patient, le sens en est amélioré. Les annotations faites de manière quantitative nous permettront de rechercher des liens ou corrélations entre les variables. Les résultats seront présentés au chapitre 8. Le PR Frequency extractor, permet d’extraire de façon quantitative, la fréquence des annotations pour chaque dossier patient. Le fichier de sortie obtenue, est un fichier « .csv » qui peut être :

1. concaténé pour regrouper les annotations des cinq années,
2. joint au tableau des données socio-démographiques,
3. intégré dans au logiciel de statistique utilisé.

Si la négation n’a été intégrée que tardivement dans le travail, nous l’avons utilisée lors de la dernière phase d’annotation de l’ensemble des corpus. Les résultats présentés au chapitre 8 prennent en compte la détection de la négation.

ID	Douleur	SigneCliniqueRespiratoire	EtatDeDependance	SoinsPalliatifs	TroubleEquilibre	DomaineEpuisement	SoinsHygiene	ActionDeCoordination
17-001	0	2	0	0	1	6	0	24
17-002	4	10	0	0	1	0	0	107
17-003	0	3	0	0	0	0	0	26
17-004	5	34	1	0	5	17	1	125
17-005	0	7	0	1	0	0	0	17
17-006	4	12	0	0	0	1	2	100
17-007	0	5	0	1	1	1	0	90
17-008	0	1	0	0	0	0	0	23
17-009	0	0	0	0	0	0	0	7
17-010	4	24	1	3	3	0	0	132
17-011	0	10	0	0	1	6	0	42
17-012	0	0	0	0	0	0	0	106
17-013	22	6	0	2	5	14	0	303
17-014	1	1	0	0	1	0	1	19
17-015	4	26	1	1	3	19	1	201
17-016	0	5	0	0	1	0	0	49
17-017	2	14	1	0	3	1	1	48
17-018	9	9	0	1	0	2	1	94
17-019	1	11	1	1	4	0	1	98
17-020	0	0	0	0	1	0	0	7
17-021	0	1	0	1	5	7	0	34
17-022	1	12	0	0	0	1	0	12
17-023	0	9	2	0	0	1	0	105
17-024	2	11	0	2	6	0	0	95
17-025	1	21	0	0	0	5	0	177
17-026	0	0	0	0	0	0	0	20
17-027	29	10	0	0	0	0	0	235
17-028	0	15	3	0	3	14	0	101
17-029	1	10	0	1	12	4	0	44
17-030	1	12	1	0	2	7	0	268
17-031	3	12	2	0	1	0	0	61
17-032	0	3	1	0	0	4	1	60
17-033	4	19	1	0	7	2	0	181
17-034	1	3	0	0	0	0	0	2
17-035	0	26	0	3	6	16	0	165
17-036	0	0	0	0	2	0	0	30
17-037	6	7	0	0	0	1	0	37
17-038	1	19	1	1	0	0	0	44

FIGURE 6.6 – Exemple d’une partie du tableau obtenu lors de l’utilisation de Frequency Extractor de ONBASAM. Les classes sont représentées en colonne et les patients sont identifiés à l’aide d’un identifiant unique.

## 6.5 ONTOPARON et ONBASAM

L’outil ONBASAM est installé depuis peu, sur des serveurs au sein du laboratoire, ce qui permet de traiter un nombre conséquent de dossiers en peu de temps. ONBASAM utilise une l’ontologie pour réaliser l’annotation. Afin que les concepts définis des différents modules soient bien repérés lors de l’annotation, il est nécessaire d’importer une version de l’ontologie incluant les concepts définis et l’ensemble des concepts inférés (cf D.1). En outre cette version de l’ontologie est limitée volontairement à l’utilisation des preflabels pour le repérage des termes dans les textes.

## 6.6 Synthèse

Tous les dossiers patients ont suivi la même séquence de traitement : un processus de pseudonymisation, une phase de correction orthographique et une phase d’annotation sémantique. Ces processus successifs nous ont permis de passer de données textuelles, à des données structurées et quantitatives. Les différentes phases d’amélioration et de modification des plugins, nous ont permis progressivement d’affiner les chaînes de

traitement. Au-delà de la fonction d'annotation sémantique, ONBASAM a permis de voir quelles étaient les classes manquantes dans l'ontologie, grâce à l'interface. La prise en compte de la négation et de l'hypothèse dans le pipeline sont arrivées tardivement dans le travail, ne permettant pas de faire une analyse quantitative chiffrée de l'efficacité du système. Un tirage aléatoire des corpus annotés intégrant la négation a été réalisé, cela nous a permis de voir de façon qualitative un gain dans l'annotation. Un travail d'amélioration de la détection de la négation et de l'hypothèse est en cours de réalisation dans le cadre d'un projet Carnot<sup>7</sup>. L'objectif à moyen terme, au LIMICS, est d'utiliser ONBASAM dans le cadre d'autres projets. Si les annotations ont été réalisées il est nécessaire d'évaluer leurs qualités. Pour cela il est nécessaire de recourir à des experts du domaine. Pour que les experts puissent réaliser cette tâche nous avons développé un outil dénommé PRONTO qui est présenté dans le chapitre suivant.

---

7. <https://www.carnot-interfaces.fr/2018/02/08/onbasam-un-outil-de-traitement-automatique-des-langues-limics/>

## Pronto outil d'évaluation

### Sommaire

---

<b>7.1</b>	<b>Présentation et objectifs de l'outil d'évaluation PRONTO</b>	<b>122</b>
7.1.1	Présentation des différentes parties	123
7.1.2	Modalité de validation et niveaux de validation	124
7.1.3	Ajout d'annotation manuelle	126
7.1.4	Métriques d'évaluation	127
<b>7.2</b>	<b>Méthodologie d'évaluation du corpus SLA</b>	<b>128</b>
7.2.1	Protocole d'évaluation	128
7.2.2	Définition des critères pour la création du corpus d'évaluation	130
<b>7.3</b>	<b>Résultats</b>	<b>131</b>
7.3.1	Résultats qualitatifs : point de vue des experts	132
7.3.2	Résultats quantitatifs	133
7.3.3	Phase d'amélioration	136
7.3.4	Limites des résultats	137

---

Le chapitre précédent a présenté l'outil d'annotation sémantique ONBASAM utilisant comme ressource lexicale ONTOPARON. Pour valider les annotations faites de manière automatique, il est nécessaire que des experts du domaine puissent évaluer la performance du système d'annotation sur un échantillon, en utilisant un outil d'évaluation facile d'usage. L'équipe du CisMef du CHU de Rouen a développé un outil l'*Extracteur de Concepts Multi-terminologiques Evaluator (ECMTE)* dont l'objectif est d'évaluer les performances de leur système d'annotation l'*Extracteur de Concepts Multi-terminologiques (ECMT)*, qui utilise les terminologies de *HeTOP* pour annoter les textes biomédicaux. Pour répondre à des besoins spécifiques du laboratoire, nous nous sommes appuyé sur l'outil *ECMTE*, pour développer une application web nommée *Pronto*. L'évaluation des annotations par les experts est un élément important permettant de juger de l'efficacité de l'outil d'annotation et secondairement d'évaluer de la pertinence ou de l'absence de concepts dans l'ontologie ONTOPARON.

Dans ce chapitre, nous présenterons l'outil PRONTO, ses fonctionnalités, les métriques utilisées, sa spécificité dans notre cas d'usage permettant aux experts d'évaluer les concepts de niveaux différents. Nous aborderons ensuite les résultats obtenus d'un point de vue qualitatif par le retour d'expérience des experts et quantitatif pour comparer notre système à l'état de l'art.

## 7.1 Présentation et objectifs de l'outil d'évaluation PRONTO

Nous avons, comme indiqué dans le chapitre 4, utilisé ONBASAM pour réaliser les corrections orthographiques sur les corpus extraits de la base du Réseau SLA IDF. Une autre chaîne de traitement dans ONBASAM, permet de réaliser les annotations sémantiques des corpus. Si les annotations sont réalisées, il est nécessaire d'évaluer la performance de cet outil d'annotation (efficacité et pertinence des annotations), pour cela nous avons développé au sein du laboratoire, une application web dénommée PRONTO qui s'est inspirée de l'outil *ECMTE* développée par l'équipe du CISMEF du CHU de Rouen. Nous avons fait le choix de créer une application interne utilisable uniquement au laboratoire, hébergée dans un serveur local, afin de garder une maîtrise dans l'utilisation et la non-divulgaration des corpus du Réseau SLA IDF. Dans le cadre de cette thèse, n'ayant pas les compétences requises pour le développement de cette application, un travail de collaboration fut mis en place avec les personnes en charge d'adapter l'*ECMTE* au cas d'usage de notre projet. Mon travail dans cette partie, s'est axé sur la réalisation d'un cahier des charges pour définir les modalités nécessaires à une utilisation optimale de l'outil par les utilisateurs et pour définir les critères de l'évaluation. L'objectif de PRONTO est d'obtenir sur une même interface les corpus annotés, les concepts auxquels sont associés les termes annotés et les modalités de validation. Pour chaque concept détecté dans le document, l'évaluateur devra indiquer son niveau de validation, à choisir parmi quatre niveaux. Une fois les évaluations réalisées par les experts, l'application va permettre au gestionnaire de projet, d'obtenir des métriques (Précision, Rappel et F-mesure) et évaluer ainsi la qualité des annotations. La partie

suivante va décrire les spécificités de l'application en fonction du rôle des utilisateurs.

### 7.1.1 Présentation des différentes parties

L'accès à l'application se fait à partir de l'url <http://134.157.234.92:8080/pronto/>. Pour accéder à l'application l'utilisateur doit se connecter au moyen d'un identifiant et d'un mot de passe unique. Deux types d'utilisateurs vont utiliser l'application : un gestionnaire de projets, chargé de la gestion du processus d'évaluation et un expert qui va réaliser l'évaluation des annotations.

#### 7.1.1.1 Partie gestionnaire

Le gestionnaire est en charge de la création et de la gestion des projets, c'est-à-dire le chargement des documents et ressources nécessaires aux experts afin de réaliser l'évaluation. Pour cela le gestionnaire doit suivre différentes phases d'un processus constitué de trois étapes.

1. **Création d'un projet.** Le gestionnaire va créer un projet en lui attribuant un nom et un descriptif. Dans notre cas d'usage, chaque projet aura pour nom, le nom de dossier d'un patient constitué de  $n$  évènements. La partie description (saisie en texte libre) permettra de donner des indications sur le projet. Pour notre usage, nous avons décidé d'indiquer, dans cette partie, le nombre d'évènements à évaluer, pour permettre aux experts de choisir les dossiers à traiter en fonction du temps dont ils disposent.
2. **Chargement des ressources.** PRONTO a besoin de l'ontologie utilisée lors de l'annotation, de l'ensemble des documents annotés et de la liste des concepts définis présents dans l'ontologie. Le chargement de l'ontologie au format « owl » dans PRONTO a deux objectifs. Le premier pour permettre à l'expert de visualiser dans l'interface de validation les définitions des concepts, le second pour permettre à l'expert de rechercher un concept dans l'ontologie lorsque celui-ci n'a pas été annoté. Le gestionnaire est le seul utilisateur à pouvoir apporter des modifications sur les projets et les ressources (ajout ou suppression de projet, de document, de l'ontologie et de la liste des concepts définis).
3. **Partage des projets avec les experts.** Le gestionnaire va définir les experts qui peuvent intervenir sur les projets. Il n'existe pas de limite quantitative sur le nombre d'experts pouvant intervenir sur un même projet.
4. **Export des métriques et des commentaires.** L'application PRONTO va permettre au gestionnaire de projet, d'exporter les métriques obtenues une fois les évaluations réalisées par les experts. Elle permet aussi d'exporter les commentaires ajoutés par les experts. Ces commentaires permettent ensuite d'enrichir ou de modifier l'ontologie.

Afin de faciliter l'utilisation et la gestion des projets par le gestionnaire, nous avons choisi d'inclure différents types de pictogrammes ayant une fonction d'information comme indiqué à la figure 7.1. Les pictogrammes indiquent l'état d'avancement de l'évaluation (évaluation non débutée, évaluation en cours, évaluation terminée).

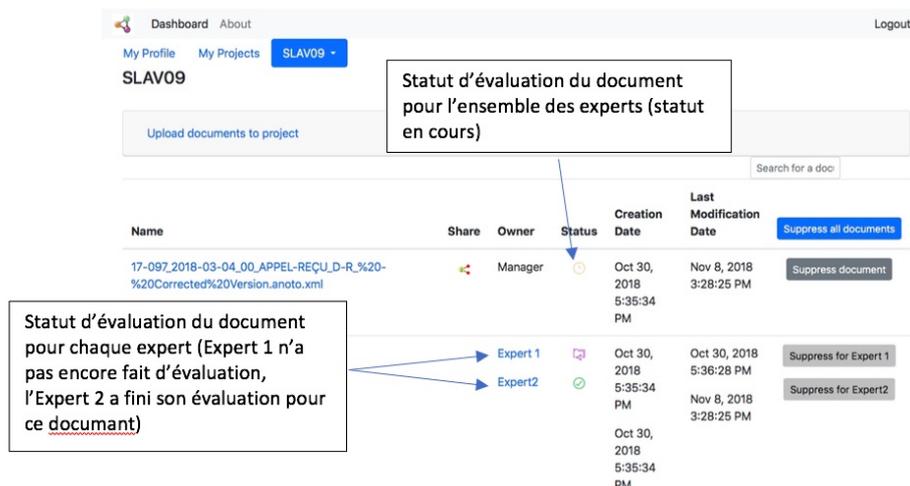


FIGURE 7.1 – Aperçu de l'interface PRONTO, côté gestionnaire de projet.

### 7.1.1.2 Partie Expert

L'expert accède, avec ces identifiants personnels, à l'ensemble des projets (dossier patient) sur lesquels il doit intervenir. En cliquant sur le dossier patient (projet), l'expert accède à tous les documents (événements) qu'il devra évaluer. Les documents sont alors présentés avec des informations : sur la date de création, la dernière date de modification, un pictogramme permet d'indiquer l'état d'évaluation (non annotés, en cours d'annotation, terminé). Les experts vont réaliser l'évaluation en attribuant pour chaque annotation faite par ONBASAM, une validation selon quatre modalités différentes. Lorsque des annotations sont absentes, ils ont la possibilité d'en créer une. Chaque expert a uniquement accès à son travail, il ne lui est pas possible de voir les commentaires, validations faites par les autres experts intervenants sur le même projet. Nous avons fait ce choix afin de ne pas induire de biais lors de l'évaluation.

## 7.1.2 Modalité de validation et niveaux de validation

La figure 7.2 illustre l'interface de PRONTO pour les experts lorsqu'ils vont initier l'évaluation d'un document. Sur la partie gauche apparaît le texte annoté, chaque annotation est surlignée permettant un repérage visuel aisé pour l'évaluateur, la partie droite de l'écran indique la liste des annotations à évaluer.

Pour chaque concept annoté et repéré par le système, quatre niveaux de validation sont possibles :

- le pictogramme vert correspond à une annotation vraie, c'est-à-dire que le terme annoté dans le texte et le concept associé sont bons, par exemple « MT » est annoté comme étant un « Médecin traitant » ;
- le pictogramme rouge correspond à une annotation fautive, c'est-à-dire que le terme annoté ou bien le concept ontologique auquel il est associé est faux. C'est

My Profile My Projects **Projet\_Version\_Finale**

Projet\_Version\_Finale

Selected document: **17-041\_2017-11-07\_00\_APPEL-REÇU\_D-R\_%20-%20Corrected%20Version.anoto.xml**

appel de l'AVS en présence de la patiente qui se plaint d'avoir bcp de glaires avec des nausées et se plaint aussi de maux de tête. la patiente doit se rendre au cabinet de kiné cet après midi 15h pour de la kiné respiratoire. --- ai conseillé d'appeler le MT pour évaluation - me tiendra informé des suites

Useful information for validating annotations

List of annotations:

Concepts left to annotate: 30

Highlighted text :	appel de l'AVS
Action De Coordination	✓ ✗ 🗨️ ?
Action De Communication De Coordination	✓ ✗ 🗨️ ?
Appel De L'auxiliaire De Vie	✓ ✗ 🗨️ ?
Highlighted text :	présence
Présence	✓ ✗ 🗨️ ?
Highlighted text :	patiente
Patient	✓ ✗ 🗨️ ?
Highlighted text :	se plaint
Plaindre	✓ ✗ 🗨️ ?
Highlighted text :	bcp
Beaucoup	✓ ✗ 🗨️ ?
Highlighted text :	glaires
État Médical	👍 🗨️ ?

FIGURE 7.2 – Interface d'évaluation des annotations sémantiques PRONTO du côté expert.

- le cas par exemple pour « ci » (du mot celui-ci) qui peut être annoté comme « carte d'invalidité », car « ci » est utilisé comme altlabel de carte d'invalidité ;
- le pictogramme du pouce vers le bas correspond à un manque de pertinence, c'est-à-dire que l'annotation est bonne, néanmoins cette dernière n'apporte pas d'information pertinente pour la compréhension du parcours de santé du patient ;
  - le point d'interrogation est utilisé lorsque l'évaluateur ne sait pas ou ne peut pas se positionner pour évaluer l'annotation.

### 7.1.2.1 Spécificités de validation des concepts définis

ONTOPARON est constitué de différents modules ontologiques. Ces derniers contiennent des concepts primaires ainsi que des concepts définis comme indiqué dans le chapitre 5. Afin de valider ou non la modélisation des concepts définis par les experts, nous avons souhaité inclure dans l'outil PRONTO la possibilité de valider les différents niveaux ontologique, à la fois les concepts de bas niveau et les concepts définis. La figure 7.2 indique que sur le premier concept annoté « *appel de l'AVS* » il est proposé 3 stratifications conceptuelles qui correspondent aux différents niveaux de l'ontologie :

- « *Action de coordination* » correspond au concept défini ;
- « *Action de communication de coordination* » correspond au concept défini du module ONTOPARONCOORD,
- « *Appel de l'auxiliaire de vie* » correspond au concept de bas niveaux du module ONTOPARONCOORD.

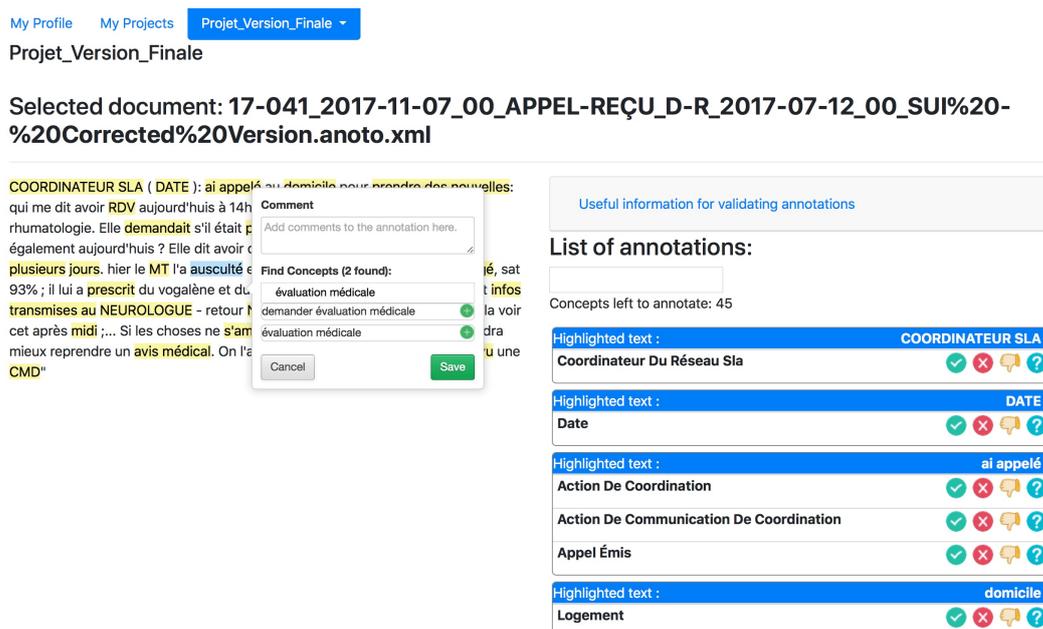


FIGURE 7.3 – Ajout d’une annotation manuelle à partir de l’outil PRONTO.

Il nous a semblé important que les experts puissent intervenir sur la validation de la modélisation ontologique. En effet, les concepts définis ont une importance dans les travaux d’analyse statistique menés au chapitre 8. Les concepts définis représentent des indicateurs ayant un intérêt clinique spécifique au sujet de l’étude du parcours de santé du patient. L’évaluation de ces concepts définis par les experts du domaine permet de valider l’utilisation de ces variables dans l’analyse de recherche de corrélation pouvant expliciter les causes de rupture de parcours.

### 7.1.3 Ajout d’annotation manuelle

Si de nombreux concepts sont déjà annotés, certains peuvent être manquants du point de vue de l’expert. Dans ce cas, l’expert a la possibilité d’ajouter une annotation manuelle. Dans l’exemple présenté dans la figure 7.3, le terme « *ausculté* », n’a pas été annoté par le système. L’évaluateur souhaite qu’il soit annoté car il correspond selon lui au concept d’une « *évaluation médicale* ». L’expert va donc surligner le terme qu’il souhaite annoter. Pour ajouter l’annotation, deux options sont possibles. Dans le premier cas l’expert souhaite associer le terme annoté à un concept présent dans l’ontologie, dans la seconde option l’expert annote le terme sans l’associer à un concept de l’ontologie.

1. Pour le premier cas, l’expert va surligner le terme et saisir le concept auquel il l’associe, dans la zone **Find Concepts** comme indiqué à la figure 7.3. Apparaîtra alors une fenêtre pop-up, avec les concepts qui lui sont proposés. Si la proposition d’un concept lui convient, il le valide par l’activation du bouton **Save**.
2. Dans le second cas, l’expert ne trouve pas de concept, dans ONTOPARON, correspondant au terme annoté. Il va alors ajouter un commentaire dans la partie

Document	Expert	Matched_term	Comment_or_Concept
105_2016-14-11_00	Expert 1	gastralgie	SigneCliniqueGeneraux
105_2016-14-11_00	Expert 2	gastralgie	douleur gastriques
105_2016-14-11_00	Expert 3	gastralgie	DouleurAbdominale
105_2016-27-09_00	Expert 2	anhédonie franche	Symptôme important
105_2016-27-09_00	Expert 2	anhédonie franche	Etat_Emotionnel_Negatif
105_2016-27-09_00	Expert 1	anhédonie franche	SigneCliniquePsychiatrique
105_2016-19-08_00	Expert 2	Fuites et asynchronismes de type auto déclenchement	probleme sur la VNI
105_2016-19-08_00	Expert 2	Fuites et asynchronismes de type auto déclenchement	ProblemeVentilation
105_2016-19-08_00	Expert 1	Pas de	négation
105_2016-19-08_00	Expert 2	Pas de	négation !
105_2016-19-08_00	Expert3	Pas de	négation
147_2018-14-05_00	Expert 2	médicaux	cr médicaux
147_2018-14-05_00	Expert 1	icaux	CompteRenduMedicaux
147_2017-18-12_00	Expert 2	point téléphonique	AppelEmis
147_2017-18-12_00	Expert 1	point téléphonique	FairePoint
147_2017-13-03_00	Expert 2	Paroxétine	DCI de l'antidépresseur
147_2017-13-03_00	Expert 1	Paroxétine	Antidépresseurs
147_2017-13-03_01	Expert 2	bruxisme	parafonction manducatrice (grincement des dents involontaire)
147_2017-13-03_01	Expert 1	bruxisme,	SigneCliniqueGeneraux
147_2017-13-03_01	Expert 1	morsure	SigneCliniqueGeneraux
147_2017-13-03_01	Expert 2	morsure joue	EtatNeurologique

FIGURE 7.4 – Extrait de la liste des annotations manuelles faites par les experts.

**Comment** pour indiquer à quel concept le terme fait référence selon lui. Ces commentaires et annotations seront enregistrés dans une liste comme indiqué à la figure 7.4.

#### 7.1.4 Métriques d'évaluation

Pour évaluer la performance des annotations et la pertinence des concepts nous utilisons 3 mesures standard, fréquemment utilisées (Manning *et al.*, 2009), la précision, le rappel et la F-mesure.

La précision permet de calculer le ratio entre les termes annotés correctement et l'ensemble des autres termes. Le rappel permet de calculer le ratio des annotations correctes sur toutes les annotations trouvées. La F-mesure combine les deux paramètres précédents et permet d'observer la performance globale du système d'annotation. Pour calculer ce résultat on utilise les formules suivantes, pour lesquelles : VP sont les vrais positifs, FP les faux positifs et FN les faux négatifs. Le tableau 7.1 explicite ces différents termes VP, VN, FN et FP etc.

$$Précision = \frac{VP}{VP + FP}$$

$$Rappel = \frac{VP}{VP + FN}$$

TABLEAU 7.1 – Catégorisation des valeurs pour le calcul des métriques de rappel, précision et F-Mesure.

	Termes pertinents	Termes non pertinents
Termes annotés	Vrais positifs (VP)	Faux positifs (FP)
Termes non annotés	Faux négatifs (FN)	Vrais négatifs (VN)

$$F - mesure = 2 * \frac{précision * rappel}{précision + rappel}$$

## 7.2 Méthodologie d'évaluation du corpus SLA

L'objectif de PRONTO est de prendre en compte les spécificités de ONTOPARON, c'est-à-dire d'indiquer les différents niveaux de l'ontologie que sont les concepts primaires et les concepts définis, permettant aux experts d'évaluer la modélisation ontologique.

Pour réaliser l'évaluation, nous avons fait le choix de solliciter des personnes ayant des champs d'expertises différents mais connexes à notre étude. Les experts qui sont intervenus dans l'évaluation sont :

- deux coordinateurs intervenant dans des réseaux se situant en Île-de-France,
- un expert ayant une expertise dans le domaine des aides techniques,
- un étudiant en médecine,
- un coordinateur du Réseau SLA IDF.

Au total, cinq experts sont intervenus dans la phase d'évaluation des annotations. Afin de proposer un environnement similaire d'évaluation aux experts, nous avons décidé de créer un protocole d'évaluation pour limiter les biais d'évaluation. Ce protocole est composé de plusieurs phases. Nous avons constitué un corpus spécifique pour les évaluations en réalisant un tirage aléatoire des événements annotés. Pour cela, nous avons défini des critères de sélection afin que l'échantillon évalué puisse être représentatif des événements présents dans la base de données du Réseau SLA IDF.

### 7.2.1 Protocole d'évaluation

Nous avons défini un protocole reproductible pour chaque session de travail avec les experts. Ce protocole est composé de trois étapes principales : (1) un temps d'information, (2) un temps d'essai et (3) un temps d'évaluation.

1. **Temps d'information** : cette étape a pour objectif de situer la phase d'évaluation des annotations sémantiques dans l'ensemble du travail de thèse. Ce temps d'information est découpé en trois séquences : (a) présentation des étapes de traitement des corpus, (b) présentation de l'ontologie, (c) présentation de l'outil PRONTO.
  - (a) **Présentation des traitements réalisés sur les corpus et des objectifs de l'évaluation.** Une présentation de la phase de pseudo-anonymisation

est faite, permettant d'expliquer la présence de mots ayant une typographie en majuscules (PATIENT, ENTOURAGE, ASSOCIATION DE PATIENT) dans le corpus regroupant les 22 catégories définies la partie 4.4.2.1 de ce manuscrit. La phase de correction orthographique est explicitée. Cela permet de justifier de la présence, parfois, de mots non contextuels à l'évènement.

- (b) **Présentation de ONTOPARON.** La deuxième séquence de ce temps d'information permet de présenter et de décrire les différents modules de l'ontologie ONTOPARON et les spécificités des concepts définis. Pour expliquer et illustrer les concepts définis nous avons fourni deux documents aux évaluateurs :
- i. Un « arbre hiérarchique statique » édité à partir de l'application Sparna<sup>1</sup> qui permet de créer une représentation des données utilisant les libellés préférés en français. Ce document permet de voir l'ensemble des concepts inférés sous les concepts définis des différents niveaux ontologiques et l'ensemble des concepts présents dans l'ontologie.
  - ii. Une « représentation graphique », obtenue à partir de l'item ontograph de PROTÉGÉ, qui permet de voir en une seule représentation les concepts inférés sous les concepts définis, comme par exemple les *actions de coordination*, ou les *actions médicales*.
- (c) **Présentation de l'outil PRONTO.** Cette troisième séquence permet de présenter l'outil PRONTO et son interface, les manipulations à réaliser pour accéder aux différents projets et aux documents. Nous avons utilisé un temps spécifique, pour expliciter les 4 niveaux de validation et la possibilité de créer des annotations, soit par la recherche dans l'ontologie, soit par la saisie de commentaire.
2. **Temps d'essai.** Une fois l'ensemble des informations données, nous avons proposé un temps libre d'apprentissage pour permettre aux experts de s'approprier l'outil. Dans les projets présents dans PRONTO, nous avons inclus un projet spécifique intitulé « Essai », sur lequel les experts pouvaient tester en situation réelle le choix des différents niveaux de validation, tout comme l'utilisation d'ajout d'annotations et de commentaires. Au-delà de l'objectif de maîtrise de l'outil, nous souhaitions que ce temps permette aux experts de formuler des questions liées aux annotations en situation pratique. Chaque expert était libre de prendre le temps qui lui semblait nécessaire pour maîtriser l'outil ; lorsqu'il se sentait à l'aise il pouvait initier seul la phase d'évaluation des corpus.
3. **Temps d'évaluation.** Si pour les autres temps, une interaction était souhaitée avec les experts, pour cette dernière phase, l'objectif était d'intervenir le moins possible afin de ne pas générer de biais dans l'évaluation. Les évaluateurs étaient libres de choisir les dossiers qu'ils souhaitaient évaluer. La seule consigne donnée lors de cette phase était de terminer tout projet initié, afin d'obtenir un score complet par dossier.

---

1. <http://labs.sparna.fr/skos-play/upload>

## 7.2.2 Définition des critères pour la création du corpus d'évaluation

L'objectif étant de valider un maximum d'annotations de différents types, nous avons mis en place une extraction aléatoire des dossiers patients à partir du logiciel JMP<sup>2</sup>. Ceci, afin d'obtenir un échantillon représentatif du corpus annoté et d'avoir un corpus sur lequel une majorité de concepts puissent apparaître. Pour créer et obtenir cet échantillon aléatoire, nous avons défini plusieurs variables intervenants dans l'extraction des dossiers.

1. **Sexe** : l'objectif était d'avoir un échantillon de dossiers d'hommes et de femmes.
2. **Répartition par niveau de complexité** : nous avons créé une répartition logarithmique en tercile des patients par niveaux de complexité, obtenant trois groupes de niveaux de complexité différents (T1 : bas niveau, T2 : niveau moyen, T3 : fort niveau de complexité). Nous avons émis l'hypothèse que le niveau de complexité défini par les coordinateurs pourrait être représentatif de problématique à identifier dans le cadre de la compréhension du parcours. Cette variable nous a donc semblé importante à intégrer dans les critères de sélection de l'échantillon afin de pouvoir illustrer des problématiques et des concepts différents.
3. **Temps passé au Réseau SLA IDF** : notre hypothèse, pour intégrer ce critère comme variable, est que les problématiques rencontrées dans les parcours des patients doivent être différentes au cours du temps. Peut-être les notions d'épuisement, de suppléance apparaissent après un certain temps dans le Réseau SLA IDF.
4. **Nombre d'évènements et de suivis** : ce critère quantitatif, est probablement un marqueur d'action de coordination. Plus il y a d'évènements plus les coordinateurs ont dû intervenir, leurs interventions ont pu porter sur différentes dimensions.
5. **Mode de vie** : il nous a semblé important de prendre en compte des personnes ayant un environnement différent. Notre hypothèse : le mode de vie de la personne (la personne vit seule, à domicile ou en institution) peut avoir un impact sur le type de problématique et de besoin rencontrés générant ainsi différents types d'action de coordination.
6. **Année d'inclusion** : ce critère nous permet d'avoir des patients inclus entre 2013 et 2017.
7. **Âge à l'inclusion** : l'hypothèse émise pour choisir ce critère est que les thématiques rencontrées au cours du parcours peuvent varier en fonction de l'âge de la personne. La thématique du travail, par exemple, n'est abordée que lorsque la personne est en activité professionnelle.
8. **Forme initiale de SLA** : la prise en compte de la forme de la SLA permet probablement d'aborder les différents signes et symptômes présents (signes respiratoires, signes moteurs, etc.).

---

2. [https://www.jmp.com/fr\\_fr/software.html](https://www.jmp.com/fr_fr/software.html)

TABLEAU 7.2 – Caractéristiques des dossiers évalués sur Pronto.

Année d'inclusion	Sexe	Âge	Forme initiale	Mode de vie	Nombre de documents	Temps au réseau	Niveau de complexité
2013	H	70	Spinale	Famille	5	413	T1
2013	H	64	Spinale	Seul	57	1218	T2
2014	F	68	Spinale	Seule	21	184	T3
2014	H	61	Spinale	Famille	26	125	T2
2014	F	70	Bulbaire	Famille	15	227	T1
2015	F	75	Spinale	Famille	16	123	T3
2016	H	59	Bulbaire	Famille	5	702	T2
2016	F	68	Spinale	Famille	15	347	T1
2016	H	56	Bulbaire	Famille	23	833	T1
2016	F	61	Bulbaire	Famille	40	612	T1
2017	H	65	Spinale	Famille	7	378	T1

Nous avons appliqué ces choix, sur la population de patients inclus entre 2013 et 2017 (931 patients). Nous avons fait le choix d'extraire des dossiers complets de patients et pas uniquement une extraction aléatoire d'évènements. Ceci afin de permettre aux évaluateurs de voir l'ensemble des évènements saisis dans le parcours des patients. La lecture chronologique des évènements doit permettre à des tiers extérieurs ne connaissant pas la situation, de comprendre les éléments intervenants au fur et à mesure dans le parcours du patient et d'avoir une vision globale de la situation. Au total sur les 931 dossiers présents dans notre base, nous avons obtenu un échantillon de 11 dossiers de patients, contenant au total 310 évènements à évaluer. Le tableau 7.2 résume les informations pour chaque dossier extrait et nous retrouvons au moins un dossier par année d'inclusion. Le critère du mode de vie est partiellement respecté puisqu'il n'y a pas de personnes vivant en institution. Les dossiers possèdent un nombre variable d'évènements allant de 5 à 57 évènements saisis. Les différentes formes initiales de début de SLA sont représentées et au moins un des dossiers est représenté par niveaux de complexité.

## 7.3 Résultats

Nous avons présenté les résultats selon deux aspects. L'aspect qualitatif, qui correspond au point de vue des experts sur l'outil PRONTO. Cette évaluation non formelle, permet d'évaluer l'ergonomie du système et d'identifier les points d'amélioration pouvant être développés pour rendre l'outil plus efficient. Le second est l'aspect quantitatif, il permet de mesurer la qualité des annotations, les erreurs et les lacunes afin d'enrichir ou modifier l'ontologie. Les évaluations ont eu lieu entre février 2018 et mai 2019. Les sessions de travail ont duré en moyenne 3 heures par experts, seul l'étudiant en médecine a eu un temps d'évaluation estimé à 25 heures. Le tableau 7.3 indique le nombre de dossiers et d'évènements évalués par expert. Au total sur l'ensemble des dossiers a

évaluer (11 dossiers et 310 évènements), 410 évènements ont été évalués par les experts. Seul un expert, est intervenu sur les 11 dossiers.

TABLEAU 7.3 – Nombre de dossiers et d'évènements évalués par expert.

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Total
Nombre d'évènements	20	52	40	68	230	410
Nombre de dossiers	2	3	1	4	11	21

### 7.3.1 Résultats qualitatifs : point de vue des experts

Les cinq experts sont intervenus à des temps différents, seuls deux experts ont travaillé sur une même session d'évaluation. Les différentes phases du protocole d'évaluation ont généré de nombreux échanges et questionnements par les experts. Sur le temps de présentation, les questions des experts ont porté essentiellement sur le choix de modélisation de l'ontologie et en particulier sur le choix de modélisation des actions de coordination. L'une des questions récurrente a été de savoir, si l'inclusion du patient dans un réseau de santé, est ou non, une action de coordination. Pour la moitié des experts coordinateurs, l'inclusion n'est pas une action de coordination, cependant pour l'autre moitié elle en fait partie car elle nécessite du temps et permet un premier recueil d'information sur la personne et sa situation.

La phase d'essai fut une phase durant laquelle de nombreuses questions furent posées par les évaluateurs, pour s'assurer qu'ils avaient bien perçu les critères et les différences entre les niveaux de validation. Cette phase leur a permis de s'approprier l'outil PRONTO, notamment en testant l'option d'ajout d'annotation avec ou sans association à un concept de l'ontologie. Le niveau de validation qui a posé le plus de questions et d'interrogations aux évaluateurs fut l'icône de la pertinence. Nous avons souhaité questionner les usagers sur leurs vécus concernant des points spécifiques :

**L'usage et l'ergonomie de l'outil PRONTO :** tous les experts ont trouvé l'utilisation de l'outil facile. Le temps d'essai proposé dans le protocole a été pour eux essentiel, afin de limiter les erreurs dans la phase d'évaluation. Le point négatif qu'ils ont mis en avant, est le temps de chargement de l'ontologie lorsqu'ils souhaitaient associer un terme à un concept de ONTOPARON.

**La perception des annotations :** les experts ont trouvé qu'un grand nombre de concepts étaient annotés et que les concepts définis étaient intéressants. Cependant tous ont estimé qu'une présentation plus approfondie sur les différents concepts définis aurait pu aider à mieux appréhender la modélisation choisie. Si certaines annotations des experts sont convergentes (annotations de signes cliniques manquants), d'autres sont différentes (annotation manuelle d'un médicament, d'une aide technique, d'une marque de complément oraux etc.). Ces différences sont probablement dues aux différentes formations initiales des experts (médecin, infirmier, ergothérapeute) et de leurs pratiques. Tous les évaluateurs ont mis en avant l'importance de prendre en compte dans l'annotation la négation et l'hypothèse.

D'autres points ont été abordés spontanément par les experts, sur le contenu des évènements, non pas pour évaluer le contenu, mais pour s'interroger sur la forme et sur leur propre manière de rédiger des « transmissions ». L'annotation par un système informatique, leur a fait prendre conscience de la nécessité de mettre en place de « bonnes pratiques » lors de la rédaction des transmissions :

- uniformiser l'utilisation des abréviations ; certaines utilisées dans le corpus n'ont pas été comprises par les experts (« fdv » pour fin de vie, ou « fv » pour feu vert etc.),
- rendre un contenu implicite, qu'un humain peut comprendre, en un contenu explicite qui puisse être annoté par un système informatique afin de ne pas perdre d'informations,
- apporter des informations quantitatives, pour quantifier les actions de coordination qui sont mises en place (« appel à de nombreux cabinets de kiné », « longue conversation », etc.).

### 7.3.2 Résultats quantitatifs

Dans cette partie nous présentons les résultats quantitatifs obtenus pour l'ensemble des dossiers évalués. PRONTO permet pour chaque projet (qui correspond à chaque dossier patient) de calculer les métriques de précision, rappel et F-mesure comme indiqué à la figure 7.5.

Sur la figure 7.5 on peut observer les résultats indiqués pour les différents niveaux de l'ontologie :

- pour les concepts primaires : « Statistics on the bottom concepts » ;
- pour les concepts définis : « Statistics on the defined concepts » ;
- il est aussi possible de connaître les métriques pour chaque module de l'ontologie en sélectionnant l'item « Statistics on ontologies ».

#### 7.3.2.1 Analyse des résultats par niveaux

La tableau 7.4 indique les scores obtenus pour les 21 dossiers annotés par les cinq experts.

Les résultats montrent que les évaluateurs ont validé en grande partie les annotations faites par ONBASAM. Cependant ces résultats sont à modérer, au regard de certains biais expliqués par (Fort *et al.*, 2009) :

1. la présence d'annotations sur un corpus, influence l'annotation des évaluateurs qui vont se focaliser sur les annotations présentes et non sur les données non annotées,
2. l'ergonomie du système utilisé est aussi un facteur à prendre en compte dans les résultats. Les fonctionnalités difficiles d'accès pour l'évaluateur ne seront pas utilisées.

Ces deux critères viennent probablement induire un biais dans nos résultats, mais il ne nous est pas possible d'en mesurer les effets.

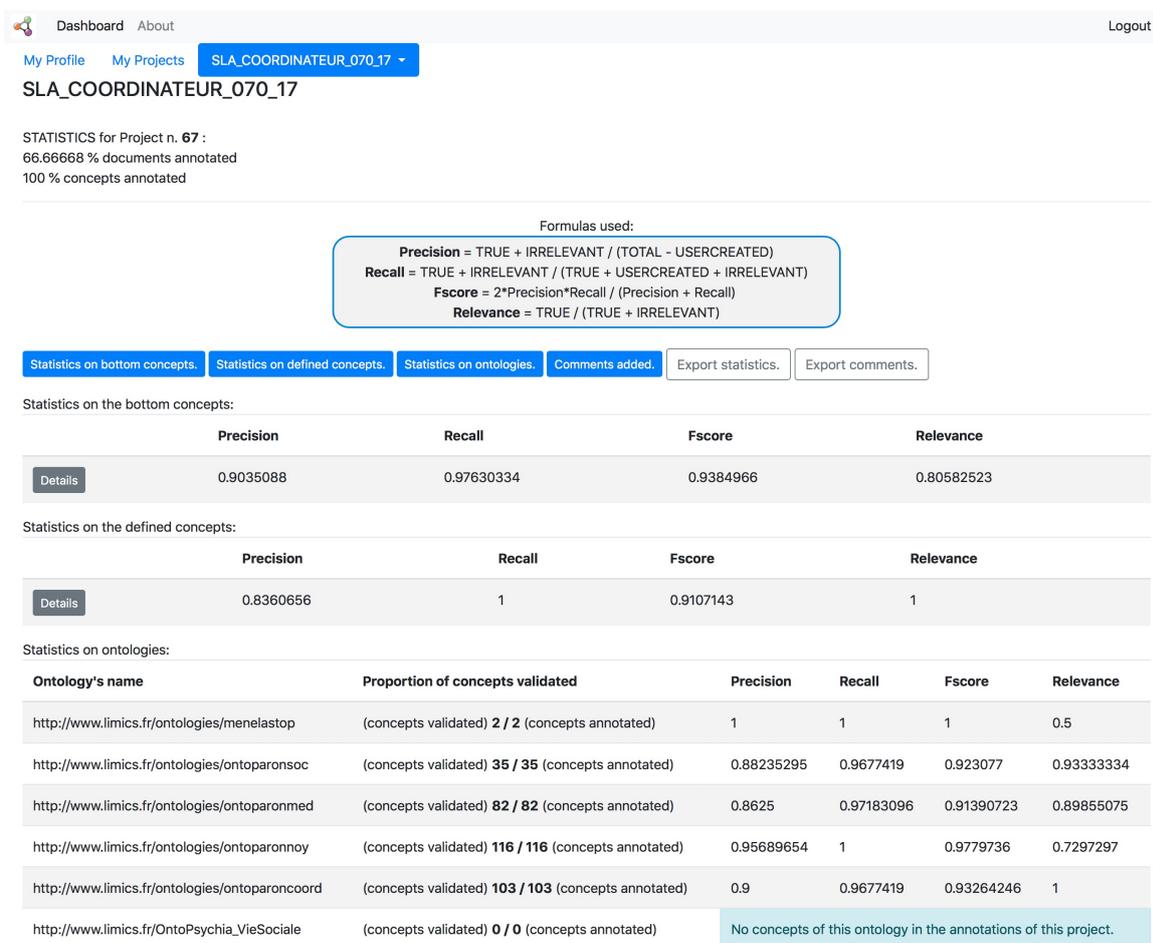


FIGURE 7.5 – Interface PRONTO des résultats quantitatifs obtenus pour un projet.

### 7.3.2.2 Analyse des annotations manuelles et des commentaires

Pour les 410 documents évalués, il a été généré par les experts 560 annotations et commentaires. Nous avons regroupé les commentaires et annotations par catégorie comme décrit ci-dessous.

TABLEAU 7.4 – Résultats quantitatifs des évaluations des annotations sémantiques par les 5 experts.

Niveaux ontologique	Précision	Rappel	F-mesure
Bottom concepts	0,85	0,98	0,91
Concepts définis	0,91	0,99	0,96
MenelasTop	0,83	0,9	0,87
ONTOPARON Social	0,92	0,98	0,95
ONTOPARON Medical	0,93	0,98	0,96
ONTOPARON Noyau	0,95	0,99	0,97
ONTOPARON Coordination	0,93	0,99	0,96

**Présence d’informations nominatives :** certains noms propres ou d’abréviations d’établissements (ex : hôpital George Pompidou : HEGP), de pays (USA), de villes n’étaient pas transformés en données anonymes. Ceci est dû au fait que nous n’avions pas intégré dans les listes utilisées lors de la phase de pseudonymisation l’ensemble des noms d’hôpitaux, de structure médico-sociale ou de cliniques présentes sur l’ensemble du territoire d’Île-de-France, ni l’ensemble des noms de villes du territoire français. Globalement, les villes situées en région Île-de-France étaient bien transformées mais pas les villes d’autres départements (Nice, Lille, etc.). Dans la liste des pays nous avons intégré États Unis mais pas USA. Nous avons complété les listes utilisées pour la pseudonymisation.

**Manque de relation entre les concepts :** certains concepts sont annotés mais les qualificatifs permettant de les qualifier ne sont pas reliés aux concepts. Par exemple dans « autonomie réduite », le concept « autonomie » et le concept « réduit » sont annotés séparément. Il en est de même pour « hypophonie majeure » ou « atteinte prédominante aux membres supérieurs » dans ce cas « atteinte » est annotée, « membres supérieurs » également mais il ne résulte pas d’annotation portant vers le concept de l’ontologie « atteinte des membres supérieurs ».

**Absence de termes dans ONTOPARON :** certains termes n’ont pas été annotés, car non présents dans l’ontologie. Ces termes sont de plusieurs catégories, soit des désignations commerciales pour dénommer des produits (poche de nutrition entérale, noms de médicaments, etc.), soit des signes cliniques et des pathologies :

- Absence de désignation des dénominations commerciales des poches de nutrition entérale « fresubin energy, sondalys energy fibre ».
- Absence de certains signes cliniques : bien qu’un grand nombre de signes et pathologies aient été inclus dans ONTOPARON, certains étaient manquants pour les experts. Par exemple « rhume, distension, gastralgie, transit semble ralenti, effet shunt, fibrillation auriculaire ou encore arythmie ».
- Absence d’annotation pour certains médicaments : de nombreux médicaments sont cités dans les corpus. Lors de la phase d’extraction des candidats termes nous en avons intégré un certain nombre (riluzole, toco, toxine botulique, charbon actif, etc.) dans l’ontologie. Cependant d’autres n’ont pas été intégrés car ayant probablement une fréquence moindre dans les corpus traités. Les experts ont repéré plusieurs médicaments non annotés, comme : « Paroxétine », « corticoïde », « prednisone », « Corticoïdes », « AVK », « laxatif », ou encore « hémigoxine ».

**Présence de fautes d’orthographe limitant l’annotation :** bien qu’une phase de correction orthographique ait été mise en place comme indiquée dans la section 4.5 de ce manuscrit. Des erreurs ont perduré dans les corpus, ne permettant pas l’annotation des concepts, par exemple « pait ente, patin te » pour « patiente », « epouse » pour « épouse » ou bien « levé personne » pour le dispositif de « lève personne ».

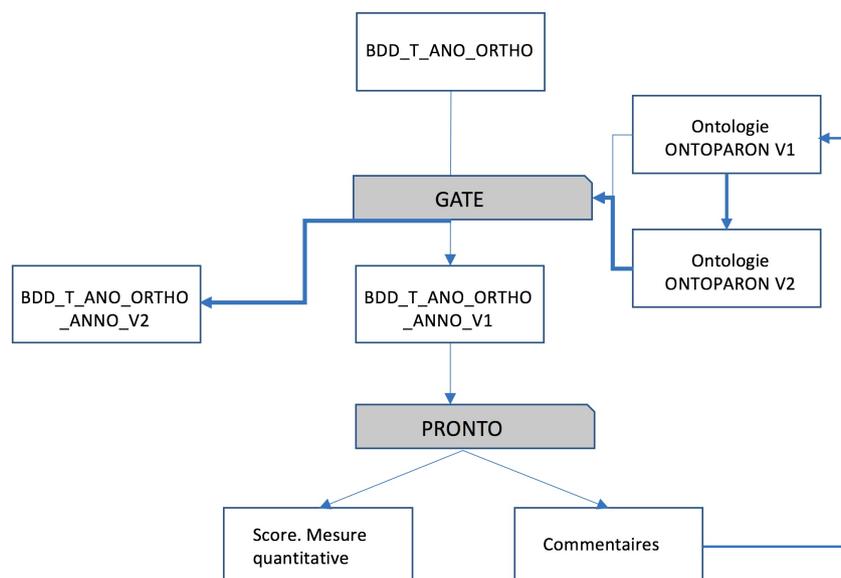


FIGURE 7.6 – Processus réalisé par PRONTO.

### 7.3.3 Phase d'amélioration

Ces évaluations nous ont permis de définir et de mettre en place des plusieurs points d'amélioration pour les outils créés :

**Modifier et corriger l'ontologie :** Suite aux annotations et commentaires des experts nous avons enrichi les différents modules d'ONTOPARON. Le concept de « poche de nutrition entérale » fut créée et les désignations commerciales furent intégrées en altlabel. Les signes cliniques annotés ont été ajoutés. Pour les médicaments, il nous semble difficile d'intégrer nous même, manuellement, l'ensemble des médicaments existants dans ONTOPARON. Cependant il nous parait plus approprié de prendre une ontologie des médicaments et de l'intégrer aux modules existants. Suite aux modifications apportées dans l'ontologie nous sommes arrivés à une version V2 de ONTOPARON.

**Compléter les listes pour la pseudonymisation :** nous avons complété les listes utilisées pour la pseudonymisation, en ajoutant aux listes initiales, l'ensemble des noms de pays, et de nombreuses abréviations d'établissement en Île-de-France.

La prise en compte des commentaires et annotations, nous a permis de passer de l'ontologie V1 utilisée pour la première annotation des 11 dossiers, à une ontologie V2 utilisée pour l'annotation des dossiers patients de notre population. La figure 7.6 illustre le processus.

### 7.3.4 Limites des résultats

La première limite de nos résultats est la taille de l'échantillon évalué. Au total 21 dossiers furent évalués par les experts, seul un dossier a été annoté par les cinq experts. Au final, peu de dossiers ont été évalués, compte tenu du nombre total de dossiers présents dans la base de données du Réseau SLA IDF. La réalisation de l'évaluation est une tâche chronophage et difficile pour les experts. Ils doivent répondre à de nombreuses questions :

- l'annotation est elle bien réalisée ?
- le terme annoté est il concordant avec le concept ?
- le terme annoté fait-il partie du concept défini auquel il est inféré ?
- y a t-il des annotations manquantes ?

Au-delà du questionnement, l'expert doit annoter et commenter les erreurs et les absences d'annotations. La recherche de concept dans ONTOPARON a été une tâche difficile, d'une part par la lenteur de l'application à télécharger l'ontologie et d'autre part pour associer le terme à un concept. Les experts ont parfois préféré faire un commentaire au fait d'associer le terme au concept présent dans l'ontologie. Les évaluations doivent être réalisées par des experts du domaine, cependant ils n'ont pu se rendre disponible qu'une seule fois, ne permettant pas de poursuivre l'évaluation de l'ensemble des dossiers. Le temps d'analyse par dossier que nous avons estimé initialement, était trop court. Nous avons mal évalué les dimensions de prise de connaissance du contexte de l'évènement (lecture, compréhension des écrits, annotations), les réflexions et questionnements de l'expert sur la modélisation de l'ontologie, et la dimension ergonomique par la lenteur du système à télécharger l'ontologie.

Au-delà des dimensions de temps et de technologie, la seconde limite pouvant expliquer les bons résultats obtenus est la provenance des corpus. En effet les corpus utilisés pour l'extraction des candidats termes lors de la construction de l'ontologie, et ceux utilisés lors de l'évaluation proviennent de la même source à savoir le Réseau SLA IDF.

Malgré ces limites, les résultats obtenus sur les différents niveaux ontologiques, ont une F-mesure de 0,90, ce qui montre que le système d'annotation ONBASAM et la modélisation choisie ont été validés par les experts. Ces résultats nous permettent d'utiliser les annotations pour l'études des parcours des patients inclus au Réseau SLA IDF. La prise en compte des commentaires des experts, a permis d'enrichir l'ontologie. Nous avons pour projet de faire une seconde phase d'évaluation des annotations, une fois la négation et l'hypothèse intégrées dans les chaînes de traitement. Cependant cela n'a pas été possible dans les délais de ce travail de thèse. Des améliorations peuvent être apportées à l'ensemble des processus, pour optimiser le système PRONTO.



Troisième partie

Résultats



# Données, analyses et résultats

## Sommaire

---

<b>8.1</b>	<b>Description des données</b>	<b>142</b>
8.1.1	Présentation des différents types de données	142
8.1.2	Exclusion de dossiers	145
8.1.3	La temporalité	146
8.1.4	Données socio-démographiques	146
8.1.5	Les données issues des annotations des corpus	147
8.1.6	Aligner les ressources aux besoins : la demande et la recherche de professionnel	151
8.1.7	Synthèse des observations	156
<b>8.2</b>	<b>Sous populations spécifiques</b>	<b>157</b>
8.2.1	Patients vivant en institution	157
8.2.2	Patients avec une trachéotomie	159
8.2.3	Patients sans besoins et demandes, et sans actions de coordination	160
8.2.4	Synthèse	161
<b>8.3</b>	<b>Analyses globales</b>	<b>161</b>
8.3.1	Liens entre les évènements et les actions de coordinations	161
8.3.2	Analyse en correspondance multiple	162
8.3.3	Le niveau de complexité	163
8.3.4	Patients avec ruptures de parcours identifiés et des difficultés de prise en charge	163
<b>8.4</b>	<b>Synthèse des analyses</b>	<b>165</b>

---

*Ce chapitre permet de décrire et analyser les données obtenues par l'annotation en les liant aux données socio-démographiques des personnes. De nombreuses analyses peuvent être faites à partir des données obtenues. Cependant nous avons décidé de nous intéresser essentiellement aux concepts définis, pour explorer les parcours des patients. La question initiale de notre recherche est de savoir quelles sont les causes de rupture de parcours de santé. La première section va décrire les données socio-démographiques et les données issues de l'annotation, pour l'ensemble de la population. La seconde section présente les différentes analyses. Une première sous-section présente les données pour trois populations spécifiques : les personnes vivant en institution, les patients ayant une trachéotomie et les patients sans besoin ni demande. La seconde sous-section présente les analyses effectuées pour la compréhension des parcours pour la population globale. Les résultats présentés dans ce chapitre ne peuvent être généralisés à l'ensemble des personnes vivant avec une SLA. Les données sont spécifiques du mode de coordination mis en place au sein du Réseau SLA IDF, et de l'ontologie ONTOPARON qui a été créé à partir des données textuelles du Réseau SLA IDF.*

## 8.1 Description des données

Comme indiqué dans le chapitre 7, nous avons soumis les annotations faites par le système ONBASAM à des experts du domaine. Les scores de F-Mesure, Précision et Rappel obtenus, nous permettent de valider la modélisation des concepts définis, ainsi que l'utilisation des annotations pour les analyses. Toutes les analyses présentées dans ce chapitre, ont été effectuées en utilisant le logiciel JMP Pro<sup>1</sup>. Les variables qualitatives ont été analysées par le test du Chi<sup>2</sup> et par l'analyse en composante multiple, les corrélations ont été analysées par le test de corrélation de Pearson. Les comparaisons ont été réalisées par le test de student. Des modèles linéaires multivariés ont été utilisés pour la recherche de relation entre les variables quantitatives. La population étudiée est constituée des patients inclus au Réseau SLA IDF entre le 2 janvier 2013 et le 31 décembre 2017. Ces années ont été choisies afin d'avoir un nombre suffisant de dossiers, pour mener les analyses. La population incluse dans notre étude est constituée de 931 patients, pour lesquels 31 260 événements ont été annotés par ONBASAM.

### 8.1.1 Présentation des différents types de données

Les données présentées dans ce chapitre sont issues de deux sources : d'une part, de la base de données du Réseau SLA IDF pour les données socio-démographiques, les données médicales et les données sur le parcours (niveau de complexité et nombre d'évènements) et d'autre part, par l'annotation sémantique réalisée à partir du système ONBASAM.

**Les données socio-démographiques**, regroupent différents types de variables : l'âge de la personne, le sexe, le statut social, etc. Différentes extractions des données

---

1. version 14.0.0 [https://www.jmp.com/fr\\_fr/software.html](https://www.jmp.com/fr_fr/software.html)

ont été réalisées au cours du travail de thèse. La dernière extraction des données et des événements a été faite le 18 juin 2019.

**Les données médicales**, apportent des renseignements sur : la forme de la pathologie (bulbaire, spinale), la date d'apparition des premiers symptômes, la date de diagnostic et la date d'inclusion au Réseau SLA IDF. Les informations sur la présence ou le refus des suppléances (VNI, gastrostomie, trachéotomie) sont également indiquées. L'information sur le positionnement des patients concernant les suppléances, résultent en grande partie de leurs choix lorsque le médecin évoque avec les patients les directives anticipées au cours des consultations de suivi au centre expert.

Ces informations sont présentées dans le tableau 8.1. En annexe F, le tableau F.1 présente les données socio-démographiques des patients, par année d'inclusion.

**Les données sur le parcours de santé.** Nous avons souhaité extraire deux données spécifiques à partir de la base du Réseau SLA IDF, qui nous semblent importantes à prendre en compte dans les analyses. Tout d'abord le nombre d'événements principaux enregistrés et le niveau de complexité.

1. **Les événements principaux et de suivi.** Le portail du Réseau SLA IDF possède un compteur indiquant le nombre de patients inclus et le nombre d'événements enregistrés sur la base. À la date du 8 août 2019 la base contient 2683 patients et 42 848 événements. Pour les 931 patients de notre étude nous avons annoté 31 260 événements. Ce différentiel important du nombre d'événements enregistrés, est expliqué par le mode de comptage du portail de la base du Réseau SLA IDF. Le portail ne comptabilise que les événements principaux et ne tient pas compte des événements de suivi. Dans notre travail nous avons individualisé chaque événement en créant un fichier « .txt » par événement. Cela permet d'identifier l'événement principal et le ou les événements de suivi qui lui sont rattachés. Les 31 260 événements sont répartis en 15 282 événements principaux et 15 978 événements de suivi. Les événements saisis dans le cadre du parcours de santé sont saisis sur la durée T3 (cf. figure 8.1).
2. **Niveau de complexité.** Le niveau de complexité nous semble, un indicateur à prendre en compte dans l'analyse des parcours patients. Lors de la saisie de l'événement principal le coordinateur attribue un score de complexité en fonction de l'évaluation de la situation. Ce score présente 3 niveaux (1 pas de complexité ; 2 complexité moyenne ; 3 grande complexité). Seul le coordinateur peut évaluer le niveau de complexité d'une situation. Une même situation peut être évaluée comme complexe chez un patient et non complexe pour un autre. Dans la méthodologie que nous avons mise en place lors de l'annotation des dossiers patients, nous avons réalisé une annotation de tous les événements. Pour pouvoir utiliser le niveau de complexité nous avons créé un « score moyen de complexité », en rapportant la somme des scores attribués au nombre d'événements principaux enregistrés.

TABLEAU 8.1 – Données socio-démographiques de la population étudiée.

	<b>Ensemble de la population</b>
<b>Âge moyen au diagnostic</b> $\pm$ ET (ans)	65,2 $\pm$ 12
<b>Âge moyen à l'inclusion</b> $\pm$ ET (ans)	65,9 $\pm$ 12,02
<b>Sexe</b> $n$ (%)	
Homme	486 (52%)
Femme	442 (48%)
<b>Forme</b> $n$ (%)	
Spinale	565 (61%)
Bulbaire	295 (32%)
Données manquantes	68 (7%)
<b>Suppléance</b> $n$ (%)	
VNI <sup>2</sup>	Refus 32 (3%) ; Présence 402 (43%)
Gastrostomie	Refus 83 (9%) ; Présence 171 (18%)
Trachéotomie	Refus 176 (19%) ; Présence 8 (1%)
<b>Délais moyen entre les premiers symptômes et le diagnostic</b> $\pm$ ET (jours)	455 $\pm$ 508
Données manquantes	79
<b>Délais moyen entre le diagnostic et l'inclusion au Réseau SLA IDF</b> $\pm$ ET (jours)	222 $\pm$ 423
Données manquantes	19
<b>Délais moyen de temps au Réseau SLA IDF</b> $\pm$ ET (jours)	548 $\pm$ 433
<b>Nombre moyen d'évènements et de suivi</b> $\pm$	33,6 $\pm$ 32
<b>Mode de sortie</b> $n$ (%)	
Déménagement	22 (2%)
Décès	676 (73%)
Choix du patient	3 (0%)
Vivant	227 (25%)
<b>Statut social</b> $n$ (%)	
Célibataire	126 (14%)
Marié	553 (60%)
Divorcé	108 (12%)
Veuf	73 (8%)
Données manquantes	69 (7%)
<b>Mode de vie</b> $n$ %	
Vit seul	166 (18%)
Vit en famille	664 (72%)
Vit en institution	17 (2%)
Données manquantes	81 (9%)
<b>Répartition par département</b> $n$ (%)	
Paris (75)	190 (20%)
Seine-et-Marne (77)	105 (11%)
Yvelines (78)	107 (12%)
Essonne (91)	107 (11%)
Hauts-de-Seine (92)	125 (13%)
Seine-St-Denis (93)	100 (11%)
Val-de-Marne (94)	101 (11%)
Val-d'Oise (95)	93 (10%)

TABLEAU 8.2 – Répartition du nombre d'évènements et de la moyenne de complexité pour l'ensemble de la population.

	Moyenne $\pm$ ET	Minimum	Maximum
<b>Évènements principaux</b>	16,4 $\pm$ 14,54	1	143
<b>Évènements de suivis</b>	17,2 $\pm$ 19,42	0	179
<b>Évènements totaux</b>	33,6 $\pm$ 32,13	1	322
<b>Niveau de complexité</b>	1,82 $\pm$ 0,38	0,33	3

Le tableau 8.2 présente la moyenne du nombre d'évènements par patient, ainsi que le score moyen de complexité, pour l'ensemble de la population.

Il est possible d'observer une grande hétérogénéité sur le nombre d'évènements saisis, certains patients n'ayant que 1 évènement alors que d'autres ont en plus de 200, la moyenne se situant à 33,6 (ET  $\pm$  32,3). Ce premier critère, est un élément important à prendre en compte dans l'analyse des parcours patients car le Réseau SLA IDF fonctionne sur sollicitation, c'est-à-dire que les patients, lorsqu'ils sont inclus ont le choix de solliciter ou non les coordinateurs. Comme pour les évènements, le score du niveau de complexité varie. Nous tenterons dans les analyses, de comprendre si ce score caractérise certains patients ou bien s'il est lié à des thématiques particulières des concepts définis (cf. 5.4.1).

**Les données d'annotation**, résultent de l'annotation des classes de l'ontologie, repérées dans les évènements et identifiées par le système d'annotation ONBASAM. Ces données sont initialement toutes quantitatives. Au vu des fréquences de certaines d'entre elles, nous avons réalisé un recodage des variables quantitatives en variables catégorielles. Il n'est pas envisagé dans cette thèse de présenter l'ensemble des annotations de chacune des classes de l'ontologie car elles sont trop nombreuses (2555).

### 8.1.2 Exclusion de dossiers

Lors de l'étude des données, certaines valeurs aberrantes sont apparues concernant le temps au Réseau SLA IDF pour 4 patients. Pour les comprendre, nous avons isolé ces 4 dossiers et les avons examinés. Au final, nous avons ciblés 3 dossiers car il est indiqué dans les évènements, que les patients présentent une Sclérose Latérale Primitive (SLP)<sup>3</sup> et non une SLA. Notre objectif étant de travailler sur les parcours des patients SLA, nous avons décidé de retirer ces dossiers de notre population car il s'agit d'une pathologie proche, mais différente. Pour le dernier patient qui présente un temps de suivi important, le diagnostic de SLA est bien posé, il n'a donc pas été exclu de la population.

Au total la population étudiée est constituée de 928 personnes.

3. [https://www.orpha.net/consor/cgi-bin/OC\\_Exp.php?Lng=FR&Expert=35689](https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Lng=FR&Expert=35689)

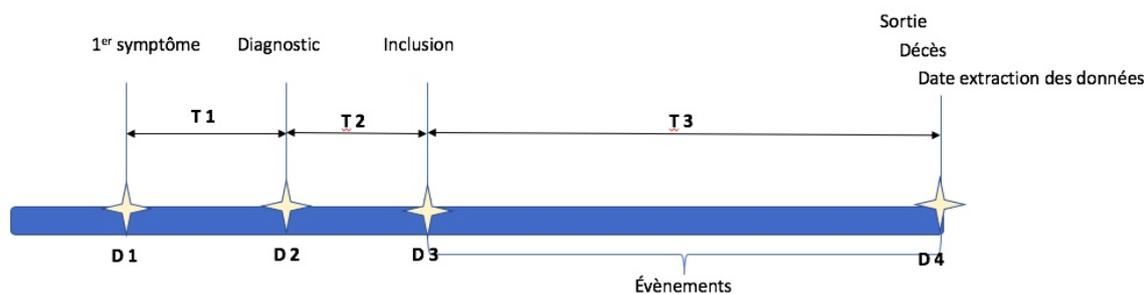


FIGURE 8.1 – Élément temporel dans le parcours des patients SLA.

### 8.1.3 La temporalité

La figure 8.1 illustre les différentes dates utilisées dans les analyses pour le calcul des différents délais. La temporalité est un élément important dans cette pathologie, en raison de la rapidité d'évolution mais aussi de la durée de survie une fois le diagnostic posé.

**La date D1** correspond à la date d'apparition des premiers symptômes. Pour 69 patients cette donnée est manquante.

**La date D2** correspond à la date du diagnostic. La donnée est manquante pour 17 patients.

**La date D3** correspond à la date d'inclusion au Réseau SLA IDF.

**La date D4** correspond à la date de sortie du réseau. La sortie peut avoir différentes modalités : choix du patient ou déménagement ; date de décès du patient ou bien date d'extraction des données. Pour les patients en vie, la date retenue est la date de la dernière extraction des données, à savoir le 18/06/2019.

A partir de ces dates, des délais ont été calculés :

1. **T 1** : délais entre la date d'apparition des symptômes et la date de diagnostic ,
2. **T 2** : délais entre la date de diagnostic et la date d'inclusion au Réseau SLA IDF,
3. **T 3** : délais entre la date d'inclusion et la date de sortie du Réseau SLA IDF. Ce qui correspond au temps d'accompagnement au Réseau SLA IDF.

### 8.1.4 Données socio-démographiques

Le tableau 8.1 présente les caractéristiques socio-démographiques des patients inclus dans notre population. En moyenne le Réseau SLA IDF, inclut 186 patients par an. La population étudiée est constituée de **52% d'hommes et de 48% de femmes** (soit un sexe-ratio de 1,09, légèrement inférieur aux données de la littérature présentées à la section 1.2.2). L'âge moyen à l'inclusion dans le Réseau SLA IDF est de **66 ans**. Les patients présentent une **forme spinale dans 61%** des cas et une **forme bulbaire dans 32%** des cas, ce qui est proche des données de la littérature. Le délai moyen

entre les symptômes inauguraux et l'annonce diagnostique est de 14,55 mois ( $\pm$  15,6 mois). L'inclusion est réalisée en moyenne 222 jours après le diagnostic.

D'un point de vue social, les personnes sont majoritairement mariées (60%), et vivent en famille (72%), 18% des personnes vivent seules et 17 personnes vivent en institution. Tous les patients vivent en région Île-de-France, le département de Paris (21%) présente une représentation plus élevée en comparaison aux autres départements de la région (10% en moyenne). Au moment de l'extraction des données, 22 patients sont sortis suite à un déménagement, 4 patients ont décidé de quitter le Réseau SLA IDF et 228 personnes sont vivantes. La sortie du Réseau SLA IDF se fait majoritairement lors du décès (73% des cas), le temps moyen d'accompagnement des patients au Réseau SLA IDF est de 552 jours. Concernant les suppléances (VNI, gastrostomie et trachéotomie), 43% des patients ont une VNI, 18% ont une gastrostomie et seul 1% ont une trachéotomie. Le refus pour la mise en place de ces suppléances est parfois exprimé, cependant, dans de nombreux dossiers, ces données sont vierges. Il n'est pas possible de savoir s'il s'agit d'une donnée manquante ou d'un non-positionnement du patient.

Dans une étude de [Cordesse \*et al.\* \(2015\)](#), les auteurs montrent que 91,5% des patients prenant du riluzole<sup>4</sup> en région Île-de-France sont inclus dans le Réseau SLA IDF. En nous basant sur les données extraites de la base de données du Réseau SLA IDF, nous pouvons dire que la population de notre étude est représentative des patients vivant avec une SLA en Île-de-France.

### 8.1.5 Les données issues des annotations des corpus

Les données obtenues par l'annotation sont présentées par module ontologique. Pour présenter ces données et leur donner un sens clinique, nous avons fait le choix de les présenter de deux façons :

1. quantitativement : c'est-à-dire telles qu'elles sont obtenues à la sortie de ONBAsAM, cela permet de voir leurs répartitions dans la population,
2. qualitativement : en les transformant en données binaires, c'est-à-dire en présence ou absence. Cela permet de repérer, par exemple, que la thématique de l'aide humaine apparaît au moins une fois pour 707 patients, soit 76 % de la population, mais que la présence de signes cognitifs dénotant une altération cognitive, n'est présente que dans 19% de la population (174 cas).

#### 8.1.5.1 L'annotation des concepts définis du module médical

Le tableau 8.3 présente les moyennes d'annotation obtenues, pour les concepts définis du module médical.

**Les concepts : État médical et Actions médicales**, ont les valeurs les plus importantes. Ce qui est cohérent avec leurs définitions. L'état médical regroupe

---

4. Médicament utilisé de façon spécifique dans la [SLA](#)

TABLEAU 8.3 – Données issues de l’annotation sémantique des concepts définis du module médical.

	<b>Ensemble de la population</b> Moyenne $\pm$ ET Présence $n$ (%) vs Absence $n$ (%)	$p$
<b>État médical</b>	47,57 $\pm$ 42,42 916 (99%) vs 12 (1%)	0.0001
<b>État d’aggravation</b>	1,83 $\pm$ 2,14 620 (67%) vs 308 (33%)	0.0001
<b>État cognitif</b>	0,49 $\pm$ 1,41 174 (19%) vs 754 (81%)	0.0001
<b>État cutané</b>	0,65 $\pm$ 1,78 254 (27%) vs 674 (73%)	0.0001
<b>État douloureux</b>	1,84 $\pm$ 3,94 441 (48%) vs 487 (52%)	0.0025
<b>État lié aux traitements</b>	5,16 $\pm$ 6,67 749 (81%) vs 179 (19%)	0.0001
<b>État moteur</b>	4,39 $\pm$ 4,40 782 (84%) vs 146 (16%)	0.0001
<b>État psychique</b>	2,62 $\pm$ 3,61 634 (68%) vs 294 (32%)	0.0001
<b>État respiratoire</b>	9,92 $\pm$ 12,03 820 (88%) vs 108 (12%)	0.0001
<b>État nutritionnel</b>	1,68 $\pm$ 2,34 560 (60%) vs 368 (40 %)	0.0001
<b>Action médicale</b>	29,05 $\pm$ 29,44 917 (99%) vs 11 (1 %)	0.0001

l’ensemble des signes et des symptômes présents dans les évènements. Les actions médicales sont liées aux soins (consultation, hospitalisation, etc.). On observe que pour 1 % de la population, il n’est fait mention d’aucune action médicale. L’hypothèse émise est que les patients ont souhaité adhérer au Réseau SLA IDF mais ne l’ont pas sollicité.

**État moteur, État respiratoire et État lié aux traitements.** Ces concepts définis ont été annotés au moins une fois dans les dossiers pour plus de 80 % de la population. Pour rappel, les concepts définis « État  $x$  », ont été créés pour permettre d’identifier la présence de signes (cliniques ou fonctionnels) dénotant une altération dans une de ces fonctions (respiratoire : dyspnée, pneumopathie, etc. ; moteur : chute, fracture, difficultés motrices, spasticité, etc.). Il est cohérent que la présence de signes d’altérations soient identifiés pour une grande majorité de la population, puisqu’ils correspondent à l’évolution naturelle de la maladie. La présence de signes, nécessite la mise en place de traitements. Le lien et l’importances d’annotation entre ces états est donc compréhensible.

**État Cognitif** ne concerne que 19% de la population, ce qui est faible en comparaison

TABLEAU 8.4 – Données issues de l’annotation sémantique des concepts définis du domaine de coordination.

	<b>Ensemble de la population</b> Moyenne $\pm$ ET	<b>Min</b>	<b>Max</b>
<b>Action de coordination</b>	74,23 $\pm$ 77,67	1	762
<b>Action de communication de coordination</b>	45,81 $\pm$ 48,9	0	480
<b>Action de communication entrante</b>	14,86 $\pm$ 16,66	0	154
<b>Action de communication sortante</b>	31,46 $\pm$ 33,37	0	329
<b>Alerte en coordination</b>	1,23 $\pm$ 1,79	0	16
<b>Aligner les ressources aux besoins</b>	10,41 $\pm$ 11,02	0	79
<b>Demandes reçues en coordination</b>	5,42 $\pm$ 6,54	0	64
<b>Entretien en coordination</b>	0,96 $\pm$ 1,67	0	28
<b>Évaluation en coordination</b>	7,30 $\pm$ 8,70	0	68

des données de la littérature qui estime que dans 50 % des cas, la SLA est associée à des troubles cognitifs et/ou psycho-comportementaux ([Viader et al., 2014](#)). Dans notre modélisation nous avons séparé les troubles du comportement, des troubles cognitifs, que nous avons liés à l’état psychique.

**État Psychique** est fréquemment repéré dans les corpus, et concerne 68 % des patients, il concerne les signes et symptômes liés à la dépression, aux troubles de l’humeur, à l’anxiété, etc..

**État douloureux**, la présence de douleurs est repérée dans 48% des cas, ce qui est légèrement inférieur à l’étude de ([Chiò et al., 2012](#)) dans laquelle 57 % des patients SLA présentent des douleurs.

**État d’aggravation**, est rencontré dans 67 % des cas. Dans les corpus il peut être noté : « aggravation importante, aggravation du patient, aggravation clinique, altération de l’état général, aggravation de la situation, etc. ». La notion d’aggravation apparaît fréquemment dans les parcours des patients et elle peut être liée à une aggravation de la pathologie ou bien à une aggravation de la situation sociale. Le positionnement de ces aggravations dans la chronologie des événements, pourrait permettre de voir quelles sont leurs conséquences sur le parcours des patients.

### 8.1.5.2 L’annotation des concepts définis du module de coordination

Nous avons souhaité observer comment se répartissaient les différentes catégories d’action de coordination. Le tableau 8.4 indique les métriques pour chacune de ces catégories.

Nous pouvons faire quelques observations :

- variabilité des moyennes des actions de coordination,

- il existe des actions de coordination plus fréquentes que d'autres. Certaines actions semblent être prédominantes, comme les actions d'aligner les ressources aux besoins, par rapport aux entretiens réalisés.

**Les actions de communication de coordination.** Ces actions ont une grande place dans les activités des coordinateurs. Si l'on fait un focus sur les deux sous-catégories, que sont **les actions de communication entrantes** et **les actions de communication sortantes** nous remarquons une différence entre les moyennes ( $p < 0.0001$ ). Les actions de communication entrantes ont une moyenne de 14,85 ( $\pm 16,65$ ) alors que les communications sortantes ont une moyenne de 31,43 ( $\pm 33,33$ ). Il y a deux fois plus d'actions de communication sortantes (vers l'extérieur du Réseau SLA IDF) que d'entrantes. Les actions de communication ne sont pas uniquement le fait de recevoir ou de transmettre une information. Le Réseau SLA IDF inclut des patients de toute la région Île-de-France, les interactions vont donc se faire majoritairement par téléphone, avec le patient et les professionnels. La réception d'une information ou d'une demande nécessite de la traiter et ensuite de la transmettre aux intervenants. La diffusion de l'information est un élément important dans la coordination (Fishman-Bosc *et al.*, 2016, Giraud *et al.*, 2017, Javel, 2011) et semble être probablement un levier de fluidité des parcours permettant à tous les acteurs et aux patients d'avoir un même niveau d'information. Au niveau des appels, il nous est possible de quantifier le nombre d'appel émis (15 802) et le nombre d'appel reçu (7341). Nous avons pu identifier que les appels reçus proviennent d'abord des patients et de leur entourage (moyenne est de  $3,13 \pm 4,36$ ) puis des structures et des professionnels du domaine médical (moyenne  $1,12 \pm 1,76$ ) et des structures et des professionnels du domaine social ( $0,11 \pm 0,43$ ). Cette répartition est retrouvée dans les appels sortants qui sont à destination des patients, puis des acteurs de santé ( $1,97 \pm 2,50$ ) et des intervenants sociaux ( $0,43 \pm 0,88$ ). Pour les transmissions il est aussi retrouver, la diffusion de l'information vers l'extérieur, 6 282 transmissions sont reçues au Réseau SLA IDF et 11 255 transmissions sont faites par les coordinateurs.

**Les actions d'aligner les ressources pour répondre aux besoins.** Ce sont les actions de recherche des structures ou des professionnels ; les préconisations (donner de l'information et délivrer des conseils) et la mise en place des ressources (la mise en place d'un séjour de répit, la mise en place d'aide humaine, etc.). Ce sont des actions très fréquentes. La mise en place de ressources, peut survenir, à la suite d'une demande formulée explicitement ou suite aux propositions des coordinateurs qui, lors de l'évaluation de la situation peuvent détecter un besoin non exprimé par le patient.

**Les demandes reçues et les évaluations en coordination.** Lorsque l'on compare les deux classes, que sont les demandes reçues en coordination (moyenne  $5,47 \pm 6,59$ ) et les actions d'aligner les ressources (moyenne  $10,46 \pm 11,09$ ) ont observé une différence significative ( $p < 0.0001$ ). L'hypothèse que nous émettons est que certaines demandes sont exprimées de façons explicites, mais que les coordinateurs réalisent des évaluations et proposent des « solutions » sans que la demande soit clairement exprimée. La notion de « proactivité » (anticipation des besoins

suite aux évaluations) expliqué par Gilles Duhamel lors de la journée d'étude « Evaluation des politiques publiques et coordination des soins » (Benamouzig *et al.*, 2015), peut expliquer ces résultats. Ces anticipations des besoins peuvent être un élément limitant les ruptures de parcours. L'annotation et la structuration de l'ontologie nous permettent d'identifier quelles sont les demandes qui sont faites. Ainsi nous pouvons identifier qu'une demande d'admission dans une structure médicale a été faite au moins une fois au cours du parcours pour 198 personnes. Les demandes sont réparties ainsi : 26 pour un Service de Soins de Rééducation, 49 une admission en USP<sup>5</sup>, 10 pour un HDJ<sup>6</sup>, 25 ont concerné une demande pour une HAD<sup>7</sup>. Pour 141 cas une demande de pch a été formulée et 76 ont concerné une demande d'APA, la mise en place de ces dossiers est essentielle lors du diagnostic. Les temps administratifs de traitement des dossiers sont longs ce qui peut avoir pour conséquence un retard dans la mise en place des aides à déployer. Dans le cadre de demande de placement nous pouvons identifier qu'une demande de maison de retraite a été formulée pour 21 patients dont la moyenne d'âge est de 69,85 ( $\pm 8,39$ ) (12 femmes et 9 hommes ; 15 présentent une forme spinale, 4 une forme bulbaire ; 3 patients ont une VNI et 1 une gastrostomie ; au niveau social 8 personnes vivent en famille et 10 seuls). Les demandes de MAS<sup>8</sup> ont été exprimées pour 5 patients (2 femmes et 3 hommes ; 4 présentent une forme spinale et 1 une forme bulbaire ; 3 patients possèdent une VNI ; 3 patients vivent en famille et 1 seul) dont la moyenne d'âge est de 54,2 ans.

### 8.1.6 Aligner les ressources aux besoins : la demande et la recherche de professionnel

Dans ONTOPARON, nous avons créé le concept de **Demande de recherche de professionnel**. Il est défini par la formulation explicite d'une demande d'aide aux coordinateurs, pour trouver un professionnel de santé (kinésithérapeute, orthophoniste, médecin traitant, etc.). Nous avons souhaité observer si la présence de cette demande était en lien à l'action de coordination qui en découle qui est la **Recherche de Professionnel** ( $r=0.28$  ;  $p<0.001$ ).

Pour cette analyse nous avons transformé les données quantitatives en données catégorielles de 2 niveaux (0 = absence de demande ou de recherche de professionnel ; 1 = présence de demande ou recherche effectuée) pour voir le nombre de patients concernés. Si dans 64 cas (7%) une «demande de recherche de professionnel» a été émise, la recherche de professionnel a concerné 209 personnes (22% des patients). Dans 145 cas, les coordinateurs ont réalisé une recherche de professionnel sans que l'on puisse relier directement la recherche effectuée à la formulation d'une demande. Cette différence peut avoir deux hypothèses :

- la demande a été émise, mais le système d'annotation ne l'a pas identifiée,

---

5. Unité de soins Palliatifs  
 6. Hôpital De Jour  
 7. Hospitalisation à domicile  
 8. Maison d'accueil spécialisée

TABLEAU 8.5 – Demande et recherche de professionnel.

<b>Demande de recherche de professionnel</b>	64
<b>Recherche de professionnel effectuée</b>	209
Recherche de kinésithérapeute	162
Recherche de médecin traitant	16
Recherche d’orthophoniste	18
Recherche d’IDE <sup>9</sup>	14
Recherche de Psychologue	9

- les coordinateurs ont évalué et identifié un manque ou un besoin, nécessitant la mise en place d’un professionnel de santé.

L’ontologie permet d’identifier de façon spécifique la catégorie de professionnel recherché. Le tableau 8.5 indique les professionnels de santé qui sont les plus recherchés par les coordinateurs. Pour certains patients, la recherche de professionnel ne concerne pas un mais plusieurs professionnels (par exemple, kinésithérapeute et IDE ; kinésithérapeute et médecin etc.). Les kinésithérapeutes sont la catégorie de professionnels paramédicaux les plus recherchés dans l’accompagnement des patients SLA.

#### 8.1.6.1 Recherche de professionnel en fonction de la densité de professionnels par territoire

Nous avons souhaité observer si la recherche de professionnel pouvait être expliquée par la densité de professionnels par département. L’hypothèse formulée est que la recherche de professionnel doit être liée à une faible présence de ces professionnels dans les territoires. Pour cela nous avons utilisé les cartes interactives de densité de professionnels, disponibles sur le portail d’accompagnement des professionnels de santé de l’ARS Île-de-France<sup>10</sup>.

La recherche de kinésithérapeute est plus importante dans l’Essonne, en Seine-Saint-Denis et à Paris et elle est moins importante dans les Yvelines et en Seine-et-Marne. En comparant ces résultats à la carte de densité des kinésithérapeutes en Île-de-France (figure F.1 en annexe F), nous ne pouvons pas expliquer le manque de kinésithérapeute par une faible densité de ces professionnels dans un territoire donné. La recherche de kinésithérapeute n’est pas en toujours en lien avec un territoire sous doté. Cela est vrai pour la Seine-Saint-Denis (territoire sous dotés), mais faux pour le département de Paris et la Seine-et-Marne.

Nous avons réalisé la même démarche, pour la recherche de médecin (plus recherché à Paris et en Seine-Saint-Denis) et la recherche d’orthophoniste (plus recherché à Paris, Essonne, Val-de-Marne et en Seine-et-Marne). Là encore, la recherche de professionnel n’est pas toujours liée à une faible densité de professionnels sur ce même territoire (figure F.2, et figure F.3 en annexe F).

10. <http://www.iledefrance.paps.sante.fr/Le-zonage-medecin.42130.0.html>

TABLEAU 8.6 – Données issues de l’annotation sémantique des concepts définis du module noyau.

	<b>Ensemble de la population</b> Moyenne $\pm$ ET Présence $n$ (%) vs Absence $n$ (%)
<b>Domaine Aide Humaine</b> Présence vs Absence	3,57 $\pm$ 4,33 707 (76 %) vs 221 (24 %)
<b>Domaine Aides techniques</b> Présence vs Absence	29 $\pm$ 42,23 859 (93 %) vs 69 (7 %)
<b>Domaine de l’épuisement</b> Présence vs Absence	3,67 $\pm$ 6,56 507 (55 %) vs 421 (45 %)
<b>Domaine besoin et demande prise en charge médicale</b> Présence vs Absence	3,74 $\pm$ 4,27 757 (82 %) vs 171 (18 %)
<b>Domaine besoin et demande prise en charge sociale</b> Présence vs Absence	3,73 $\pm$ 4,53 724 (78 %) vs 204 (22 %)
<b>Domaine Problème social</b> Présence vs Absence	3,10 $\pm$ 4,12 695 (75 %) vs 233 (25 %)
<b>Domaine accessibilité</b> Présence vs Absence	2,11 $\pm$ 5,32 468 (50 %) vs 460 (50 %)

L’hypothèse que nous faisons, est que les demandes doivent probablement intervenir lorsque le patient rencontre plus de difficultés dans ces déplacements, ne lui permettant plus de se rendre en cabinet. La poursuite des soins nécessite, alors, le déplacement du professionnel au domicile. La densité de professionnel n’explique pas la recherche qui est faite. Nous faisons l’hypothèse que les professionnels peuvent être présents sur un territoire donné mais qu’ils ne réalisent pas, ou refusent de faire des prises en charge à domicile.

Si les recherches ont été effectuées, il ne nous est pas possible, actuellement, de quantifier le temps nécessaire aux coordinateurs pour mener ces démarches (identifier les professionnels sur le territoire, les solliciter, exposer la situation du patient et la pathologie, les relancer si pas de réponse, etc.), ni de savoir si ces recherches aboutissent à la mise en place du professionnel. La lecture de certains événements nous ont permis de voir que dans certains départements comme à Paris, les coordinateurs peuvent solliciter jusqu’à 15 cabinets de kinésithérapeutes. Cette tâche peut être chronophage, il peut être difficile pour des patients ou leur entourage, de rechercher un professionnel qui puisse venir au domicile, ils sollicitent alors le Réseau SLA IDF pour intervenir.

### 8.1.6.2 L’annotation des concepts définis du module noyau

Le tableau 8.6, indique les valeurs quantitatives obtenues, pour les concepts définis de l’ontologie ONTOPARON. Ces concepts définis représentent les grandes thématiques intervenant dans le parcours patients.

Nous pouvons faire quelques observations suite aux données obtenues, selon le do-

maine :

**Le domaine des aides techniques**, est la thématique qui va concerner le plus de patient (93%), alors que le problème d'accessibilité ne concerne que 50% des patients. Ceci peut s'expliquer par :

1. la SLA provoque progressivement des incapacités physiques qui nécessitent la mise en place de compensations, quelle que soit sa forme initiale. Dans les formes bulbaires, la nécessité de mettre en place des outils de communication peut apparaître rapidement. Dans les formes spinales l'atteinte progressive au niveau des membres inférieurs, par exemple, nécessite une graduation des aides techniques de déplacement (béquille, releveur, rollator, fauteuil roulant manuel et fauteuil roulant électrique) ;
2. la présence d'une coordinatrice ergothérapeute, au sein du Réseau SLA IDF peut être un facteur favorisant les préconisations, les conseils et l'orientation des patients. L'expertise de cette professionnelle permet, d'avoir une réponse rapide ou bien d'être orienté vers une personne (revendeur, association de patients, etc.) ou une structure ressource (centre d'essais, plateforme nouvelle technologie, etc.). Les besoins en aides techniques nécessitent une évaluation fréquente par des experts dans le domaine et une expertise sur les différents dispositifs (Moutet, 2006). En effet les besoins vont évoluer au cours du temps et vont concerner tous les domaines de vie quotidienne : la mobilité, le positionnement et la communication (Hobson & McDermott, 2016).

**Domaine des problèmes d'accessibilité.** Un patient sur deux va rencontrer des difficultés d'accessibilité dans son environnement. L'évaluation et la préconisation d'aménagement sur le lieu de vie ou lieu de travail, nécessite de faire une évaluation en situation (sur place). Le nombre de patients inclus au Réseau SLA IDF étant important, il n'est pas possible pour l'ergothérapeute de réaliser toutes ces évaluations. Lors de l'identification de problème d'accessibilité, il est fréquent que la coordinatrice recherche une structure, ou un professionnel de proximité pour effectuer cette évaluation.

**Domaine des besoins et demandes de prise en charge médicale et domaine des besoins et demandes de prise en charge sociale.** Pour ces deux thématiques, la moyenne est proche pour l'ensemble de la population (3,73), cependant rapportée au nombre de patients concernés, les aspects médicaux vont concerner 82 % de la population et les aspects sociaux 78 %. La première observation est que pour certains patients, les demandes et besoins dans ces deux domaines sont absents. L'importance des aspects médicaux est cohérent du fait de la pathologie. Pour les aspects sociaux, nous avons émis deux hypothèses : a) les coordinateurs sont clairement identifiés comme personnes ressources pouvant aider dans les questions et démarches liées aux aspects sociaux (mise en place de dossier de prestations "apa ou pch" par exemple) ; b) lors d'appels ayant pour thématique principale, des questions d'ordre médical, le questionnement et l'évaluation de la situation par le coordinateur, permettent d'identifier des besoins dans la thématique sociale.

**Domaine de l'aide humaine et problème social.** Ces deux thématiques sont proches en termes de personnes concernées, respectivement 76% et 75% de la population. Nous pouvons faire l'hypothèse que la présence de problèmes sociaux, comme l'isolement social ou des difficultés familiales nécessitent la mise en place d'aides humaines, pour intervenir auprès du patient.

**Domaine de l'épuisement.** Il concerne 54% de la population soit 507 patients, c'est-à-dire que pour un patient sur deux, une situation d'épuisement est rencontrée au cours du parcours. La SLA, par sa nature progressive implique une dépendance croissante auprès des aidants, qu'ils soient des aidants professionnels ou aidants familiaux [Brulletti et al. \(2015\)](#). La présence d'un épuisement de l'aidant dans la SLA est reconnu dans de nombreuses études, il est lié à l'état cognitif et comportemental du patient ([Burke et al., 2015](#)) Si l'épuisement est présent dans plus de 50% de la population, il existe des patients pour lesquels l'épuisement est absent.

Ces premiers résultats montrent que les thématiques que nous avons choisies et définies grâce aux concepts définis, sont fréquentes dans les parcours de santé des patients. Le repérage de ces items dans les parcours, fait sens avec la description des conséquences de la SLA de la première partie de ce manuscrit. La diminution progressive des capacités motrices nécessite le recours à des compensations (humaines ou techniques) qui doivent être mises en place et financées. Cependant pour tous ces domaines, on observe que certaines personnes ne rencontrent pas les mêmes problématiques ou n'ont pas les mêmes demandes.

### 8.1.6.3 L'analyse des concepts définis du module noyau par les données socio-démographiques

Après la description des concepts définis ci-dessus nous avons souhaité analyser ces données avec les données socio-démographiques (sexe, âge à l'inclusion, statut social, mode de vie, département et les données de parcours (forme de la pathologie, les délais "T1", "T2" et "T3" cf. [8.1](#)), en utilisant un modèle de régression linéaire où  $X$  est le domaine des concepts définis et  $Y$  les variables socio-démographiques et données de parcours.

Le domaine des aides techniques va concerner préférentiellement les patients les plus jeunes ( $p < 0.001$ ), elle augmente avec le temps de présence au Réseau SLA IDF ( $p < 0.001$ ) et est plus importante dans les formes spinales que dans les formes bulbaires ( $p < 0.0095$ ). Ces résultats sont compréhensibles en prenant en compte les caractéristiques de l'atteinte spinale qui touche principalement les membres, limitant le patient dans les gestes de vie quotidienne et nécessite la mise en place d'aides techniques.

Les besoins et les demandes de prise en charge médicale sont liées au sexe ( $p = 0.003$ ) plus fréquemment annotés chez les femmes que chez les hommes ; à l'âge ( $p = 0.0033$ ) plus nombreuses avec l'avancé en âge ; elles progressent avec le temps de présence au Réseau SLA IDF ( $p < 0.001$ ) ; le mode de vie ( $p = 0.0143$ ) est aussi un indicateur avec

des moyennes plus importantes pour les personnes vivant en famille (en comparaison des personnes qui vivent seules). L'hypothèse qui peut-être faite est que lorsque le patient vit en famille, avec la présence d'un tiers à domicile (enfant, parent, frère, etc.), ce dernier sollicite le réseau à l'apparition de nouveaux symptômes ou difficultés.

Le domaine de l'aide humaine est lié à la forme ( $p = 0.0017$ ) et à l'âge ( $p = 0.007$ ) au temps au Réseau SLA IDF. Les personnes présentant une forme spinale présentent des besoins et des demande plus fréquents que les personnes présentant une forme bulbaire ( $p = 0.0017$ ), l'atteinte initialement motrice dans la forme spinale permet d'expliquer cette différence. Les besoins en aide humaines varient en fonction de l'âge à l'inclusion ( $r=-0.25$ ;  $p<0.0001$ ), ils sont moins importants pour les personnes les plus âgées. Le recours aux aides humaines nécessitent un financement, le reste à charge étant plus important pour les plus de 60 ans, comme indiqué dans le rapport de l'IRDES [Penneau et al. \(2019\)](#), les personnes peuvent décider de ne pas y recourir.

Les besoins et demandes de prise en charge sociale sont liées : à l'âge ( $p<0.001$ ) ils sont plus importants pour les personnes les plus jeunes ; augmentent avec le temps de présence au Réseau SLA IDF ( $p<0.001$ ) ; bien que la forme et le département aient des p-value faibles (respectivement  $p= 0.0253$  et  $p= 0.032$ ) le test de student met en avant des différences pour chacune de ces variables. Ainsi les besoins et demandes sont plus importantes pour les personnes présentant des formes spinales et pour les personnes vivants à Paris. Dans les formes spinales, les atteintes motrices vont limiter progressivement l'indépendance des personnes dans la réalisation des activités de vie quotidienne. La mise en place d'aides techniques et d'aides humaines vont être nécessaire, ce qui est en lien avec les résultats rencontrés ci-dessus. La mise en place de ces aides nécessitent des financements pour lesquels des démarches administratives doivent être réalisées (dossier apa, dossier pch), ce qui explique que le concept des besoins et demandes sociales soit plus fréquemment annoté pour les personnes ayant une forme spinale.

Pour le domaine de l'épuisement, la variable du temps au Réseau SLA IDF ( $p=0.0045$ ) du mode de vie ( $p=0.0169$ ) et du statut social ( $p =0.0138$ ) sont identifiées. Le test de student montre que la moyenne de l'épuisement est plus important pour les personnes célibataires que chez les personnes mariées ( $p = 0.0276$ ), et plus fréquent pour les personnes qui vivent en familles que les personnes qui vivent seules ( $p = 0.0169$ ). L'épuisement va apparaître au cours du temps et doit être liée à l'augmentation progressive des déficiences des personnes.

### 8.1.7 Synthèse des observations

Nous avons décrit les données socio-démographiques et les données issues des annotations. Nous pouvons voir de nombreuses disparités sur ces données. La prise en compte des moyennes et des écarts-types apporte des informations globales. En transformant certaines de ces données quantitatives en données qualitatives, il est possible

d'observer combien de patients sont concernés par ces thématiques. Ces premiers résultats nous permettent de dire que tous les patients n'ont pas le même parcours, les mêmes besoins ni les mêmes demandes et les actions de coordination qui en résultent sont différentes.

Lors de l'analyse des concepts définis (des grandes thématiques sociales, médicales, de l'aide humaine, etc.) ont observé que toutes ces variables sont liées au temps de suivi au Réseau SLA IDF. Ceci peut s'expliquer par la formulation des demandes par les patients ou professionnels qui ont clairement identifié le rôle du Réseau SLA IDF, de plus l'accompagnement peut permettre la détection et l'identification des besoins par les coordinateurs au cours du suivi. L'évolution naturelle de la SLA font que les besoins vont évoluer dans le temps, plus les personnes solliciteront le Réseau SLA IDF, plus ces thématiques seront annotés dans les corpus.

Lorsque l'on évoque la notion de parcours, la notion temporelle est implicitement liée. Dans notre approche et dans les outils développés, les événements d'un patient sont annotés individuellement, cependant le fichier de sortie de ONBASAM va générer une annotation globale de tous les événements du patient. Ce premier travail, d'analyse dans une grande population peut ensuite permettre un travail plus fin où les annotations sont placées chronologiquement. La chronologie des faits doit apporter plus d'informations sur les séquences d'apparition des problématiques. Il serait alors possible d'observer si des demandes, des besoins ou des problèmes apparaissent à des temps spécifiques dans le parcours.

## 8.2 **Sous populations spécifiques : patients vivant en institution, patients tracheotomisés, patients sans besoins ni demandes**

Lors de l'analyse des données d'annotations et des données socio-démographiques, nous avons identifié trois catégories de patients spécifiques : les patients vivant en institution, les patients ayant une trachéotomie et les patients sans demandes et besoins médicaux et sociaux.

### 8.2.1 **Patients vivant en institution**

Nous avons souhaité isoler, les patients vivant en institution. **17 des patients** de la population étudiée vivent en maison de retraite. Nous avons séparé ces personnes de la population globale, en faisant l'hypothèse que le fait de vivre dans une institution médicalisée ne devait pas créer les mêmes besoins ni les mêmes demandes dans le parcours de santé. La présence de professionnels médicaux et paramédicaux permet d'avoir un recours aux soins nécessaires. Le nombre de 17 patients inclus dans cet échantillon ne nous permet pas de faire des tests statistiques robustes. Il ne sera présenté que des observations sur cet échantillon.

Dans cet échantillon il y a **12 femmes et 5 hommes**, l'âge moyen à l'inclusion dans le Réseau SLA IDF est de **79 ans** ( $\pm 6,7$ ). Le délai moyen du diagnostic à

l'inclusion est de 177 jours ( $\pm 104$ ) et le temps au Réseau SLA IDF est en moyenne de 242 jours ( $\pm 224$ ). La moyenne d'évènements par patient est de **22,41** ( $\pm 22,60$ ) et le nombre d'action de coordination est en moyenne de **50,76** ( $\pm 43,36$ ).

Dans cette population il n'a pas été repéré, par l'annotation, de demande ou de recherche de professionnel, le fait de vivre en institution doit permettre d'accéder aux professionnels de santé. La recherche de structure est annotée, et elle concerne la recherche de réseaux de soins palliatifs, et la recherche de séjour de répit. Le concept défini le plus annoté est le domaine des besoins et demandes médicales. Les structures semblent bien identifier le Réseau SLA IDF comme ressource pouvant donner des indications, des conseils ou faire le lien avec le centre expert sur les problématiques médicales. Bien que les personnes travaillant en institution soient des professionnels de santé, il leur est nécessaire d'avoir recours à des personnes ressources, pour accompagner les patients SLA.

Deux des domaines, nous ont interpellé et nous avons souhaité les regarder en particulier : le domaine des aides techniques qui est présent dans 100% des cas et celui de l'épuisement de l'équipe présents dans 64% des cas.

#### 8.2.1.1 Les aides techniques en institution

La question des aides techniques est abordée pour tous les patients, les aides techniques concernent principalement : les aides à la mobilité (essentiellement les fauteuils roulants) dans 85% des cas, les aides techniques de transferts et le contrôle de l'environnement dans 64%, les produits pour la communication pour 53% des patients et les aides techniques liées au confort et au positionnement dans 41% des cas. Ces données bien qu'issues d'un petit échantillon suivent une répartition à peu près identique aux besoins d'aides techniques identifiés dans l'étude de [Funke \*et al.\* \(2018\)](#) effectuée en Allemagne.

La spécificité des aides techniques, notamment pour la communication et pour les déplacements, nécessite des évaluations fréquentes par la progression parfois rapide des déficiences. La mise en place de ces dispositifs nécessite une expertise, ce qui peut expliquer que les professionnels travaillant en institution sollicitent les coordinateurs du Réseau SLA IDF pour être conseillé sur les dispositifs existants et les mieux adaptés aux besoins des personnes.

#### 8.2.1.2 L'épuisement en institution

Nous avons été surpris d'observer que le domaine de l'épuisement était présent pour 11 patients (64%). Lors de l'analyse des données, nous avons remarqué que l'épuisement était associé à l'annotation des classes de problème de comportement et de difficulté de prise en charge. Comme pour la population générale, la présence d'un épuisement va être en lien avec les actions de coordination de recherche de séjour de répit. Pour la suite des analyses ces patients ont été exclus de la population.

Si pour 17 personnes il est identifié qu'elles vivent en institution, nous avons pu identifier que pour 21 patients (non inclus dans ceux vivant en institution), une demande de maison de retraite a été faite au Réseau SLA IDF. Au niveau des caractéristiques

TABLEAU 8.7 – Comparaison des moyennes d’annotations des concepts définis entre la population globale et les patients trachéotomisés.

Concepts définis	Population général	Patients trachéotomisés	<i>p</i>
Action de coordination	73,58 ( $\pm$ 74,9)	191 ( $\pm$ 233)	0,001
Domaine besoin et prise en charge sociale	3,67 ( $\pm$ 4,43)	7,12 ( $\pm$ 8,79)	0,005
Domaine besoin et demande de prise en charge médicale	3,73 ( $\pm$ 4,22)	7,9 ( $\pm$ 9)	0,005
Domaine de l’aide humaine	3,6 ( $\pm$ 4,3)	7,62 ( $\pm$ 9,66)	0,008
Domaine de l’épuisement	3,65 ( $\pm$ 6,52)	5 ( $\pm$ 7,10)	0,5
Domaine des aides techniques	28,7 ( $\pm$ 41,6)	70,4 ( $\pm$ 92,2)	0,004

socio-démographiques des personnes pour lesquelles une demande de placement a été faite au cours du parcours, l’âge moyen est de 70 ans ( $\pm$  8,4) à l’inclusion, 15 personnes présentent une forme spinale et 4 une forme bulbaire, il s’agit majoritairement de femme 12 (et de 9 hommes), 10 de ces personnes vivent seules et 8 en famille. La présence d’un isolement social semble être corrélé à une demande d’entrée en institution ( $p=0.0017$ ).

### 8.2.2 Patients avec une trachéotomie

Dans la population étudiée seules 8 personnes ont une trachéotomie. Les patients trachéotomisés ont en moyenne **56 ans** ( $\pm$  14,02), le temps d’accompagnement au Réseau SLA IDF est de 753,25 jours ( $\pm$  591). La moyenne d’évènement est de **86** ( $\pm$  96) et le nombre d’actions de coordinations est en moyenne de **191,25** ( $\pm$  233,58). Ce groupe est constitué de 5 hommes et de 3 femmes. Tous les patients vivent en famille, au niveau du statut social : 2 personnes sont célibataires, 1 personne est divorcée et 4 sont mariées (pour l’une des personnes l’information sur le statut social est manquante).

Les personnes vivants avec une trachéotomie présentent un nombre important d’évènements et d’actions de coordination par rapport à la population générale. La durée d’accompagnement au Réseau SLA IDF est plus importante, puisqu’il est de 753 jours pour les patients ayant une trachéotomie contre 548 jours en moyenne pour la population générale. Ce qui est concordant avec les données de la littérature indiquant que la présence d’une trachéotomie est associée à une augmentation de la durée de survie dans la SLA. Nous avons comparé les moyennes des concepts définis, entre les personnes ayant une trachéotomie et la population totale. Le tableau 8.7 présente ces résultats.

On observe des différences significatives entre les deux populations. Seul le domaine de l’épuisement n’est pas significatif sur les données quantitatives. L’observation de l’épuisement de manière qualitative, indique que l’épuisement est présent pour 6 des 8 patients (75% des cas) ; ce qui est supérieur à la population totale. Si des différences sont observées, entre les deux populations, pour les concepts définis, il en est de même pour les actions de coordination. La moyenne des actions de coordination est supérieure à

celle rencontrée dans la population. L'hypothèse que nous pouvons formuler, est que les patients vivant avec une trachéotomie nécessitent plus de soins et d'accompagnement dans leur parcours. La taille de cette population limite les analyses qui peuvent être faites.

### 8.2.3 Patients sans besoins et demandes, et sans actions de coordination

Nous avons pu observer que certains patients ne présentent pas de demandes ni de besoins, dans deux des domaines principaux (social et médical). Nous avons réalisé un regroupement des patients pour lesquels le « Domaine de demande et besoin de prise en charge médicale et sociale » était à 0 et qui avaient au maximum 3 actions de coordination<sup>11</sup>.

Ce regroupement a permis d'identifier 27 patients dont les données sont indiquées au tableau F.2 en annexe. La moyenne d'âge lors de l'inclusion est de **68 ans** (médiane à 69 ans). Les personnes sont majoritairement mariées 73% et vivent en famille dans 78% des cas, seules 6 personnes ont une suppléance (5 une VNI et 1 une gastrostomie). La moyenne d'évènement et de suivi est de **3,63** (médiane à 3), contre 33,6 pour la population totale. Ce sont donc des personnes sans activité de coordination du Réseau SLA IDF.

Pour ces patients le temps d'accompagnement au réseau est de **546 jours** ( $\pm 476$  avec un minimum de 8 jours et un maximum de 1744 jours). Ce temps d'accompagnement nous a semblé important en comparaison du nombre d'évènements. Nous avons observé la date du dernier évènement saisi. 12 des patients « vivant » au moment de l'analyse, ont été inclus avant 2016 et pour 6 d'entre eux, le dernier évènement saisi date de 2017. En indiquant la date du dernier évènement saisi sur la base, le temps au Réseau SLA IDF passe de 546 à 284 jours ( $\pm 389$ ).

Au-delà des données des actions de coordination et des évènements, nous avons calculé la moyenne des signes cliniques et des actions médicales. L'hypothèse que nous pouvons formuler est que les patients peuvent être inclus au Réseau SLA IDF mais ne ressentent pas le besoin de le solliciter. Ils peuvent cependant continuer de bénéficier du suivi au centre expert SLA.

Les actions médicales ont une moyenne de 3,59 ( $\pm 5,71$ ) ; les signes cliniques une moyenne de 4,93 ( $\pm 8,7$ ). Les données indiquent que le Réseau SLA IDF reçoit des informations sur l'état de santé des patients, mais ils n'interviennent pas dans les parcours. Pour valider ces informations nous avons réalisé une lecture des évènements qui montre :

- les évènements saisis évoquent : le processus d'inclusion (présentation du réseau, entretien avec le patient, appel au médecin) et la saisie de différents comptes rendus médicaux (consultation, cmd).

---

11. Le choix de 3 actions de coordination correspond à une activité minimale de coordination. Il s'agit de : a) la présentation du réseau lors de la consultation du retour d'annonce, b) la réception des documents d'inclusion, c) l'appel au médecin pour l'informer de l'inclusion du patient dans le Réseau SLA IDF.

- les dates de saisies des derniers évènements, sont pour certaines anciennes (2017). Certains patients sont : soit perdus de vue, soit décédés et le Réseau SLA IDF n'a pas eu l'information, pour mettre à jour la base.
- la présence d'actions médicales et de signes cliniques, présents dans les comptes rendus, indiquent que les personnes bénéficient du suivi au centre expert, sans avoir recours à la coordination. Il n'est pas possible de dire que ces personnes soient en rupture de parcours.

La taille de cet échantillon est trop petite pour faire des conclusions sur les raisons qui font que certains patients n'aient pas de demandes ni de besoins, comme les autres patients. Il est simplement possible d'observer que par rapport à la population totale, les patients sont en moyenne plus âgés et que dans 78% des cas ils vivent en famille. Le mode de vie peut avoir un impact sur les besoins en coordination.

#### 8.2.4 Synthèse

Bien que ces trois populations de patients soient faibles en nombre de personnes, les données indiquent que :

- la présence d'une trachéotomie va nécessiter une forte activité de coordination, tout comme la formulation de demandes et de besoins dans le domaine médical et social probablement pour suivre l'évolution de la pathologie et adapter les aides au domicile,
- l'âge, le mode de vie (en institution ou en famille), sont des critères qui ont un impact sur les types de besoins et des demandes formulées,
- l'épuisement est une problématique fréquente dans la SLA, qui se retrouve en institution et à domicile.

### 8.3 Analyses globales

#### 8.3.1 Liens entre les évènements et les actions de coordinations

Dans la description des données, nous avons observé la variation du nombre d'évènements et des actions de coordination. L'hypothèse est que ces deux variables doivent être corrélées, des sollicitations doivent engendrer des actions de coordination. Il existe une forte corrélation ( $r=0.96$  ;  $p<0.001$ ) entre les actions de coordination et le nombre d'évènements par patients. De plus la Plus il y aura de sollicitations, ou d'informations reçues, plus les coordinateurs du Réseau SLA IDF pourront intervenir dans le parcours des patients. Si les patients ou les professionnels ne sollicitent pas le Réseau SLA IDF pour intervenir, les patients peuvent ne pas bénéficier de coordination.

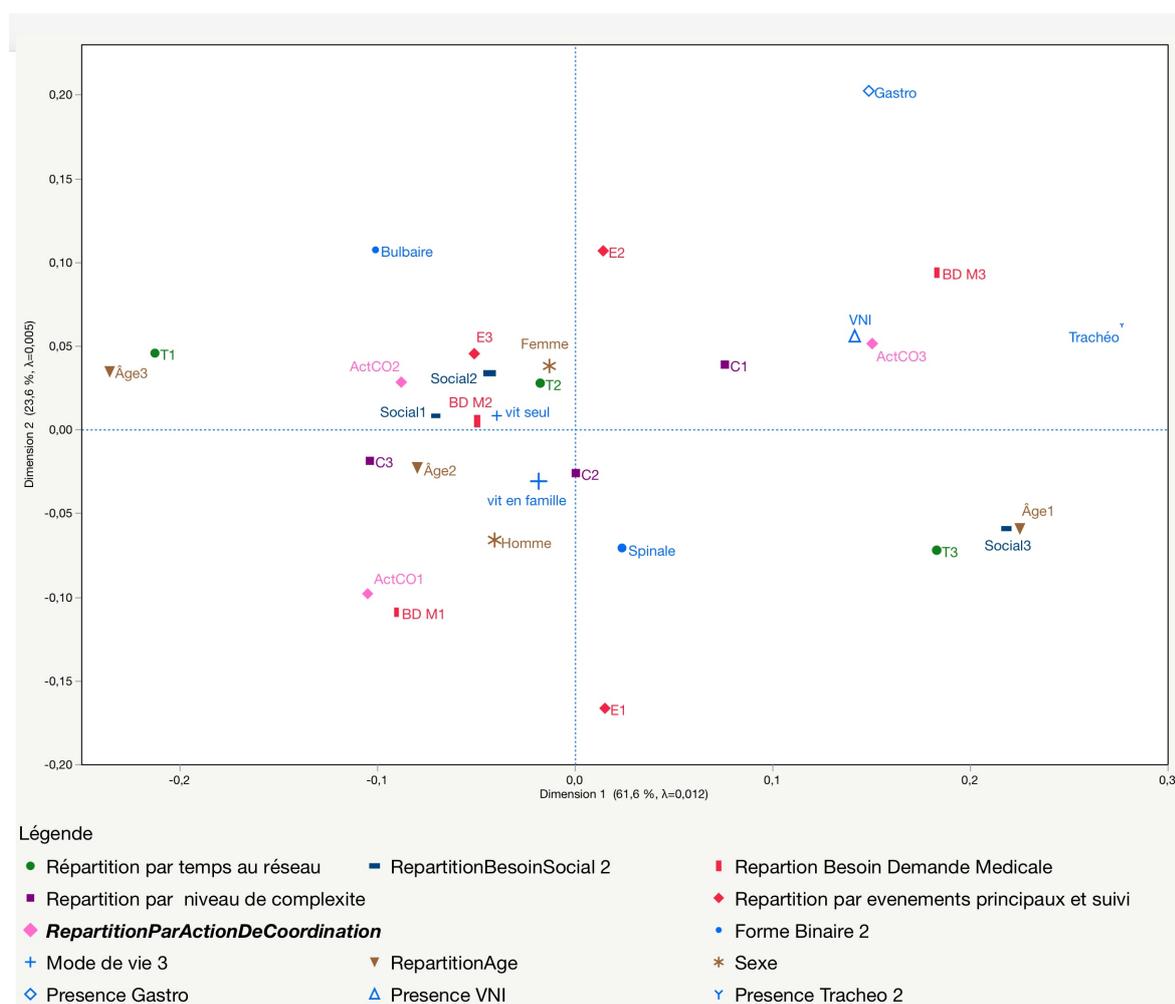


FIGURE 8.2 – Analyse en correspondance multiple montrant les associations entre les données issues de l’annotation et les caractéristiques des patients.

### 8.3.2 Analyse en correspondance multiple

Nous avons réalisé des classements en tertiles des patients selon différentes variables : a) l’âge, b) le nombre d’actions de coordination, c) le score de complexité, c) les demandes et besoins médicaux et sociaux et d) le temps au Réseau SLA IDF. Le tableau F.3 en annexe, indique les moyennes et les écarts types de chacun des groupes.

Pour comprendre les liens entre ces variables, nous avons réalisé une analyse des correspondances multiples (Roux & Rouanet, 2010), en intégrant dans le modèle, les caractéristiques du patient (âge, la forme de la SLA, la présence de suppléances et le mode de vie) ainsi que les données d’activités de parcours (temps au Réseau SLA IDF, nombre d’action de coordination niveau moyen de complexité, demande et besoins médicaux et sociaux). L’analyse en correspondance multiple est illustrée en figure 8.2, utilisant deux dimensions chacune expliquant respectivement 61,6% et 23,6% de l’inertie totale.

Différentes observations peuvent être formulées.

**Lien entre les besoins et demandes médicales (BD M) et les actions de coordination (ActCO).** Le groupe pour lequel il y a peu d'actions de coordination est proche du groupe ayant peu de besoins et demandes médicales (ActCO1 et BD M1). Le groupe pour lequel il y a plus d'actions de coordination est proche du groupe ayant plus de besoins et demandes médicales (BD M3).

**L'âge des personnes est lié à la durée d'accompagnement et à la thématique sociale.**

Il semble que les personnes les plus jeunes, ont plus de besoins et de demandes dans la thématique sociale. Les patients les plus âgés (Âge 3) ont un temps d'accompagnement plus court (T1), et le groupe de patients les plus jeunes (Âge1) ont un temps d'accompagnement plus long (T3). L'âge moyen des personnes est de 78,17 ans (pour le groupe Âge 3), alors qu'il est de 52,06 ans dans le premier tertile (Âge 1). Ces données peuvent s'expliquer sur le critère que l'âge élevé est un facteur de mauvais pronostic lié à un processus dégénératif plus rapide ([Husami et al., 2018](#)) Au-delà du temps d'accompagnement au Réseau SLA IDF, il semble que l'âge de la personne, va être lié à la formulation des besoins et demandes sociales. Ainsi les personnes plus jeunes vont avoir au cours de leurs parcours plus de besoins et de demandes au niveau social. Cela peut s'expliquer par :

- la poursuite d'une activité professionnelle, qui nécessite des adaptations sur le lieu de travail,
- la mise en place du dossier [MDPH](#) et la ré-évaluation des plans de compensation du handicap,
- la mise en place progressive d'aides humaines et d'aides techniques.

**Le mode de vie, le sexe et la forme de la pathologie** Ces trois caractéristiques sont regroupées au centre du graphique et ne permettent pas de tirer de conclusion.

### 8.3.3 Le niveau de complexité

L'analyse en correspondance multiple ne nous permet pas d'apporter des éléments de compréhension sur le niveau de complexité. La complexité n'est indiquée que pour l'évènement principal, peut-être, l'attribution d'un score pour chaque évènement de suivi permettrait de mieux identifier les thématiques du parcours qui créées de la complexité dans la prise en charge.

### 8.3.4 Patients avec ruptures de parcours identifiés et des difficultés de prise en charge

Les ruptures de parcours peuvent être analysées selon deux aspects :

**La rupture de parcours de soins**, correspond à une non continuité ou l'absence de soins et/ou la présence d'hospitalisations évitables (hospitalisation non-programmée). Pour identifier si une hospitalisation annotée dans les corpus est une hospitalisation évitable, l'intervention d'un médecin est nécessaire. L'évolution de la SLA,

nécessite des temps d'hospitalisations notamment lors de la mise en place de suppléance, ou bien lors d'apparition d'épisode d'aggravation respiratoire, qui peuvent nécessiter une hospitalisation en urgence (non-programmé), mais qui est nécessaire. Si l'annotation des corpus, permet de repérer les hospitalisations, il ne nous est pas possible d'identifier si ces hospitalisations sont ou non des hospitalisations évitables.

**La rupture de parcours de santé**, va prendre en compte la dimension médicale et médico-sociale. Certaines classes de l'ontologie permettent d'identifier clairement la présence de « rupture de parcours », et les « difficultés de prise en charge ». Ainsi la classe de l'ontologie « Difficulté de prise en charge médico-sociale » subsume les classes : absence de prise en charge, arrêt de la prise en charge, difficulté de prise en charge, échec de mise en place, refus de prise en charge, rupture de prise en charge, urgente prise en charge.

Les patients pour lesquels ces classes sont annotées sont donc en risque ou en rupture de parcours.

Nous avons identifié 247 patients pour lesquels ces difficultés de prise en charge sont annotées et pour 139 d'entre eux la rupture de prise en charge est identifiée. La population concernée par ces difficultés et ces ruptures est composée de 121 femmes et 126 hommes, 70% des patients présentent une forme spinale et 30% une forme bulbaire. Les patients vivent à 84% en famille et dans 14% des cas seuls, les personnes sont principalement mariées 63%, (14% célibataires, 15% divorcés et 7% veufs). L'âge moyen des personnes présentant des difficultés est de 64,4 ans ( $\pm 11,9$ ) et de 66,09 ( $\pm 11,9$ ) pour ceux ne présentant pas de difficultés. Nous avons souhaité identifier si des facteurs spécifiques étaient liés à la présence de ces difficultés de prise en charge. Les résultats obtenus montrent que les difficultés de prise en charge sont liées à la présence de :

1. l'expression d'un refus par les personnes ( $p < 0.001$ ),
2. à la présence de signes démontrant une altération de l'état cognitif ( $p < 0.001$ ),
3. à une altération des fonctions motrices ( $p = 0.0106$ ),
4. à la mise en place de structure médico-sociale comme un SSIAD, une HAD-R ou un SAMSAH<sup>12</sup> ( $p = 0.009$ ),
5. à la recherche d'un séjour de répit ( $p = 0,0273$ ),
6. à la mise en place de prestations sociales (apa, pch<sup>13</sup>).

L'expression de refus du patient lors du parcours semble être en lien avec une difficulté ou une rupture de prise en charge. Ce refus peut être dû à des altérations dans le jugement de la personne ou bien dans des difficultés de comportement qui peuvent produire un arrêt de la prise en charge par les intervenants. Pour expliquer la présence d'altérations motrices qui soient liées aux difficultés de prise en charge et qui prennent en compte le refus du patient, l'hypothèse que nous pouvons faire (avec

---

12. Service de Soins Infirmier à Domicile, Hospitalisation à Domicile de Réadaptation, Service d'Accompagnement Médico-Social pour Adultes Handicapés

13. Allocation Personnalisée d'Autonomie, Prestation de Compensation du Handicap

l'expérience des lectures des évènements) et que la diminution des capacités motrices nécessite la mise en place de matériel pour réaliser les soins d'hygiène (mise en place de lève-personne, de chaise de douche, ou de lit médicalisé pour réaliser la toilette au lit) et la mise en place d'aides humaines. Parfois les patients et les familles sont dans le refus d'utiliser ces aides, les intervenants refusent alors la poursuite de la prise en charge si les aides ne sont pas mises en place car la réalisation des soins est alors trop difficile ou chronophage. Il est intéressant de voir que la présence de ces difficultés est liée à la mise en place de structure médico-sociale, ainsi que la recherche de séjour de répit. Lors de l'identification de problèmes, les coordinateurs peuvent proposer la mise en place d'un relais de type SSIAD pour pallier les difficultés, ou bien un séjour de répit. Le séjour de répit permet de faire un point de la prise en charge et de rechercher les intervenants pouvant agir au domicile. La mise en place de structure nécessite le recours à des financements ce qui peut expliquer que la mise en place de prestations sociales de types apa et pch soient nécessaires.

## 8.4 Synthèse des analyses

Les données obtenues par l'annotation sont nombreuses. Nous avons souhaité les analyser en prenant comme point de départ les concepts définis et les données socio-démographiques. Les données obtenues semblent concordantes avec les données présentes dans la littérature, notamment, sur le besoin en aides (techniques et humaines), l'identification d'un épuisement des aidants (professionnels ou aidant naturel) et la présence de signes cliniques démontrant une altération des fonctions (respiratoires, nutritionnelles, motrices etc.) au cours du temps. L'annotation des corpus permet d'identifier que les besoins, les demandes et les actions de coordination s'expriment différemment dans les parcours des patients. La fréquence des annotations dans les concepts définis permet de mettre en avant l'intervention du réseau de coordination sur les dimensions médicales, médico-sociales et sociales.

Si pour des patients la rupture est identifiée à travers le repérage des annotations, des actions de coordination peuvent être mise en place pour limiter l'impact de ces ruptures. Cependant il est possible que des patients en rupture de parcours, ne soient pas repérés car le Réseau SLA IDF ne reçoit pas d'informations permettant d'identifier ces personnes. Les actions de coordination sont corrélées au nombre d'évènements saisis dans la base de données, une sollicitation peut permettre aux coordinateurs d'effectuer une évaluation globale et détecter des besoins, sans que des demandes aient été exprimées. Le suivi régulier, par les coordinateurs, au cours du parcours va permettre de proposer et d'orienter les patients, l'entourage et les professionnels. Les données obtenues résultent des annotations faites par le système, la structuration syntaxique et grammaticale étant variée dans les corpus il est possible que des termes n'aient pas été annotés car non reconnu. Pour améliorer l'annotation des termes l'une des pistes qui puissent être envisagées et de mener une réflexion sur les modalités de saisies des évènements. En améliorant les données textuelles initiales, le système d'annotation sémantique peut être plus performant. Bien que chacune des classes de l'ontologie ait été annotée au moins une fois, certaines classes présentent un nombre d'annotation

faible (de 1 à 3 pour l'ensemble du corpus) ce qui pose la question de la granularité de l'ontologie.

Si l'annotation par l'ontologie permet de repérer des termes et de décrire les éléments intervenants dans les parcours, nous ne pouvons pas répondre à la question initiale des causes de rupture de parcours dans la SLA. Nous pensons que les données obtenues peuvent le permettre mais elles nécessitent la mise en place d'une collaboration entre les coordinateurs, les médecins neurologues et des statisticiens afin de déterminer les indicateurs définissant les ruptures de parcours. Une fois ces indicateurs définis il est possible de faire des clusters de patients et rechercher dans les annotations la corrélation entre ces indicateurs.

Le nombre de dossiers patients inclu dans ces résultats est de 928 dossiers, il peut être envisagé dans un second temps de faire l'annotation de l'ensemble des dossiers patients du [Réseau SLA IDF](#) (2684) pour obtenir une puissance statistique plus forte et valider les hypothèses faites dans ce travail.

# Conclusion et perspectives

## Problématique

La Sclérose Latérale Amyotrophique est une pathologie neuro-dégénérative, pour laquelle a été mis en place en région Île-de-France, un réseau de coordination ville hôpital le [Réseau SLA IDF](#), dont l'objectif est d'accompagner les patients dans leurs parcours de santé. Le Réseau SLA IDF dispose d'une base événementielle (base textuelle) sur laquelle sont saisies l'ensemble des besoins et des demandes, ainsi que les actions de coordination qui sont réalisées lors du parcours de santé des patients. Le parcours de santé prend en compte la globalité de la personne, à la fois la dimension médicale, la dimension sociale et la dimension socio-environnementale. L'objectif du travail mené est d'analyser ces données, pour comprendre les causes de rupture de parcours. La compréhension des causes de rupture doit permettre d'anticiper les besoins, améliorer les parcours afin de limiter les hospitalisations évitables et améliorer la qualité de vie des patients. Les données saisies sur la base événementielle sont essentiellement textuelles. Pour exploiter ces textes, nous avons mis en place un processus, décomposé en quatre étapes. La première étape a été la réalisation de la pseudonymisation et le traitement orthographique des corpus. La seconde fut la création d'un système d'organisation des connaissances utilisées pour l'annotation sémantique des données, pour cela nous avons développé une ontologie modulaire : ONTOPARON. La troisième étape a été l'annotation des données, pour cela nous avons développé un outil d'annotation ONBASAM. Pour évaluer les annotations faites automatiquement, ainsi que le choix de modélisation des concepts définis, un protocole d'évaluation a été mis en place. Cinq experts extérieurs au projet de recherche ont réalisé l'évaluation de 410 événements, en utilisant l'outil d'évaluation développé PRONTO. Les résultats de l'évaluation (F-Mesure  $\geq 0,90$ ) nous ont permis de valider les annotations et la modélisation choisie. La dernière étape de notre processus de traitement des données fut l'extraction des annotations, de 931 dossiers patients, par un processus d'extraction de ONBASAM. Ces données d'annotations ont été mises en lien avec les données socio-démographiques des patients. Des analyses des données ont été réalisées en utilisant un logiciel statistique. Ce processus nous a permis de passer de données brutes à des données semi-structurés qui peuvent être exploitées et analysées. Si toutes les étapes du processus ont pu être mises en place, cela n'a été rendu possible que par la mise en place de collaboration

avec divers intervenants.

## L'ontologie ONTOPARON

Une ontologie permet de définir un vocabulaire commun et une représentation du domaine étudié, elle permet de partager de l'information entre des humains et d'être utilisée par une machine. Pour construire l'ontologie nous avons suivi la méthodologie ARCHONTE. La thématique de l'étude des parcours de santé, nous a fait passer rapidement d'une ontologie monolithique à une ontologie modulaire. D'une part pour faciliter le partage avec les experts du domaine, d'autre part pour sa facilité de manipulation et de mise à jour. L'ontologie ONTOPARON est composée de cinq modules :

- **un module noyau** qui modélise les concepts de haut niveau, permettant de subsumer les concepts des différents modules ontologiques. Nous avons utilisé, une ontologie noyau développée et utilisée au sein du laboratoire LIMICS, Onto-Ménélas,
- **un module médical** ONTOPARONMED, qui modélise les connaissances du domaine de la SLA, par les aspects médicaux comme les signes, les symptômes spécifiques de la pathologie et l'ensemble des traitements qui peuvent être mis en place comme la ventilation non invasive (VNI), la trachéotomie, etc.,
- **un module socio-environnemental** ONTOPARONCOORD, qui modélise les aspects sociaux et environnementaux. Les aspects sociaux sont liés aux spécificités du système social et médico-social français. Dans ce module les structures sociales, médico-sociales ainsi que les prestations dont peut bénéficier le patient y sont modélisés. Les aspects environnementaux, vont concerner le lieu de vie de la personne, son domicile, son entourage etc.,
- **un module de coordination** ONTOPARONCOORD, modélise et conceptualise les activités de coordination qui peuvent être mises en place par les coordinateurs du Réseau SLA IDF. Les actions de coordination sont de différentes natures, elles peuvent être des actions de communication, des actions de conseil ou d'orientation, etc.,
- **un module de consolidation**, dont l'objectif est d'importer les différents modules du domaine. Ce module ne modélise aucune connaissance. il permet, une fois l'ensemble des modules importés, d'utiliser le raisonneur de Protégé, qui permet d'inférer sous les concepts définis l'ensemble des classes qui possède une même relation.

Au total l'ontologie modulaire, est constituée de 2555 classes. Nous avons décidé de créer des concepts définis représentant une thématique spécifique qui est importante dans l'analyse des parcours de santé des patients. L'utilisation des concepts définis, permet d'avoir une approche globale et de regarder des problématiques spécifiques. 47 concepts définis ont été créés dans les différents modules ontologiques. 7 concepts définis ont spécifiquement été utilisés et étudiés lors des analyses :

- le domaine des besoins et des demandes de prise en charge médicales,

- 
- le domaine des besoins et des demandes de prise en charge sociales,
  - le domaine de l'épuisement,
  - le domaine des aides techniques,
  - le problème d'accessibilité,
  - le domaine des problèmes sociaux,
  - les actions de coordination.

## Les apports et les contraintes de la modularité

La modularité présente de nombreux avantages, comme une gestion plus aisée par une taille des modules plus petite et une thématique spécifique (médicale, sociale, etc.) par module, facilitant les échanges avec les experts. La modularité rend possible l'import de  $n$  module permettant d'enrichir l'ontologie. Lors de la phase d'évaluation par les experts du domaine, de nombreux médicaments n'ont pas été annotés. Pour répondre à ce besoin d'annotation des médicaments, il peut être envisagé d'importer, dans ONTOPARON un module ontologique modélisant l'ensemble des médicaments. Si la modularité présente de nombreux avantages, elle possède aussi des inconvénients. La modification d'une classe doit se faire sur le module ontologique et non sur l'ontologie important tous les modules, il est donc nécessaire d'être vigilant lors de la réalisation de modification. Nous avons décidé de modéliser les connaissances de chaque domaine par module, cependant cette séparation impose de faire des choix, qui peuvent être discutés. Nous avons décidé, par exemple, de positionner les aides techniques et les services de soins infirmiers à domicile (SSIAD) dans le module socio-environnemental. Ces concepts pourraient être modélisés dans le module médical car ils résultent d'un besoin qui a pour source un problème médical. Cependant dans notre travail, l'étude des parcours de santé et la possibilité de réutiliser les modules ontologiques pour d'autres pathologies nous ont renforcé dans le choix de cette modélisation.

## L'évaluation de l'annotation et de la modélisation des concepts définis

Le processus d'évaluation que nous avons mis en place, et l'outil que nous avons développé (PRONTO) nous ont permis d'obtenir des résultats sur deux niveaux. Tout d'abord sur les performances de ONBASAM, pour l'annotation sémantique des termes présents dans les corpus textuels, ensuite sur la validation de la modélisation des concepts définis. Le travail d'évaluation avec les experts, a mis en avant des perceptions différentes des experts, sur les définitions de certains concepts comme les actions de coordination ou le concept processus social. Cependant les scores montrent une validation des experts pour l'annotation et pour les concepts définis, même si des améliorations peuvent être apportées.

Une fois l'évaluation réalisée, nous avons apporté des modifications à l'ontologie en prenant en compte les commentaires et les annotations faites par les experts. Nous avons procédé ensuite à l'annotation de 931 dossiers, des patients inclus au Réseau SLA IDF entre 2013 et 2017.

(N°211) Demande d'inclusion NB: Patient parkinsonien , évoluant depuis plus de 10ans. Amorce d'un déclin intellectuel, **risque de chute** élevé et présence dyskinésies+++  
 Vient d'être vu en consultation pour une **baisse** du traitement a été faite : **nécessité** de vérifier **tolérance** dans quelques jours.

Face à la venue du patient en hôpital de semaine pour rééquilibrage du traitement antiparkinsonien j'informe l'équipe médicale des éléments principaux du parcours de soins.  
 Je rencontre l'épouse et le patient à plusieurs reprises sur le temps de l'hospitalisation pour vérifier la bonne compréhension des examens réalisés et des modifications thérapeutiques proposées.

Lors de l'un des entretiens réalisés durant l'hospitalisation, ENTOURAGE m'informe qu'elle ne m'a pas recontactée pour l'orthophoniste car souhaitait attendre effet EXELON (introduit en novembre 2014) et l'**avis NEUROLOGUE** (revu en consultation en DATE) pour confirmer cette indication. Nous évoquons également les difficultés liés au **maintien à domicile**, **les chutes** et les besoins du patients étant de plus en plus importants.

Je propose à PATIENT et ENTOURAGE de rechercher une orthophoniste à domicile. Je propose une augmentation des heures d'aides au domicile, ce qui est **refusé par le patient** et l'épouse. Je propose de rechercher un accueil de jour et en explique le fonctionnement. PATIENT et ENTOURAGE vont y réfléchir et me tiendront au courant. Je leur réexplique qu'une demande Trajectoire est en cours pour un séjour en rééducation et de répit. NB: ENTOURAGE souhaite un centre à proximité de leur domicile car ne pourra s'y déplacer qu'en transports en communs.

Face à l'accord de l'épouse pour que je les aide à trouver une orthophoniste à domicile

Recherche orthophoniste à domicile. Appel de 5 cabinets d'orthophonie et transmission à l'épouse des coordonnées de PROFESSIONNEL ADRESSE VILLE Tél : N° TÉLÉPHONE.  
 Je me mets à la disposition de PROFESSIONNEL s'il souhaite avoir davantage d'information sur le parcours de soins de PATIENT.

Appel de l'épouse qui me signale que le patient a été très perturbé lors du retour à domicile DATE: confus++ et désorienté. Le MT est passé au domicile le DATE : pas de préconisations particulières. Face à l'**aggravation de l'état confusionnel**, ENTOURAGE a arrêté le LEPONEX 1/2cp le soir (introduit durant l'hospitalisation) il y a une semaine mais difficultés persistent. A noter également un grand nombre de **chutes**, non traumatiques d'après l'épouse. ENTOURAGE semble épuisée+++ NB : ortho a bien été débuté avec PROFESSIONNEL.

- PriseEnChargeParamedicale
- Probleme
- Processus
- ProcessusMedicauxLiesHospitalisation
- ProcessusSocial
- ProcessusSortieHospitalisation
- ProfessionnelDeSante
- Professionnel
- ProfessionnelParamedical
- QualitatifObjetMetaFonctionnel
- Qualitatifs
- RecevoirAlerte
- RechercheDeProfessionnel
- RechercheDeRessourceCoordination
- RechercherOrtho
- Reeducation
- RefusDuPatient
- Refuser
- Rencontrer
- Repit
- RetourADomicile
- RisqueDeChuter
- SejourDeRepit
- Semaine
- Sentence
- SigneAppareilLocoMoteur
- SigneClinique
- SigneCliniqueNeurologique

FIGURE 8.3 – Exemple d'évènement de patients suivis dans un réseau parkinson, annotés avec ONBASAM et l'ontologie ONTOPARON.

## L'utilisation d'ONTOPARON pour d'autres pathologies

La SLA a pour particularité de présenter de nombreux points communs avec d'autres pathologies neurologiques et neuro-dégénératives comme la maladie de Parkinson dans les signes et les symptômes moteurs ou la présence de troubles de déglutition, avec la maladie d'Alzheimer pour les troubles cognitifs notamment, ou avec les personnes ayant eu un traumatisme crânien, etc.. ONTOPARON a été développé pour annoter les données du corpus du Réseau SLA IDF. Cependant nous avons fait l'hypothèse qu'ONTOPARON pouvait être utilisée, pour l'annotation de données de coordination d'autres pathologies. Pour vérifier cette hypothèse, nous avons utilisé les outils développés dans ce travail de thèse, à savoir ONTOPARON et ONBASAM pour annoter des corpus (pris de manière aléatoire) issus de la coordination de parcours de patient présentant une maladie de Parkinson et de Traumatisme crânien. Les figures ?? illustrent les annotations obtenues.

On remarque que ONTOPARON permet d'annoter ces évènements, issus de coordination de parcours pour d'autres pathologies que la SLA. Cependant l'ontologie présente des limites dans ce cadre. Si les aspects sociaux et les actions de coordination sont repérés, il existe des manques pour le domaine médical dans les évènements du patient ayant une maladie de Parkinson, certains médicaments et signes ne sont pas annotés (dyskinésies, traitement antiparkinsonien, etc.). Il en est de même pour les évènements de la personne ayant eu un traumatisme crânien, ou par exemple les termes de

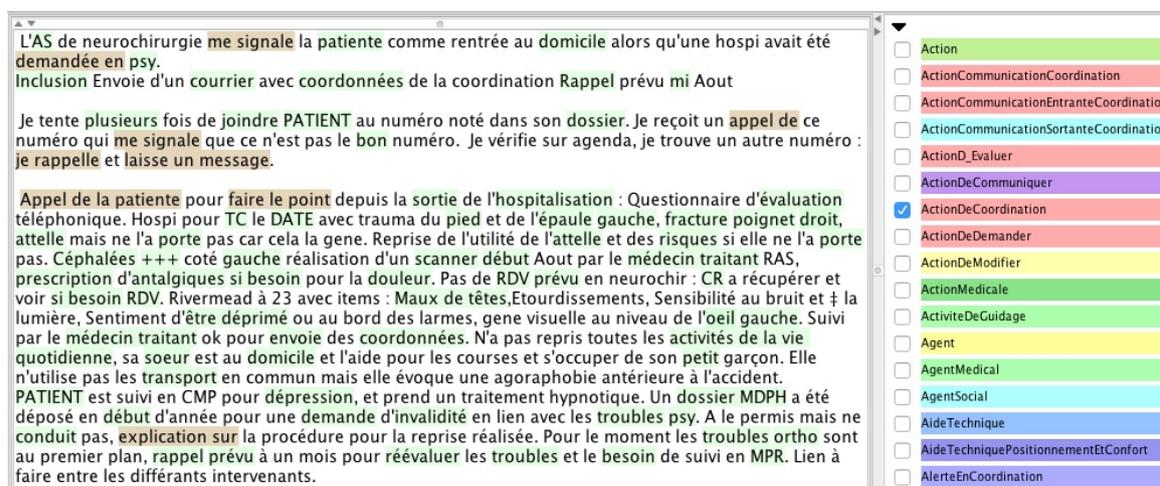


FIGURE 8.4 – Exemple d'évènement de patients suivis dans un réseau traumatisé crânien, annotés avec ONBASAM et l'ontologie ONTOPARON.

neurochirurgie ou de CMP<sup>14</sup> ne sont pas repérés. Pour traiter ces corpus deux options peuvent être envisagées, soit enrichir le module médical ONTOPARONMED, avec les signes et symptômes des autres pathologies, ou bien créer un module spécifique par pathologie. L'annotation des médicaments a présenté les mêmes limites dans le cadre des évènements de la SLA ou de la maladie de Parkinson. Les patients suivis dans le cadre des maladies neurologiques, peuvent présenter des maladies intercurrentes, ou présenter une polyopathie qui nécessite la mise en place de plusieurs traitements médicamenteux. Ceci renforce notre idée, d'importer un module spécifique des médicaments, qui serait utilisable pour toutes les pathologies. Une ontologie des médicaments est en cours de validation au sein du LIMICS Au-delà du module médical, il est tout à fait possible d'envisager, dans le cadre de la coordination de parcours d'autres pathologies (cancer, diabète, /etc), d'utiliser uniquement le module socio-environnemental et de coordination, pour identifier les situations d'épuisement, de problématiques sociales et les actions de coordination qui sont réalisées. Ceci peut permettre, par exemple, d'identifier si la coordination est dépendante ou non des pathologies rencontrées.

## Les outils ONBASAM et PRONTO

Deux outils ont été créés au cours de ce travail de thèse, grâce aux collaborations mises en place.

Le premier, est un outil d'annotation sémantique ONBASAM. Pour le développer nous avons utilisé les ressources disponibles en open source dans GATE et les avons adaptées à nos besoins. Différentes chaînes de traitement ont été développées : la première chaîne permet de corriger les erreurs d'orthographe présentes dans les corpus lors de la saisie des évènements. La seconde permet l'annotation sémantique des corpus, elle utilise l'ontologie ONTOPARON comme ressource lexicale pour réaliser l'annotation. Il a

14. Centre Médico-Psychologique

été possible, lors de la dernière phase de développement d'ONBASAM, d'intégrer l'algorithme ConText qui permet de détecter les modalités de négation et d'hypothèse, qui sont importantes à prendre en compte. L'export des annotations, de manière quantitative est rendue possible par un Processing Ressource (PR), qui permet d'obtenir sur un même fichier (excel) la somme des annotations par classe de l'ontologie et organisée par dossier patient. Ce fichier peut être, ensuite, importé dans un logiciel d'analyse statistique, permettant de lier les données d'annotations aux données socio-démographiques des patients. Nous avons réalisé, au cours de ce travail de thèse, l'annotation de 931 dossiers patients soit 31 260 évènements. Des évaluations sur la détection de la négation et de l'hypothèse doivent être menées afin d'évaluer la performance du système et de le comparer aux résultats d'autres systèmes d'annotation. Depuis peu ONBASAM est installé sur des serveurs au sein du laboratoire, il va être utilisé à court terme pour d'autres projets et pour d'autres pathologies, notamment pour un projet portant sur la psychiatrie.

Le second outil développé PRONTO, est un outil d'évaluation des annotations. C'est un système intégrant les corpus annotés par ONBASAM. Il permet d'avoir sur une même interface les annotations et les concepts auxquels sont associés les termes des corpus. Son objectif est de permettre à des experts d'évaluer les annotations. Le développement de PRONTO s'est basé sur l'ECMTE du CHU de Rouen, pour lequel nous avons apporté des modifications. Notre souhait été que PRONTO puisse permettre l'évaluation des concepts primitifs mais aussi des concepts définis. Lors de la présence d'erreurs ou d'absences d'annotations, l'expert a la possibilité de créer une annotation en liant le terme à un concept de l'ontologie ou bien de mettre un commentaire. Ces évaluations permettent d'enrichir et de modifier l'ontologie, en prenant en compte les annotations commentaires des experts. Les résultats obtenus lors de l'évaluation ont permis d'obtenir une F-mesure de 0,96 pour les concepts définis, et de 0,91 pour les concepts primitifs.

Des améliorations techniques peuvent être apportées, comme la vitesse de chargement de l'ontologie lorsque l'expert souhaite associer un terme à un concept de l'ontologie ou augmenter la zone d'affichage des validations, cependant à ce jour, ces outils sont fonctionnels. Tout comme ONBASAM, PRONTO est un outil qui va être utilisé dans d'autres projets du laboratoire LIMICS.

## **L'analyse des éléments de parcours et de la coordination**

L'annotation des corpus nous permet d'obtenir de nombreuses informations. Les concepts définis rendent possibles une observation des parcours patients à travers des thématiques spécifiques comme : la détection de situation d'épuisement des aidants, la thématique des aides humaines, les difficultés liées à des problématiques sociales et d'identifier les demandes et les besoins dans le domaine médical et social, etc.. Pour les concepts médicaux, en particulier les états (état respiratoire, état nutritionnel, etc.)

que nous avons définis, ils sont concordants avec les données de la littérature.

Les données obtenues nous permettent de valider que les patients présentent des parcours qui sont différents, par la variabilité du nombre d'évènements créés mais aussi par le nombre d'actions de coordination qui sont réalisés. Certains patients vont avoir de nombreux évènements et actions de coordination, mais d'autres ne présentent pas de besoin et de demande sans être obligatoirement en rupture de parcours.

## Les actions de coordinations

À travers les annotations il est possible de quantifier et de qualifier les actions de coordination qui sont réalisées. Si les actions de coordination que nous avons définies comme action de communication sont importantes d'un point de vue quantitatif, elles mettent en avant l'importance de diffusion de l'information auprès des patients et de leur entourage mais aussi auprès des professionnels intervenants dans les parcours. En effet, il existe une différence significative entre les actions de communication qui arrivent au Réseau SLA IDF et celles qui en sortent. Les coordinateurs agissent sur le parcours de santé et pas uniquement sur le parcours de soins, leurs interventions se situent à la fois dans le domaine médical, social et médico-social. Les actions de coordination mises en place, peuvent répondre à une demande clairement formulée, comme par exemple dans le cadre d'une demande de recherche de professionnel et la recherche qui s'ensuit. Cependant, certaines actions de mise en place de ressources ne répondent pas à ce schéma. L'hypothèse formulée est que le coordinateur détecte un besoin, réalise une évaluation et propose un dispositif adapté à mettre en place (aide technique, séjour de répit, mise en place d'aide humaine, évaluation médicale, etc.). Une étude plus approfondie sur les ressources mises en place par les coordinateurs permettrait d'identifier les domaines pour lesquels les patients rencontrent le plus de difficultés.

## Les ruptures de parcours

La notion de « rupture de parcours » est une problématique complexe, car si l'on évoque les parcours de soins ou les parcours de santé les indicateurs ne sont pas les mêmes et pas définis clairement dans la littérature. Dans le cadre des parcours de soins, les indicateurs identifiés dans la littérature pour définir ces ruptures sont l'absence, la non continuité des soins ou la présence d'hospitalisations évitables. Dans le cadre des parcours de santé, il n'existe pas de définition précise ni d'indicateur identifié pour expliciter ce qu'est une rupture. Dans les évènements saisis par les coordinateurs, il est parfois indiqué la présence de ruptures, l'absence ou l'arrêt de la prise en charge. Nous avons pu repérer par l'annotation la présence de ces difficultés pour 247 patients, les variables en lien avec la présence de ces ruptures sont l'expression d'un refus du patient ( $p < 0.001$ ) la présence de signes cliniques démontrant une altération cognitive ou la présence de troubles psychique ( $p = 0.007$ ;  $p = 0.01$ ) et la présence d'altérations de la fonction motrice ( $p = 0.003$ ). Le refus peut être exprimé sur le refus de soins, le refus de mise en place d'aides qu'elles soient humaines ou techniques.

## Synthèse

Les résultats obtenus sont liés à l'annotation qui a été réalisée par le système ONBASAM, et au mode de coordination mis en place au Réseau SLA IDF, ces résultats ne peuvent être généralisés aux parcours de tous les patients vivants avec une SLA en France.

Le travail de description et d'analyse des résultats présenté dans ce manuscrit, n'est qu'une infime partie du travail qui peut être réalisé sur l'ensemble des données obtenues. Des pistes peuvent être envisagées pour poursuivre ces analyses :

- la prise en compte de la temporalité pour placer les éléments chronologiquement, pour identifier si des périodes sont plus complexes que d'autres et si des séquences sont prédictives de difficultés,
- inclure lors de la saisie des événements un niveau de complexité pour chaque événement de suivi,
- définir les critères qui représentent une rupture de parcours de santé, pour identifier d'une part le nombre de patients concernés et d'autres si des séquences spécifiques apparaissent ou bien s'il est possible d'identifier des « déclencheurs » de situations problématiques. Pour cela des méthodes d'apprentissage peuvent être envisagées pour rechercher des motifs séquentiels, sur le modèle exposé dans la thèse de [Pinaire \(2017\)](#). Pour cela un travail de collaboration doit être mis en place entre les coordinateurs, les neurologues et des statisticiens habitués à l'utilisation de ces méthodes,
- une approche par patient, ou groupe de patients présentant les mêmes caractéristiques et non par population comme nous l'avons fait dans ce travail.

L'annotation des corpus, dépend des connaissances modélisées dans l'ontologie mais aussi des données textuelles initiales des corpus. Lors de la phase d'évaluation des annotations, les experts de coordination ont mis en avant la présence d'éléments implicites qui sont sous-jacents à la lecture des événements, mais non annotés car non identifiables par le système. L'annotation est sensible aux termes et à la construction grammaticale et syntaxique utilisée dans les corpus, ainsi la même information peut-être exprimée de  $n$  manières qui ne sera pas annotée de la même manière (Appel hier de l'épouse ; Hier appel reçu de l'épouse ; L'épouse a appelé hier, etc.), ne donnant pas les mêmes résultats. L'une des améliorations envisagées pour le système ONBASAM, est l'intégration et la prise en compte des relations sémantiques définies dans l'ontologie dans le processus d'annotation. Ceci devrait permettre d'obtenir des triplets plus informatifs que la simple annotation des classes de l'ontologie. Ceci permettrait de repérer par exemple, quels sont les intervenants demandant une hospitalisation du patient, le patient lui-même, l'entourage ou bien un intervenant, ou si des structures ou des professionnels refusent la prise en charge des patients.

# Liste des acronymes

- AAH** Allocation Adulte Handicapé. [21](#)
- ACTP** Allocation Compensatrice Tierce Personne. [35](#)
- AIVQ** Activités Instrumentales de la Vie Quotidienne. [20](#)
- ALD** Affections Longue Durée. [20](#)
- AMM** Autorisation de Mise sur le Marché. [16](#), [17](#)
- APA** Allocation Personnalisée d'Autonomie. [20](#), [22](#), [23](#), [35](#), [104](#)
- ARS** Agence Régionale de Santé. [27](#), [150](#), [215](#)
- ARSLA** Association pour la Recherche sur la Sclérose Latérale Amyotrophique et autres maladies du Motoneurone. [24](#)
- AT** Aides Techniques. [12](#), [20](#)
- AVQ** Activité de Vie Quotidienne. [17](#), [20](#)
- BioTex** BIOMedical Term EXtraction. [ix](#), [xvii](#), [58](#), [85–89](#), [91](#)
- CDAPH** Commissions Départementales des Personnes Handicapées. [21](#)
- CIF** Classification Internationale du Fonctionnement, du Handicap et de la Santé. [52](#), [53](#), [93](#)
- CIM** Classification Internationale des Maladies. [44](#)
- CLIC** Centre Local d'Information et de Coordination. [27](#)
- CMI** Carte Mobilité Inclusion. [21](#)
- CNIL** Commission Nationale de l'Informatique et des Libertés. [69](#), [70](#), [72](#), [88](#)
- CNSA** Caisse Nationale de Solidarité pour l'Autonomie. [28](#)
- DA** Directives Anticipées. [15](#), [16](#)
- DARK** Data Annotation using Rules and Knowledge. [71](#)
- DFT** Démence Fronto-Temporale. [16](#)
- DHOS** Direction de l'Hospitalisation et de l'Organisation des Soins. [24](#)

- DRESS** Direction de la recherche, des études, de l'évaluation et des statistiques. 21, 22
- ECMT** Extracteur de Concepts Multi-terminologiques. 120
- ECMTE** Extracteur de Concepts Multi-terminologiques Evaluator. 120
- EHPAD** Etablissement d'Hébergement pour Personnes Âgées Dépendantes. 27
- ENMG** Electro-neuro-myogramme. 14
- FAM** Foyer d'Accueil Médicalisé. 21
- FILSLAN** Filière de santé maladies rares Sclérose Latérale Amyotrophique et maladies du neurone moteur. 12, 24, 25
- HAD** Hospitalisation à Domicile. 27
- HAS** Haute Autorité de Santé. 1, 11–14, 25, 29
- HeTOP** Health Terminology Ontology Portal. 93, 111, 120
- JAPE** Java Annotation Patterns Engine. 80
- LPP** Liste des Produits et Prestations. 43
- MAIA** Méthodes d'Accompagnement à l'Autonomie. 20
- MDPH** Maison Départementale des Personnes Handicapées. 20, 21, 23, 27, 38, 104, 161
- MeSH** Medical Subject Headings. 43
- NCCO** Nurse Care Coordination Ontology. 93
- NGAP** Nomenclature Générale des Actes Professionnels. 43
- NMC** Neurones Moteurs Centraux. 14
- NMP** Neurones Moteurs Périphériques. 14
- OMS** Organisation Mondiale de la Santé. 44, 52
- PCH** Prestation de Compensation du Handicap. 20–23, 35, 104
- PMSI** Programme de Médicalisation du Système d'Informations. 38
- PNDS** Protocole National de Diagnostic et de Soins. 14, 25, 26
- PPC** Plan Personnalisé de Compensation. 21
- Réseau SLA IDF** Réseau SLA Île-de-France. vii, viii, 2, 3, 5, 10–12, 25, 27, 31, 32, 35, 40, 68, 77, 79, 87, 110, 111, 120, 140, 148, 164, 165, 172, 207
- RGPD** Règlement Général sur la Protection des Données. 68–70
- RQTH** Reconnaissance Qualité de Travailleur Handicapé. 21

- SAMSAH** Services d'Accompagnement Médico-Social pour Adultes Handicapés. [21](#), [35](#), [104](#)
- SLA** Sclérose latérale Amyotrophique. [ix](#), [xiv](#), [5](#), [11–21](#), [23](#), [24](#), [26](#), [31](#), [40](#), [86](#), [88](#), [89](#), [101](#), [107](#), [119](#), [126](#), [127](#), [145](#), [153](#), [166](#), [192](#)
- SNIRAM** Système National d'Information Interrégimes de l'Assurance Maladie. [34](#), [38](#)
- SOC** Système d'Organisation des Connaissances. [viii](#), [39](#), [41](#), [42](#), [56](#), [90](#)
- SSIAD** Services de Soins Infirmiers A Domicile. [27](#), [35](#), [106](#), [107](#)
- SSR** Soins de Suite et de Réadaptation. [27](#)
- TALN** Traitement Automatique de la langue Naturelle. [57](#), [58](#), [78](#), [79](#), [87](#)
- UMLS** Unified Medical Language System. [93](#)
- USLD** Unité de soins de longue durée. [27](#)
- VNI** Ventilation Non Invasive. [14](#), [24](#), [35](#), [141](#), [158](#)



## Bibliographie

- ABACHA A. B. & ZWEIGENBAUM P. (2010). Annotation et interrogation sémantiques de textes médicaux. In *Atelier Web Sémantique*, p.10.
- ABDELNOUR M. & CORDESSE V. (2008). Le réseau SLA Île-de-France, un partenariat multidisciplinaire. *Soins*, **53**(726), 53–55.
- ABDOUNI S. (2018). *Fin 2017, les départements ont attribué 2 millions de prestations d'aide sociale aux personnes âgées ou handicapées*. Rapport interne 1091, DRESS.
- AIMÉ X. (2015). Eléments de réflexion sur l'utilisation de corpus pour la construction d'ontologies. In *IC2015*, Rennes, France.
- AIMÉ X., GEORGE S. & HORNING J. (2014). VETIVOC, une ressource terminologique modulaire du domaine du textile, de la mode et de l'habillement. *Revue d'intelligence artificielle*, **28**(6), 689–728.
- AIMÉ X., GEORGES S. & HORNING J. (2015). Vetivoc : a modular ontology for the fashion, textile, and clothing domain. *Applied ontology*.
- ALCIMED, MORINI M., BONNIN V., PUPIN A.-C. & LETELLIER M. (2009). *Évaluation des dispositifs mucoviscidose et sclérose latérale amyotrophique*. Rapport interne, Rapport DHOS/ministère de la Santé, Paris.
- ANESM (2018). *Pratiques de coopération et de coordination du parcours de la personne en situation de handicap*. Recommandations de bonnes pratiques professionnelles, Agence nationale de l'évaluation et de la qualité des établissements et services sociaux et médico-sociaux, Saint-Denis.
- ATRASH A., ABEL M.-H. & MOULIN C. (2014). Ontologie Modulaire pour la Collaboration. In *25es Journées francophones d'Ingénierie des Connaissances*, p. 8 – 11, Clermont Ferrand, France.
- AUSSENAC-GILLES N. (2005). *Méthodes ascendantes pour l'ingénierie des connaissances*. PhD thesis, Université Toulouse III-Paul SABATIER.
- AUSSENAC-GILLES N. (2012). Donner du sens à des documents semi-structurés : de la construction d'ontologies à l'annotation sémantique. In L. CALDERAN, P. LAURENT, H. LOWINGER & J. MILLET, Eds., *Le document numérique à l'heure du web*, Sciences et techniques de l'information, p. 105–140. ADBS. Chapitre 05 : Donner du sens à des documents semi-structurés : de la construction d'ontologies à l'annotation sémantique.
- BACHIMONT B. (2000). Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en Ingénierie des connaissances. In EYROLLES, Ed., *Ingénierie des connaissances : évolution récentes et nouveaux défis*, Collection technique et scientifique des télécommunications, chapter 19, p. 305–323. Eyrolles, eyrolles edition.

- BACHIMONT B., ISAAC A. & TRONCY R. (2002). Semantic Commitment for Designing Ontologies : A Proposal. In G. GOOS, J. HARTMANIS, J. VAN LEEUWEN, A. GÓMEZ-PÉREZ & V. R. BENJAMINS, Eds., *Knowledge Engineering and Knowledge Management : Ontologies and the Semantic Web*, volume 2473, p. 114–121. Berlin, Heidelberg : Springer Berlin Heidelberg.
- BANEYX A. (2007). *Construire une ontologie de la Pneumologie. Aspects théoriques, modèles et expérimentations*. Phd thesis, Université Pierre et Marie Curie - Paris VI.
- BAO J. & HONAVAR V. G. (2006). Adapt OWL as a modular ontology language. In *Proceedings of the OWLED\*06 Workshop on OWL : Experiences and Directions, Athens, Georgia, USA, November 10-11, 2006*.
- BARANES M. (2012). Vers la correction automatique de textes bruités : Architecture générale et détermination de la langue d'un mot inconnu. In *RECITAL'2012 - Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues*, p. 95–108, Grenoble, France.
- BEN ABBÈS S., SCHEUERMANN A., MEILENDER T. & D'AQUIN M. (2012). Characterizing Modular Ontologies. In *International Conference on Formal Ontologies in Information Systems (FOIS)*, p. 13–25, Graz, Austria.
- BENAMOUZIG D., BERGERON H., BLOCH M.-A., BOURGUEIL Y., HÉNAUT L. & TOUZET C. (2015). *Évaluation des politiques publiques et coordination des soins*. Rapport interne, idres ; LIEPP Sciences Po, Paris.
- BENBRIKA S., DESGRANGES B., EUSTACHE F. & VIADER F. (2019). Cognitive, Emotional and Psychological Manifestations in Amyotrophic Lateral Sclerosis at Baseline and Overtime : A Review. *Front. Neurosci.*, **13**.
- BENSIMON G., LACOMBLEZ L. & MEININGER V. (1994). A Controlled Trial of Riluzole in Amyotrophic Lateral Sclerosis. *New England Journal of Medicine*, **330**(9), 585–591.
- BLANC F.-X., BANEYX A., CHARLET J. & HOUSSET B. (2010). Représentation des connaissances en pneumologie : l'ontologie doit pouvoir aider au codage. <http://www.em-premium.com/data/revues/07618425/v27i7/S0761842510003141/>.
- BLOCH M.-A. & HÉNAUT L. (2014). *Coordination et parcours. : La dynamique du monde sanitaire, social et médico-social*. Dunod.
- BOENTERT M. (2019). Sleep disturbances in patients with amyotrophic lateral sclerosis : current perspectives. *Nat Sci Sleep*, **11**, 97–111.
- BOUQUET B., DUBÉCHOT P. & JAEGER M. (2017). Introduction. *Vie sociale*, n° **18**(2), 7–11.

- BOURIGAULT D. & AUSSENAC-GILLES N. (2003). Construction d'ontologies à partir de textes. In *Actes de la 10ème conférence sur le Traitement Automatique des Langues Naturelles*, p. 27–47, Batz-sur-Mer, France : Association pour le Traitement Automatique des Langues.
- BOURIGAULT D., AUSSENAC-GILLES N. & CHARLET J. (2004). Construction de ressources terminologiques ou ontologiques à partir de textes : un cadre unificateur pour trois études de cas. In J.-M. PIERREL & M. SLODZIAN, Eds., *Techniques informatiques et structuration de terminologies*. Hermès. *Revue d'intelligence artificielle*, 18(1).
- BOUTELOUP C. (2006). Quels sont les moyens de suppléance de la fonction alimentaire et leurs indications ? *Revue Neurologique*, **162**, 309–319.
- BROCQ H. & BIOY A. (2006). Une réflexion sur la prise en charge des patients atteints de Sclérose Latérale Amyotrophique. *Médecine Palliative : Soins de Support - Accompagnement - Éthique*, **5**(6), 337–342.
- BRULETTI G., COMINI L., SCALVINI S., MORINI R., LUISA A., PANERONI M. & VITACCA M. (2015). A two-year longitudinal study on strain and needs in caregivers of advanced ALS patients. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, **16**(3-4), 187–195.
- BRUN N., FERNANDEZ-CURIEL S. & SAOUT C. (2016). Logique de parcours et coordination des soins. *Soins*, **61**(806), 50–52.
- BURKE T., ELAMIN M., GALVIN M., HARDIMAN O. & PENDER N. (2015). Caregiver burden in amyotrophic lateral sclerosis : a cross-sectional investigation of predictors. *Journal of Neurology*, **262**(6), 1526–1532.
- CABOT C. (2017). *Recherche d'information clinique dans le Dossier Patient Informatisé : modélisation, implantation et évaluation*. Phd thesis, Université de Rouen, Rouen.
- CAMPANA-SALORT E. (2006). Comment dit-on le diagnostic de Sclérose Latérale Amyotrophique (SLA) ? *Revue Neurologique*, **162**, 113–121.
- CARDOSO S., AIME X., MEININGER V., GRABLI D., COHEN K. B. & CHARLET J. (2018). De l'intérêt des ontologies modulaires. Application à la modélisation de la prise en charge de la SLA. In S. RANWEZ, Ed., *29es Journées Francophones d'Ingénierie des Connaissances, IC 2018*, 29es Journées Francophones d'Ingénierie des Connaissances, IC 2018, p. 121–128, Nancy, France : AFIA.
- CARDOSO S., MELO MORA L. F., JAULENT M.-C., AIMÉ X., GRABLI D., MEININGER V. & CHARLET J. (2017). Ontologie et TALN : l'anonymisation au service du repérage conceptuel dans le contexte de la SLA. In C. ROUSSEY, Ed., *28es Journées francophones d'Ingénierie des Connaissances IC 2017*, Actes IC 2017 28es Journées francophones d'Ingénierie des Connaissances, p. 98–103, Caen, France.

- CHARLET J. (2002). *L'ingénierie des connaissances : développements, résultats et perspectives pour la gestion des connaissances médicales*. PhD thesis, Université Pierre et Marie Curie - Paris VI. Habilitation à Diriger des Recherches.
- CHARLET J., BACHIMONT B. & JAULENT M.-C. (2006). Building medical ontologies by terminology extraction from texts : An experiment for the intensive care units. *Computers in Biology and Medicine*, **36**(7-8), 857–870.
- CHARLET J., BACHIMONT B. & TRONCY R. (2004). Ontologies pour le Web sémantique. In *Revue I3*, page 31p, p.31.
- CHIÒ A., CANOSA A., GALLO S., MOGLIA C., ILARDI A., CAMMAROSANO S., PAPURELLO D. & CALVO A. (2012). Pain in amyotrophic lateral sclerosis : a population-based controlled study. *European Journal of Neurology*, **19**(4), 551–555.
- CHIÒ A., LOGROSCINO G., TRAYNOR B., COLLINS J., SIMEONE J., GOLDSTEIN L. & WHITE L. (2013). Global Epidemiology of Amyotrophic Lateral Sclerosis : A Systematic Review of the Published Literature. *Neuroepidemiology*, **41**(2), 118–130.
- CINTAS P. (2006). Quels sont les traitements médicamenteux symptomatiques ? *Revue Neurologique*, **162**, 235–243.
- CLAVELOU P. & GUY N. (2006). [Symptomatic treatments in amyotrophic lateral sclerosis]. *Rev. Neurol. (Paris)*, **162 Spec No 2**, 4S228–224S234.
- CNSA (2012). *Promouvoir la continuité des parcours de vie*. Rapport interne, Caisse nationale de solidarité pour l'autonomie, Paris.
- CNSA (2017). *Accès à l'aide humaine : élément 1 de la prestation de compensation du handicap*. Guide d'appui aux pratiques des maisons départementales des personnes handicapées, CNSA.
- CORCIA P. & BOURDILLON F. (2008). *Evaluation du centre de référence national de la prise en charge de la sclérose latérale amyotrophique (SLA) de Paris*. Rapport interne, Société Française de santé publique.
- CORCIA P., PRADAT P.-F., SALACHAS F., BRUNETEAU G., FORESTIER N. L., SEILHEAN D., HAUW J.-J. & MEININGER V. (2008). Causes of death in a post-mortem series of ALS patients. *Amyotrophic Lateral Sclerosis*, **9**(1), 59–62.
- CORDESSE V., JAMETAL T., GUY C., LEFEBVRE S., ROUSSEL M., RUGGERI J., SCHIMMEL P., HOLSTEIN J. & MEININGER V. (2013). Analyse du parcours de santé au cours des maladies neurologiques handicapantes et évolutives. *Revue Neurologique*, **169**(6–7), 476–484.
- CORDESSE V., SIDOROK F., SCHIMMEL P., HOLSTEIN J. & MEININGER V. (2015). Coordinated care affects hospitalization and prognosis in amyotrophic lateral sclerosis : a cohort study. *BMC Health Services Research*, **15**(1), 134.

- CORDIER A. (2014). Les parcours de soins et de santé sont-ils une évolution inéluctable de notre système de santé? *adsp : Actualité et dossier en santé publique*, p.3.
- CORR B., FROST E., TRAYNOR B. & HARDIMAN O. (1998). Service provision for patients with ALS/MND : A cost-effective multidisciplinary approach. *Journal of the Neurological Sciences*, **160**, S141–S145.
- COURATIER P., CORCIA P., LAUTRETTE G., NICOL M., PREUX P.-M. & MARIN B. (2016). Epidemiology of amyotrophic lateral sclerosis : A review of literature. *Revue Neurologique*, **172**(1), 37–45.
- COURATIER P., MARIN B., LAUTRETTE G., NICOL M. & PREUX P.-M. (2014). Épidémiologie, spectre clinique de la SLA et diagnostics différentiels. *La Presse Médicale*, **43**(5), 538–548.
- CUNNINGHAM H., TABLAN V., ROBERTS A. & BONTCHEVA K. (2013). Getting More Out of Biomedical Documents with GATE’s Full Lifecycle Open Source Text Analytics. *PLoS Computational Biology*, **9**(2), e1002854.
- DANEL-BRUNAUD V., PEREZ T., JUST N. & DESTÉE A. (2008). Surveillance et traitement des troubles respiratoires associés à la sclérose latérale amyotrophique. *Revue Neurologique*, **161**, 480–485.
- DANEL-BRUNAUD V., TOUZET L., CHEVALIER L., MOREAU C., DEVOS D., VANDOOOLAEGHE S. & DEFEBVRE L. (2017). Ethical considerations and palliative care in patients with amyotrophic lateral sclerosis : A review. *Revue Neurologique*, **173**(5), 300–307.
- DES COMPTES C. (2018). *Le rapport public annuel 2018*. Le suivi des recommandations Tome II, Cour des comptes.
- DESNUELLE C. (2016). Structures et réseaux de prise en charge de la sclérose latérale amyotrophique. *La Revue du praticien*, **66**(5), 569.
- DESPRÉS S. & SZULMAN S. (2009). Choisir une ontologie noyau en vue de sa réutilisation Une stratégie d’évaluation. In L. BELLATRÈCHE, G. KASSEL & P. THIRAN, Eds., *Journées Francophone sur les Ontologies*, p. 131–139, Poitiers, France.
- DHOMBRES F., JOUANNIC J.-M., JAULENT M.-C. & CHARLET J. (2010). Choix méthodologiques pour la construction d’une ontologie de domaine en médecine prénatale. In S. DESPRES, Ed., *21èmes Journées Francophones d’Ingénierie des Connaissances*, p. 171–182, Nîmes, France : Ecole des Mines d’Alès.
- DRAME K. (2014). *Contribution à la construction d’ontologies et à la recherche d’information : application au domaine médical*. Phd thesis, Université de Bordeaux.
- DRAMÉ K., DIALLO G., DELVA F., DARTIGUES J. F., MOUILLET E., SALAMON R. & MOUGIN F. (2014). Reuse of termino-ontological resources and text corpora

- for building a multilingual domain ontology : an application to Alzheimer's disease. *J Biomed Inform*, **48**, 171–182.
- ERMINE J.-L. (2000). *Les systèmes de connaissances*. Hermes Science Publication.
- FISHMAN-BOSC A., LEVEAU E., CRELEROT-KLOPFENSTEIN S., GENTILE S. & COLSON S. (2016). Clarification de concept : la coordination appliquée au domaine des soins. *Revue francophone internationale de recherche infirmière*, **2**, 77–86.
- FONTENELLE T. (2006). Les nouveaux outils de correction linguistique de microsoft. In *Actes de la 13ème conférence sur le Traitement Automatique des Langues Naturelles*, p. 3–19, Leuven, Belgique : Association pour le Traitement Automatique des Langues.
- FORT K., EHRMANN M. & NAZARENKO A. (2009). Vers une méthodologie d'annotation des entités nommées en corpus? In *Traitement Automatique des Langues Naturelles 2009*, Senlis, France.
- FOUFI V., GAUDET-BLAVIGNAC C., CHEVRIER R. & LOVIS C. (2017). De-Identification of Medical Narrative Data. *Stud Health Technol Inform*, **244**, 23–27.
- FUNKE A., SPITTEL S., GREHL T., GROSSKREUTZ J., KETTEMANN D., PETRI S., WEYEN U., WEYDT P., DORST J., LUDOLPH A. C., BAUM P., OBERSTADT M., JORDAN B., HERMANN A., WOLF J., BOENTERT M., WALTER B., GAJEWSKI N., MAIER A., MÜNCH C. & MEYER T. (2018). Provision of assistive technology devices among people with ALS in Germany : a platform-case management approach. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, **19**(5-6), 342–350.
- FUSSELLIER M. (2006). Témoignage du vécu de la prise en charge d'un patient atteint de Sclérose Latérale Amyotrophique et de son entourage. *Revue Neurologique*, **162**, 209–214.
- GIRAUD J., THEVENET M., HADDAD R., BRUERE I., LEVEQUE S., MION M. & RIEUTORD A. (2017). Un outil d'amélioration de la transition hôpital-ville chez les personnes âgées. *Soins Gériatrie*, **22**(127), 34–39.
- GÓMEZ-PÉREZ A. (1999). Ontological engineering : A state of the art. *Expert Update : Knowledge Based Systems and Applied Artificial Intelligence*, **2**(3), 33–43. Ontology Engineering Group? OEG.
- GONZALEZ-BERMEJO J., HURBAULT A., COUPÉ C., MEININGER V. & SIMILOWSKI T. (2010). Soins palliatifs et ventilation mécanique dans la sclérose latérale amyotrophique (SLA). *Médecine Palliative : Soins de Support - Accompagnement - Éthique*, **9**(6), 309–317.
- GRAU B. C., PARSIA B., SIRIN E. & KALYANPUR A. (2006). Modularity and Web Ontologies. In *KR*.

- GROSJEAN J. (2014). *Modélisation, réalisation et évaluation d'un portail multi-terminologique multi-discipline, multi-lingue (3M) dans le cadre de la Plateforme d'Indexation Régionale (PlaIR)*. Phd thesis, Rouen.
- GROSJEAN J., SOUALMIA L. F., BOUARECH K., JONQUET C. & DARMONI S. J. (2014). Comparing BioPortal and HeTOP : towards a unique biomedical ontology portal? In *IWBBIO : International Work-Conference on Bioinformatics and Biomedical Engineering*, Granada, Spain.
- GROUIN C. (2013). *Anonymisation de documents cliniques : performances et limites des méthodes symboliques et par apprentissage statistique*. Phd thesis, Université Pierre et Marie Curie - Paris VI.
- GRUBER T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, **43**(5), 907–928.
- GUARINO N. (1997). Semantic matching : Formal ontological distinctions for information organization, extraction, and integration. In M. T. PAZIENZA, Ed., *Information Extraction A Multidisciplinary Approach to an Emerging Information Technology*, p. 139–170, Berlin, Heidelberg : Springer Berlin Heidelberg.
- GUARINO N. & GIARETTA P. (1995). Ontologies and knowledge bases : towards a terminological clarification. In *Towards Very Large Knowledge Bases*, p. 25–32. Amsterdam : IOS Press, ios press edition.
- HAMIDOU B. (2015). *Epidemiology of amyotrophic lateral sclerosis : Risk factors, incidence and phenotypes*. Phd thesis, Université de Limoges.
- HARKEMA H., DOWLING J. N., THORNBLADE T. & CHAPMAN W. W. (2009). Context : An Algorithm for Determining Negation, Experienter, and Temporal Status from Clinical Reports. *J Biomed Inform*, **42**(5), 839–851.
- HERMET P. (2018). *Parcours de soins pneumologique des patients atteints de sclérose latérale amyotrophique au CHU de Rouen*. Phd thesis, Université de Rouen.
- HJØRLAND B. (2008). What is Knowledge Organization (KO)? *KNOWLEDGE ORGANIZATION*, **35**(2-3), 86–101.
- HOBSON E. V. & MCDERMOTT C. J. (2016). Supportive and symptomatic management of amyotrophic lateral sclerosis. *Nat Rev Neurol*, **12**(9), 526–538.
- HODGE G. (2000). *Systems of knowledge organization for digital libraries : beyond traditional authority files*. Washington, DC : Digital Library Federation, Council on Library and Information Resources. OCLC : 248396976.
- HUDON M. & HADI W. M. E. (2010). Organisation des connaissances et des ressources documentaires. *Les Cahiers du numérique*, **Vol. 6**(3), 9–38.

- HUSSAMI A., KATELL B., LEMESLE-MARTIN M., BEJOT Y., GIROUD M. & JACQUIN-PIQUES A. (2018). Facteurs pronostiques de la sclérose latérale amyotrophique à la première consultation dans un centre tertiaire : expérience du centre SLA de Dijon de 2007 à 2016. *Revue Neurologique*, **174**, S52.
- HÉNAUT L. & BLOCH M.-A. (2016). Coordonnateurs de parcours, quelles dynamiques professionnelles pour les infirmiers ? *Soins*, **61**(806), 20–22.
- JAVEL P. (2011). Coordonnateur de réseau de santé : une préfiguration d'un nouveau métier ? *Soins Cadres*, **20**(78S), 21–22.
- JENSEN M., COX A. P., CHAUDHRY N., NG M., SULE D., DUNCAN W., RAY P., WEINSTOCK-GUTTMAN B., SMITH B., RUTTENBERG A., SZIGETI K. & DIEHL A. D. (2013). The neurological disease ontology. *Journal of Biomedical Semantics*, **4**, 42.
- JONQUET C., ANNANE A., BOUARECH K., EMONET V. & MELZI S. (2016). SIFR BioPortal : Un portail ouvert et générique d'ontologies et de terminologies biomédicales françaises au service de l'annotation sémantique. In *JFIM : Journées Francophones d'Informatique Médicale, e-health pour tous*, Genève, Switzerland.
- KIERNAN M. C., VUCIC S., CHEAH B. C., TURNER M. R., EISEN A., HARDIMAN O., BURRELL J. R. & ZOING M. C. (2011). Amyotrophic lateral sclerosis. *The Lancet*, **377**(9769), 942–955.
- LAVERGNE T., CAPPÉ O. & YVON F. (2010). Practical Very Large Scale CRFs. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, ACL '10*, p. 504–513, Stroudsburg, PA, USA : Association for Computational Linguistics.
- LEHÉRICEY G., FORESTIER N. L., DUPUIS L., GONZALEZ-BERMEJO J., MEININGER V. & PRADAT P.-F. (2012). Prise en charge nutritionnelle dans la sclérose latérale amyotrophique : un enjeu médical et éthique. *La Presse Médicale*, **41**(6P1), 560–574.
- LEMAIRE B. (2008). Limites de la lemmatisation pour l'extraction de significations. In *9e Journées internationales d'Analyse Statistique des Données Textuelles*, p. 725–732, Lyon, France.
- LOGROSCINO G., TRAYNOR B. J., HARDIMAN O., CHIO A., MITCHELL D., SWINGLER R. J., MILLUL A., BENN E., BEGHI E. & FOR EURALS (2010). Incidence of amyotrophic lateral sclerosis in Europe. *Journal of Neurology, Neurosurgery & Psychiatry*, **81**(4), 385–390.
- LOSSIO-VENTURA J. A., JONQUET C., ROCHE M. & TEISSEIRE M. (2014a). Bio-medical Terminology Extraction : A new combination of Statistical and Web Mining Approaches. In *JADT : Journées d'Analyse statistique des Données Textuelles*, p. 421–432, Paris, France.

- LOSSIO-VENTURA J. A., JONQUET C., ROCHE M. & TEISSEIRE M. (2014b). BIO-TEX : A system for biomedical terminology extraction, ranking, and validation. In *Proceedings of the 2014 International Conference on Posters & Demonstrations Track-Volume 1272*, p. 157–160 : CEUR-WS. org.
- LUNA J., LOGROSCINO G., COURATIER P. & MARIN B. (2017). Current issues in ALS epidemiology : Variation of ALS occurrence between populations and physical activity as a risk factor. *Revue Neurologique*, **173**(5), 244–253.
- LÓPEZ F. (1999). Overview Of Methodologies For Building Ontologies. In *Ontologies and Problem-Solving Methods (KRR5)*, p.13, Stockholm.
- LÓPEZ M. F., PÉREZ A. G. & AMAYA M. D. R. (2000). Ontology's Crossed Life Cycles. In R. DIENG & O. CORBY, Eds., *Knowledge Engineering and Knowledge Management Methods, Models, and Tools*, Lecture Notes in Computer Science, p. 65–79 : Springer Berlin Heidelberg.
- MA Y., AUDIBERT L. & NAZARENKO A. (2009). Ontologies étendues pour l'annotation sémantique. In *20es Journées Francophones d'Ingénierie des Connaissances*, p. 205–216, Hammamet, Tunisie, Tunisia.
- MANNING C. D., RAGHAVAN P. & SCHÜTZE H. (2009). *Introduction to Information Retrieval*. Cambridge University Press.
- MARIN B. (2011). *Epidemiology of Amyotrophic Lateral Sclerosis : incidence, mortality and nutritional prognosis factors of survival*. Phd thesis, Ecole Doctorale n° 523 Gay Lussac "Sciences pour l'environnement".
- MARIN B., FONTANA A., ARCUTI S., COPETTI M., BOUMÉDIENE F., COURATIER P., BEGHI E., PREUX P. M. & LOGROSCINO G. (2018). Age-specific ALS incidence : a dose–response meta-analysis. *European Journal of Epidemiology*, **33**(7), 621–634.
- MBOKO G. (2009). La coordination des acteurs, maillon indispensable du maintien à domicile. *Soins Gériatrie*, **14**(75), 18–21.
- MCDONALD K., SCHULTZ E., ALBIN L., PINEDA N., LONHART J., SUNDARAM V., SMITH-SPANGLER C., BRUSTROM J., MALCOLM E., ROHN L. & DAVIES S. (2014). *Care Coordination Atlas Version 4*. Number 14-0037- EF in Agency for Healthcare Research and Quality. Rockville,MD : U.S. Department of Health and Human Services, ahrq publication edition.
- MCDONALD K. M., SUNDARAM V., BRAVATA D. M., LEWIS R., LIN N., KRAFT S. A., MCKINNON M., PAGUNTALAN H. & OWENS D. K. (2007). *Closing the Quality Gap : A Critical Analysis of Quality Improvement Strategies (Vol. 7 : Care Coordination)*. AHRQ Technical Reviews. Rockville (MD) : Agency for Healthcare Research and Quality (US).

- MC ELHINEY M. C., RABKIN J. G., GORDON P. H., GOETZ R. & MITSUMOTO H. (2009). Prevalence of fatigue and depression in ALS patients and change over time. *Journal of Neurology, Neurosurgery & Psychiatry*, **80**(10), 1146–1149.
- MEININGER V., FONTAINE B., GRABLI D., BASDEVANT A., CLEMENT K., VIGNOT M. & CORDESSE V. (2014). Quelle place pour un nouveau métier dans la santé : la coordination des soins ? *La Presse Médicale*, **43**(10, Part 1), 1031–1033.
- MENKEN M., STUCKENSCHMIDT H., (VRIJE H. W., STOILOS G., TZOUVARAS V., HELLAS T., FENSEL C. P. D. & LEGER C. P. A. (2005). D2.1.3.1 report on modularization of ontologies coordinated by stefano spaccapietra (ecole polytechnique fédérale de lausanne) with contributions from :.
- MEYSTRE S. M., FRIEDLIN F. J., SOUTH B. R., SHEN S. & SAMORE M. H. (2010). Automatic de-identification of textual documents in the electronic health record : a review of recent research. *BMC Medical Research Methodology*, **10**(1).
- MOUTET F. (2006). Quelle est la prise en charge de la perte d'autonomie et de ses conséquences sociales ? *Revue Neurologique*, **162**, 275–277.
- MUSEN M. A. (2015). The ProtÉGÉ Project : A Look Back and a Look Forward. *AI Matters*, **1**(4), 4–12.
- OGDEN C. K. & RICHARDS I. A. (1923). *The Meaning of meaning : a study of the influence of language upon thought of the science of symbolism*. New York : Harcourt.
- PARENT C. & SPACCAPIETRA S. (2009). An Overview of Modularity. In H. STUCKENSCHMIDT, C. PARENT & S. SPACCAPIETRA, Eds., *Modular Ontologies : Concepts, Theories and Techniques for Knowledge Modularization*, Lecture Notes in Computer Science, p. 5–23. Berlin, Heidelberg : Springer Berlin Heidelberg.
- PATHAK J., JOHNSON T. M. & CHUTE C. G. (2009). Survey of modular ontology techniques and their applications in the biomedical domain. *Integr Comput Aided Eng*, **16**(3), 225–242.
- PENNEAU A., PICHETTI S. & ESPAGNACQ M. (2019). *Dépenses et restes à charge sanitaires des personnes en situation de handicap avant et après 60 ans*. Number 571 in Les rapports de l'IRDES. IDRES, idres edition.
- PINAIRE J. (2017). *Exploring trajectories of patients via medico-economic databases : application to myocardial infarction*. Theses, Université Montpellier.
- PIVETEAU D., ACEF S., DEBRABANT F.-X., JAFFRE D. & PERRIN A. (2014). <<Zéro sans solution>> : Le devoir collectif de permettre un parcours de vie sans rupture, pour les personnes en situation de handicap et pour leurs proches. Rapport interne, Ministère des Affaires sociales et de la Santé.

- POPEJOY L. L., KHALILIA M. A., POPESCU M., GALAMBOS C., LYONS V., RANTZ M., HICKS L. & STETZER F. (2015). Quantifying care coordination using natural language processing and domain-specific ontology. *J Am Med Inform Assoc*, **22**(e1), e93–e103.
- POUGET J. (2013). Les centres dédiés à la sclérose latérale amyotrophique ont-ils changé les pratiques et la prise en charge? *Revue Neurologique*.
- PRADAT P.-F., CORCIA P. & MEININGER V. (2016). Sclérose latérale amyotrophique. *EMC - Neurologie*, **13**(2), 1–15.
- PRETE L. (2016). *Management of patients with amyotrophic lateral sclerosis in French palliative care unit : review of the situation and difficulties*. Phd thesis, Université de Lorraine.
- RAMOS-GORAND M. (2016). *Le Non-recours à l'APA à domicile vu par les professionnels de terrain. Entre contraintes et expression du choix des personnes âgées*, volume 10 of *Dossiers de la DREES*. DREES.
- RIANCHO J., GIL-BEA F., SANTURTUN A. & LÓPEZ DE MUNAÍN A. (2019). Amyotrophic lateral sclerosis : a complex syndrome that needs an integrated research approach. *Neural Regeneration Research*, **14**(2), 193.
- RICHARD M. (2017). *Apport de la modélisation ontologique pour le partage des connaissances en psychiatrie*. Phd thesis, Université Pierre et Marie Curie - Paris VI.
- RICHARD M., AIMÉ X., KREBS M.-O. & CHARLET J. (2015). Enrich classifications in psychiatry with textual data : an ontology for psychiatry including social concepts. *Stud Health Technol Inform*, **210**, 221–223.
- ROCHE C. (2005). Terminologie et ontologie. *Langages*, n° **157**(1), 48–62.
- ROUSSEY C., PINET F., KANG M. A. & CORCHO O. (2011). An Introduction to Ontologies and Ontology Engineering. In *Ontologies in Urban Development Projects*, volume 1, p. 9–38. London : Springer London.
- ROUX B. L. & ROUANET H. (2010). *Multiple Correspondence Analysis*. SAGE. Google-Books-ID : GWsHakQGEHsC.
- SORIANI M.-H. & DESNUELLE C. (2009). Épidémiologie de la SLA. *Revue Neurologique*, **165**(8-9), 627–640.
- SORIANI M.-H. & DESNUELLE C. (2017). Care management in amyotrophic lateral sclerosis. *Revue Neurologique*, **173**(5), 288–299.
- TALBOT K. (2009). Motor neuron disease. *Practical Neurology*, **9**(5), 303–309.

- TALON-CHRÉTIEN M.-C., JAKOVENKO F., FRANZIN-GARREC M., RÉVY M., NERICH B., ROUCOULES B., JAKOVENKO F., GARREC M. F., RÉVY M., NERICH B. & ROUCOULES B. (2016). La coordination des intervenants au domicile du patient. *Soins*, **61**(807), 28–32.
- TISSAOUI A., AUSSENAC-GILLES N., LAUBLET P. & HERNANDEZ N. (2013). Evonto : un outil d'évolution de ressource termino-ontologique pour l'annotation sémantique. *Technique et Science Informatiques*, **32**, 817–840.
- TORNY F., LACOSTE M., NGUYEN J. M., TYMOCZKO-NGUYEN M. E. & COURATIER P. (2006). Étude des causes du retard au diagnostic de la sclérose latérale amyotrophique. *Revue Neurologique*, **162**(5), 617–622.
- USCHOLD M. & KING M. (1995). Towards a methodology for building ontologies. In *In Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95*.
- UZUNER O., LUO Y. & SZOLOVITS P. (2007). Evaluating the State-of-the-Art in Automatic De-identification. *J Am Med Inform Assoc*, **14**(5), 550–563.
- VANDENBUSSCHE P.-Y. (2011). *Définition d'un cadre formel de représentation des Systèmes d'Organisation de la Connaissance*. Phd thesis, Université Pierre et Marie Curie - Paris VI.
- VIADER F., MONDOU A., DESGRANGES B. & CARLUER L. (2014). Troubles cognitifs et SLA. Au-delà de l'atteinte motrice. *La Presse Médicale*, **43**(5), 587–594.
- WOLFARTH C., PONTON C. & TOTEREAU C. (2017). Apports du TAL à la constitution et à l'exploitation d'un corpus scolaire au travers du développement d'un outil d'annotation orthographique. *Corpus*.
- YOUNESI E., MALHOTRA A., GÜNDEL M., SCORDIS P., KODAMULLIL A. T., PAGE M., MÜLLER B., SPRINGSTUBBE S., WÜLLNER U., SCHELLER D. & HOFMANN-APITIUS M. (2015). PDON : Parkinson's disease ontology for representation and modeling of the Parkinson's disease knowledge domain. *Theoretical Biology and Medical Modelling*, **12**(1).
- ZACKLAD M. (2010). Évaluation des systèmes d'organisation des connaissances. *Les Cahiers du numérique*, **Vol. 6**(3), 133–166.
- ZACKLAD M. & GIBOIN A. (2010). Introduction. *Document numérique*, **Vol. 13**(2), 7–12.
- ZWEIGENBAUM P. (1999). Encoder l'information médicale : des terminologies aux systèmes de représentation des connaissances. *Innovation Stratégique en Information de Santé*, p. 27–47.

# Quatrième partie

## Annexes



## Publications

CARDOSO, S., MELO MORA, L. F., GRABLI, D., MEININGER, V., & CHARLET, J. (2016). Les ontologies pour aider à comprendre les parcours de santé dans le cadre des maladies neurodégénératives. Montpellier, France. Retrieved 2019-08- 14, from <https://ic2016.sciencesconf.org/conference/ic2016/pages/IASante.pdf>.

CARDOSO, S., MELO MORA, L. F., JAULENT, M.-C., AIMÉ, X., GRABLI, D., MEININGER, V., & CHARLET, J. (2017, jul). Ontologie et TALN : l'anonymisation au service du repérage conceptuel dans le contexte de la SLA. In C. Roussey (Ed.), *28es Journées francophones d'Ingénierie des Connaissances IC 2017*, Actes IC 2017 (pp. 98–103). Caen, France.

CARDOSO, S., AIMÉ, X., MEININGER, V., GRABLI, D., COHEN, K. B., & CHARLET, J. (2018, jul). De l'intérêt des ontologies modulaires. Application à la modélisation de la prise en charge de la SLA. In S. Ranwez (Ed.), *29es Journées Francophones d'Ingénierie des Connaissances, IC 2018*, Actes IC 2018 (pp. 121– 128). Nancy, France : AFIA.

CARDOSO, S., AIMÉ, X., MEININGER, V., GRABLI, D., MELO MORA LUIS, F., BRETONNEL, C. K., & CHARLET, J. (2018). A Modular Ontology for Modeling Service Provision in a Communication Network for Coordination of Care. *Studies in Health Technology and Informatics*, 890–894.

CARDOSO, S., AIMÉ, X., MEININGER, V., GRABLI, D., MENETON, P., & CHARLET, J. (2019, jul). Using Equivalent Classes of an Ontology to Understand Care Pathway in Amyotrophic Lateral Sclerosis. *Stud Health Technol Inform*, 262, 93–96.



# Annexe **B**

## Information globales

### **B.1 Les aides techniques dans la SLA**

Représentation graphique des aides techniques qui peuvent être proposées aux patients durant leurs parcours.

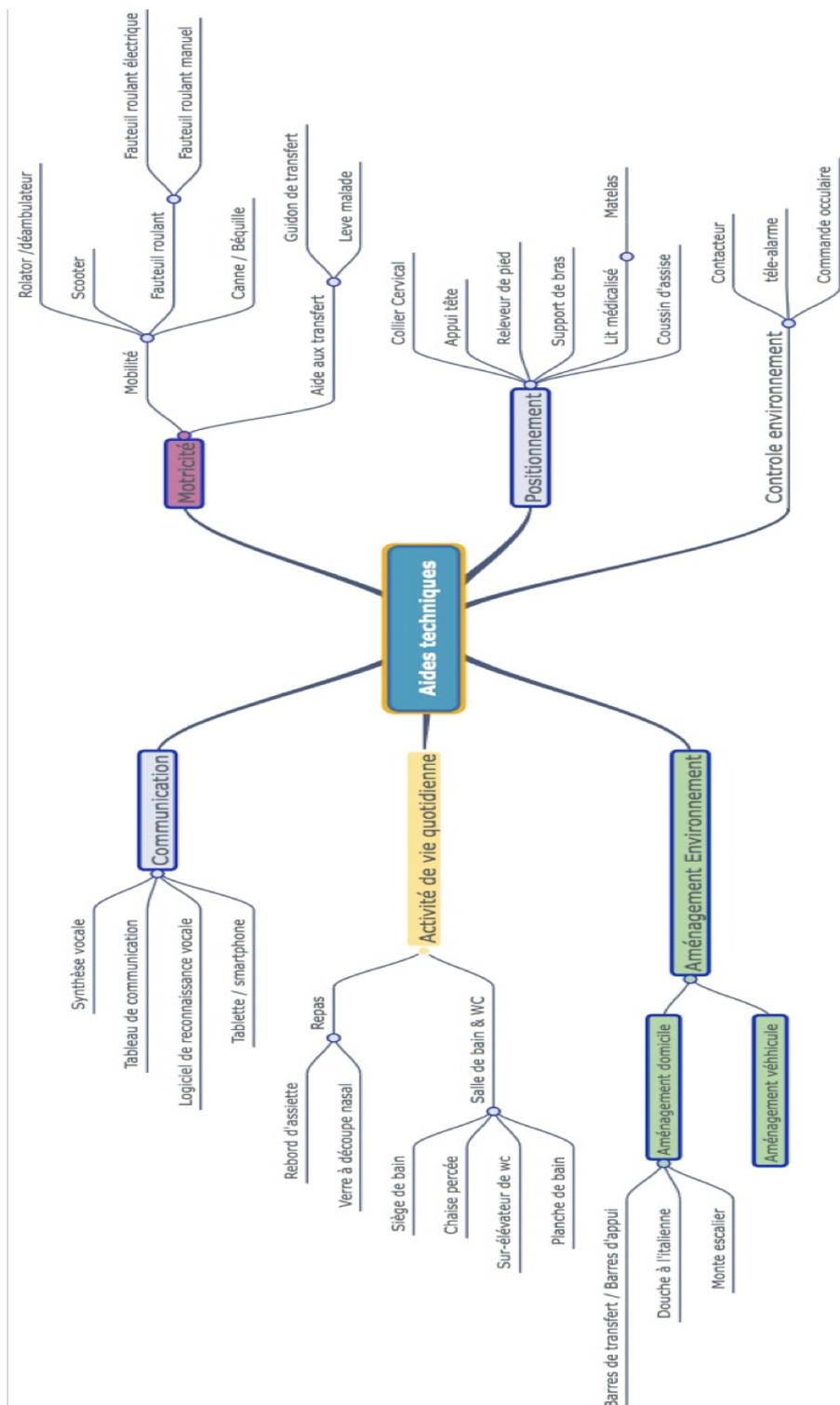


FIGURE B.1 – Exemples d’aides techniques pouvant être proposées dans la SLA.

## **B.2 Captures d'écrans des volets de recueil d'informations de la base du Réseau SLA IDF**

Portail de la coordination des centres et des réseaux de prise en charge de la Sclérose Latérale Amyotrophique

Portail SLA Pro - Accès Membres

User Dernière connexion le 10/01/2019 Déconnexion

Nouveau patient Patients actifs Archives Alertes Bilans & Exports

Patient : 473 Sélectionner

Fiche Médical Social Ergo. Psy. Intervenants Situ. Événements

**Références du dossier**

Numéro d'inclusion 158/17

Date d'inclusion [redacted] /mm/aaaa

Neurologue référent [redacted]

**État civil du patient**

Genre  M.  Mme  Mlle

Nom [redacted]

Prénom [redacted]

Date de naissance [redacted] /mm/aaaa

Régime d'assurance sociale Général

Caisse d'affiliation [redacted]

**Adresse principale**

Adresse [redacted]

Adresse (suite) [redacted]

Ville [redacted]

Code postal [redacted]

**Adresse secondaire**

Adresse [redacted]

Adresse (suite) [redacted]

Ville [redacted]

Code postal [redacted]

**Contact**

Téléphone [redacted]

Téléphone secondaire [redacted]

Téléphone portable [redacted]

Fax [redacted]

E-mail [redacted]

**Personne de confiance**

Nature [redacted]

Genre  M.  Mme  Mlle

Nom [redacted]

Prénom [redacted]

Adresse [redacted]

Adresse (suite) [redacted]

Ville [redacted]

Code postal [redacted]

Téléphone [redacted]

Téléphone portable [redacted]

E-mail [redacted]

**Personne à prévenir**

Nature [redacted]

Genre  M.  Mme  Mlle

Nom [redacted]

Prénom [redacted]

Adresse [redacted]

Adresse (suite) [redacted]

Ville [redacted]

Code postal [redacted]

Téléphone [redacted]

FIGURE B.2 – Capture d'écran de la fiche d'information patient sur la base du Réseau SLA IDF.

Portail de la coordination des centres et des réseaux de prise en charge de la Sclérose Latérale Amyotrophique

SLA PRO

Portail SLA Pro - Accès Membres

User Dernière connexion le 10/01/2019 Déconnexion

Nouveau patient Patients actifs Archives Alertes Bilans & Exports

Patient : 473 Sélectionner

Fiche Médical Social Ergo. Psy. Intervenants Situ. Évènements

Début de la maladie [ ] j/mm/aaaa

Date de diagnostic [ ] j/mm/aaaa

Mode début membre supérieur

État actuel atteinte mortice + respi + bulb  
10h vni/fr

**Situation actuelle**

Situation Inclu dans ce réseau depuis le [ ]

**VNI**

État oui

Date [ ] j/mm/aaaa

Discussion non

**Gastrostomie**

État non

Discussion non

**Chambre implantable**

Etat non

Discussion non

**Trachéotomie**

État non

Discussion oui

Date [ ] j/mm/aaaa

Status refusé

Enregistrer Annuler

© Copyright 2019 Réseau SLA, all rights reserved.

FIGURE B.3 – Capture d'écran de l'interface de la base du Réseau SLA IDF pour le volet médical.

Portail de la coordination des centres et des réseaux de prise en charge de la Sclérose Latérale Amyotrophique

**SLA PRO** Portail SLA Pro - Accès Membres

Dernière connexion le 10/01/2019 Déconnexion

Nouveau patient Patients actifs Archives Alertes Bilans & Exports

Patient : 473 Sélectionner

Fiche Médical Social Ergo. Psy. Intervenants Situ. Évènements

**Situation familiale**

Situation maritale

Situation familiale

100% sécurité sociale  non  en cours  oui

Remarques  
A 3 enfants  
Vit en EHPAD

Carte d'invalidité  non  en cours  oui

Carte européenne de stationnement  non  en cours  oui

**Prestations de maintien à domicile**

ACTP  non  en cours  oui

Prestation de compensation  non  en cours  oui

Majoration pour tierce personne  non  en cours  oui

APA  non  en cours  oui

Aide complémentaire (aide DDASS)  non  en cours  oui

**Signalement**

Date

A qui?

Remarques

Organisation actuelle au domicile

Détail des aides et coordonnées

Enregistrer Annuler

© Copyright 2019 Réseau SLA, all rights reserved.

FIGURE B.4 – Capture d'écran de l'interface de la base du Réseau SLA IDF pour le volet social.

Portail de la coordination des centres et des réseaux de prise en charge de la Sclérose Latérale Amyotrophique

SLA PRO

Portail SLA Pro - Accès Membres

User Dernière connexion le 10/01/2019 Déconnexion

Nouveau patient Patients actifs Archives Alertes Bilans & Exports

Patient : 473 Sélectionner

Fiche Médical Social Ergo. Psy. **Intervenants** Situ. Évènements

Ajouter un intervenant de proximité

**Kinésithérapeute :**

Adresse: [redacted]  
 Adresse (suite): [redacted]  
 Code postal: [redacted]  
 Téléphone: [redacted]  
 Fax: [redacted]

Ville: PARIS  
 Téléphone portable: [redacted]  
 E-mail: [redacted]

Modifier Supprimer

**Médecin traitant :**

Adresse: [redacted]  
 Adresse (suite): [redacted]  
 Code postal: [redacted]  
 Téléphone: [redacted]  
 Fax: [redacted]

Ville: PARIS  
 Téléphone portable: [redacted]  
 E-mail: [redacted]

Modifier Supprimer

**Ortophoniste :**

Adresse: 2 x par semaine à domicile  
 Adresse (suite): [redacted]  
 Code postal: [redacted]  
 Téléphone: [redacted]  
 Fax: [redacted]

Ville: PARIS  
 Téléphone portable: [redacted]  
 E-mail: [redacted]

Modifier Supprimer

**Prestataire de Services :**

Adresse: Installation Cough assist+  
 Adresse (suite): aspirateur mucosités  
 Code postal: [redacted]  
 Téléphone: [redacted]  
 Fax: [redacted]

Ville: [redacted]  
 Téléphone portable: [redacted]  
 E-mail: [redacted]

Modifier Supprimer

**Prestataire de Services Nutrition**

Adresse: IDE Coordinatrice  
 Adresse (suite): [redacted]  
 Code postal: [redacted]  
 Téléphone: [redacted]  
 Fax: [redacted]

Ville: [redacted]  
 Téléphone portable: [redacted]  
 E-mail: [redacted]

Modifier Supprimer

**Laboratoire d'analyses :**

Adresse: [redacted]  
 Adresse (suite): [redacted]  
 Code postal: [redacted]  
 Téléphone: [redacted]  
 Fax: [redacted]

Ville: [redacted]  
 Téléphone portable: [redacted]  
 E-mail: [redacted]

Modifier Supprimer

**Infirmière : IDE Libéral**

Adresse: [redacted]  
 Adresse (suite): [redacted]  
 Code postal: [redacted]  
 Téléphone: [redacted]  
 Fax: [redacted]

Ville: [redacted]  
 Téléphone portable: [redacted]  
 E-mail: [redacted]

Modifier Supprimer

© Copyright 2019 Réseau SLA, all rights reserved.

FIGURE B.5 – Capture d'écran de la fiche intervenant, qui permet de recueillir l'ensemble des informations sur les professionnels intervenant auprès du patient.



Annexe **C**

Requêtes SPARQL utilisées dans CORESE

```
=====
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
=====
```

```
# Liste les concepts qui ont plusieurs prefLabels en français
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?prefLabel ?concept (COUNT(?prefLabel) AS ?nbPrefLabel)
  WHERE{
    ?concept a owl:Class.
    ?concept skos:prefLabel ?prefLabel.
    FILTER(LANG(?prefLabel)="fr").
  }
GROUP BY ?concept
HAVING(?nbPrefLabel>=2)
```

```
# Liste les concepts qui n'ont pas de prefLabel
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?concept
WHERE{ ?concept a owl:Class.
  FILTER NOT EXISTS {?concept skos:prefLabel ?prefLabel}
}
```

```
# Liste les concepts qui ont des preflabel mais pas en français
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT DISTINCT ?concept
WHERE{
  ?concept a owl:Class.
  ?concept skos:prefLabel ?prefLabel.
  FILTER NOT EXISTS {
    ?concept skos:prefLabel ?prefLabel2.
    FILTER(LANG(?prefLabel2)="fr")
  }
}
```

```
# Liste les prefLabels en anglais
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?concept ?prefLabel
  WHERE{
    ?concept a owl:Class.
    ?concept skos:prefLabel ?prefLabel.
    FILTER(LANG(?prefLabel)="en").
  }
GROUP BY ?concept
```

```
# Compte les prefLabels en anglais
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT (COUNT(?prefLabel) AS ?nbPrefLabel)
WHERE{
  ?concept a owl:Class.
  ?concept skos:prefLabel ?prefLabel.
  FILTER(LANG(?prefLabel)="en").
}
```

```
# Liste les concepts qui ont des preflabel mais pas en anglais
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT DISTINCT ?concept
WHERE{
  ?concept a owl:Class.
  ?concept skos:prefLabel ?prefLabel.
  FILTER NOT EXISTS {
    ?concept skos:prefLabel ?prefLabel2.
    FILTER(LANG(?prefLabel2)="en")
  }
}
```

# Annexe **D**

## Classes, relations, concepts définis et notions thématiques

### D.1 Les classes et les relations formelles

Nous avons dans la partie 2.2.2, les principaux concepts permettant de représenter les éléments de connaissance dans une ontologie. En termes formels, quand on manipule les connaissances dans un éditeur d'ontologie comme Protégé, l'identité des concepts manipulés change un peu par rapport au travail de modélisation :

- les concepts sont des classes (*class*);
- les relations sémantiques sont des relations binaires (*object property*);
- les propriétés attachées aux classes sont des axiomes (*axiom*) formels qui traduisent des restrictions (*restriction*) attachées à la classe par des relations. Ainsi, la classe `EpuisementDeLaPriseEnCharge` est dans la thématique de l'épuisement et cela est représenté par la restriction « `aPourThematique some ThematiqueEpuisement` » construit avec la relation sémantique `a_PourThematique` et la classe « `ThematiqueEpuisement` ».

### D.2 OWL-DL et les concepts définis

Nous reprenons dans cette section les éléments de langages des ontologies formelles, à l'origine exprimés dans [http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4\\_v1\\_3.pdf](http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf) du site<sup>1</sup> et repris en français dans [https://www.researchgate.net/publication/318208041\\_Creation\\_d'une\\_ontologie\\_en\\_OWL\\_avec\\_Protege\\_45](https://www.researchgate.net/publication/318208041_Creation_d'une_ontologie_en_OWL_avec_Protege_45). Comme rappelé en partie 2.2.2.4, OWL permet d'exprimer des ontologies et OWL-DL est le sous-langage idéal pour décrire et gérer nos ontologies. L'une des principales caractéristiques des ontologies décrites à l'aide de OWL-DL est qu'elles peuvent être traitées par un raisonneur. L'un des principaux services offerts par

---

1. <http://owl.cs.manchester.ac.uk/publications/talks-and-tutorials/protg-owl-tutorial/>

un raisonneur consiste à vérifier si une classe est une sous-classe d'une autre classe<sup>2</sup>. En effectuant de tels tests sur toutes les classes d'une ontologie, il est possible pour un raisonneur de calculer la hiérarchie de classes d'ontologies inférées. Un autre service standard offert par les raisonneurs est la vérification de la cohérence. Sur la base de la description (conditions) d'une classe, le raisonneur peut vérifier s'il est possible que la classe ait des instances. Une classe est réputée incohérente si elle ne peut avoir aucune instance.

### D.2.1 Conditions nécessaires et suffisantes (classes primitives et définies)

Les classes que nous créons dans une première étape, utilisent uniquement les conditions nécessaires pour les décrire. Les conditions nécessaires peuvent être lues comme suit : « Si un élément appartient à cette classe, il est nécessaire de remplir ces conditions ». Dans notre domaine, on a ainsi défini par condition nécessaire la classe `Epuisement-DeLaPriseEnCharge` (Cf. supra). C'est une classe primitive. Avec les seules conditions nécessaires, nous ne pouvons pas dire : « Si quelque chose remplit ces conditions, alors il doit être membre de cette classe ». Dans notre domaine, en poursuivant notre exemple, le `DomaineEpuisement` est un `ObjetAbstrait` totalement défini par la condition nécessaire et suffisante « `ObjetAbstrait and (aPourThematique some ThematiqueEpuisement)` » et on peut dire que tout objet qui satisfait la condition énoncée est une sous-classe de `DomaineEpuisement`. Une classe qui ne possède que les conditions nécessaires est appelée classe primitive. Une classe qui a au moins un ensemble de conditions nécessaires et suffisantes est appelée classe définie.

### D.2.2 Classes et raisonneurs

Ainsi, le raisonneur ne peut classer automatiquement les classes que sous des classes définies - c'est-à-dire des classes avec au moins un ensemble de conditions nécessaires et suffisantes. A l'inverse, une classe primitive, définie par les seules conditions nécessaires laisse ouverte d'autres conditions possibles et ne permet que d'autres classes se positionnent dessous automatiquement avec le raisonneur. Pouvoir utiliser un raisonneur pour calculer automatiquement la hiérarchie des classes est l'un des principaux avantages de la construction d'une ontologie à l'aide du sous-langage OWL-DL. En effet, lors de la construction de très grandes ontologies (contenant plusieurs milliers de classes), l'utilisation d'un raisonneur pour calculer les relations sous-classe-superclasse entre classes devient presque vitale. Sans un raisonneur, il est très difficile de conserver de grandes ontologies dans un état maintenable et logiquement correct. Dans les cas où les ontologies peuvent avoir des classes comportant plusieurs super-classes (héritage multiple), il est presque toujours préférable de construire la hiérarchie des classes sous la forme d'un simple arbre. Les classes de la hiérarchie affirmée (hiérarchie construite

---

2. Connu sous le nom de test de subsumption - les descriptions des classes (conditions) sont utilisées pour déterminer s'il existe une relation super-classe/sous-classe.

manuellement) n'ont à ce moment pas plus d'une superclasse. Le calcul et la maintenance de l'héritage multiple sont le travail du raisonneur. Cette technique permet de maintenir l'ontologie dans un état modifiable et maintenable. Cela favorise non seulement la réutilisation de l'ontologie par d'autres ontologies et applications, mais minimise également les erreurs humaines inhérentes au maintien d'une hiérarchie à héritages multiples. Notons que cette contrainte n'est pas toujours facile à respecter.

## D.3 Les concepts définis de ONTOPARON

Les concepts définis ou *equivalent class* sont définis par des conditions suffisantes et nécessaires. La figure D.1 illustre les notions thématiques que nous avons créé dans ONTOPARONNOY. Ces thématiques font référence à des domaines d'intérêt cliniques pour comprendre la répartition des fréquences d'annotations dans les dossiers patients.

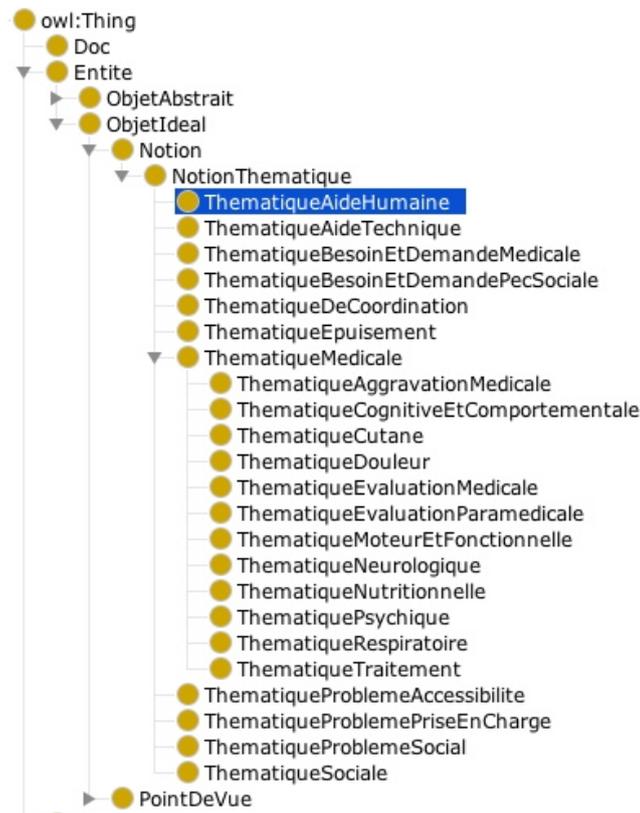


FIGURE D.1 – Notion thématique présente dans ONTOPARONNOY.

### D.3.1 Exemple du concept définis : Domaine de l'épuisement

Dans cet exemple le concept définis, *DomaineEpuisement* est un objet abstrait qui a `pour_thematique some ThematiqueEpuisement` comme indiqué à la figure D.2.

L'utilisation du raisonneur permet d'inférer l'ensemble des concepts qui sont liés par la même relation comme illustrer à la figure D.3



FIGURE D.2 – Définition du concept défini : Domaine de l'épuisement.

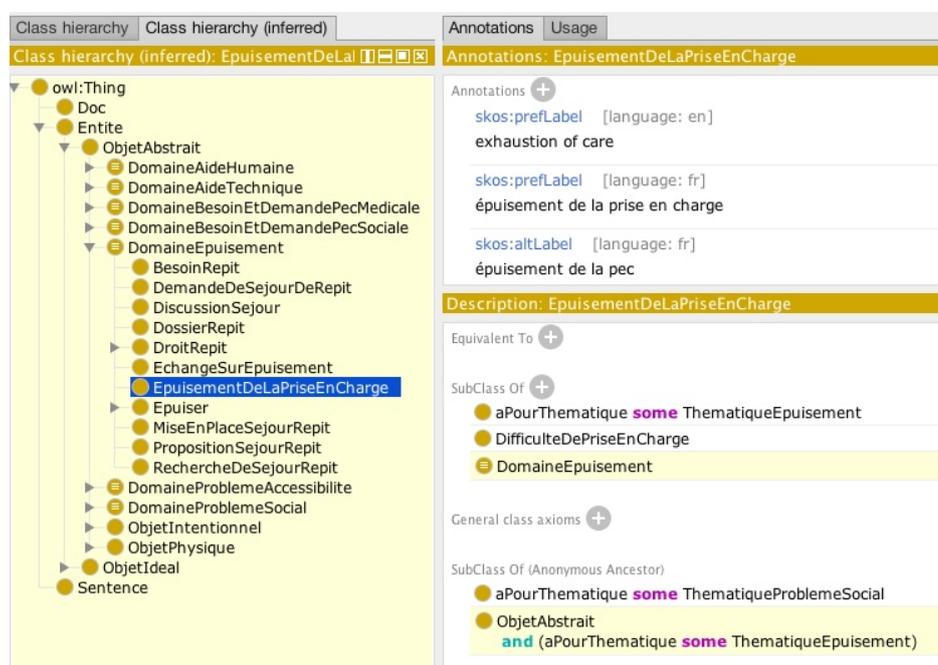


FIGURE D.3 – Ensemble des classes inférées après utilisation du raisonneur HerMiT, sous le concept définis du Domaine de l'épuisement.

### D.3.1.1 Concept définis présents dans le module médical

Plusieurs concepts définis sont présents dans ce module, d'une part les « États  $n$  » décrits dans la partie 5.4.3, mais aussi les actions et les évaluation comme indiqués ci-dessous.

**Actions médicales :** sont les actions qui sont réalisées par un professionnel médical ou paramédical. On y trouve les prises en charge paramédicale en kinésithérapie, orthophonie, etc. L'avis médical émis par un médecin spécialiste, l'ensemble des examens (d'imagerie ou de biologie)etc..

1. Évaluation médicale : regroupe l'ensemble des consultations, et les examens cliniques, que sont les bilans (bilans moteur, bilan douleur).
2. Évaluation paramédicale : regroupe les bilans réalisés fait par des professionnels paramédicaux comme un bilan ergothérapeutique, un bilan kinésithérapeutique ou bien encore un bilan nutritionnel, etc..



# Annexe **E**

## Exemple d'annotation des corpus du Réseau SLA IDF fait par d'autres systèmes d'annotation.

Exemples d'annotations réalisés sur des corpus du [Réseau SLA IDF](#), produit par deux systèmes existants que sont : le SIFR Annotator<sup>1</sup> de l'équipe du LIRMM<sup>2</sup> et l'ECMT<sup>3</sup> du CHU de Rouen.

### **E.1 Annotation réalisée par le SIFR Annotator du LIRMM**

- 
1. <http://bioportal.lirmm.fr/annotator>
  2. Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier
  3. <http://cispro.chu-rouen.fr/CISMeFecmtservice/REST>

## Annexe E : Exemple d'annotation des corpus du Réseau SLA IDF fait par d'autres systèmes d'annotation.

FIGURE E.1 – Capture d'écran de l'interface du SIFR.

Annotations total results 10 (direct 10 / ancestor 0 / mapping 0)

CLASS <small>filter</small>	ONTOLOGY <small>filter</small>	TYPE <small>filter</small>	CONTEXT	MATCHED CLASS <small>filter</small>	MATCHED ONTOLOGY <small>filter</small>	NEGATION	EXPERIENCER	TEMPORALITY
situation	OntoPneumo	direct	... sur la <b>situation</b> . - Au niveau ...	situation	OntoPneumo	AFFIRMED	PATIENT	RECENT
situation	OntoPneumo	direct	Appel de l'ergo de la MDPH pour me faire un retour sur la situation. - Au niveau du logement, vit dans un appartement ( bailleur privé ) avec accès par ascenseur exigü. Salle de bain avec baignoire et WC séparé. Conseille un logement. Doivent faire une demande d'autorisation au bailleur pour la réalisation de travaux. ( le mari avait acheté du matériel pour faire les travaux lui même, mais le projet n'était pas adapté et il devait ramener le matériel ) - Au niveau des aides humaines, a fait une PCH U. Intervention pourra commencer dès le 01. 04. Souhaite des intervenants extérieurs. --- - ... sur la <b>situation</b> à COORDINATEUR SLA.	situation	OntoPneumo	AFFIRMED	PATIENT	RECENT
Logement	Medical Subject Headings, version française	direct	Appel de l'ergo de la MDPH pour me faire un retour sur la situation. ... niveau du <b>logement</b> , vit dans un ...	Logement	Medical Subject Headings, version française	AFFIRMED	PATIENT	RECENT
baignoire	Systematized Nomenclature of MEDicine, version française	direct	Appel de l'ergo de la MDPH pour me faire un retour sur la situation. ... bain avec <b>baignoire</b> et WC séparé. ...	baignoire	Systematized Nomenclature of MEDicine, version française	AFFIRMED	PATIENT	RECENT
séparé	Systematized Nomenclature of MEDicine, version française	direct	Appel de l'ergo de la MDPH pour me faire un retour sur la situation. ... et WC <b>séparé</b> . Conseille un logement. ...	séparé	Systematized Nomenclature of MEDicine, version française	AFFIRMED	PATIENT	RECENT
séparation	Systematized Nomenclature of MEDicine, version française	direct	Appel de l'ergo de la MDPH pour me faire un retour sur la situation. ... et WC <b>séparé</b> . Conseille un logement. ...	séparation	Systematized Nomenclature of MEDicine, version française	AFFIRMED	PATIENT	RECENT
Divorce	Medical Subject Headings, version française	direct	Appel de l'ergo de la MDPH pour me faire un retour sur la situation. ... et WC <b>séparé</b> . Conseille un logement. ...	Divorce	Medical Subject Headings, version française	AFFIRMED	PATIENT	RECENT
Accomplissement	Medical Subject Headings, version française	direct	Appel de l'ergo de la MDPH pour me faire un retour sur la situation. ... pour la <b>réalisation</b> de travaux. (...)	Accomplissement	Medical Subject Headings, version française	AFFIRMED	PATIENT	RECENT
adaptation	Systematized Nomenclature of MEDicine, version française	direct	Appel de l'ergo de la MDPH pour me faire un retour sur la situation. ... n'était pas <b>adapté</b> et il devait ...	adaptation	Systematized Nomenclature of MEDicine, version française	NEGATED	PATIENT	RECENT

FIGURE E.2 – Résultats des annotations réalisées par le SIFR.

## E.2 Annotations réalisées par l'ECMTv3

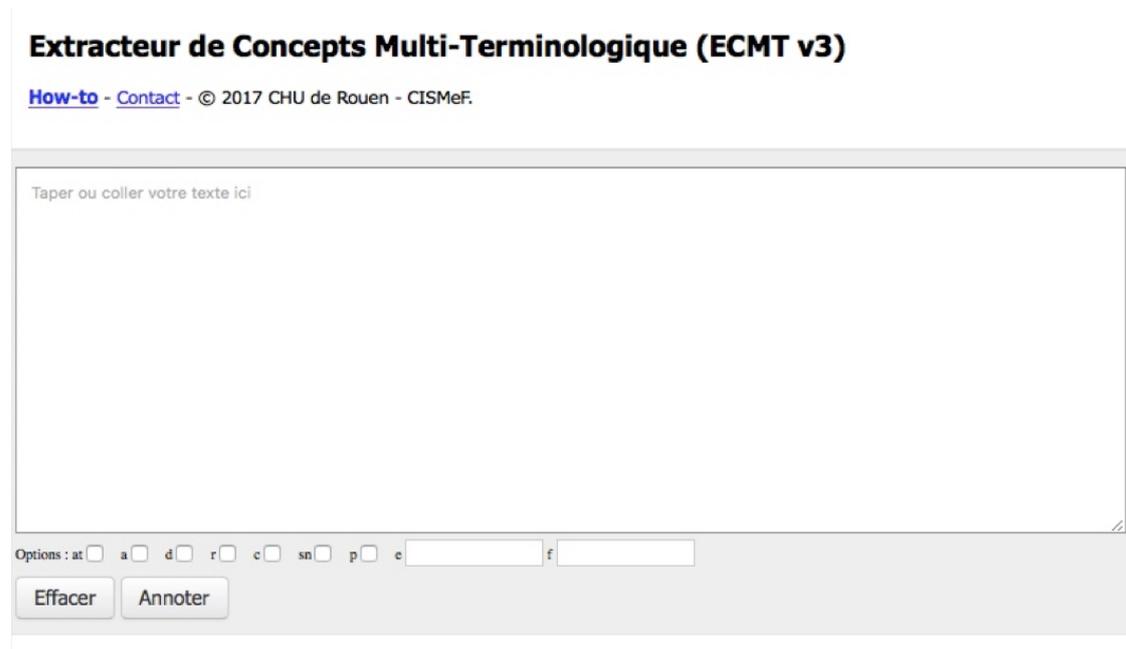


FIGURE E.3 – Capture d'écran de l'interface de l'ECMTv3.

# Extracteur de Concepts Multi-Terminologique (ECMT v3)



**How-to - Contact** - © 2017 CHU de Rouen - CISMeF.

appel reçu de MR ENTOURAGE pour m'informer comme nous en avions convenu a cette date s'il n'avait pas de nouvelles pour le séjour de répit --- - Appel fait HÔPITAL: la demande médicale est acceptée, mais la place n'est pas encore possible a définir ( séjour de répit fin DATE, il est trop tôt, il faudra appeler début DATE ) messages laisse a HÔPITAL et a HÔPITAL pour me tenir informée attente d e leur rappel  
Effacer

1 phrases annotées en 7 ms. 48 codes distincts identifiés.

## Codes identifiés

([SCT\\_CO\\_420942008](#))

- **catégorie(s)** : produit intellectuel (UML\_ST\_T170);

**Acceptation** ([ICN\\_F\\_10000329](#))

**Acceptation des soins par le patient** ([MSH\\_D\\_010342](#))

- **catégorie(s)** : activité de soins médicaux (UML\_ST\_T058);

**Attente** ([ICN\\_F\\_10023679](#))

- **catégorie(s)** : concept ou idée (UML\_ST\_T078);

**Attente** ([NCI\\_CO\\_C25359](#))

- **catégorie(s)** : concept ou idée (UML\_ST\_T078);

**Attente** ([SCT\\_CO\\_385635001](#))

- **catégorie(s)** : concept ou idée (UML\_ST\_T078);

**Fait** ([NCI\\_CO\\_C61036](#))

- **catégorie(s)** : Fait (UML\_ST\_T033);

**Hôpitaux** ([MSH\\_M\\_0010590](#))

- **catégorie(s)** : organisation lié aux services médicaux (UML\_ST\_T093); objet manufacturé (UML\_ST\_T073);

**Information** ([NCI\\_CO\\_C48191](#))

- **catégorie(s)** : concept ou idée (UML\_ST\_T078);

**Informé (procédure)** ([SCT\\_CO\\_223458004](#))

- **catégorie(s)** : activité de soins médicaux (UML\_ST\_T058);

**Possible** ([NCI\\_CO\\_C53258](#)) pas

nouveau ([NCI\\_CO\\_C25586](#)) pas

- **catégorie(s)** : concept qualitatif (UML\_ST\_T080);

nouveau ([SCT\\_CO\\_7147002](#)) pas

- **catégorie(s)** : concept qualitatif (UML\_ST\_T080);

protocole APEL ([MSH\\_SC\\_045350](#))

- **catégorie(s)** : procédure thérapeutique ou préventive (UML\_ST\_T061);

rappel ([NCI\\_CO\\_C37922](#))

- **catégorie(s)** : concept fonctionnel (UML\_ST\_T169);

rappel ([NCI\\_CO\\_C53609](#))

- **catégorie(s)** : activité gouvernementale ou réglementaire (UML\_ST\_T064);

rappel ([NCI\\_CO\\_C54270](#))

- **catégorie(s)** : entité conceptuelle (UML\_ST\_T077);

rappel mnésique ([MSH\\_D\\_011939](#))

- **catégorie(s)** : enseignement et éducation (CIS\_MT\_53); processus mental (UML\_ST\_T041);

tenir ([NCI\\_CO\\_C62290](#))

- **catégorie(s)** : activité (UML\_ST\_T052);

tennis ([MSH\\_D\\_013715](#))

- **catégorie(s)** : médecine du sport (CIS\_MT\_18); activité quotidienne ou de loisirs (UML\_ST\_T056);

Terme	Ter.	Code	CUI	Cond.	Ctxt.
	SCT	420942008			
Acceptation	ICN	10000329			
Acceptation des soins par le patient	MSH	D010342			
Attente	ICN	10023679			
Attente	NCI	C25359			
Attente	SCT	385635001			
Fait	NCI	C61036			
Hôpitaux	MSH	M0010590			
Information	NCI	C48191			
Informé (procédure)	SCT	223458004			
Possible	NCI	C53258			pas
Possible	NCI	C82490			pas
Possible	SCT	371930009			pas
Tennis	MSH	M0021163			

accepté	NCI C62559	
accepté	SCT 385645004	
actualités	MSH D016433	pas
appeler	NCI C64924	
appellation	SCT 27468005	
attente	SCT 410517006	
comme	NCI C64938	
comme - fragment d'instruction posologique (valeur qualificative)	SCT 420503003	
date de début	SCT 298059007	
date de fin	NCI C68617	
demande	NCI C48312	
demande pour	SCT 103320006	
défini	SCT 255545003	pas
faire	NCI C86066	
fait	SCT 385640009	
fait	SCT 385658003	
hôpital	NCI C16696	
hôpital	SCT 22232009	
hôpitaux	MSH D006761	
informé	NCI C37896	
messenger	SCT 40838000	
médical	NCI C25261	
médical	SCT 74188005	
médicalisation	MSH M0568535	
médicalisation	MSH D062528	
nouveau	NCI C25586	pas
nouveau	SCT 7147002	pas
protocole APEL	MSH C045350	
rappel	NCI C37922	
rappel	NCI C53609	
rappel	NCI C54270	
rappel mnésique	MSH D011939	
tenir	NCI C62290	
tennis	MSH D013715	

# Annexe **F**

## Données liées aux analyses

### **F.1 Données des patients par année d'inclusion**

Le tableau [F.1](#) regroupe l'ensemble des données patients inclus entre 2013 et 2017 par année d'inclusion.

TABLEAU F.1 – Données de la population par année d'inclusion au Réseau SLA IDF.

	2013	2014	2015	2016	2017
<b>Âge moyen à l'inclusion ± ET (ans)</b>					
Moyenne	65,6 ± 11,44	66,4 ± 12,87	66,1 ± 11,39	66,3 ± 12,9	65,05 ± 11,37
Minimum	26	20	26	22	23
Maximum	90	91	90	88	85
<b>Sexe <i>n</i> (%)</b>					
Homme	110 (55%)	100 (53%)	79 (46%)	104 (53%)	94 (52%)
Femme	89 (45%)	86 (46%)	90 (53%)	93 (47%)	86 (48%)
Sexe ratio	1,23	1,16	0,88	1,12	1,09
<b>Forme <i>n</i> (%)</b>					
Spinale	117 (59%)	104 (56%)	89 (53%)	133 (68%)	124 (69%)
Bulbaire	72 (36%)	61 (33%)	56 (33%)	60 (30%)	47 (26%)
Donnée manquantes	10 (5%)	21 (11%)	24 (14%)	4 (2%)	9 (5%)
<b>Délais moyen de temps au Réseau SLA IDF ± ET (jours)</b>					
Moyenne	585,8 ± 543,3	569,6 ± 493,6	525 ± 428,9	587,9 ± 400	485,6 ± 272,4
Minimum	6	6	7	10	6
Maximum	2335	1950	2062	1251	887
<b>Nombre moyen d'évènements et de suivi ±</b>					
Moyenne	39,9 ± 33,6	39,5 ± 41,5	30,9 ± 28,8	29,5 ± 22,4	30,7 ± 30,3
Minimum	1	1	1	2	2
Maximum	166	322	151	126	164
<b>Mode de sortie <i>n</i> (%)</b>					
Déménagement	9 (5%)	5 (3%)	5 (3%)	1 (1%)	2 (1%)
Décès	175 (88%)	155 (83%)	124 (73%)	127 (64%)	93 (52%)
Choix du patient	1 (1%)	0	3 (2%)	0	0
Vivant	14 (7%)	26 (14%)	37 (22%)	69 (35%)	85 (47%)

## F.2 Carte de densité des professionnels en région Île-de-France

Pour comprendre si les recherches de professionnels, effectuées par les coordinateurs étaient en rapport avec des territoires spécifiques nous avons souhaité observer la densité des professionnels sur les départements de la région île-de-France. Ces cartes sont issues du portail d'accompagnement des professionnels de santé<sup>1</sup> et du site santégraphie<sup>2</sup> de l'ARS Île-de-France.

## F.3 Densité des kinésithérapeutes

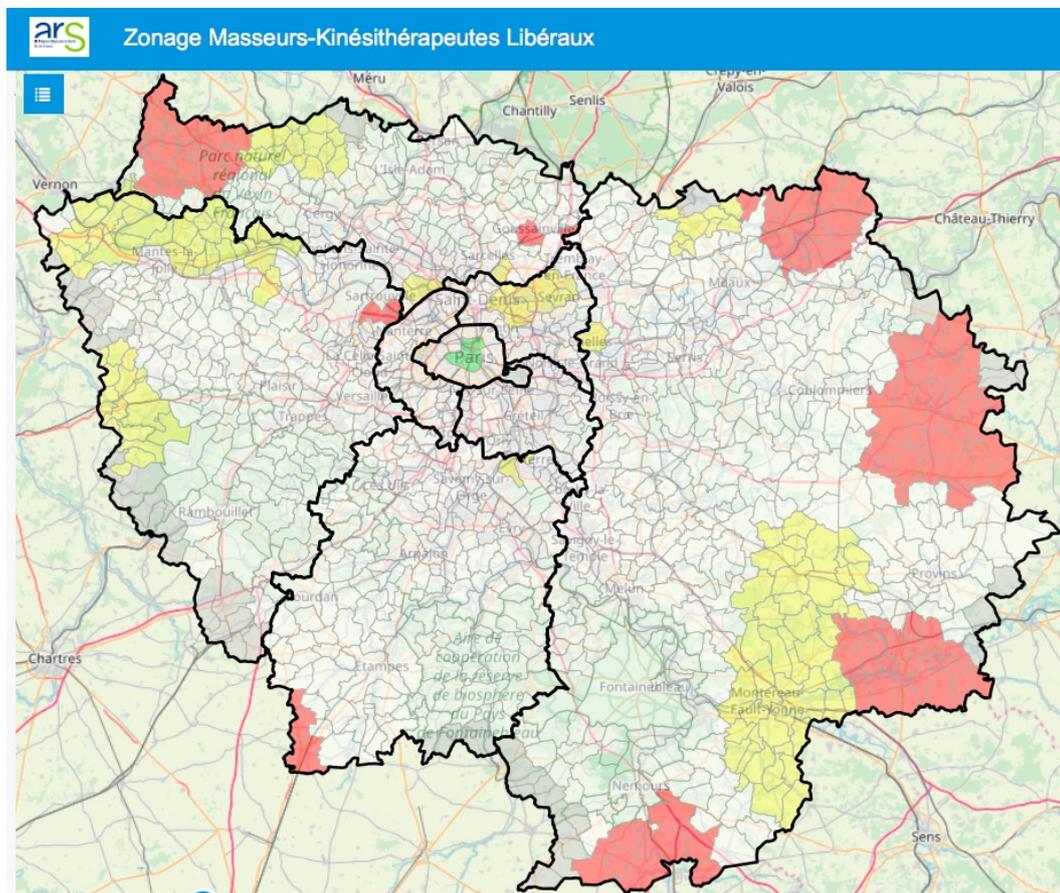


FIGURE F.1 – Densité des kinésithérapeutes en Île-de-France. Les zones "sous dotées" sont représentées en jaune, les zones "très sous dotées" sont représentées en rouge.

1. <http://www.iledefrance.paps.sante.fr/Le-zonage-medecin.42130.0.html>
2. <https://santegrappie.fr/accueil/accueil>

## F.4 Densité des médecins

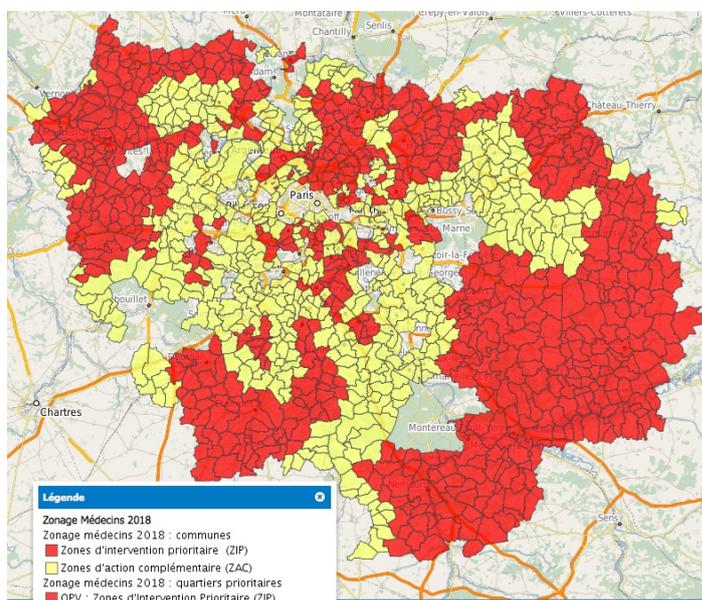


FIGURE F.2 – Densité des médecins en Île-de-France.

## F.5 Densité des orthophonistes

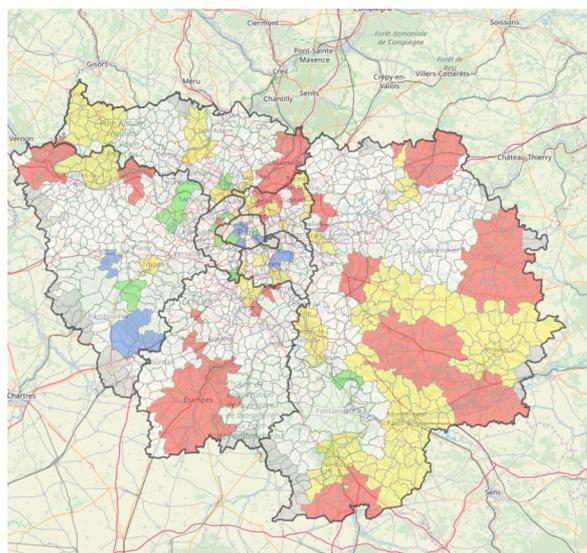


FIGURE F.3 – Densité des orthophonistes en Île-de-France. Les zones rouges sont "très sous dotées", les jaunes sont "sous dotées", les blanches "intermédiaires", les vertes "très dotées" et les bleus "sur dotées".

## **F.6 Données socio-démographiques des patients sans besoins ni demandes**

Le tableau [F.2](#) présente les caractéristiques socio-démographiques des patients pour lesquels, il ne fut annoté aucun besoin ni demande de prise en charge dans le domaine médical et social, et qui ont bénéficié de 3 actions de coordination, correspondant au processus d'inclusion.

TABLEAU F.2 – Caractéristiques des patients sans demandes ni besoins dans le domaine social et médical.

	<b>Patient</b>
<b>Âge moyen au diagnostic</b> $\pm$ ET (ans)	66,8 $\pm$ 15 ; médiane 69,5
<b>Âge moyen à l'inclusion</b> $\pm$ ET (ans)	68,10 $\pm$ 15,8 ; médiane 71
<b>Sexe</b> <i>n</i>	
Homme	14 (52%)
Femme	13 (48%)
<b>Forme</b> <i>n</i>	
Spinale	13
Bulbaire	11
Donnée manquantes	3
<b>Suppléance</b> <i>n</i> (%)	
VNI <sup>3</sup>	5
Gastrostomie	1
Trachéotomie	0
<b>Délais moyen entre le diagnostic et l'inclusion au Réseau SLA IDF</b> $\pm$ ET (jours)	186,4 $\pm$ 172 ; médiane 131
Données manquantes	1
<b>Délais moyen de temps au Réseau SLA IDF</b> $\pm$ ET (jours)	284,5 $\pm$ 389 ; médiane 168
<b>Nombre moyen d'évènements et de suivis</b> $\pm$	3,6 $\pm$ 2,2 ; médiane 3
<b>Mode de sortie</b> <i>n</i>	
Décès	14
Choix du patient	1
Vivant	12
<b>Statut social</b> <i>n</i>	
Célibataire	2
Marié	19
Divorcé	2
Veuf	3
Données manquantes	1
<b>Mode de vie</b> <i>n</i> %	
Vit seul	5 (19%)
Vit en famille	21( 81%)
Données manquantes	1
<b>Répartition par département</b> <i>n</i>	
Paris (75)	6
Seine-et-Marne (77)	1
Yvelines (78)	1
Essonne (91)	6
Hauts-de-Seine (92)	3
Seine-St-Denis (93)	3
Val-de-Marne (94)	2
Val-d'Oise (95)	5

## F.7 Classement en tertile

Le tableau F.3 présente les moyennes des variables pour les classement en tertile de chacun des groupes.

TABLEAU F.3 – Moyenne par classement en Tertile.

Classement par âge <i>n</i>	Âge 1 52,06 ± 8,12 299	Âge 2 66,9 ± 3 320	Âge 3 78,17 ± 4,5 292
Classement par action de coordination <i>n</i>	ActCO 1 15,7 ± 3 292	ActCO 2 51,17 ± 2,9 311	ActCO 3 154,1 ± 3 308
Classement par temps au réseau <i>n</i>	T1 139,90 ± 13,2 294	T2 461,2 ± 12,3 307	T3 1044,8 ± 12,3 310
Classement par besoins et demandes médicales <i>n</i>	BD M1 1 ± 0,2 166	BD M2 3,2 ± 0,1 367	BD M3 9,9 ± 0,2 211
Classement par besoins et demandes sociales <i>n</i>	Social 1 1 ± 0,2 153	Social 2 3,2 ± 0,1 360	Social 3 10,4 ± 0,2 200
Classement par complexité <i>n</i>	C1 1,43 299	C2 1,79 307	C3 2,22 305



## Résumé

Les patients atteints de Sclérose Latérale Amyotrophique nécessitent un accompagnement pluridisciplinaire au cours de leur parcours. La pathologie provoque de nombreuses incapacités et situations de handicaps. Cette prise en charge complexe peut amener à des situations de rupture de parcours par l'absence, l'arrêt ou des difficultés de prise en charge, cependant les causes de ces ruptures ne sont pas connues. Le réseau de coordination ville hôpital SLA Île-de-France dispose d'une base textuelle de coordination, sur laquelle les besoins et les demandes des patients sont décrits tout au long de leurs parcours. À cette fin, nous proposons d'analyser cette base pour en extraire de la connaissance et décrire les parcours patients. Pour cela nous utilisons un outil d'annotation sémantique qui utilise comme ressource une ontologie du domaine que nous avons créé ONTOPARON. ONTOPARON est une ontologie modulaire composée d'un module médical, d'un module socio-environnemental, d'un module de coordination, d'un module noyau et d'un module de consolidation. Ces modules ontologiques permettent de prendre en compte les dimensions intervenant dans les parcours des patients. Des concepts définis, rendant compte de thématiques importantes comme l'épuisement de l'aidant ou bien encore la présence de problèmes sociaux ont été créés pour détecter les difficultés rencontrées. L'annotation de 931 dossiers patients a permis de voir que tous les patients ne présentent pas les mêmes besoins ni les mêmes demandes, certaines thématiques s'expriment différemment en fonction de l'âge, de la forme de pathologie ou du mode de vie des personnes. Ce travail de thèse s'accompagne de la mise en place d'un outil d'annotation sémantique ONBASAM et d'un outil d'évaluation d'annotation PRONTO.

**Mots-clés :** informatique médicale, ontologie modulaire, concepts définis, sclérose latérale amyotrophie, parcours de soins.

---

## Abstract

Patients with Amyotrophic Lateral Sclerosis (ALS) require multidisciplinary support during their course. Pathology causes many disabilities and disabilities. This complex management can lead to situations of course failure through absence, cessation or difficulties of management, however the causes of these breaks are not known. The coordination network SLA Île-de-France has a textual coordination base, on which the needs and demands of patients are described throughout their journey. To this end, we propose to analyze this basis in order to extract knowledge and describe the patient pathways. For this we use a semantic annotation tool that uses as a resource an ontology of the domain that we have created ONTOPARON. ONTOPARON is a modular ontology composed of a medical module, a socio-environmental module, a coordination module, a core module and a consolidation module. These ontological modules make it possible to take into account all the dimensions involved in patients' journeys. Defined concepts, reporting on important themes such as the exhaustion of the caregiver or even the presence of social problems were created to detect the difficulties encountered. The annotation of 931 patient files has shown that not all patients have the same needs or the same demands, some themes are expressed differently depending on the age, the pathology or the lifestyle. This thesis work is accompanied by the implementation of a semantic annotation tool ONBASAM and an annotation evaluation tool PRONTO.

**Keywords :** medical informatics, modular ontology, defined concepts, ALS, care pathway.

