



République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la
recherche scientifique

Université Larbi Tébessi - Tébessa



Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie
Département : Mathématiques et Informatique

Mémoire de fin d'étude
Pour l'obtention du diplôme de *MASTER*
Domaine : Mathématiques et Informatique
Filière : Informatique
Option : Systèmes d'information

Thème

**Ontologie pour la traçabilité des conduites
thérapeutiques en oncologie médicale**

Présenté Par :

Abbassi Oumaima

Devant le jury :

Dr. Amroune Mohamed	MCA	Université Larbi Tébessi	Président
Dr. Bourougaa Salima	MCB	Université Larbi Tébessi	Examinateur
Dr. Gasmi Mohamed	MAB	Université Larbi Tébessi	Encadreur
Dr. Bendhib Issam	MCB	Université Larbi Tébessi	Co-Encadreur

Date de soutenance : 08/07/2019

Remerciement

Je voudrais tout d'abord exprimer mes plus profonds remerciements à mes encadreurs Dr. Bendhib Issam & Dr. Gasmi Mohamed pour ses accord d'être mes directeurs de mémoire et de ses disponibilité et ses aide pendant toute la préparation de ce travail.

Je tiens aussi à remercier tous les membres de jury : Dr. Amroune Mohamed et Dr. Bourougaa-Tria Salima, pour leur disponibilité et acceptation d'examiner et de rapporter mon travail.

Je remercie ainsi tous les enseignements de département Mathématique et informatique pour ses conseils et son encouragement.

Je remercie tous mes collègues de l'université de Tébessa, pour leurs encouragements et précieuses orientations pendant toute la période de l'élaboration de ce travail.

Je ne saurais oublier de remercier ma mère ainsi que mes sœurs pour leur soutien moral, leurs encouragements et leur patience durant les étapes de réalisation de ce travail.

Enfin, Que tous ceux qui directement ou indirectement m'ont apporté leur aide, trouvent ici l'expression de mes sincères remerciements.

Dédicace

A mes parents et ma grand-mère,

A mes frères, et toute ma famille,

A tous mes enseignants,

A mes collègues,

A mon Marie,

A mes amis.

الملخص

في علم الأورام الطبية، فإن اختيار السلوك العلاجي هو القرار الحاسم للطبيب ليكون قادرًا على إنقاذ المريض. يعتمد هذا الاختيار دائمًا على التاريخ العلاجي للمريض.

يوفر دمج الأنطولوجيات في المجال الطبي مفردات مشتركة بالإضافة إلى وصف لمعنى مصطلحات مجال ما والعلاقات التي يحافظون عليها ، ويمكن أن يكون هذا التكامل قابلاً للاستغلال عن طريق المعلوماتية ، كما أنه يقدم مساهمة أساسية لإعادة استخدام الموارد الطبية أو تبادل المعلومات.

في هذا العمل ، نحن مهتمون بتنفيذ علم الأنطولوجيا في مجال الأورام (سرطان القولون) الذي يسمح بالحفاظ على تتبع جميع السلوكيات العلاجية من أجل تسهيل عمل مختلف الجهات الطبية الفاعلة في البحث عن المعلومات وأخذها القرار.

الكلمات المفتاحية: التتبع ، الإدارة العلاجية ، علم الوجود ، علم الأورام.

RÉSUMÉ

En oncologie médicale, le choix de la conduite thérapeutique est la décision cruciale du médecin pour pouvoir sauver le patient. Ce choix dépend toujours de l'historique thérapeutique du patient.

L'intégration des ontologies dans le domaine médical fournit un vocabulaire commun ainsi qu'une description de la signification des termes d'un domaine et des relations qu'ils entretiennent, cette intégration peut être exploitable de manière informatique, elle apporte aussi une contribution primordiale à la réutilisation des ressources médicales ou à l'échange d'informations.

Dans ce mémoire nous nous sommes intéressées à implémenter une ontologie dans le domaine d'oncologie (cancer de colon) qui permet de garder la traçabilité de toutes conduite thérapeutique afin de faciliter le travail des différents acteurs médicaux dans la recherche d'information et la prise de décision.

Mots-clés : traçabilité, conduite thérapeutique, ontologie, oncologie.

ABSTRACT

In medical oncology, the choice of therapeutic behavior is the crucial decision of the physician to be able to save the patient. This choice always depends on the therapeutic history of the patient.

The integration of ontologies in the medical field provides a common vocabulary as well as a description of the meaning of the terms of a domain and the relations that they maintain, this integration can be exploitable in a computer way, it also brings a primordial contribution the reuse of medical resources or the exchange of information.

In this thesis we are interested in implementing an ontology in the field of oncology (colon cancer) that allows to keep the traceability of all therapeutic conduct in order to facilitate the work of different medical actors in the information search and the taking of decision.

Keywords: traceability, therapeutic management, ontology, oncolog

Table des matières

<i>Introduction générale</i>	1
Chapiter 1 Les ontologies	3
1.1 Historique	4
1.2 Définitions	4
1.3 Composants d'une ontologie	6
1.4 Types d'Ontologie.....	7
1.4.1 Objet de conceptualisation.....	7
1.4.2 Niveau de complétudes	10
1.5 Cycle de vie d'une ontologie	10
1.6 Construction d'ontologie.....	11
1.7 Les Formalismes pour la représentation	16
1.8 Les langages de représentation.....	18
1.9 Outils de développements d'Ontologies.....	22
1.10 Conclusion	25
Chapiter 2 L'ontologie médicale	26
2.1 L'ontologie Médicale	27
2.1.1 Terminologies.....	27
2.1.2 Ontologies Vs Terminologies	27
2.1.3 Construction et édition d'ontologies	28
2.1.4 Ressources terminologiques et ontologiques en médecine	30
2.1.5 Enrichissement, maintenance et évolution d'ontologies	35
2.2 L'Oncologie	35
2.2.1 Le cancer	37
2.2.2 Les causes du cancer	37
2.2.3 Les principaux cancers.....	38
2.2.4 Les traitements.....	38
2.2.5 Les effets du cancer et de son traitement	39
2.2.6 La prévention	39
2.3 Le Cancer du colon.....	40
2.3.1 Facteurs de risque.....	40
2.3.2 Le dépistage.....	40
2.3.3 Stades du cancer	41

2.3.4	Traitements	41
2.3.5	Suivi	41
2.4	Ontologies dans le domaine oncologie	42
2.5	Conclusion	43
Chapiter 3	Conception d'une ontologie de traçabilité.....	45
1	Processus de développements d'ontologie	46
2	Description d'ontologie carcinome de colon	48
2.1	Conception de l'ontologie du domaine carcinome de colon.....	48
2.1.1	Domaine de l'ontologie.....	48
2.1.2	Les besoins examinés par cette ontologie	48
2.2	Processus d'affinement de l'ontologie	48
2.3	Respect des principes de construction	49
2.4	Présentation de cette ontologie	49
2.4.1	Les axes primitives de l'arbre ontologique construire	51
2.4.1.1	Classification scientifique de carcinome de colon	51
2.4.1.2	Taxon	51
2.5	Les propriétés des classe de cette ontologie.....	52
2.5.1	propriété d'objet.....	52
2.5.2	propriété de type de donnée.....	52
2.6	représentation hiérarchique des concepts	52
2.7	Glossaire des concepts	54
2.8	La représentation graphique de carcinome du colon.....	55
3	Traitements	57
3.1	Description d'ontologie de Traitements	57
3.2	Construction d'une ontologie de traitements	58
3.2.1	La Méthode METHONTOLOGY	58
3.2.2	Spécification	58
3.2.3	Conceptualisation	59
3.2.4	Construction de diagramme de relations binaires	59
3.2.5	Construction de visualisation d'ontologie	60
3.2.6	Représentation hiérarchique des classes binaires	61
4	Définition de l'ontologie dédiée à la Traçabilité	62
4.1	Traçabilité dans le domaine médicale	62
4.2	Ontologie de traçabilité.....	63
4.3	La visualisation d'ontologie de traçabilité	65
5	Conclusion	68
	Conclusion générale.....	70

LISTE DES FIGURES

Figure 1.1 Typologie classé par objet de conceptualisation	8
Figure 1.2 Classification de Lassila et Mc Guinness	8
Figure 1.3 Différents types d'ontologies].....	10
Figure 1.4 Cycle de vie d'une ontologie.	11
Figure 1.5 Processus général de construction d'une ontologie	12
Figure 1.6 Utilisation des ontologies.....	13
Figure 1.7 La pyramide des langages du Web sémantique.	22
Figure 2.1 Les différents sous-domaines intégrés à l'UMLS	33
Figure 2.2 Exemple (NF2 et protéines et maladies associées dans l'UMLS)	33
Figure 2.3 Description d'un terme dans GO.....	34
Figure 2.4 Comment une cellule devient cancer.	37
Figure 3.1 Classification du cancer à partir d'OD.	51
Figure 3.2 représentation hiérarchique des concepts principaux de carcinome du colon	53
Figure 3.3 représentation de visualisation OWL des concepts principaux de carcinome du colon..	54
Figure 3.4 représentation de graphe des concepts principaux de carcinome du colon.....	56
Figure 3.5 Graphe de relations binaires.....	60
Figure 3.6 la visualisation d'ontologie de traitement.....	60
Figure 3.7 représentation hiérarchique de classes principales de traitements.....	61
Figure 3.8 Représentation de l'information dans l'ontologie de traçabilité.....	64
Figure 3.9 Diagramme conceptuel de l'ontologie de traçabilité.....	65
Figure 3.10 Visualisation de traçabilité.....	66
Figure 3.11 diagramme de relations binaires.....	67

LISTE DES TABLEAUX

Tableau 1.1 Les principaux moteurs d'inférence	23
Tableau 2.1 Les différentes catégories ou domaines du MeSH	31
Tableau 2.2 Les onze axes de la SNOMED	32
Tableau 3.1 Définition des concepts	54
Tableau 3.2 Définition concepts principale d'ontologie de traitements.....	57
Tableau 3.3 Table des relations binaires	61

GLOSSAIRE

- **Asymptomatique** : ne présentant pas de signe clinique permettant le diagnostic de la maladie.
- **Comorbidités**: pathologies associées.
- **Incidence** : nombre de nouveaux cas d'une maladie dans une population, pendant un temps donné.
- **Métastase** : localisation à distance d'une tumeur cancéreuse propagée par voie sanguine ou lymphatique.
- **Réunion de concertation pluridisciplinaire (RCP)** : réunion régulière entre professionnels de santé, au cours de laquelle se discutent la situation d'un patient, les traitements possibles en fonction des dernières études scientifiques, l'analyse des bénéfices et les risques encourus ainsi que l'évaluation de la qualité de vie qui va en résulter.
- **Test OC Sensor** : test immunologique de dépistage de recherche de sang occulte dans les selles, avec lecture automatisée permettant une évaluation quantitative.

Introduction Générale

INTRODUCTION GÉNÉRALE

Comme dans d'autres pays du monde, l'Algérie s'intéresse à l'oncologie et cherche toujours à se tenir au courant de ses propres développements, comme en témoignent les efforts déployés, dont on sait qu'ils augmentent sans cesse pour lutter contre les tumeurs et réduire le taux de mortalité en suivant les différentes méthodes de traitement que les spécialistes travaillent à développer dans le monde.

Lors de la quatrième journée internationale d'oncologie tenue l'année passée (2018) à Oran et organisées par le Chuo Docteur Benzerdjeb, en collaboration avec l'Agence thématique de recherches en sciences de la santé (ATRSS), ont été l'occasion de débattre des cancers digestifs, particulièrement les cancers métastatiques et les tumeurs neuroendocrines.

Ont participé à ce rendez-vous des spécialistes référents dans les divers domaines de l'oncologie digestive entre gastroentérologie, chirurgie, biologie moléculaire, entre autres spécialités, venus des différentes régions du pays, de France, ainsi que de Belgique. Le constat premier est la "nette augmentation" de ce type de cancer estimé à 5% chaque année de par le monde. En Algérie, son incidence est de 10%. [1]

Pour le Pr Abdelkader Bousahba, le président de ces journées, cette augmentation des cas de cancer colorectal "est liée au changement du mode alimentaire ainsi qu'à d'autres facteurs de risques sur lesquels se penche la recherche à l'heure actuelle". Pour lui, il s'agit en premier lieu de faire le point sur les actions prioritaires à mener sur le terrain dans le cadre de la feuille de route que constitue le Plan national cancer (PNC) 2015-2019 et de fournir une revue des plus importantes avancées réalisées dans la prise en charge des cancers digestifs, en général, et du cancer colorectal et des tumeurs neuroendocrines [1], en particulier le cancer de colon est l'un des plus difficiles et des plus difficiles à détecter, et le patient atteint de cancer du côlon compte plusieurs étapes de traitement, ce qui peut constituer un obstacle pour les experts sur le terrain, car il n'est pas facile de procéder à un dépistage exhaustif et que le patient doit tout garder pendant le traitement. Dès le début et être à la découverte de la tumeur jusqu'à la mort du patient.

Ce processus dérange les patients et les médecins en raison de l'importance de conserver toutes les connaissances et les informations liées au traitement du patient parce que le traitement du cancer est lié en série et systématique et tout défaut ou déficience dans les étapes précédentes affecte négativement la vie du patient Dans les informations de santé du patient.

Certaines des questions que nous viennent à l'esprit comme :

Comment le patient peut-il conserver toutes ses données médicales thérapeutiques?

Comment le médecin peut-il identifier les traitements antérieurs subis par le patient?

Est-il possible de fournir un moyen efficace permettant au patient de conserver ses données thérapeutiques et les résultats de chaque traitement, qu'il soit pharmaceutique, chimique ou chirurgical?

Peut-on résoudre le problème du suivi thérapeutique pour les patients atteints de cancer?

Pour répondre à ces questions, nous nous proposons de modéliser les connaissances sous forme d'ontologie afin de prendre en considération la traçabilité thérapeutique dans le domaine d'oncologie (cancer de colon) afin de consigner et de conserver chaque traitement, sa cause, son résultat, sa durée et sa quantité... etc.

C'est ce que nous essayons d'expliquer dans ce travail lié à l'ontologie médicale, reconnue pour son développement et sa grande attention par les spécialistes et les utilisateurs, ainsi que par les médecins,...

Le mémoire est organisé en plusieurs chapitres, le chapitre un donne les généralités sur l'ontologie à savoir : les définitions et concepts, les composants d'une ontologie, les différents types, le cycle de vie d'ontologie, aussi nous expliquons comment nous pouvons construire une ontologie avec ces formalismes, langage et outils.

Dans le deuxième chapitre nous présentons l'ontologie dans le domaine médical en trois sections, la première explique la différence entre les terminologies et les ontologies aussi nous avons cité les méthodes les plus utilisées pour la construction d'une ontologie acceptable, aussi bien les ressources terminologiques et ontologiques médicales. La deuxième section est pour l'oncologie et ses définitions et spécialités. On a choisi le cancer de colon comme domaine de travail ou on le définit dans la troisième section.

Nous sommes concentrés sur la conception de notre travail et le processus de suivi pour le développement de notre ontologie de traçabilité dans le troisième chapitre. On a essayé de définir nos 3 ontologies, la 1^{ère} qui concerne l'oncologie (cancer de colon), la 2^{ème} l'ontologie thérapeutique et on terminera par la définition de l'ontologie noyau de notre travail qui est l'ontologie de traçabilité.

Finalement, le mémoire se termine par une conclusion et des perspectives très objectives.

Chapitre 1
LES ONTOLOGIES

LES ONTOLOGIES

L'ingénierie de connaissances (IC) a longtemps été considérée comme le domaine de prédilection du développement d'expertise en conception de système à base de connaissances. La démarche consiste à ajouter des métadonnées aux connaissances qui décrivent leurs contenus et leurs fonctionnalités, celles-ci doivent s'appuyer sur des ontologies afin de pouvoir être partagées et munies d'interprétations opérationnelles.

Les ontologies constituent l'une des bases les plus importantes pour identifier le vocabulaire commun et les règles relatives à un domaine particulier, ainsi que d'établir des relations entre le vocabulaire et les règles.

Dans ce chapitre, nous allons révéler les différentes définitions. Nous allons montrer aussi la place d'ontologie dans les systèmes à bases de connaissances. Ensuite, nous montrerons le processus et les méthodologies servant leur construction, ainsi que les principaux formalismes et les outils utilisés pour servir leur représentation et développement, à savoir les langages de spécification, les moteurs d'inférences, les langages d'interrogation, et les éditeurs d'ontologie.

1.1 Historique

Le terme ontologie est traduit en origine grec et implique l'existence, la logique et toute science. Il a été introduit pour la première fois en 1613, dans le dictionnaire écrit par Rudolf Goclenius. Le terme a été utilisé pour la première fois au 18ème siècle par Christian von Wolfe (1679-1754).

La difficulté de définir le champ d'application de l'ontologie réside non seulement dans la nouveauté du terme, mais aussi dans les doutes qui accompagnent le mot existence depuis qu'il a été utilisé par Parménide.

Platon a essayé de rechercher l'existence constante et immortelle dans les idéaux, Bien qu'Aristote n'approuve pas l'abstraction des idéaux par Platon, il lie son existence à la connaissance, mais sa définition de l'existence devient la porte d'entrée de toutes les sciences possibles.

Le terme « ontologie » est utilisé depuis le début des années 1990, et son champ d'application s'élargit considérablement.

Introduit en Intelligence Artificielle (IA) il y a 23 ans, le terme d'ontologie est cependant usité en philosophie depuis le XIXème siècle. Dans ce domaine, l'ontologie est une étude de l'être en tant qu'être, c'est-à-dire, une étude des propriétés générales de ce qui existe.

C'est à l'occasion de l'émergence de l'Ingénierie des Connaissances que les ontologies sont apparues en IA, comme réponses aux problématiques de représentation et de manipulation des connaissances au sein des systèmes informatiques.

1.2 Définitions

Les ontologies, à l'origine c'est une branche de la philosophie qui s'intéresse à la nature et à l'organisation de la réalité, correspondent à la partie de la métaphysique qui s'intéresse à l'être en tant qu'être, par opposition aux philosophies secondes qui s'intéressent à l'étude des manifestations de l'être. En informatique, la littérature fournit un tas de définitions du mot ontologie. Ces définitions, dans leur diversité, offrent des points de vue à la fois différents et complémentaires. Dans ce qui suit, nous décrivons quelques-unes.

Définition 1 :

« Une ontologie définit les termes et les relations de base du vocabulaire d'un domaine ainsi que les règles qui indiquent comment combiner les termes et les relations de façon à pouvoir étendre le

vocabulaire ». [1]

Définition 2 :

« *Une ontologie est une spécification explicite d'une conceptualisation* ». [1]

Cette définition se base sur deux dimensions :

Une ontologie est la conceptualisation d'un domaine, c'est-à-dire un choix de comment on définit un domaine. La spécification de cette conceptualisation, c'est-à-dire sa description formelle.

La définition de Gruber est la plus utilisée dans la littérature. Elle a été légèrement modifiée par Borst.

Définition 3 :

« *Une ontologie est une spécification explicite et formelle d'une conceptualisation partagée* » [1]

Le terme « **conceptualisation** » réfère dans cette définition à une abstraction d'un phénomène du monde, obtenue en identifiant les concepts appropriés à ce phénomène.

Le terme « **formelle** » indique que les ontologies sont interprétables par la machine.

Cependant, « **spécification explicite** » signifie que les concepts de l'ontologie et les contraintes liées à leur usage sont définis de façon déclarative. Enfin, le terme «partagé» signifie que l'ontologie capture la connaissance consensuelle. Mais cette définition laisse la porte ouverte à de nombreuses définitions. [1]

Définition 4 :

« *Une ontologie est une description formelle d'entités et leurs propriétés, relations, contraintes, comportement* » [1]

Cette définition de Grüninger est simplifiée dans [1] où une ontologie est définie comme un ensemble de définitions de concepts et leurs relations. A ne pas confondre avec un modèle qui est un ensemble d'instances de ces concepts.

Définition 5 :

« *Sont des spécifications partielles et formelles d'une conceptualisation commune* » [1]

En 1997, Guarino accentue l'ambiguïté du terme conceptualisation qui doit être pris dans son sens intuitif. La spécification des ontologies est partielle, car une conceptualisation ne peut pas toujours être entièrement formalisée dans un cadre logique, du fait d'ambiguïtés ou du fait qu'aucune représentation de leur sémantique n'existe dans le langage de représentation d'ontologies choisi. « Commune » renvoie à l'idée qu'une ontologie rend compte d'un savoir consensuel, c'est-à-dire qu'elle n'est pas l'objet d'un individu, mais qu'elle est reconnue par un groupe. [1]

A partir de ces définitions nous constatant que les experts se diffèrent dans la définition de l'ontologie, mais ils sont d'accord qu'il s'agit d'une science large qui définit les termes de base et les relations du vocabulaire d'un domaine donné, ainsi que la création de règles qui indiquent comment combiner les termes et permettant de lier les termes et les relations pour étendre l'expansion du vocabulaire.

1.3 Composants d'une ontologie

Les ontologies produisent un vocabulaire commun d'un domaine et définissent, de façon plus ou moins formelle, la signification des termes et des relations entre eux. Les connaissances intégrées dans les ontologies sont formalisées en mettant en jeu cinq types de composants concepts, relations, fonctions, axiomes, instances. [1]

❖ Concepts

Ils sont appelés aussi termes ou classes de l'ontologie. Un concept est un constituant de la pensée (un principe, une idée, une notion abstraite) sémantiquement évaluable et communicable. L'ensemble des propriétés d'un concept constitue sa compréhension ou son intention et l'ensemble des êtres qu'il englobe, son extension. [1]

- Selon un concept se définit à trois niveaux : Un concept est une signification. Sa place dans un système de significations permet de le comprendre, de le distinguer et de le différencier par rapport à d'autres concepts. Un concept est une construction.
- Selon, ces concepts peuvent être classifiés selon plusieurs dimensions : [1]
 - Niveau d'abstraction (concret ou abstrait).
 - Atomicité (élémentaire ou composée).
 - Niveau de réalité (réel ou fictif).

❖ Relations

Représentent un type d'interaction, ou bien des associations existant entre les concepts d'un domaine. Elles se définissent formellement à partir d'un produit de n concepts :

$$R : C_1 \times C_2 \times \dots \times C_{n-1}(C)$$

sous-classe-de (Spécialisation, généralisation), partie-de (agrégation ou composition), associée-à, instance-de sont des exemples de relations binaires.

En résumé les relations sont les liens que les concepts peuvent avoir entre eux.

❖ Fonctions

Ce sont des cas particuliers de relations dans lesquelles le Nième élément de la relation est défini de manière unique à partir des n-1 premiers. Formellement, les fonctions sont définies par des expressions :

$$F : C_1 \times C_2 \times \dots \times C_{n-1} \rightarrow C_n.$$

❖ Axiomes

Ils ont pour objectif de représenter des concepts et des relations dans un langage logique afin de représenter leur sémantique. Ils représentent les intentions des concepts et des relations du domaine et, de manière générale, les connaissances n'ayant pas un caractère strictement terminologique [1]. L'utilisation des axiomes sert à définir le sens des entités, mettre des restrictions sur la valeur des attributs, examiner la conformité des informations spécifiées ou en déduire de nouvelles. [1]

❖ Instances

Elles constituent la définition extensionnelle de l'ontologie ; ces objets véhiculent les connaissances (statiques, factuelles) d'un domaine d'application [1] Par exemple : (Clio est une instance du concept voiture).

1.4 Types d'Ontologie

Il existe plusieurs types d'ontologie dans la littérature, mais cette section nous allons présenter les plus généralement utilisés.

1.4.1 Objet de conceptualisation

Les ontologies sont classifiées selon leur objet de conceptualisation (but de leur utilisation) de la façon suivante : [1]

1. Supérieure/Haut niveau,
2. Domaine,
3. Tâche,
4. Application.

- **Ontologie de haut niveau** : décrit des concepts très généraux comme l'espace, le temps, la matière, les objets, les événements, les actions, etc. Ces concepts ne dépendent pas d'un problème ou d'un domaine particulier, et doivent être, du moins en théorie, consensuels à de grandes communautés d'utilisateurs. [1].

- **Ontologie de domaine** : Contrairement aux ontologies de haut niveau, les ontologies de domaine sont plus spécifiques. Elles synthétisent les connaissances spécifiques à un domaine particulier. Elles décrivent le vocabulaire ayant trait à un domaine générique (ex. l'enseignement, la médecine...), notamment en spécialisant les concepts d'une ontologie de haut niveau. [1]
- **Ontologie de tâches** : Ce type d'ontologies est utilisé pour conceptualiser des tâches spécifiques dans les systèmes, telles que les tâches de diagnostic, de planification, de conception, de configuration, de tutorat. Soit tout ce qui concerne la résolution de problèmes. Ce type d'ontologies décrit le vocabulaire concernant une tâche générique (ex. enseigné, diagnostiqué...), notamment en spécialisant les concepts d'une ontologie de haut niveau. [1]
- **Ontologie d'application** : Cette ontologie est la plus spécifique, elle contient des concepts dépendants d'un domaine et d'une tâche particuliers, qui sont généralement subsumés par des concepts de ces deux ontologies. Ces concepts correspondent souvent aux rôles joués par les entités du domaine lors de l'exécution d'une certaine activité. [1]

Les ontologies génériques ou ontologies de haut niveau (super ontology) expriment des conceptualisations valables dans différents domaines de valeur relativement générale comme les notions d'objets, de propriété, de valeur, d'état, ou encore des concepts de temps d'espace d'événements. Elles sont prévues pour être utilisées dans des situations diverses, et pour servir une large communauté d'utilisateurs. [3]



Figure 1.1 Typologie classé par objet de conceptualisation [5]

Niveau de formalisme de représentation :

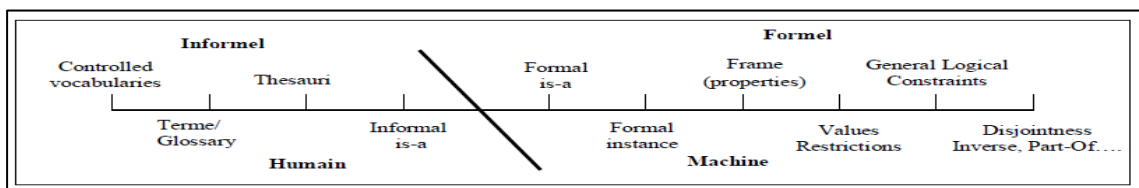


Figure 1.2 Classification de Lassila et Mc Guinness [9]

Ce type d'ontologies regroupe les concepts utilisés pour formaliser les connaissances.

Parmi les ontologies de représentation, on trouve des ontologies qui vont décrire les notions utilisées dans toutes les ontologies pour spécifier les connaissances, telles que les substances, les concepts, les relations etc. Par exemple, la «Frame-Ontology » est une ontologie de représentation.

Elle définit de manière formelle les concepts utilisés principalement dans les langages à base de frames : classes, sous-classes, attributs, valeurs, relations et axiomes.

Les ontologies de représentation sont indépendantes des différents domaines de connaissances, puisqu'elles décrivent des primitives cognitives communes aux différents domaines. [3]

Selon le niveau du formalisme de représentation, [1] propos une classification comprenant les principales catégories :

1. **Informelles** : ontologies opérationnelles dans un langage naturel (sémantique ouverte). Ainsi cette ontologie est très explicite et facile à comprendre. Cependant ce formalisme ne respecte pas la relation de subsomption connue aussi sous le nom de relation taxonomique. Ainsi c'est difficile de l'utiliser dans le système informatique.
2. **Semi-informelles** : utilisation d'un langage naturel structuré et limité pour décrit la connaissance. Elle spécifie ainsi la langue naturelle contrôlée et structurée.
3. **Semi-formelles** : (langage artificiel défini formellement.) Cette ontologie exprime la connaissance formelle non interprétée par la machine. Elle est faite pour les langages artificiels qui sont bien définis formellement.
4. **Formelles** : utilisation d'un langage artificiel contenant une sémantique formelle, ainsi que des théorèmes et des preuves de propriétés telles la robustesse l'exhaustivité. [1] Les langage utilisé est artificiel et les informations sont lisibles par la machine. [5]
5. **Le vocabulaire contrôlé** : C'est une liste finie de termes définis par un groupe de personnes ou une communauté. La signification des termes n'est pas forcément définie et il n'y a pas d'organisation logique entre les termes. Les catalogues sont des exemples de cette catégorie. [9]
6. **Le glossaire** : C'est une liste de termes avec leurs significations spécifiées en langage naturel. Cette représentation apporte plus d'informations car une personne peut lire la définition, cependant elle n'est pas interprétable par l'ordinateur. [9]
7. **Le thésaurus** : Il fournit une sémantique supplémentaire entre les termes tels que la relation de synonymie. Mais, il n'offre pas la structure hiérarchique explicite. [9]

1.4.2 Niveau de complétudes

Bachimont utilisait le terme engagement pour modéliser les connaissances. Il exploitait l'engagement sémantique pour le sens linguistiques des concepts et l'engagement ontologique pour le sens formel des concepts ainsi que l'engagement computationnel pour l'exploitation des concepts. Trois types d'ontologie sont associés à cette ontologie. [5]

a) Ontologie régionale :

Cette ontologie est représentée comme étant un arbre de concepts sémantique. Elle permet de définir le sens d'une unité linguistique sans utiliser un concept, ni une référence ni un linguistique lui-même. Elle permet de contester l'usage du paradigme différentiel.

b) Ontologie référentielle :

C'est une sorte d'ontologie formelle constituée par des concepts formels appelés aussi concepts référentiels. Ainsi cette ontologie s'appuie sur l'analyse de la fonction référentielle qui n'est autre que l'utilisation des lexiques pour le couple « termes/libellé ».

c) Ontologie computationnelle :

Dans le but de générer des inférences, cette ontologie utilise les concepts computationnels à travers les différents types d'opérations possibles.

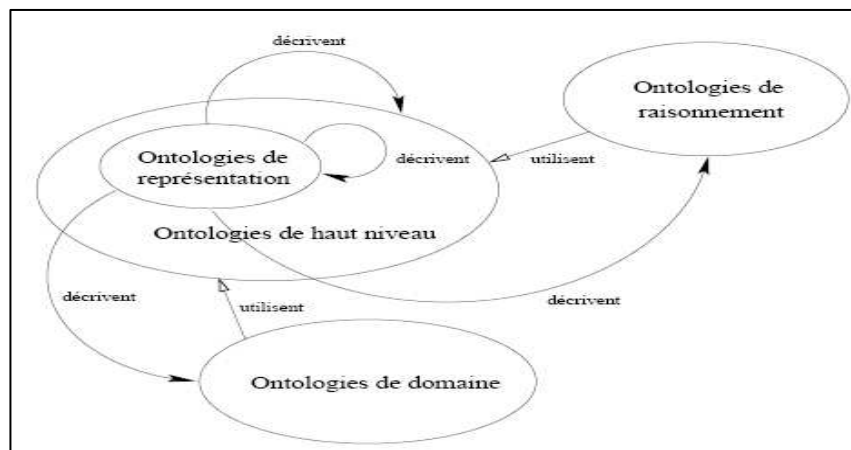


Figure 1.3 Différents types d'ontologies [1]

1.5 Cycle de vie d'une ontologie

Les ontologies étant destinées à être utilisées comme des composants logiciels dans des systèmes répondant à des objectifs opérationnels différents, leur développement doit s'appuyer sur les mêmes principes que ceux appliqués en génie logiciel. Ainsi, les ontologies doivent être considérées comme des objets techniques évolutifs et possédants un cycle de vie qui nécessite

d'être spécifié. Dans ce contexte, les activités liées aux ontologies sont d'une part des activités de gestion incluant la planification, le contrôle, et la garantie de qualité, et d'autre part des activités orientées développement regroupant. Les activités de pré-développement, de développement et de post-développement ;[7]

Un cycle de vie inspiré du génie logiciel (voir figure 1.4), comprend une étape initiale d'évaluation des besoins, qui se transforme en idée, la concrétisation de l'idée qui se traduit par la conception qui est diffusée pour son utilisation.

Vient ensuite l'étape de l'évaluation qui donne naissance, le plus souvent à une étape d'évolution. et de maintenance du modèle. Une réévaluation de l'ontologie et des besoins devra se faire après chaque utilisation significative. L'ontologie peut être étendue et, si nécessaire, en partie reconstruite. La validation du modèle de connaissances est au centre du processus et se fait de manière itérative.

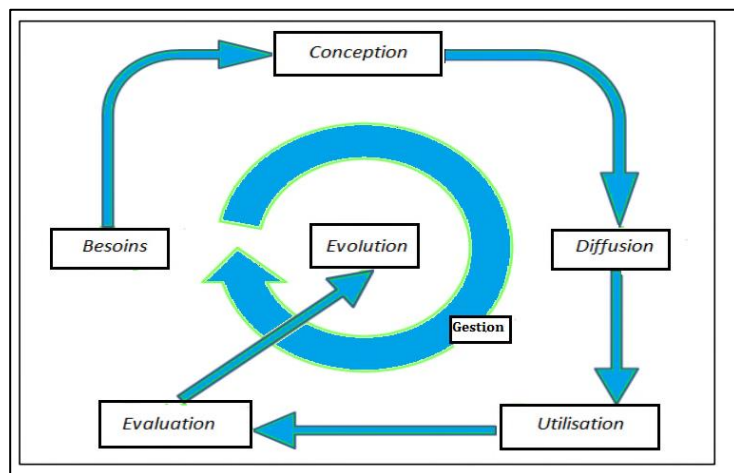


Figure 1.4 Cycle de vie d'une ontologie. [7]

1.6 Construction d'ontologie

L'activité de construction est la partie centrale du cycle de vie d'une ontologie. Le processus de leur construction est une collaboration qui réunit des experts du domaine de connaissance, des ingénieurs de la connaissance, voire les futurs utilisateurs de l'ontologie. Cette collaboration ne peut être fructueuse que si les objectifs du processus ont été clairement définis. [7]

La figure suivante présente les différentes étapes permettant de passer des données brutes à l'ontologie opérationnelle.

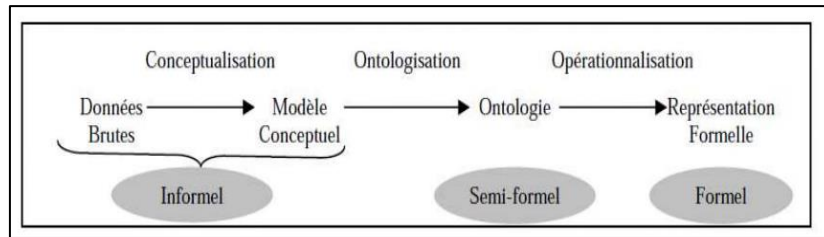


Figure 1.5 Processus général de construction d'une ontologie [7]

- **Conceptualisation** : La première étape de construction, appelée conceptualisation, permet d'aboutir à un modèle informel, donc sémantiquement ambiguë et généralement exprimé en langage naturel. Cette étape, consiste, à partir des données brutes, à dégager les concepts et les relations entre ces concepts permettant de décrire de manière informelle les entités cognitives du domaine. Ainsi, le modèle obtenu consiste en un ensemble de termes désignant les entités du domaine de connaissances (concepts, relations, propriétés des concepts et des relations, etc.).
- **Ontologisation** : L'ontologisation consiste en une formalisation partielle, sans perte d'information, du modèle conceptuel obtenu dans l'étape précédente. Ce qui permet de faciliter sa représentation ultérieure dans un langage complètement formel et opérationnel. Le modèle obtenu est souvent qualifié de semi-formel (car certaines connaissances ne peuvent pas être totalement formalisées). [8]
- **Opérationnalisation** : On appelle ontologie opérationnelle une ontologie exprimée dans un langage opérationnel et dotée d'une sémantique opérationnelle. [9] L'opérationnalisation de l'ontologie consiste donc en la spécification informatique des opérations applicables aux concepts dans un langage opérationnel (i.e. doté des services permettant de mettre en œuvre des raisonnements). L'utilisation opérationnelle d'une ontologie suppose sa représentation dans un langage formel mais aussi opérationnel, i.e. offrant des mécanismes de raisonnements adaptés aux manipulations de connaissances envisagées. Cette étape devra donc intégrer des outils permettant d'opérationnaliser l'ontologie. . [9]
- **Pourquoi quelqu'un voudrait-il développer une ontologie ?**

Certaines des raisons sont :

- ✓ Analyser la connaissance du domaine,
- ✓ Rendre explicites les hypothèses de domaine,
- ✓ Permettre la réutilisation de la connaissance du domaine,
- ✓ Séparer la connaissance du domaine de la connaissance opérationnelle,
- ✓ Partager une compréhension commune de la structure de l'information entre des personnes ou des agents logiciels.

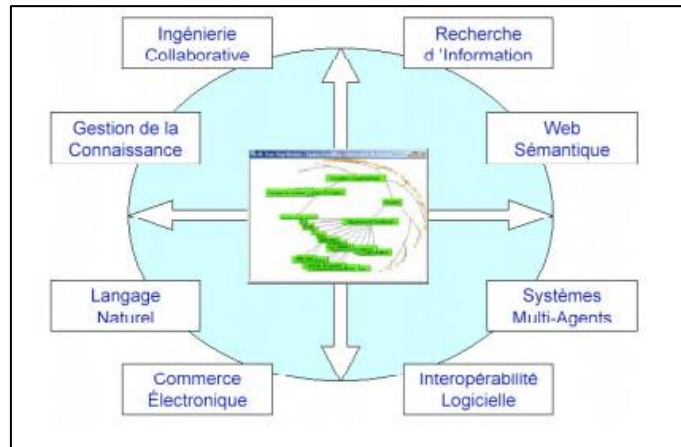


Figure 1.6 Utilisation des ontologies [7]

Les méthodes et les méthodologies recensées permettent la construction d'ontologies à partir de zéro c.-à-d. à partir des données brutes, ou par réutilisation d'autres ontologies, la réingénierie, l'intégration ou fusion avec d'autres ontologies, la construction collaborative ainsi que l'évolution des ontologies construites. [9]

- **Les méthodes et méthodologies pour la construction d'ontologies en partant de zéro**

Dans cette catégorie, la première a été la méthodologie CYC [10], élaborée dans le cadre du développement de l'ontologie CYC. Cette méthodologie propose deux étapes de construction :

- l'extraction manuelle de la connaissance commune implicite dans les différentes sources ;
- l'utilisation des techniques de Traitement Automatique de la Langue Naturelle (TALN) et d'acquisition de connaissances pour générer de nouvelles connaissances à partir de celles acquises à l'étape précédente.

- **Les méthodes pour la réingénierie d'ontologies**

Reconstruire et lier un modèle conceptuel d'une ontologie déjà implémentée à une autre en cours d'implémentation représente le processus de réingénierie d'ontologies. En adaptant une technique de réingénierie de schémas à une ontologie de domaine Gomez-Perez , donne une méthode représentative de ce processus.[11]

- **Les méthodes pour la construction coopérative d'ontologies**

Les ontologies sont souvent utilisées par plusieurs praticiens, elles doivent donc faire l'objet d'un consensus et être acceptée par sa communauté d'utilisateurs. Leur construction est une construction coopérative. Les méthodes de cette catégorie adoptent donc une approche collaborative de construction incluant l'intervention de personnes localisées à des endroits différents. Un certain nombre de problèmes que pose la construction d'ontologies dans un contexte distribué : la gestion de l'interaction et la communication entre les différentes

personnes ; le contrôle de l'accès aux données ; la reconnaissance de l'attribution des droits d'auteurs ; la détection et la correction d'erreurs ont été identifiés par Euzenat [12].

Comme exemple de ces méthodes citons celle de Farquhar et ses collègues, utilisée pour la construction d'ontologies pour l'intégration de données et la méthode élaborée dans le cadre de l'initiative KA, utilisée pour la construction d'ontologies pour l'annotation sémantique de documents. [11]

Le partage d'une compréhension commune de la structure de l'information entre des personnes ou des agents logiciels est l'un des objectifs les plus communs du développement des ontologies (Musen, 1992 ; Gruber, 1993).[4]

Nous allons décrire les plus importants [1].

- Approche de construction d'ontologie de domaine à partir de grandes ontologies (SENSUS, Cyc, AKT,...) : Construire des ontologies de domaine à partir des grandes ontologies déjà existantes.
- Méthode de Uschold et King's : Basé sur l'expérience acquise lors du développement de l'ontologie d'Entreprise, the enterprise ontology, ont proposé une méthodologie générale, inspirée de leur expérience de construction d'ontologies dans le domaine de la gestion des entreprises [1].
- La méthodologie On-To-Knowledge¹ (OTK) : La méthodologie On-To-Knowledge propose de construire une ontologie en tenant compte de comment l'ontologien va être utilisée par l'application plus tard (cette méthodologie est très dépendante de l'application).
- Methontology : Cette méthodologie a été développée au sein du groupe d'ontologie à l'université polytechnique de Madrid. élaborée en 1998 par A. GOMEZ-PEREZ, couvre tout le cycle de vie d'une ontologie. [1]

- **Les autres méthodes de développement d'ontologies**

Ils existent d'autres méthodes de développement d'ontologies utilisées dans d'autres domaines. Certaines sont basées sur des études comparatives des méthodes de développement d'ontologie à partir de zéro. Dans ce groupe, nous pouvons citer entre autres:

- **La méthode Cyc dans le domaine de la microélectronique ;**
- **Les méthodes d'Uschold et King et de Gruninger**, dont toutes les trois cherchent à modéliser le domaine de l'entreprise. Elles sont moins détaillées et n'ont été testées que dans le domaine des affaires (business) ; [9]

- **La méthode Kactus** , proposée dans le domaine de l'intelligence artificielle ; [9]
- **La méthode Sensus**, proposée par l'équipe de Swartout , est utilisée pour construire le squelette d'une ontologie de domaine à partir d'une grande ontologie. [9]

D'autres méthodes, qui ne s'occupent que d'une activité précise du processus de développement existent. On peut les classer dans les catégories suivantes:

- **Les méthodes d'étude des ontologies ou «ontologies learning».** Cette catégorie de méthodes s'intéresse à l'activité de l'acquisition de connaissance. L'acquisition de la connaissance pour la construction des ontologies exige toujours beaucoup de temps et de ressources, malgré que l'un des principaux objectifs des ontologies est de focaliser et réduire le niveau de l'acquisition de la connaissance. Ces méthodes sont utilisées pour créer une nouvelle ontologie à partir de zéro, enrichir une ontologie existante avec de nouveaux termes, et acquérir la connaissance pour des tâches.
- **Les méthodes d'alignement et de fusion des ontologies.** Les ontologies ont pour but de capturer la connaissance consensuelle d'un domaine donné d'une manière générique et formelle pour être réutilisée et partagée à travers des applications et par des groupes de personnes. De cette définition nous pourrions incorrectement déduire qu'il y a seulement une ontologie pour un domaine. En fait, nous pouvons trouver dans la littérature plusieurs ontologies qui modélisent, de différentes manières, le même type de connaissance ou de domaine. Noy et Musen, distinguent deux approches pour unifier les terminologies des ontologies : alignement d'ontologie et fusionnement d'ontologie. Les méthodes d'alignement d'ontologie établissent différents types de correspondances entre les ontologies, par conséquent cette option préserve les ontologies originales. [9]

Plusieurs méthodes d'alignement d'ontologie ont été proposées, comme celle utilisée par Anchor PROMPT [9]. Cependant, les méthodes de fusion d'ontologies proposent de produire une ontologie unique à partir des ontologies originales. Le processus de fusionnement exige habituellement d'établir des correspondances entre les ontologies afin de les fusionner. Etant donné l'état actuel des choses et dans le contexte du Web Sémantique, il est plus approprié d'établir des correspondances ontologiques entre les ontologies existantes sur le même sujet que de prétendre construire un modèle unifié de la connaissance pour un tel sujet à partir de zéro. Dans cet ordre d'idées, nous pouvons citer la méthodologie OIGNONS [13], les méthodes FCA-Merge et PROMPT pour fusionner des ontologies. [9]

- **Les méthodes d'évaluation des ontologies** sont des ontologies qui ne peuvent être réutilisées par d'autres ontologies ou être utilisées par des applications sans une évaluation préalable de leur contenu d'un point de vue technique. Comme directives pour évaluer des ontologies, les premiers travaux ont été proposés par l'équipe de Gomez-Perez ; et ont

continué sur l'évaluation de la connaissance taxonomique. L'équipe de Guarino a développé OntoClean [9], qui est une méthode pour analyser et nettoyer la taxonomie d'une ontologie existante au moyen d'un ensemble de principes basés sur la philosophie. [9]

1.7 Les Formalismes pour la représentation

Les ontologies sont apparues en informatique, plus précisément en Ingénierie des connaissances, dans le cadre des démarches d'acquisition des connaissances pour les systèmes à base de connaissances (SBC). Les SBC proposaient alors de spécifier, d'un côté des connaissances du domaine modélisé, et de l'autre, des connaissances de raisonnement qui manipule et utilise ces connaissances du domaine. L'idée de cette séparation modulaire était de construire mieux et plus rapidement des SBC en réutilisant le plus possible des composants génériques, que ce soit au niveau du raisonnement ou des connaissances du domaine. [1]

Représenter des connaissances propres à un domaine particulier consiste à décrire et à coder les entités de ce domaine de manière à ce qu'une machine puisse les manipuler afin de raisonner. Comme alternative à la logique classique, l'IA a proposé divers formalismes de représentation d'ontologies.

❖ Les frames

Le formalisme frames est introduit par M. Minsky. [1]

Dans ce formalisme, la structure de données enregistrées représente une situation et un objet. L'idée est de collecter toutes les informations nécessaires concernant une situation et de les mettre dans une place appelée frame. Certains auteurs, par exemple Hayes, ont critiqué l'absence d'une sémantique formelle dans ce formalisme et ont montré qu'une fois convenablement formalisé, ce formalisme est une variante syntaxique de la logique de prédicat du premier ordre L.P.P.O.

Le modèle des Frames a été initialement proposé comme langage de représentation d'ontologies par T. GRUBER. Le principe de ce modèle est de décomposer les connaissances en classes (frames) qui représentent les concepts du domaine. À un frame est rattaché un certain nombre d'attributs (slots), chaque attribut peut prendre ses valeurs parmi un ensemble de facettes (facets). Une autre façon de présenter ces attributs est de les considérer comme des relations binaires entre classes dont le premier argument est appelé domaine (Domain) et le deuxième est appelé portée (range). [1]

❖ Les graphes conceptuels

Le modèle des Graphes Conceptuels (GC), introduit par J. SOWA [1] au début des années 80, est un modèle opérationnel de représentation de connaissances, qui appartient à la famille des réseaux sémantiques.

Le formalisme réseaux sémantique est développé par M. Quillian pour représenter la sémantique du langage naturel . [1].

Ce formalisme a une représentation de graphe dont les noeuds représentent des concepts et des individus. Ces noeuds sont connectés par des arcs étiquetés. Il y a deux sortes d'arcs : les arcs de propriété qui affectent les propriétés à des concepts ou a des individus et les arcs IS-A qui introduisent les relations hiérarchiques entre des concepts ou entre des individus . [1]

Le modèle des GCs est mathématiquement fondé sur la logique et la théorie des graphes. Cependant, pour raisonner à l'aide de GC, deux approches peuvent être distinguées :

- (1) raisonner à l'aide de la logique en considérant les GCs comme une interface graphique,
- (2) considérer les GCs comme un modèle de représentation à part entière disposant de ses propres mécanismes de raisonnement fondés sur la théorie des graphes.

Le modèle des GCs se décompose en deux parties [1] :

Une partie terminologique dédiée au vocabulaire conceptuel des connaissances à représenter, c'est-à-dire les types de concepts, les types de relations et les instances des types de concepts.

Cette partie correspond à la représentation du modèle conceptuel mais intègre également des connaissances sur la hiérarchisation des types de concepts et de relations.

Une partie assertionnelle dédiée à la représentation des assertions du domaine de connaissance étudié.

❖ Les logiques de description

L'objectif principal des LDs consiste à pouvoir raisonner efficacement pour minimiser les temps de réponse. Par conséquent, la communauté scientifique a publié de nombreuses recherches qui portent sur l'étude du rapport expressivité/performance des différentes LDs [1].

Le développement des LDs fut fortement influencé par les travaux sur la logique des prédicats, les schémas (frames) [1] et les réseaux sémantiques. Des correspondances existent entre les LDs et ces formalismes [1]. Les premiers travaux sur les LD commencèrent au début des années 1980 avec des systèmes à base de connaissances tels que KL-ONE, BACK et LOOM [1].

Dans les années 1990, une nouvelle classe d'algorithmes est apparue : les algorithmes de vérification de satisfiabilité à base de tableaux. Ces derniers raisonnent sur des LD dites expressives ou très expressives, mais en temps exponentiel. Cependant, en pratique, le comportement des algorithmes est souvent acceptable [1]. L'expressivité accrue a ouvert la porte à de nouvelles applications telles que le Web sémantique [1]. Le terme logiques de description expressive (LDE) désigne l'ensemble des LDs qui ont émergé pendant cette période.

Les deux niveaux de description

La modélisation des connaissances d'un domaine avec les LDs se réalise en deux niveaux. Le premier, le niveau terminologique ou TBox, décrit les connaissances générales d'un domaine alors que le second, le niveau assertionnel ou ABox, représente une instanciation spécifique.

Une TBox comprend la définition des concepts et des rôles, alors qu'une ABox décrit les individus en les nommant et en spécifiant en termes de concepts et de rôles, des assertions qui portent sur ces individus nommés.

✓ **Le niveau terminologique (TBox)**

Les concepts atomiques et rôles atomiques constituent les entités élémentaires d'une TBox.

Les noms débutant par une lettre majuscule désignent les concepts, alors que ceux débutant par une lettre minuscule dénomment les rôles (par exemple : les concepts Femelle, Mâle, Homme et Femme, et le rôle relationParentEnfant).

✓ **Le niveau assertionnel (ABox)**

Une ABox contient un ensemble d'assertions sur les individus. Chaque ABox doit être associée à une TBox, car les assertions s'expriment en terme de concepts et de rôles de la TBox. Une ABox désigne des individus caractérisés par des assertions d'individus nommés.

Une assertion de rôle, de la forme $R(a, b)$ indique que pour cette ABox qu'il existe un individu nommé a qui est en relation avec un individu nommé b par le rôle R (défini dans la TBox associée).

L'inférence

L'inférence s'effectue au niveau terminologique ou assertionnel (factuel) :

- ✓ L'inférence au niveau terminologique comprend quatre principaux problèmes [1] : la satisfiabilité, la Subsumption, l'équivalence, et la disjonction.
- ✓ L'inférence au niveau assertionnel comprend quatre principaux problèmes aussi [1] : la Cohérence, la vérification d'instance, La vérification de rôle, et le problème de récupération.

1.8 Les langages de représentation

Les langages de représentation d'ontologies qui existent aujourd'hui sont divers et variés. Cet état de choses suscite beaucoup d'interrogations lorsque vient le moment d'en choisir un pour construire une nouvelle ontologie...

Plus le langage est riche plus le support de raisonnement devient inefficace et un langage supporté par des raisonneurs assez efficaces aussi un langage capable d'exprimer de vastes classes d'ontologies et de connaissances.

Dans cette section nous parlerons des langues les plus utilisées comme : XML ; RDF ; OWL...

❖ RDF

RDF [1] est un langage pour la représentation de méta-données à propos des ressources. Le modèle RDF permet cette représentation par des assertions sous la forme d'un triplet (ressource, propriété, valeur), ou encore (sujet, prédicat, objet) :

- **Ressources** : les ressources sont tous les objets décrits par RDF. Généralement, ces ressources peuvent être aussi bien des pages Web que tout objet ou personne du monde réel. Les ressources sont alors identifiées par leur URI (Uniform Resource Identifier).
- **Propriétés** : une propriété est un attribut, un aspect, une caractéristique qui s'applique à une ressource. Il peut également s'agir d'une mise en relation avec une autre ressource.
- **Valeurs** : les valeurs en question sont les valeurs particulières que prend la propriété. La valeur pouvant être une autre ressource ou bien un littéral.

• Objectif de RDF

Le développement de RDF par le W3C a été entre autres motivé par la perspective des applications suivantes [6] :

- manipulation et classification des métadonnées Web, afin de fournir des informations sur les ressources Web et les systèmes qui les utilisent.
- développement de modèles d'information ouverts plutôt que fixés pour certaines applications (par exemple les activités de planification, de description de processus organisationnels, d'annotation de ressources web, etc.)
- optimisation de la coopération entre applications, en permettant de combiner les données de plusieurs applications, pour générer de nouvelles informations.

❖ RDF(S)

RDFS [1] est un langage permettant de définir des schémas de méta-données. Il définit le sens, les caractéristiques et les relations d'un ensemble de propriétés. La principale nouvelle notion est la distinction entre une classe (concept d'une ontologie) et une instance (individu d'une ontologie). Quelques notions définies sont : (rdfs : Class), (rdfs : subclassOf), (rdfs : domain), et (rdfs : range).[1]

❖ DAML-OIL

DAML (The DARPA Agent Markup Language. <http://www.daml.org/>.) est un langage qui a comme but de fournir les fondations pour la génération suivante du Web sémantique. Comme RDFS, ce langage n'est pas assez expressif relativement aux exigences du Web sémantique, un nouveau langage nommé DAML-ONT a été développé en tant qu'extension de RDF avec les capacités d'un langage de représentation du savoir.

En même temps, nouveau langage nommé OIL a été développé par un groupe des chercheurs, pour le même but. Ce langage a une syntaxe basée sur RDF et il est explicitement construit pour que sa sémantique puisse être spécifiée à travers une description logique très expressive, la logique de description de type SHIQ.

DAML+OIL est la combinaison de ces deux langages. Il hérite des avantages de ces deux langages. En conséquence, DAML+OIL est un langage très expressif et lisible par la machine ainsi que par un être humain avec une syntaxe basée sur RDF.

❖ OWL

OWL [W3C04c] est un langage fondé sur la syntaxe RDF/XML et héritier des travaux de DAML+OIL. OWL introduit l'aspect sémantique qui manque RDF, et offre, par ses primitives plus riches, au machine une capacité d'interprétation plus grande que celle de RDF et RDFS.

OWL est, tout comme RDF, un langage XML profitant de l'universalité syntaxique de XML. Fondé sur la syntaxe de RDF/XML, OWL offre un moyen d'écrire des ontologies web. OWL se différencie du couple RDF/RDFS en ceci que, contrairement à RDF, il est justement un langage d'ontologies. SI RDF et RDFS apportent à l'utilisateur la capacité de décrire des classes (avec des constructeurs) et des propriétés, OWL intègre, en plus, des outils de comparaison des propriétés et des classes : identité, équivalence, contraire, cardinalité, symétrie, transitivité, disjonction, etc.

Ainsi, OWL offre aux machines une plus grande capacité d'interprétation du contenu web que RDF et RDFS, grâce à un vocabulaire plus large et à une vraie sémantique formelle. Plus un outil est complet, plus il est, en général, complexe. C'est cet écueil qu'a voulu éviter le groupe de travail WebOnt du W3C en dotant OWL de trois sous-langages OWL Lite, OWL DL et OWL Full, qui offrent des capacités d'expression croissantes et, naturellement, destinés à des communautés différentes d'utilisateurs, chacun est une extension par rapport à son prédécesseur plus simple.

- **OWL Lite** : est le sous langage le plus simple, il répond à des besoins de hiérarchie de classification et de fonctionnalités de contraintes simples.
- **OWL DL** : plus complexe que OWL Lite, permettant une expressivité bien plus importante. Il est fondé sur la logique de description SHIN(D). Garantit la complétude

des raisonnements (calculabilité des inférences) et leur décidabilité (leur calcul se fait en une durée finie).

- **OWL Full** : est la version la plus complexe d'OWL mais également celle qui permet le plus haut niveau d'expressivité. OWL Full est destiné aux situations où il est plus important d'avoir un haut niveau de capacité de description. Il a l'avantage de la compatibilité complète avec RDF/RDFS, mais l'inconvénient est qu'il ne garantit pas la complétude et la décidabilité des calculs liés à l'ontologie. OWL Full offre cependant des mécanismes intéressants, comme par exemple la possibilité d'étendre le vocabulaire par défaut de OWL. Il existe entre ces trois sous langages une dépendance de nature hiérarchique : toute ontologie OWL Lite valide est également une ontologie OWL DL valide, et toute ontologie OWL DL valide est également une ontologie OWL Full valide. [6]

❖ RDQL

RDQL (RDF Data Query Language) [W3C04d] est un langage d'interrogation de données définies en RDF. Ce langage n'est pas standardisé, il existe de nombreuses implémentations.

Sa syntaxe est très proche de SQL (Structured Query Language). Les clauses de RDQL sont :

- **SELECT** définit la liste des variables que l'on désire obtenir. Une variable est composée de caractères alphanumériques et commence par un '?'.
▪ **FROM** définit l'emplacement des documents RDF utilisés pour la requête.
▪ **WHERE** définit le triplet RDF (sujet - prédicat - objet), les éléments de ce triplet sont décrits soit par les valeurs de l'ontologie interrogée soit par des variables.
▪ **AND** définit les restrictions booléennes de la requête. Une restriction booléenne est constituée de valeurs ou variables composés à l'aide d'opérateurs.

❖ SPARQL

SPARQL [W3C06] est une amélioration de RDQL, ce langage est en cours de standardisation au niveau du W3C. SPARQL offre une syntaxe quasi identique à RDQL, mais y ajoute notamment les opérateurs UNION et OPTIONAL dans la clause WHERE. L'opérateur UNION définit la disjonction de triplets RDF.

❖ nRQL

nRQL (new Racerpro query Language) [Haarslev01] est le langage d'interrogation de Racer. Comme RDQL et SPARQL, nRQL est basé sur la recherche de graphes RDF. Sa syntaxe est proche des deux autres, sauf pour sa notation préfixée des opérateurs.

1.9 Outils de développements d'Ontologies

Les outils de développement d'ontologies qui existent dans le marché aujourd'hui sont divers et variés. La chose qui ne mène à poser plusieurs questions: L'outil offre-t-il une assistance au développement ? L'outil dispose-t-il d'un moteur d'inférence ? Quels langages d'ontologies l'outil supporte-t-il ? L'outil permet-il d'importer/exporter des ontologies ?

L'outil offre t-il un support à la réutilisation d'ontologies existantes ? L'outil permet-il de documenter les ontologies construites ? L'outil offre-t-il un support graphique à la construction des ontologies ? L'outil est-il stable, convivial, « mature » ?

Les réponses à toutes ces questions pourraient s'avérer décisives dans le choix de l'un ou l'autre outil. Dans cette section nous présentons les principaux outils disponibles.[1]

- ❖ **Langages de spécification d'ontologies :** Ces dernières années, le domaine de l'ontologie a évolué, notamment en ce qui concerne le développement de nombreux langages de l'ontologie dans le contexte du Web sémantique. Certains d'entre eux sont basés sur la syntaxe de XML, tels que XOL, SHOE, OML, RDF, et RDF Schéma. Les deux derniers sont créés par des groupes de travail du W3C. Trois autres langages sont établis sur RDF(S) pour améliorer ses caractéristiques : OIL, DAML+OIL et OWL.

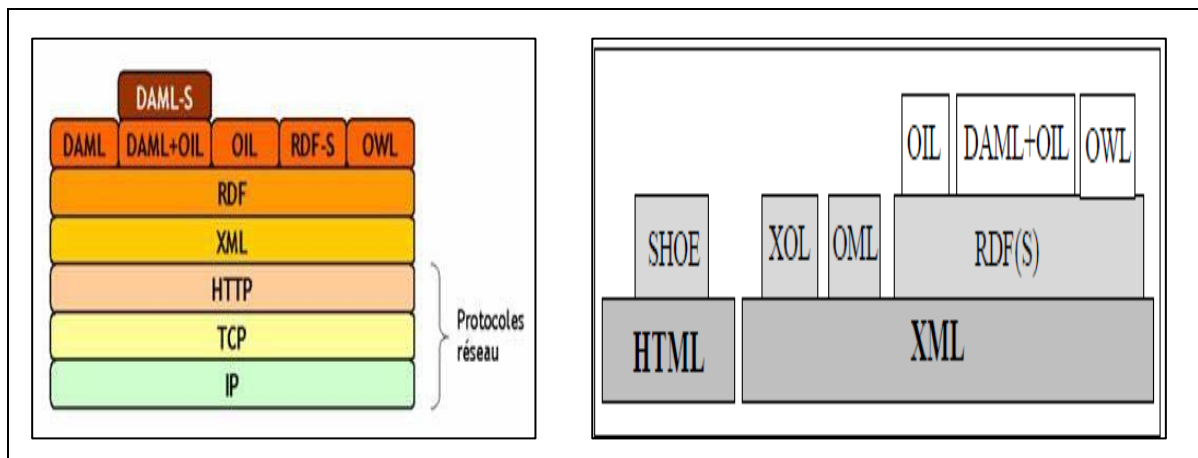


Figure 1.7 La pyramide des langages du Web sémantique.

- ❖ **Les moteurs d'inférences :** La plupart de ces moteurs acceptent en entrée des fichiers OWL et sont conçus pour raisonner sur les logiques de descriptions. Une fois l'ontologie chargée, ces moteurs effectuent les inférences sur la TBox et la ABox. Le tableau ci-dessous dresse une comparaison des principaux moteurs d'inférence pour les logiques de description expressives : FaCT[1], Racer [1], Pellet[1], FaCT++[1], F-OWL, Surnia, et Hoolet.

Moteur	Logique de Description	Implantation	Inférence	API Java	OWL	Décidabilité
Racer	SHIQ	C++	TBox/ABox	oui	OWL DL	Oui
Pellet	SHIN(D), SHON	Java	TBox/ABox	natif	OWL DL	Oui
FaCT	SHIQ, SHF	Common Lisp	TBox	oui	OWL DL	Oui
FaCT++	SHIF	C++	TBox	oui	OWL -Lite	Oui
Surnia	Logique prédicats	Python	TBox/ABox	non	OWL Full	Non
Hoolet	Logique prédicats	Java	TBox/ABox	oui	OWL DL	Non
F-OWL	SHIQ	Java	TBox/ABox	oui	OWL Full	Non

Tableau 1.1 Les principaux moteurs d'inférence [1]

Racer par exemple est le moteur d'inférence le plus connu et l'un des plus utilisés dans le domaine pour ces performances et sa stabilité. Il est commercialisé par Racer Systems GmbH & Co. KG, fondé en 2004 par Volker Haarslev, Kay Hidde, Ralf Möller et Michael Wessel qui travaillaient à l'université de Hambourg. Racer travaille sur les ontologies modélisées par son langage, mais il accepte des ontologies décrites en RDF ou OWL, ces dernières étant traduites vers le langage utilisé par Racer. Ce moteur d'inférence possède également son propre langage de requête nRQL (new Racerpro query Language) pour interroger l'ontologie sur la TBox et ABox.

- Racer possède quelques avantages :
 - ✓ Racer permet l'utilisation de règles SWRL.
 - ✓ Racer permet l'ajout d'assertions et d'individus dans les ABox après le chargement de l'ontologie.
- Les inconvénients de Racer sont :
 - ✓ Racer suppose que tous les propriétés sur les data types sont fonctionnelles (pas de valeurs multiples pour un data type property).
 - ✓ Racer ne permet pas l'utilisation de type de données utilisateur (type défini par l'utilisateur), et il n'existe pas de version libre d'utilisation. Cependant il est possible d'obtenir une licence gratuite dans le cadre de la recherche scientifique.
- ❖ **Langages d'interrogation d'ontologies** : Dans cette section, nous présentons les trois langages d'interrogation basés sur la reconnaissance de graphe RDF : RDQL, SPARQL et nRQL. Le principe est de décrire un graphe RDF à l'aide de variables. Les résultats d'une

requête étant les valeurs des variables pour lesquelles il existe un graphe RDF dans l'ontologie correspondant au graphe défini par la requête. RDQL et SPARQL sont les langages d'interrogation utilisables sur Pellet et nRQL est le langage d'interrogation de Racer.

❖ **Editeur d'ontologies** : L'ontologie connaît un grand développement et une augmentation du nombre d'éditeurs et de chercheurs dans ce domaine. Les éditeurs varient en fonction de l'utilisation de différentes formalités et de leur présentation avec différentes caractéristiques. Parmi ces outils on trouve : OIEd, OntoEdit, , Web ode, DOE, Protégé2000... voici la description de quelques-uns : [1]

1. **OIEd** : a été conçu pour éditer des ontologies dans le langage de représentation OIL, il est souvent considéré comme une simple interface de la logique de description SHIQ. Cet éditeur offre également les services d'un raisonneur, FaCT qui permet de tester la satisfiabilité des définitions de classes et de découvrir des subsumptions restées implicites dans l'ontologie. L'outil dispose de mécanismes pour la classification et le contrôle de la cohérence des ontologies. La version 3.4 d'OIEd est gratuite et disponible sur le site web d'OIEd.
2. **OntoEdit** : est également un environnement de construction d'ontologies indépendant de tout formalisme. Des outils graphiques dédiés à la visualisation d'ontologies sont inclus dans l'environnement. ONTOEDIT intègre un serveur destiné à l'édition d'une ontologie par plusieurs utilisateurs. Un contrôle de la cohérence de l'ontologie est assuré à travers la gestion des ordres d'édition.
3. **Ontolingua** : de l'Université Stanford. Le serveur Ontolingua est le plus connu des environnements de construction d'ontologies en langage Ontolingua. Il consiste en un ensemble d'environnements et de services qui supportent la construction en coopération d'ontologies, entre des groupes séparés géographiquement. Il supporte plusieurs langages et dispose de traducteurs permettant de passer de l'un à l'autre.
4. **ONTOSAURUS** : de l'Information Science Institute de l'Université de Southern California. Ontosaurus consiste en un serveur utilisant LOOM comme langage de représentation des connaissances, et en un serveur de navigation créant dynamiquement des pages HTML qui affichent la hiérarchie de l'ontologie; le serveur utilise des formulaires HTML pour permettre à l'utilisateur d'éditer l'ontologie. Des traducteurs du LOOM en Ontolingua, KIF, KRSS et C++, ont été développés.
5. **Protégé2000** : est une interface modulaire, développée au Stanford Medical Informatics de l'Université de Stanford⁷, permettant l'édition, la visualisation, le contrôle (vérification des contraintes) d'ontologies, l'extraction d'ontologies à partir de sources textuelles, et la fusion semi-automatique d'ontologies. Le modèle de connaissances de

Protégé2000 est issu du modèle des frames et contient des classes (concepts), des slots (propriétés) et des facettes (valeurs des propriétés et contraintes), ainsi que des instances des classes et des propriétés. De nombreux plug-ins sont disponibles ou peuvent être ajoutés par l'utilisateur.

1.10 Conclusion

Dans ce chapitre, nous avons essayé d'éclaircir la notion d'ontologie. Nous avons commencé par ces origines philosophiques en arrivant à définir son sens en ingénierie des connaissances et son cycle de vie.

Ensuite, nous avons étudié la manière de concevoir et de réaliser une ontologie en énumérant ses composants et en présentant les méthodes et les outils pour leur construction.

Chapitre 2
L'ontologie médicale

L'ONTOLOGIE MEDICALE

La santé est un domaine très particulier qui rejoint tout le monde si nouveau nés jusqu'à la personne mourante.

C'est un domaine qui a beaucoup d'innovation et beaucoup de recherches mais qui aussi beaucoup de problèmes, C'est pourquoi les spécialistes Ils essaient se trouvent des solutions en intégrant des systèmes d'information dans le domaine médical afin de réduire ou d'éliminer certains de ces problèmes.

Depuis de nombreuses années, la médecine a produit de nombreuses terminologies pour des applications diverses de tous types classification, thésaurus et plus récemment des ontologies. (L'ontologie médicale est l'étude de ce qui en médecine et du processus de leur formation. Elle s'intéresse à la genèse des entités médicales : les maladies, les signes cliniques, les syndromes cliniques, les symptômes, les lésions, les syndromes lésionnels, les anomalies biologiques et les anomalies radiologiques.)

Le web sémantique réclame par définition de construire des ontologies permettant par leur caractère formel, d'automatiser un certain nombre de tâches liées principalement à la recherche d'information et à la classification, nous montrons que la médecine a des spécificités qui contraignent l'usage des ontologies et leur construction. [14]

Ces spécificités amènent à prendre en compte le caractère linguistique des thésaurus médicaux et à les utiliser en coordination avec les ontologies pour exploiter les informations médicales. En complètement, la construction des ontologies peut prendre en compte l'expression l'linguistique des connaissances en utilisant les textes générés par les professionnels de santé durant leur activité. [14]

2.1 L'ontologie Médicale

2.1.1 Terminologies

Les ontologies terminologiques sont définies comme étant des ontologies dont les catégories ne sont pas spécifiées par des axiomes et des définitions. Un exemple est Word Net [49] dont les catégories sont partiellement spécifiées par des relations comme l'hyponymie, qui détermine les positions relatives des concepts, sans pour autant la définir de façon complète. La plupart des organisations établissent des conventions de nommage, des classifications et des standards. Ces éléments sont un début de formalisation. [16]

2.1.2 Ontologies Vs Terminologies

Une distinction entre terminologie et ontologie est apportée par C. Roche qui stipule que « ...parler de terminologie c'est parler des mots et du sens des mots... L'étude du sens d'un mot relève de la linguistique... ». Dans une terminologie on s'intéresse aux mots et aux relations entre eux ; la relation structurante de base est la relation d'hyponymie et son inverse l'hyponymie, tandis que dans une ontologie, on s'intéresse à la notion de concept et aux relations entre eux. [16]

Cependant, ontologies et terminologies servent au même objectif ; fournir à une communauté d'utilisateurs une conceptualisation partagée d'une partie spécifique du monde, dans le but de faciliter une communication efficace d'une connaissance complexe. Gamper fournit quatre critères pour distinguer les deux notions [16] :

- Le cadre formel de leur définition : le sens des termes est défini dans un texte libre en langue naturelle. L'interprétation correcte du sens véhiculé dépend de l'utilisateur. Les ontologies spécifient explicitement la connaissance conceptuelle en utilisant un langage formel avec une sémantique claire, qui permet une interprétation non ambiguë des termes ;
- Le support computationnel : La plupart des terminologies utilisées actuellement fournissent très peu de sémantique pour une représentation explicite aussi bien de la connaissance que pour la maintenance des données. Pour les ontologies, quant à elles et grâce à l'utilisation de langages de représentation formels comme les Logiques de Description (DLs), il est possible de vérifier leur consistance, d'inférer de nouvelles connaissances, etc ;
- les utilisateurs : les terminologies sont utilisées pour véhiculer l'information entre les humains, tandis que les ontologies sont principalement développées pour le partage de connaissances entre l'être et la machine.
- l'usage de la langue naturelle : la terminologie s'occupe du transfert de la connaissance comme une activité linguistique, c'est-à-dire l'exploration de la langue naturelle afin d'identifier tous les termes utilisés par les humains pour faire référence aux concepts sous-jacents.

La question de savoir si une ontologie donnée doit être représentée de façon formelle ou informelle est directement liée à l'objectif de cette ontologie .Si le but est d'assurer la communication entre des personnes, une description en langue naturelle peut suffire. Par contre, si elle doit assurer le rôle d'un format d'échange dans un environnement d'interopérabilité entre programmes, il est nécessaire qu'elle soit formelle afin que la communication soit cohérente et que la vérification automatique de cette cohérence soit possible. [16]

2.1.3 Construction et édition d'ontologies

La construction d'une ontologie n'est pas une activité aisée, d'autant plus qu'il n'existe pas aujourd'hui une méthodologie communément admise, comme c'est le cas dans le domaine des bases de données par exemple, avec la démarche Entité/Association. Le processus de construction intègre souvent un expert du domaine. Une fois construite, l'ontologie doit être claire, cohérente, compréhensible, facile à utiliser et extensible [17].

Nous pouvons classer les travaux sur la construction des ontologies comme suit [16] :

- les méthodes et méthodologies pour la construction d'ontologies en partant de zéro ;
- les méthodes pour la réingénierie d'ontologies ;
- les méthodes de construction coopérative d'ontologies.

Les spécialistes ont fourni une description détaillée des méthodes et des outils permettant d'établir l'ontologie. Korsho, Ali et Murshed ainsi que leurs collègues[18],[19].Corcho et ses collègues ont fait une revue des principaux outils et langages ainsi que des méthodologies pour construire des ontologies, tandis que Murshed et Ali ont proposé des critères d'évaluation et des techniques de classement permettant aux développeurs d'ontologies de choisir le bon outil de construction.[16]

2.1.3.1 Méthodes et méthodologie de construction d'ontologie

Les ontologies sont devenues un moyen indispensable pour représenter et exploiter les données et les connaissances d'un domaine, et plus particulièrement celles du domaine médical. Plusieurs systèmes terminologiques ont été proposés et développés : des vocabulaires contrôlés pour annoter des gènes et classer les documents, et des thesaurus pour guider et faciliter la recherche d'informations. Néanmoins, le succès de ces systèmes est limité en raison de leur dépendance à des cas et des tâches spécifiques et de l'absence de possibilités de raisonnement. [15]

Une ontologie est toujours liée à une méthodologie de construction, à un outil de construction et avec un langage de représentation d'ontologie.

Ce processus de développement est un processus complexe Pour cela, il est nécessaire d'utiliser des méthodes ou méthodologies pour seconder le processus de construction des ontologies.

2.1.3.2 La méthodologie METHONTOLOGY

Cette méthodologie a été développée par le groupe d'Ontologie à l'Université Polytechnique de Madrid. METHONTOLOGY prend ses racines dans les activités principales identifiées par le processus de développement logiciel et dans les méthodologies d'ingénierie de connaissance.

Cette méthodologie inclue : l'identification du processus de développement, un cycle de vie basé sur des prototypes évolutifs, et les techniques pour effectuer chaque tâche dans les activités de gestion, de développement et de support.

METHONTOLOGY a été proposée pour la construction d'ontologie par la FIPA (Foundation for Intelligent Physical Agents), qui favorise l'interopérabilité à travers les applications [15].

2.1.3.3 La méthodologie On-To-Knowledge

On-To-Knowledge recommande un procédé itératif de développement, et comporte quatre phases principales : une phase de spécification de condition, une phase d'amélioration, une phase d'évaluation et une phase d'application et d'évolution. On-To-Knowledge propose l'acquisition des connaissances en spécialisant une ontologie générique.

METHONTOLOGY inclut une méthode de réingénierie pour résoudre certains des problèmes liés à la construction d'une ontologie par la réutilisation d'une autre ontologie.

Elle propose de construire l'ontologie en tenant compte de la manière dont elle sera utilisée dans d'autres applications. Par conséquent, les ontologies développées avec cette méthodologie sont fortement dépendantes de l'application [21].

2.1.3.4 La méthode Ontology Development 101

Ontology Development 101 a été développée à l'Université de Stanford, elle cherche à construire des ontologies formelles par la reprise et l'adaptation des ontologies déjà existantes, et propose de suivre les démarches ci-après [22] :

- Déterminer le domaine et la portée de l'ontologie ;
- Considérer la réutilisation des ontologies existantes ;
- Énumérer les termes les plus importants dans l'ontologie ;
- Définir les classes et hiérarchie des classes ;
- Définir les propriétés des classes ;
- Définir les facettes des attributs ;
- Construire les instances.

2.1.3.5 ARCHONTE

B. Bachimont s'est basé sur la sémantique différentielle pour proposer la méthodologie ARCHONTE (ARCHitecture for ONTological Elaborating).

Selon cette méthodologie, la construction d'une ontologie passe par trois étapes principales :

1. choisir les termes pertinents du domaine et normaliser leurs sens puis justifier la place de chaque concept dans la hiérarchie ontologique en précisant les relations de similarités et de différences que chaque concept entretient avec ses concepts frères et son concept père.
2. formaliser les connaissances, ce qui implique par exemple d'ajouter des propriétés à des concepts, des axiomes, de contraindre les domaines d'une relation, etc.
3. représenter l'ontologie dans un langage formel de représentation des connaissances.

Quel que soit la méthodologie adoptée, le processus de construction d'une ontologie est une collaboration qui réunit des experts du domaine de connaissance, des ingénieurs de la connaissance, voire les futurs utilisateurs de l'ontologie [15].

2.1.4 Ressources terminologiques et ontologiques en médecine

Les ressources terminologiques ou ontologiques en médecine ont la spécificité d'être nombreuses, d'abord parce qu'il existe de nombreux thésaurus, ensuite parce que les réflexions et les développements sur les ontologies sont très avancés dans le milieu médical.

Il existe dans le domaine médical un grand nombre de ressources terminologiques et ontologiques (RTO) construites pour répondre à des besoins précis et divers connaissances. Dans cette partie, nous présentons quelques projets de construction d'ontologies. [17]

2.1.4.1 CIM

CIM (Classification internationale des maladies) permet le codage des maladies, des traumatismes et de l'ensemble des motifs de recours aux services de santé. Elle est publiée par l'Organisation Mondiale de la Santé et est utilisée à travers le monde pour enregistrer les causes de morbidité et de mortalité, à des fins diverses, parmi lesquelles le financement et l'organisation des services de santé ont pris ces dernières années une part croissante.

Elle bénéficie d'une remise à niveau régulière, la version la plus récente étant la 10-ième révision (publiée en 1993). Il s'agit d'une classification monoaxiale avec 21 chapitres principaux dont 17 concernent des maladies et 4 concernent les signes et résultats anormaux, les causes de traumatismes, d'empoisonnement ou de morbidité, l'état de santé et les facteurs de recours aux soins. Les catégories de maladies sont définies en fonction d'un caractère commun qui peut être l'étiologie par exemple (1 = maladies infectieuses, lettres A et B), la topographie (9 = maladies de l'appareil circulatoire, lettre I). Le chapitre des maladies infectieuses est le plus gros et le plus détaillé parce que ces maladies sont la première cause mondiale de morbidité et de mortalité [17].

2.1.4.2 MeSH

Le MeSH est un thésaurus médical conçu par la NLM (National Library of Medicine) aux États-Unis. Il est utilisé en particulier par PubMed (l'interface de la NLM à la base de données

bibliographiques MEDLINE qui couvre tous les domaines médicaux) pour l'indexation et la recherche de publications scientifiques.

Il compte 24 767 descripteurs dans sa version de 2008. Les descripteurs MeSH sont organisés en 16 catégories : la catégorie A pour les termes anatomiques, la catégorie B pour les organismes, la catégorie C pour les maladies, etc. Chaque catégorie est subdivisée en sous-catégories.

A l'intérieur de chaque catégorie, les descripteurs sont structurés hiérarchiquement, du plus général au plus spécifique [17].

[A]	Anatomie
[B]	Organisme
[C]	Maladie
[D]	Produits chimiques et pharmaceutiques
[E]	Techniques analytiques, diagnostiques, thérapeutiques et équipements
[F]	Psychiatrie et psychologie
[G]	Phénomènes et processus
[H]	Disciplines et professions
[I]	Anthropologie, enseignement, sociologie, et phénomènes sociaux
[J]	Technologie, industrie et agriculture
[K]	Science humaines
[L]	Sciences de l'information
[M]	Individus
[N]	Santé
[V]	Caractéristiques d'une publication
[Z]	Lieux géographiques

Tableau 2.1 Les différentes catégories ou domaines du MeSH

2.1.4.3 CISMef

l'équipe CISMef du Centre Hospitalier Universitaire de Rouen a initié le projet CISMef (Catalogue et Index des Sites MEDicaux Francophones), Ce catalogue indexe les principaux sites et documents francophones de qualité médicale contrôlée.

En décembre 2007, il a dépassé les 41300 ressources indexées avec une moyenne de 80 nouvelles ressources par semaine.

Cette liste de sites contient un classement thématique, en particulier des spécialités médicales, un classement alphabétique, et un accès par type de ressources. Depuis juin 2000, l'outil associé, Doc'CISMef, permet d'effectuer des recherches dans le catalogue de ressources, et offre des possibilités de recherches plus étendues.

CISMef utilise deux outils standards pour organiser l'information : le thésaurus MeSH (Medical Subject Headings), utilisé notamment par la base de données bibliographique Medline, et

le format de métadonnées du Dublin Core [17].

2.1.4.4 La SNOMED

La SNOMED est une nomenclature pluri-axiale couvrant tous les champs de la médecine et de la dentisterie humaines, ainsi que de la médecine vétérinaire. SNOMEDCT (SNOMED Clinical Terms) représente la dernière version de la nomenclature mais seule la version SNOMED 3.5 (appelée également SNOMED International) a été traduite en français. La SNOMED 3.5 comporte 11 axes (Tableau 2.2)

Axe	Nom de l'axe
A	Agent physiques
C	Produits chimiques
D	Diagnostics
F	Fonctions
G	Qualificatifs
J	Métiers
L	Organismes vivants
M	Morphologie
P	Procédure
S	Contexte Social
T	Topographie

Tableau 2.2 Les onze axes de la SNOMED

Dans chaque axe, les concepts sont représentés par une série de termes au sein de laquelle on peut distinguer une formulation préférée et des synonymes de diverses natures syntaxiques. La version française comporte 97 485 concepts désignés par 144796 termes. Par ailleurs, chaque axe représente une hiérarchie simple de concepts qui peuvent représenter une combinaison de concepts. La recherche en informatique médicale a montré que la SNOMED est la terminologie la plus adaptée à l'indexation des informations du dossier patient. Cependant, elle contient des éléments non pertinents à l'indexation.

Ce sont les éléments de l'axe G contenant les qualificatifs et termes de relations qui n'ont pas de sens lorsqu'ils ne sont pas reliés aux autres termes SNOMED par exemple : "sans", "disponible", etc [17].

2.1.4.5 UMLS

UMLS est un recueil de plusieurs vocabulaires dans le domaine médical, il fournit une structure de cartographie de vocabulaires et permet la traduction entre des différents systèmes de terminologies. UMLS a été mis en place dans le but d'améliorer l'accès à l'information médicale à partir de sources diverses : bases de données bibliographiques, bases de données d'enregistrements cliniques et bases de connaissances médicales. Un des moyens d'UMLS est alors de définir un vocabulaire médical de base, un « métathésaurus » qui reprend et dédouble les termes de

l'ensemble des 95 ressources terminologiques qu'il inclut (MeSH, SNOMED ...). Ce métathésaurus propose une description hiérarchique des connaissances médicales utilisées dans divers documents et systèmes à base de connaissances.

L'intérêt d'UMLS réside dans sa grande couverture du domaine médical (1 276 301 concepts dans la version A du 1er trimestre 2006) et dans sa disponibilité.

Chaque concept UMLS a un identifiant unique, le CUI (Concept Unique Identifier). A chaque concept est associé un ensemble de termes dans différents lexiques. Chaque CUI a dans chaque langue un terme préféré unique appelé SUI (String Unique Identifier).

Chaque SUI est lié à un ou plusieurs termes selon ses différentes variations lexicales, qui sont les LUI (Lexique Unique Identifier). Les SUI dans les différentes langues sont nécessaires pour la communication, mais le vrai identifiant du concept est son CUI [17].

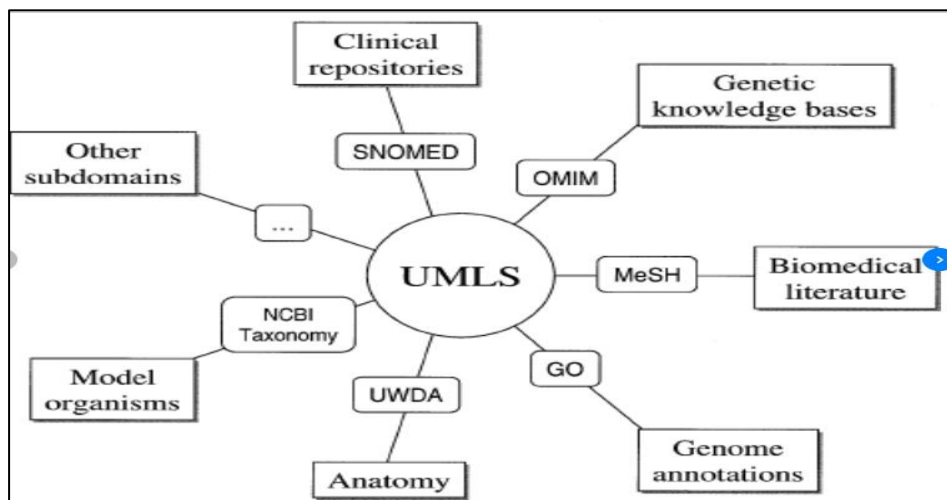


Figure 2.1 Les différents sous-domaines intégrés à l'UMLS.[24]

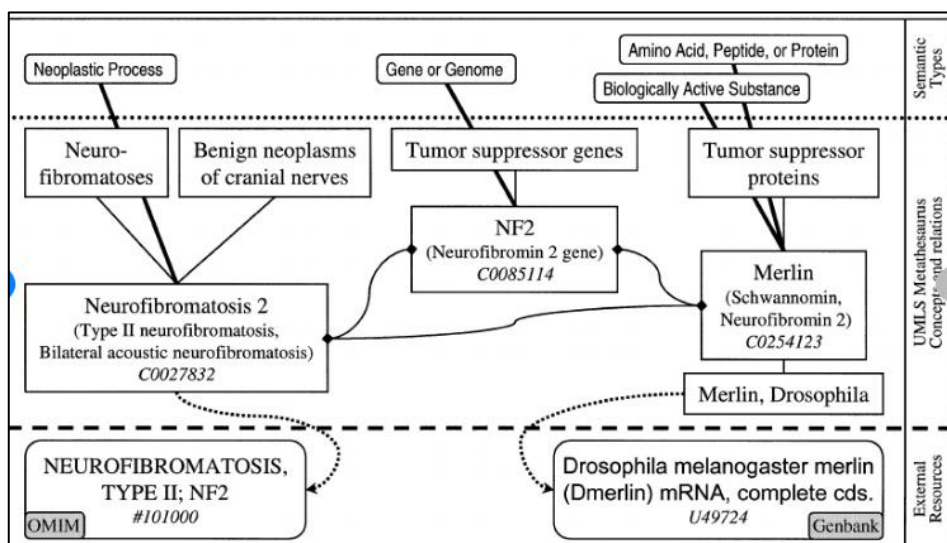


Figure 2.2 : Exemple (NF2 et protéines et maladies associées dans l'UMLS (représentation partielle)) [24]

2.1.4.6 FMA

La FMA (Foundational Model of Anatomy) est une ontologie de référence dans le domaine de l'anatomie. Elle vise à représenter les entités anatomiques et les relations nécessaires pour la modélisation symbolique de la structure phénotypique du corps humain sous une forme qui soit compréhensible par l'homme et qui soit également traitable par une machine.

Les entités anatomiques sont représentées dans FMA, allant des macromolécules biologiques aux cellules, tissus, organes, systèmes d'organes, les majeures parties du corps, y compris le corps entier. Elle contient actuellement autour de 75 000 entités (concepts) anatomiques et plus de 120 000 termes. [17]

2.1.4.7 Ontologie de gènes – GO

L'ontologie de gènes (GO - Gene Ontology) est une ressource terminologique destinée à structurer la description des gènes et des produits géniques dans le cadre d'une ontologie commune à toutes les espèces. Ce projet, qui s'inscrit dans la démarche plus large d'Open Biomédical Ontologies (OBO) regroupant d'autres projets bioinformatiques dans le domaine biomédical, poursuit trois objectifs [17] :

- gérer et enrichir son vocabulaire contrôlé décrivant les gènes et leurs produits,
- gérer les annotations, c'est-à-dire les informations rattachées aux gènes et à leurs produits,
- fournir les outils permettant d'accéder aux informations structurées dans le cadre du projet.

id :	GO :0000016
name :	lactase activity
namespace :	molecular_function
def :	"Catalysis of the reaction : lactose + H2O = D-glucose + D-galactose." [EC :3.2.1.108]
synonym :	"lactase-phlorizin hydrolase activity" BROAD [EC :3.2.1.108]
synonym :	"lactose galactohydrolase activity" EXACT [EC :3.2.1.108]
xref :	EC :3.2.1.108
xref :	MetaCyc :LACTASE-RXN
xref :	Reactome :20536
is_a :	GO :0004553 ! hydrolase activity, hydrolyzing O-glycosyl compounds

Figure 2.3 Description d'un terme dans GO

2.1.4.8 GALEN

GALEN (General Architecture for Language, Encyclopedia and Nomenclature) est un projet européen qui avait pour but de proposer des terminologies réutilisables et partageables pour le domaine médical.

GALEN utilise un formalisme appelé GRAIL (Galen Representation and Integration Language) qui permet de saisir la connaissance terminologique dans le domaine médical.

Ce formalisme est hautement génératif et permet de définir des concepts complexes, composés

de concepts plus élémentaires. Tous les concepts, et les relations qui les lient, sont représentés indépendamment du langage dans lequel ils sont exprimés.

La version actuelle de GALEN comprend une hiérarchie assez riche de concepts (~ 25000 concepts) ainsi qu'un ensemble de relations associatives permettant de définir des structures complexes [17].

2.1.4.9 MENELAS

MENELAS est un projet européen son but était la conception et l'implémentation d'un système de pilotage capable d'accéder à des rapports médicaux rédigés en langage naturel dans trois langues : l'anglais, le français et le néerlandais. Ce système devait pouvoir analyser le contenu de rapports médicaux (comptes rendus d'hospitalisation ou CRH) et l'archiver dans une base de données sous la forme d'un ensemble de structures conceptuelles (graphes conceptuels). Ces structures, qui constituent la représentation de chaque CRH, devaient pouvoir ensuite être consultées pour accéder à des informations spécifiques contenues dans le CRH. Une partie des informations était encodée à l'aide de nomenclatures internationales, ce qui permettait leur échange à partir de CRH écrits en différentes langues. Le projet a été confronté aux problèmes habituels de la compréhension de textes en langage naturel.

Cette ontologie couvrant le domaine des maladies coronariennes comporte plus de 1800 concepts et 300 relations [17].

2.1.5 Enrichissement, maintenance et évolution d'ontologies

Une ontologie dans son cycle de vie est appelée à évoluer afin de prendre en compte les spécificités des nouveaux besoins qui peuvent apparaître. Ce changement peut être de natures très différentes ; concernant la structure conceptuelle, l'ajout de nouveaux concepts, de nouvelles définitions, de nouveaux termes, etc. L'enrichissement d'une ontologie par de nouveaux termes pour désigner les concepts est particulièrement souhaité dans le cadre de l'utilisation d'ontologies pour la recherche d'information [16].

2.2 L'Oncologie

L'oncologie, aussi appelée cancérologie ou carcinologie, est la spécialité médicale chargée de l'étude, du diagnostic et du traitement des cancers, conséquences d'une multiplication anarchique de certaines cellules du corps humain. Elle s'intéresse à tous les types de tumeurs cancéreuses, qu'elles soient solides ou non, localisées ou étendues. Le principal but de l'oncologie est de comprendre au mieux les causes possibles des cancers et leur processus de développement afin de pouvoir les prévenir et les guérir avec plus d'efficacité. [25]

Le terme de « cancer » est utilisé depuis l'antiquité pour désigner toute excroissance anormale

comme des tumeurs, des ulcérations... (Hippocrate a déjà introduit le terme de « karcinos » pour désigner ces processus expansifs)

Le terme est utilisé plus particulièrement pour les maladies oncologiques (« oncos » masse, grosseur) depuis les connaissances scientifiques de la maladie. [26]

Au Moyen-Âge les chirurgiens décrivirent assez précisément des tumeurs comme le cancer du rectum ou du sein dans le cadre des autopsies.

C'est, au fur et à mesure des études avec notamment la naissance de l'anatomopathologie et de la découverte du système lymphatique, que les scientifiques ont commencé à comprendre le principe de l'évolution de la maladie : le point de départ est locorégionale puis la dissémination se fait par voie lymphatique.

Néanmoins, cette maladie restait incurable malgré les traitements (chirurgie, onguents, saignées). [26]

La découverte des rayons X en 1895 et de la radioactivité en 1896 révolutionna les pratiques thérapeutiques. Le développement de la radiothérapie s'est alors opéré en un peu plus d'un siècle, sur deux fronts parallèles. Cette dualité reposant sur la distinction entre les rayonnements d'origine nucléaire (radioactivité), et ceux d'origine extra nucléaire produits par des appareils. [26]

- La radiothérapie externe au Radium
- La radiothérapie dite conventionnelle (basse énergie)
- La télé cobalthérapie
- Les accélérateurs de particules (haute énergie)
- La curiethérapie

Et parallèlement : [26]

- Pour les traitements à visée locorégionale : la chirurgie
- Pour les traitements à visée générale : chimiothérapie, hormonothérapie, immunothérapie.
- La recherche permet et continue de progresser sur plusieurs fronts : [26]
- L'oncologie médicale ou la connaissance de la maladie par le biais de l'expérience médicale et des études en laboratoire.
- La biologie cellulaire ou une meilleure compréhension de la cancérogenèse pour trouver des pistes de contrôle de la maladie.

- Les thérapies pour tendre vers toujours plus d'efficacité et de tolérance.

2.2.1 Le cancer

Le cancer commence lorsque les cellules d'une partie du corps échappent à tout contrôle. Il existe plusieurs types de cancer, mais ils commencent tous à cause de la croissance incontrôlée des cellules anormales. [28]

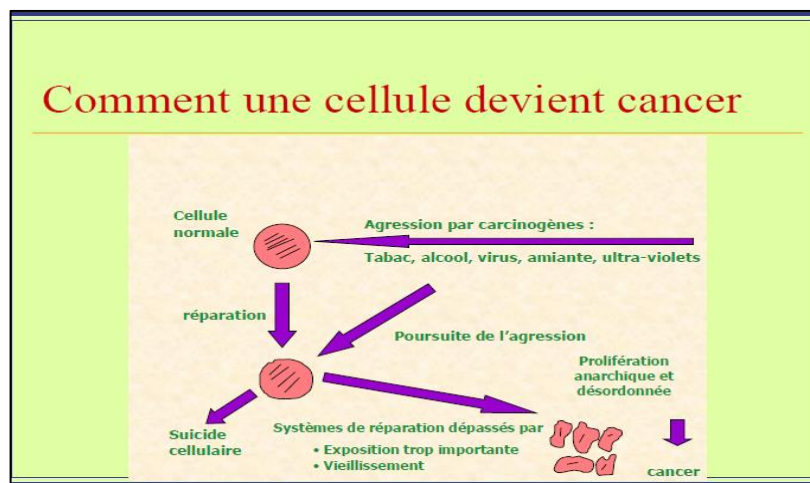


Figure 2.4 Comment une cellule devient cancer.[28]

Le cancer est une prolifération anarchique de cellules aboutissant à la formation d'une tumeur qui a tendance à envahir les tissus environnants et à se disséminer loin de son site initial.

Le cancer va donc se présenter comme une maladie générale d'évolution complexe, dominée par la variabilité dans le temps et dans l'espace de relations particulières entre les cellules cancéreuses et celles de l'hôte qui les contient, sinon qui les héberge. [29]

Le mot « cancer » désigne plus de 200 maladies. Chacune d'entre elles porte un nom différent : cancer du poumon, cancer du côlon, leucémie, etc. Tous les cancers sont différents les uns des autres, mais ils ont une chose en commun : ils s'attaquent aux cellules.

2.2.2 Les causes du cancer

On sait que plus de la moitié des cancers surviennent après l'âge de 65 ans. La consommation de tabac est aussi la cause de plusieurs cancers dont ceux du poumon, de la gorge, de l'estomac, de la vessie et du rein. Comme société, nous avons également développé de très mauvaises habitudes alimentaires.

De plus, comme nous ne faisons pas assez d'activité physique, notre état de santé général est moins bon, ce qui pourrait favoriser l'apparition de plusieurs maladies, dont le cancer.

Par ailleurs, nous vivons dans un environnement où nous sommes en contact tous les jours avec un grand nombre de polluants qui pourraient être cancérigènes : produits chimiques pour détruire

les insectes ou les mauvaises herbes, fumées d'usine, gaz d'échappement des autos, etc.

De la même façon, nos conditions de vie, comme le chômage et la pauvreté, peuvent avoir un effet sur la santé.

En fait, il est encore plus probable qu'une combinaison de ces différents facteurs cause la plupart des cancers. [30]

2.2.3 Les principaux cancers

Le cancer est une maladie qui touche tous les organismes vivants, incluant les insectes et les plantes. [30] Il y a plus de 200 types de cancers. Toutefois, certains sont beaucoup plus fréquents que d'autres :

- Le cancer de la peau
- Le cancer du poumon
- Le cancer du sein
- Le cancer de la prostate
- Les cancers du côlon et du rectum

Les cancers du côlon et du rectum touchent les hommes autant que les femmes, surtout après 50 ans. Ce sont des cancers qui se traitent assez bien quand ils ne sont pas trop avancés. [30]

C'est ce que nous allons expliquer au plus tard.

2.2.4 Les traitements

Lorsqu'un médecin pense qu'une personne a des signes de cancer, il l'envoie passer une série d'examens. C'est à la suite de ces examens qu'on pourra poser un diagnostic, c'est à dire identifier la maladie.

C'est seulement à la suite du diagnostic que l'oncologue, un médecin spécialiste du cancer, pourra proposer un traitement.

Il faut savoir qu'il se fait beaucoup de recherche sur le cancer et que les traitements sont de plus en plus efficaces. Aujourd'hui, le cancer n'est plus une condamnation à mort.

On peut guérir environ la moitié des cancers.

Il existe plusieurs traitements pour le cancer, les trois principaux sont les suivants : la chirurgie, la chimiothérapie et la radiothérapie. [30]

- La chirurgie : est une opération qui consiste à enlever une tumeur en tout ou en partie. En

effet, selon le type de cancer et l'endroit où la tumeur est située, il peut être possible de l'enlever au complet. Dans ce type d'opération, on ôte souvent un peu de tissu sain par mesure de prudence. Si la maladie est plus avancée, on peut enlever une partie de la tumeur, ce qui permet de retarder son développement ou de réduire la douleur.

- La chimiothérapie : consiste à donner des médicaments à la personne atteinte afin de tuer les cellules cancéreuses. Ces médicaments peuvent prendre plusieurs formes : liquides, pilules ou injections. Ce type de traitement peut durer plusieurs mois.
- La radiothérapie : consiste à traiter le cancer au moyen de rayons X très puissants ou de substances radioactives. On concentre ces rayons directement sur les tumeurs afin de détruire les cellules cancéreuses. [30]

2.2.5 Les effets du cancer et de son traitement

Malgré l'efficacité grandissante des traitements, le cancer reste une maladie difficile à combattre. Les traitements ne détruisent pas seulement les cellules cancéreuses, mais aussi certaines cellules saines.

Aussi, les personnes qui ont le cancer sont souvent aux prises avec un grand nombre de désagréments qui peuvent découler de la maladie elle-même, de son traitement ou de la combinaison des deux.

Ces désagréments sont nombreux, mais voici les plus courants : (La fatigue, L'anémie La douleur, Les nausées et les vomissements, La perte de cheveux et de poils, La perte de poids). [30]

2.2.6 La prévention

Comme la prévention est toujours le meilleur moyen de rester en santé, voici sept conseils inspirés de la Société canadienne du cancer [30] .

- Ne fumez pas et évitez la fumée des autres.
- Ayez un régime alimentaire varié : peu d'aliments gras (beurre, charcuterie) et beaucoup de fibres (légumes, fruits, céréales).
- De plus, limitez votre consommation d'alcool.
- Ne prenez pas trop de soleil.
- Faites de l'activité physique.
- Suivez les recommandations des professionnels de la santé pour dépister les différents types de cancer.
- Si votre santé vous semble moins bonne, allez voir votre médecin ou votre dentiste.

- Suivez les consignes de sécurité lorsque vous manipulez des matières dangereuses, que ce soit à la maison ou au travail.

2.3 Le Cancer du colon

Le côlon, qui mesure environ 1,5 mètres, est la partie de l'intestin qui fait suite à l'intestin grêle. Le cancer du côlon est une maladie de la muqueuse, qui tapisse l'intérieur du côlon ou du rectum pour le cancer rectal. Les tumeurs colorectales se développent généralement à partir de la muqueuse tapissant les parois du côlon et du rectum.

Le cancer colorectal met plusieurs années à se développer, de 9 à 10 ans. Il se forme à partir de la transformation d'une tumeur bénigne, un polype (excroissance charnue) apparu sur la muqueuse.

Le cancer colorectal représente, par sa fréquence, le troisième cancer chez l'homme et le second chez la femme. Sa fréquence augmente après 45 ans. Il est découvert dans 73 % des cas au niveau du côlon et de la région recto-sigmoïdienne, et dans 27 % des cas dans le rectum. L'âge médian du diagnostic est de 72 ans chez l'homme et de 75 ans chez la femme.

La baisse de la mortalité par cancer colorectal résulte des progrès réalisés dans la détection de la maladie grâce à un diagnostic plus précoce, à l'amélioration de la prise en charge thérapeutique et à une diminution de la mortalité opératoire (Bouvier AM2009).

2.3.1 Facteurs de risque

- La consommation importante de viande (sauf la volaille), le tabac et la sédentarité augmentent le risque de cancer colorectal.
- Un régime riche en fibres et une activité physique régulière ont un effet protecteur établi.
- Un dépistage systématique précoce est proposé en cas d'antécédents familiaux ou chez les porteurs d'anomalies génétiques spécifiques.

2.3.2 Le dépistage

Le test de dépistage du cancer du côlon et du rectum est basé sur la découverte de sang microscopique (non vu à l'oeil nu) dans les selles. En effet, très souvent, les gros polypes et les cancers de l'intestin saignent sur un mode discret et non visible.

Ce saignement microscopique pourra être repéré grâce à un test de laboratoire : le test OC Sensor. La pratique du dépistage a montré qu'elle pouvait diminuer la mortalité par cancer colorectal : entre 2005 et 2012 la mortalité due à ces cancers a diminuée de 1,5% par an chez l'homme et de 1,1% chez la femme (Binder-Foucard, 2013)[31].

2.3.3 Stades du cancer

Le cancer du côlon peut rester longtemps asymptomatique*, Certains symptômes peuvent également attirer l'attention et conduire à une consultation médicale.

Le facteur le plus important pour décider du traitement et donner un pronostic est le stade de la maladie au moment du diagnostic.

Trois critères (TNM) permettent d'évaluer l'étendue de la maladie :

- la taille et la profondeur de la tumeur (T).
- L'atteinte ou non des ganglions et leur nombre en cas d'atteinte (N).
- La présence ou non de métastases (M).

Ces critères sont combinés pour donner un stade global dont va dépendre la stratégie thérapeutique.

Selon l'emplacement de la tumeur les symptômes sont plus ou moins précoces.

Les complications sont souvent le signe d'une maladie avancée.

Le pronostic est amélioré par un diagnostic et une prise en charge précoces. [31]

2.3.4 Traitements

Pour traiter le cancer colorectal, plusieurs thérapeutiques sont employées. On évite une opération mutilante chaque fois que cela est possible.

Le choix du traitement dépend de la localisation du cancer et de son étendue, de l'état général du patient et des comorbidités * éventuelles.

La prise en charge proposée sera définie à l'issue d'une réunion de concertation pluridisciplinaire (RCP), après l'analyse par plusieurs spécialistes de l'état du patient et de l'incidence* potentielle des différents traitements envisageables (bénéfices attendus, risques et effets secondaires). [31]

2.3.5 Suivi

Le suivi a pour objectif de vérifier la survenue ou non de récurrences, d'évaluer et de traiter les séquelles et les complications en relation avec la maladie et/ou les traitements.

Classiquement, le patient est suivi tous les 3 mois pendant 3 ans puis tous les 6 mois pendant 2ans puis tous les 5 ans.

Pendant ce suivi, le patient bénéficiera d'un soutien psychologique et social autant que

nécessaire.

Le malade est suivi régulièrement tous les ans [31]:

- pendant 5 ans ce suivi inclut un examen clinique, une radiographie thoracique, une échographie abdominale,
- puis tous les 5 ans, une coloscopie totale est réalisée.

L'intérêt éventuel de la surveillance par scanner est en cours d'évaluation. La poursuite de la surveillance clinique est recommandée entre 5 et 10 ans après le diagnostic.

2.4 Ontologies dans le domaine oncologie

Les informations issues du domaine médical sont de grande taille, chose qui est due principalement à la diversité, non encore cernée, de tumeur et des humains. Structurer cette abondante quantité d'information est maintenant nécessaire pour pouvoir extraire les données pertinentes en un temps acceptable ce qui permettra de mieux connaître notre présent pour pouvoir anticiper et bien préparer notre futur. Les ontologies offre une solution pour organiser et partager ces connaissances, offrant ainsi un outil automatique capturant à la fois données et sémantique.

La liste des ontologies disponibles dans le domaine d'oncologie est illimité pour cela nous cité quelques un :

- Germ cell cancer (cancer des cellules germinales).
- Malignant adenoma (adénome malin).
- Carcinoma Carcinome).
- Melanoma (mélanome)

2.5 Conclusion

Le cancer est encore une maladie terrible, mais une maladie qu'on apprend de plus en plus à connaître.

La médecine évolue à toute vitesse et les traitements deviennent de plus en plus efficaces. De plus, on comprend mieux comment cette maladie est liée aux habitudes et aux conditions de vie. Le cancer colorectal a largement bénéficié des progrès thérapeutiques tant chirurgicaux, radiothérapeutiques que médicamenteux, permettant d'éviter souvent une mutilation et de prévenir des récurrences. [31]

Nous avons vu dans ce chapitre que le traitement de la tumeur cancéreuse nécessite de nombreuses analyses, discussions médicales et médicaments, ainsi que de nombreux tests, révélations et ordonnances du médicament nécessaire, ainsi que le nombre de médecins diagnostiqués et, bien sûr, chaque médecin demande tous les documents du patient liés à tous les suivis médicaux et Ou perdu. [31]

Aussi nous avons présenté les méthodes et les méthodologies les plus utilisées pour seconder le processus de construction d'une ontologie. Et décrit les principales ressources terminologiques les plus utilisées dans le domaine médical.

Ceci a conduit à l'identification des trois méthodologies de base pour le développement des thérapeutiques oncologiques, ce que nous expliquerons dans le chapitre suivant.

Chapitre 3

Conception d'une ontologie de traçabilité

Conception d'une ontologie de traçabilité

La traçabilité thérapeutique de cancer reste un problème un peu difficile qui attire les chercheurs. la difficulté revient à plusieurs paramètres, dont on cite quelques un :

- *Les masses des consultations et des traitements.*
- *Le taux des fichiers médicaux (les analyses, les fiches des traitements ...).*
- *La possibilité de changement du médecin et des cliniques.*

Tous ces paramètres et bien d'autres des obstacles à surmonter afin d'obtenir la trace nécessaire pendant la période de traitement.

Dans ce chapitre, nous exposons le processus de construire notre ontologie de traçabilité thérapeutique, dédié à traiter tous la durée de traitements de malade, nous commencerons, en premier lieu, par présenter la méthode utilisée et les étapes principale de construction.

1. Processus de développement d'ontologies

Notre ontologie de traçabilité dédiée à la conduite thérapeutique est basée principalement sur le développement de deux ontologies dérivées, la première concerne l'ontologie des maladies oncologiques et bien précisément les cancers du côlon, cette ontologie va être développée à partir d'une ontologie existante nommée « carcinome du colon » à partir de « disease ontology (ontologie des maladies) » [32]. La deuxième utilisée dans notre travail est l'ontologie qui concerne les différents types de traitement, cette ontologie doit être construite à partir de zéro vu le manque des travaux connexes dans la thérapie oncologique.

Pour l'ontologie qui concerne les maladies cancéreuses on va la définir à partir d'une ontologie existante (la fameuse ontologie « Disease ») et on va la réduire juste au « carcinome du colon ». Pour ce faire on a choisi la méthode ARCHONTE (ARCHitecture for ONTOlogical Elaborating) qui le considère la plus adaptée pour la construction d'une ontologie à partir de l'existant. Cette méthode fournit les étapes suivantes :

1. choix des termes pertinents du domaine, normalisation sémantique précisant les relations de similarités et de différences que chacun des concepts entretient avec ses frères et son père (principes de la sémantique différentielle) ;
2. formalisation des connaissances ou engagement ontologique : construction d'une ontologie référentielle et ajout des propriétés et annotations, définition des domaines et codomaines des relations ;
3. opérationnalisation de l'ontologie dans un langage de représentation des connaissances.

Elle a été appliquée en utilisant la Classification Internationale des Maladies (CIM-10) de l'Organisation Mondiale de la Santé, la CCAM (Classification Commune des Actes Médicaux), le FMA, et la SNOMED.

On va entrer autres détaillé : METHONTOLOGY proposé par le laboratoire d'Intelligence Artificielle de l'Université Technique de Madrid (UPM). Cette méthodologie s'applique dans le cadre du développement à partir de zéro d'ontologie Elle couvre tout le cycle de vie d'une ontologie. Elle comporte un ensemble établi de principes, de processus, de pratiques, de méthodes et d'activités utilisés pour le design, le développement, l'évaluation et l'implémentation.

Methontology est supportée par la plateforme ODE (Ontology Development Environment) qui a donné suite à WebODE. Beaucoup d'ontologies ont été développées en utilisant cette méthodologie : CHEMICALS (dans le domaine des éléments chimiques), la Reference Ontologie, etc. De plus, c'est une recommandation de la Fondation for Intelligent Physique Agents (FIPA).

Cette méthodologie propose un processus de développement qui s'appuie sur la construction d'un modèle conceptuel robuste et sur la détermination claire et consistante des exigences de l'ontologie à construire par :

- l'identification du processus de développement de l'ontologie (qui consiste à identifier les concepts, les relations du domaine d'intérêt et à les représenter explicitement dans un modèle formel. Cette phase prévoit également l'exploitation éventuelle d'ontologies existantes),
- un cycle de vie basé sur des prototypes en évolution,
- des techniques particulières pour réaliser chaque activité,
-

2. Description d'ontologie Carcinome du colon

2.1 Conception de l'ontologie du domaine de carcinome du colon

2.1.1 Domaine de l'ontologie

L'oncologie est un vaste domaine caractérisé par sa difficulté et son développement rapide, que ce soit au niveau du traitement ou des méthodes de découverte ou du recours à d'autres sciences pour contribuer à l'éradication des tumeurs.

En raison de sa complexité, il nécessite des explications détaillées et méticuleuses, et c'est ce que les experts ont jugé nécessaire de traiter le cancer du côlon comme l'une des tumeurs les plus graves subies par l'humanité.

L'oncologie a suscité un vif intérêt dans divers domaines, notamment la bioinformatique, qui souhaite étudier et faciliter son traitement et s'efforcer de développer des mécanismes pour le combattre.

2.1.2 Les besoins examinés par cette ontologie

Le cancer du côlon se caractérise par ses parties multiples et précises, qui doivent être expliquées et classées par type et statut, ce qui nous a amenés à nous appuyer sur cette ontologie pour expliquer les détails de cette partie du cancer du côlon.

On s'intéresse dans cette ontologie de carcinome du colon, comment il est classé, la diversité des classifications appliquées, sur les taxons qui sont la base de certaine classification, ainsi que le type de reproduction adopté et la structure cellulaire de chaque espèce du colon.

2.2 Processus d'affinement de l'ontologie

On ne peut pratiquement dissocier les étapes de construction d'une ontologie, car il s'agit d'un processus non linéaire, plusieurs allers-retours sont souvent nécessaires lors du développement d'une ontologie, pour les raisons suivantes:

- il n'est pas possible de savoir dès le départ, que les termes collectés sont suffisant pour répondre à l'objectif pour lequel l'ontologie a été construite. Il est possible d'ajouter des

nouveaux termes lorsque cela s'avère nécessaire, comme il est possible, aussi, de retirer des termes jugés inutiles. [33]

- il n'est pas toujours facile de prédire qu'un terme va jouer le rôle d'une classe ou celui d'un attribut, plusieurs modifications seraient effectuées dans ce sens. [33]

Pour représenter l'ontologie choisie, il est nécessaire de construire:

- Une liste de concepts.
- Une liste d'attributs.
- Une liste de relations
- Une représentation hiérarchique des concepts.

2.3 Respect des principes de construction

- ❖ **Clarté et objectivité:** pour répondre à ce principe, tous les termes utilisés dans cette ontologie ont été associés à des définitions. [33]
- ❖ **Complétude:** pour répondre à ce principe, les définitions des concepts et des relations de cette ontologie ont été associés à des conditions nécessaires, d'autres ont été associés à des conditions nécessaires et suffisantes, dans la mesure de pouvoir définir ces conditions. [33]
- ❖ **Extensibilité ontologique maximale:** la définition d'un terme n'explique que le terme lui-même, sa définition ne peut être la même que celle d'un terme plus général, ou d'un terme plus spécialisé. [33]
- ❖ **Principe de distinction ontologique:** les concepts dans cette ontologie sont suffisamment disjoints. [33]

2.4 Présentation de cette ontologie

Une ontologie est composée d'un vocabulaire partagé. Un premier travail consiste donc à construire un corpus, qui rassemble le vocabulaire employé dans le domaine. [33] A partir de « **DO**¹ » on a extrait les concepts de cette ontologie.

Disease Ontology intègre sémantiquement les vocabulaires maladie et médical par le biais d'une cartographie croisée étendue des termes DO, MeSH, ICD, le thésaurus de NCI, SNOMED et OMIM.

¹DO : Disease ontology(ontologie des maladies) a été développée comme une ontologie normalisée des maladies humaines dans le but de fournir à la communauté biomédicale des descriptions cohérentes, réutilisables et durables des termes de maladie humaine, des caractéristiques de phénotype et des concepts de maladie du vocabulaire médical connexes, grâce aux efforts de collaboration de chercheurs de la Northwestern University, Centre for Genetic Medicine et l'Institut des sciences du génome de l'École de médecine de l'Université du Maryland.

Pour cette présentation nous nous sommes basés sur la méthodologie ARCHONTE (ARCHitecture for ONTOlogical Elaborating), proposée par B. Bachimont [33] [11], B. Bachimont propose de contraindre l'ingénieur des connaissances à un "engagement sémantique", c'est-à-dire à expliciter clairement le sens de chacun des concepts de l'ontologie, en introduisant une "normalisation sémantique" :

Les primitives nécessaires à la représentation des connaissances doivent être modélisées à partir des données empiriques dont on dispose, à savoir l'expression linguistique des connaissances. Le travail de modélisation doit s'effectuer à partir de documents attestés dans la pratique d'un domaine et rassemblés en un corpus. Le corpus est constitué de documents produits dans le contexte où le problème à résoudre se pose [33].

❖ Normalisation sémantique et engagement sémantique

L'étape de normalisation sémantique a pour objectif de rendre explicite le sens des expressions linguistiques. Il s'agit, par ce processus, d'en faire des primitives du domaine, c'est-à-dire d'identifier les notions élémentaires à partir desquelles l'ensemble des connaissances du domaine sont construites. B. Bachimont propose de se baser sur la sémantique différentielle, présentée dans les travaux de F. Rastier [35], pour réaliser cette étape. Cette théorie attribue un sens aux termes grâce à la définition de traits sémantiques génériques et spécifiques. [15]

❖ Formalisation des connaissances et engagement ontologique

La seconde étape permet de formaliser les connaissances du domaine à représenter. Il s'agit de définir des concepts, et non plus des notions, selon une sémantique formelle et extensionnelle. [15]

Les concepts sont exprimés dans un langage de représentation des connaissances et leur sens est décontextualisé. Selon la sémantique extensionnelle, les concepts sont liés à un ensemble de ce Chapitre. Traitement du corpus et conceptualisation référents dans le monde, i.e., à un ensemble d'objets du domaine. Cet ensemble est appelé l'extension du concept. A ce stade, des opérations ensemblistes, telles que la réunion ou l'intersection, peuvent être utilisées pour composer de nouveaux sens et donc de nouveaux concepts formels. [15]

Cette étape permet également de formaliser les relations qui existent entre les concepts. [15]

2.4.1 Les axes primitifs de l'arbre ontologique construit

2.4.1.1 Classification scientifique de carcinome du colon

▪ Méthode

Le cancer colon est classé scientifiquement par les spécialistes dans des catégories qui se chevauchent et qui sont étroitement liées (Carcinome du côlon récurrent \longrightarrow Cecum Carcinoma), en fonction du type de la partie touchée, des glandes et des cellules touchées.

Nous verrons les classifications du cancer du côlon dans la section annotation.

▪ Intérêts

Dans la classification scientifique, l'espèce est l'unité de base, le groupe élémentaire.

2.4.1.2 Taxon

En taxinomie, un taxon est une entité conceptuelle qui est censée regrouper tous les types de cancer.

Dans cette ontologie, on n'a présenté que le carcinome de colon.

La classification du cancer propose une hiérarchie codifiée en des étapes qui sont présentent dans la figure suivante :

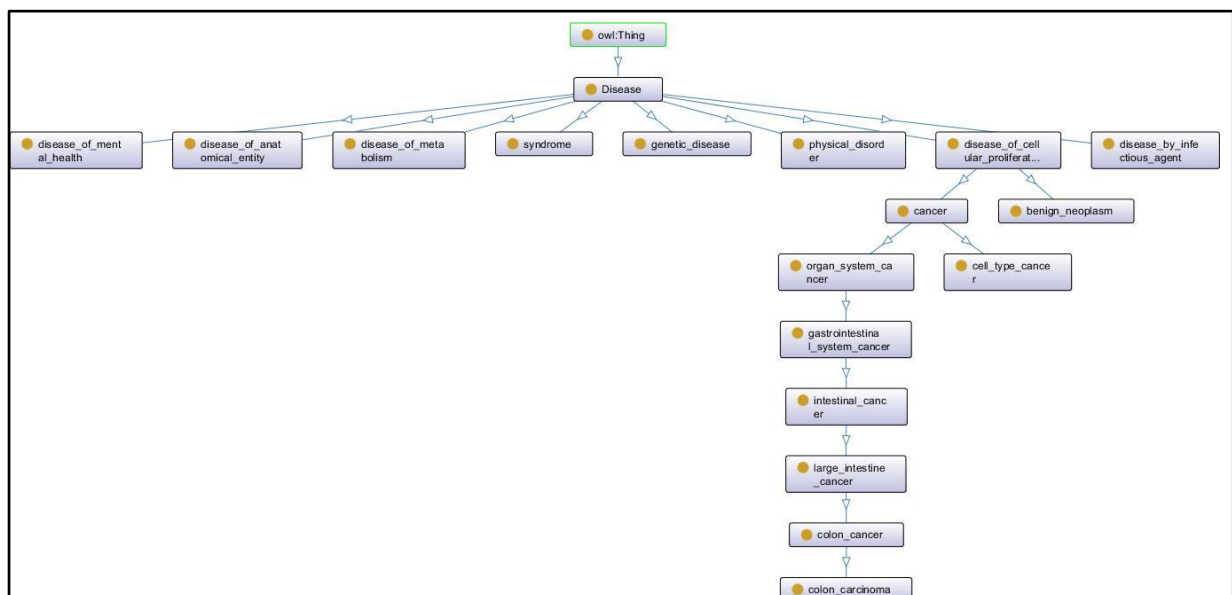


Figure 3.1 : Classification du cancer à partir d'OD

2.5 Les propriétés des classes de cette ontologie

Les classes seules ne fourniront pas assez d'informations pour assurer la compétence de l'étape 1. Après avoir défini l'ensemble des classes, on devait décrire la structure interne des concepts. [33]

2.5.1 Propriété d'objet

Sont des propriétés qui relient les instances d'une classe aux instances d'une autre classe. Chaque propriété dans l'ontologie relie les classes entre elles. Ces dernières font partie d'un domaine. [33]

2.5.2 Propriété de type de données

Dans cette ontologie, nous nous intéressons à la classification scientifique et globale des éléments et stades du cancer du côlon et à son ordre conformément aux normes internationales. Il est difficile de faire attention à toutes les étapes et à toutes ses composantes, nous avons donc limité notre attention à la partie carcinome du colon.

2.6 Représentation hiérarchique des concepts

L'ontologie choisit, couvre un bon nombre de termes, c'est pour cette raison que la hiérarchie de classe ci-dessous ne représente que les classes principales.

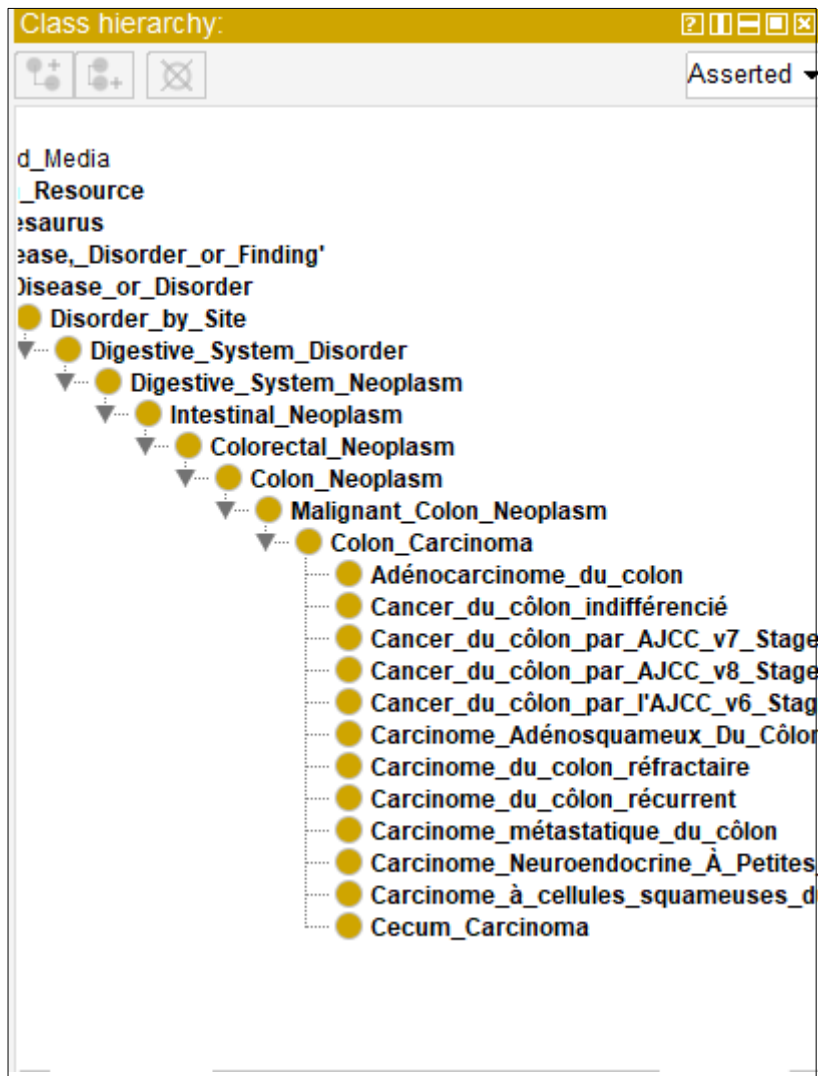


Figure 3.2 : représentation hiérarchique des concepts principaux de carcinome du colon.

❖ La visualisation OWL Pour les concepts principaux

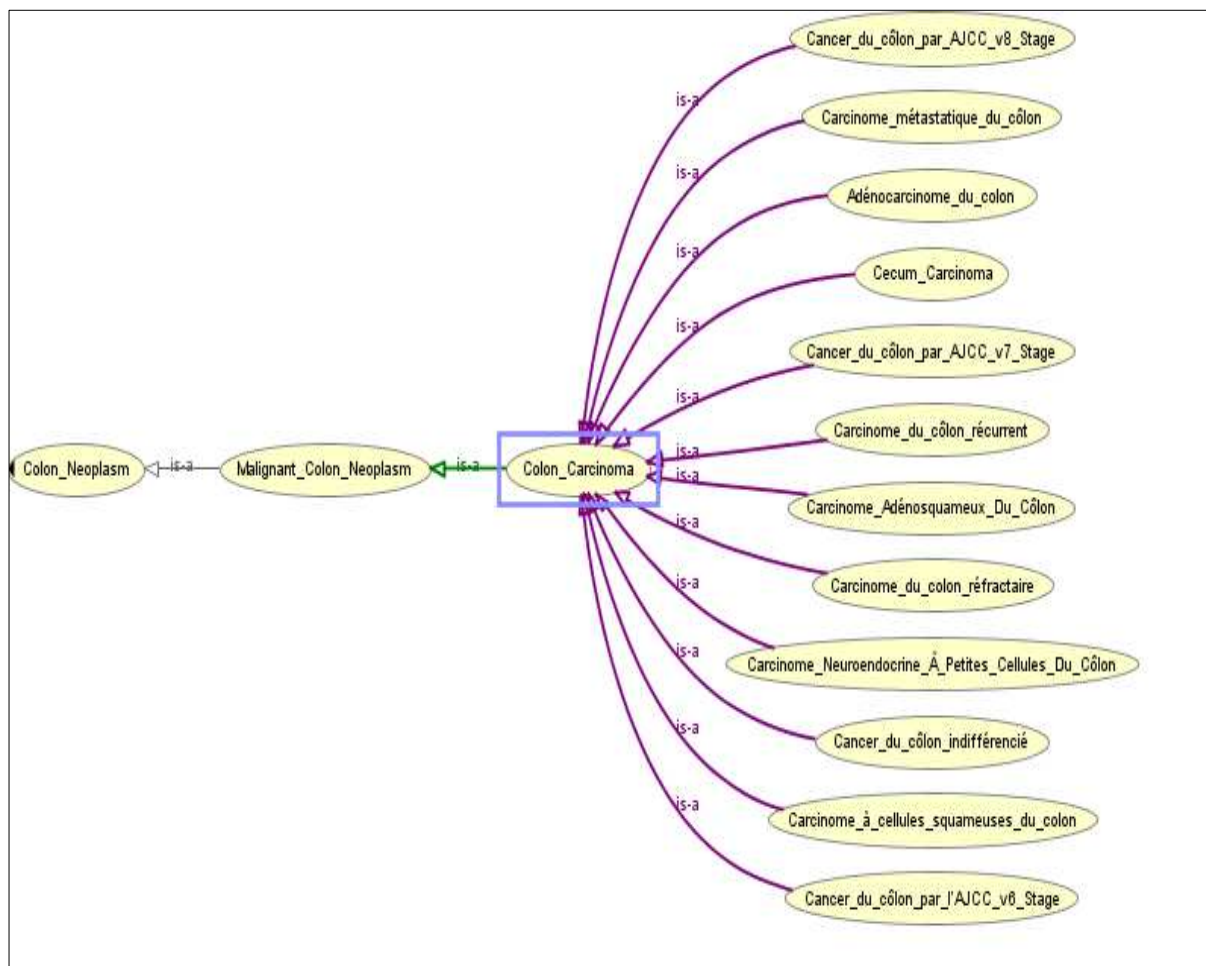


Figure 3.3 : représentation de visualisation OWL des concepts principaux de carcinome du colon.

2.7 Glossaire des concepts

Dans cette section nous le tab 1 illustre le glossaire de termes de carcinome du colon. Les termes recensés sont ceux en commun avec le domaine médicale et ceux propre au domaine considéré.

Nom	Description
Carcinome du colon	Cancer qui se forme dans les tissus du côlon (la partie la plus longue du gros intestin). La plupart des cancers du côlon sont des adénocarcinomes (cancers qui commencent par des cellules qui produisent et libèrent du mucus et d'autres fluides).
Cecum Carcinoma	Une tumeur épithéliale maligne qui provient du caecum et envahit la muqueuse musculieuse dans la sous-muqueuse. La grande majorité sont des adénocarcinomes.
Carcinome du colon récurrent	Réapparition du carcinome du côlon après une période de rémission.
Carcinome du colon non différencié	Une tumeur épithéliale maligne invasive qui provient du côlon. Il n'existe aucune preuve morphologique,

	immunophénotypique ou biologique moléculaire de différenciation glandulaire ou squameuse.
Adénocarcinome du colon	Adénocarcinome issu du côlon. Il est plus fréquent chez les populations ayant un régime de type occidental et chez les patients ayant des antécédents de maladie inflammatoire chronique de l'intestin.
Carcinome adénoquameux du colon	Un carcinome du côlon inhabituel caractérisé par la présence de composants carcinomateux glandulaires et squameux. Les deux composants carcinomateux peuvent être mélangés à l'intérieur de la tumeur, ou les deux peuvent apparaître séparément dans des zones différentes.
Cancer du côlon par stade v6 de l'AJCC	terme faisant référence à la stadification du cancer du côlon selon la 6ème édition du Comité mixte américain sur le cancer.
Cancer du côlon par stade AJCC v7	terme faisant référence à la stadification du cancer du côlon selon la 7ème édition de l'American Joint Committee on Cancer.
Cancer du côlon par stade v8 de l'AJCC	terme faisant référence à la stadification du cancer du côlon selon la 8ème édition de l'American Joint Committee on Cancer. Ce système de classification s'applique aux adénocarcinomes, aux carcinomes neuroendocriniens de haut grade et aux carcinomes épidermoïdes du côlon. (8e éd. de l'AJCC)
Carcinome neuroendocrinien à petites cellules du côlon	Un carcinome agressif, de haut grade et peu différencié avec différenciation neuroendocrinienne qui provient du côlon. Il est caractérisé par la présence de petites cellules malignes.
Carcinome à cellules squameuses du colon	Un carcinome du colon très rare caractérisé par la présence d'un infiltrat malin des cellules malpighiennes.
Carcinome métastatique du côlon	Un carcinome qui provient du côlon et s'est métastasé à un autre site anatomique.
Carcinome du colon réfractaire	Cancer du colon résistant au traitement.

Tab 3.1 : Définition des concepts

2.8 La représentation graphique de carcinome du colon

La figure suivante représenter le graphe des concepts qui sont définis dans le tableau précédent.

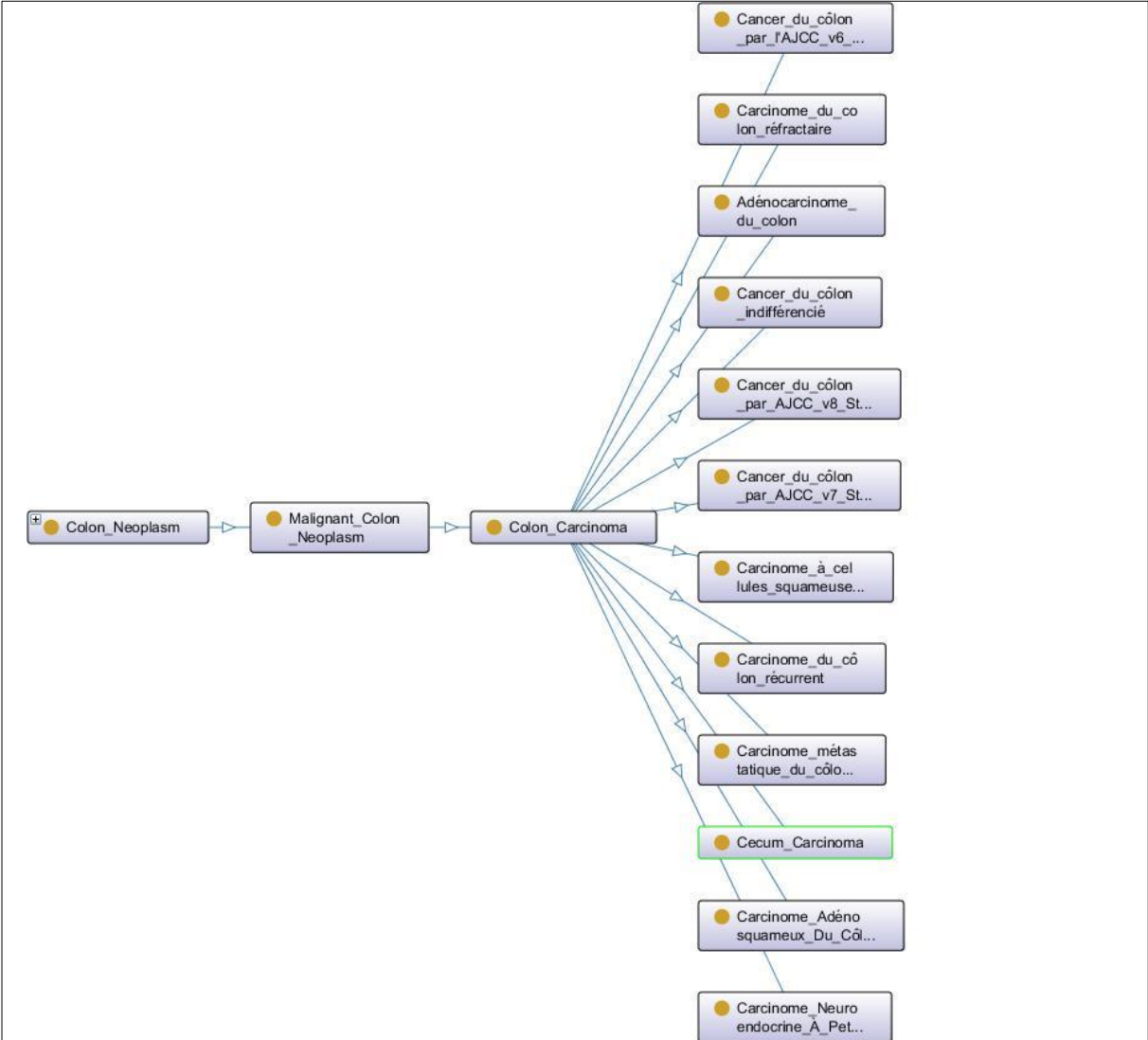


Figure 3.4 : représentation de graphe des concepts principaux de carcinome du colon.

3 Traitements

3.1 Description d'ontologie de Traitements

Nous avons également expliqué à l'avance la nécessité d'une ontologie de traitements thérapeutiques nous permettant de suivre le processus de traitement sur une longue période en raison de la difficulté et de la sensibilité de cette étape.

Dans cette section, nous expliquerons les moyens de traiter le cancer du côlon et les relations entre chaque traitement. Cette ontologie de traitement est une ontologie construite à partir de méthode de construction à zéro .Cet axe représente les concepts de type traitement ou thérapie d'une maladie, par exemple : chimiothérapie, radiothérapie, chirurgie ...

Le tableau suivants présenter la description des concepts :

concept	Description
Traitement	une thérapie, ou plus généralement la thérapeutique, est un ensemble de mesures appliquées par un professionnel de la santé (ou thérapeute) à une personne vis-à-vis d'une maladie, afin de l'aider à en guérir, de soulager ses symptômes, ou encore d'en prévenir l'apparition.
Thérapie ciblée	"biothérapies" désignent des médicaments dirigés contre des cibles moléculaires : récepteurs, gènes ou protéines impliquées dans les voies de signalisation intracellulaires jouant un rôle dans la transformation des cellules en cellules cancéreuses ou dans le développement des tumeurs malignes. Par opposition aux médicaments de chimiothérapie traditionnelle qui s'opposent, globalement, à la multiplication des cellules, les médicaments de chimiothérapie ciblée visent les mécanismes intimes de la cancérisation des cellules.
La chimiothérapie	Les protocoles usuels de chimiothérapie sont préconisés en traitement adjuvant et dans les formes avancées de la maladie pour contrôler efficacement la maladie. En cas de métastase(s) hépatique(s), ces traitements permettent, dans un nombre non négligeable de cas, de les rendre résécables.

LA CHIRURGIE	la partie de la thérapeutique qui implique des opérations internes ou des manœuvres externes ¹ sur les tissus, notamment par incision et suture. Un chirurgien est un professionnel de la santé habilité à pratiquer la chirurgie (médecin spécialiste, chirurgien-dentiste, vétérinaire). Un acte médical pratiqué par un chirurgien est une opération chirurgicale.
---------------------	--

Tab 3.2 : Définition concepts principale d'ontologie de traitements

3.2 Construction d'une ontologie de traitements

3.2.1 La Méthode METHONTOLOGY

METHONTOLOGY [36] se veut être une approche intégrée s'inscrivant dans le cadre dans un processus de gestion de projet complet, allant la spécification des besoins jusqu'à la phase réalisation et maintenance. Ce processus est composé des étapes suivantes:

1. Spécification : déterminer l'utilisation future de l'ontologie;
2. Conceptualisation : obtenir un modèle du domaine au niveau des connaissances;
3. Formalisation : transformation du modèle conceptuel en modèle formel;
4. Intégration : réutilisation d'autres ontologies;
5. Implémentation : construction d'un modèle opératoire utilisable par un ordinateur;
6. Maintenance : mise à jour de l'ontologie en cas de besoin.

METHONTOLOGY permet de caractériser l'ontologie au niveau des connaissances et insiste sur la nécessité de travailler à partir de représentations intermédiaires des connaissances lors de la phase de conceptualisation.

3.2.2 Spécification :

Une ontologie ne peut être construite qu'après la phase de spécification. Il s'agit d'établir un document informel de spécification de besoins. Au niveau de ce document, nous décrivons l'ontologie à construire à travers les cinq aspects suivants :

- **Le domaine de connaissance :**

Domaine de prise de décision carcinome du côlon.

- **L'objectif :**

Suivi complet et précis de tous les stades du traitement du cancer colon à ses débuts, sa propagation ou son élimination.

- **Les utilisateurs futurs de l'ontologie :**

Patientes, les médecins, Professionnels de santé.

- **Les sources d'informations :**

NCIthesaurus, NCI Terme Browser.

- **La portée de l'ontologie :**

Patientes, Professionnels de santé, médecin, fiche clinique ...

3.2.3 Conceptualisation :

Dans cette étape on distingue les principales tâches suivantes :

- **Construction du glossaire de termes**

Construire un glossaire de termes est la première tâche à effectuer dans l'étape de conceptualisation. Il recueille et décrit tous les termes (concepts, instances, attributs, relations entre les concepts, etc.) qui sont utiles et potentiellement utilisables dans le domaine où nous spécifions leurs descriptions détaillées et non ambiguës dans un langage naturel. Le tableau (3.2) illustre quelque terme de glossaire de TraçCCOnto. Les termes recensés sont ceux en commun avec le domaine médicale et ceux propre au domaine considéré.

- **Classification des concepts en hiérarchies de concepts**

Cette tâche consiste à définir les relations taxonomiques entre les concepts définis dans le glossaire de termes. [37]

En outre, METHONTOLOGY propose d'utiliser quatre relations taxonomiques qui sont sous-classe-de (Subclass-Of), décomposition-disjointe (Disjoint-Decomposition), décomposition-exhaustive (Exhaustive-Decomposition), et partition (Partition).

- Un concept C1 est sous classe d'un autre concept C2 si et seulement si chaque instance de C1 est aussi une instance de C2. [37]
- Une décomposition disjointe d'un concept C'est un ensemble de sous classes de C qui n'ont pas des instances en commun et ne couvrent pas C c.à.d. il peut exister des instances du concept C qui sont des instances d'aucun concept de la décomposition.
- Une décomposition exhaustive d'un concept C'est un ensemble de sous classes de C qui couvrent C et peuvent avoir des instances en commun et des sous classes. [37]

Autrement dit, il ne peut pas exister des instances du concept C qui ne sont pas des instances d'au moins un concept de la décomposition. [37]

- Une partition d'un concept C'est l'ensemble de sous classes de C qui ne partagent pas des instances en commun et couvrent C. [37]

3.2.4 Construction de diagramme de relations binaires

Une relation binaire permet de relier deux concepts entre eux (un concept source et un concept cible). « Si R est une relation entre deux concepts C1 et C2 alors pour tout couple d'instances des concepts C1 et C2, il existe une relation de type R qui lie deux instances de C1 et C2 ». Cette tâche permet de représenter d'une manière graphique les différentes relations qui existent entre les divers concepts que se soit de même ou de différentes hiérarchies. [37]

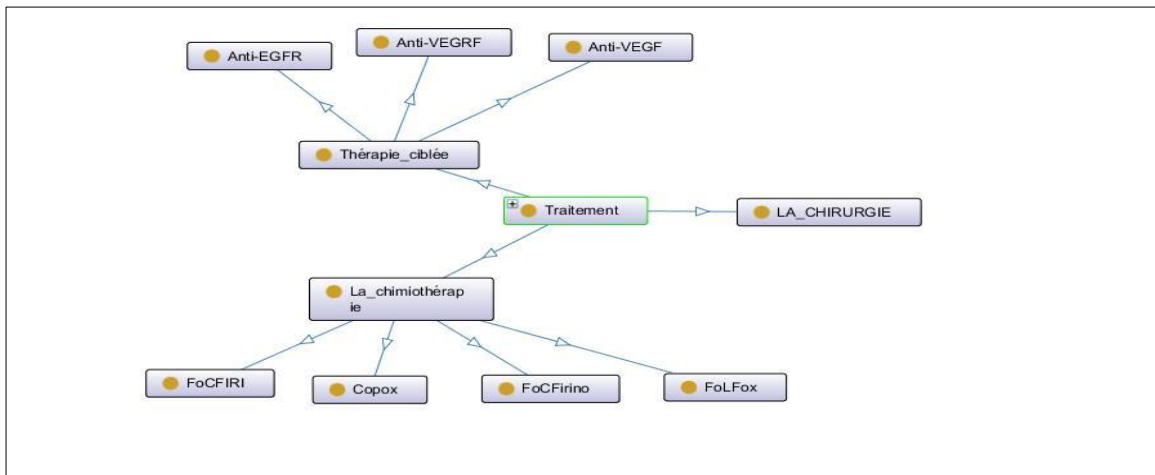


Figure 3.5 : Graphe de relations binaires.

3.2.5 Construction de visualisation d'ontologie

La visualisation permet de nous valider les relations « is-a » de chaque classe.

La figure suivante présente cette visualisation :

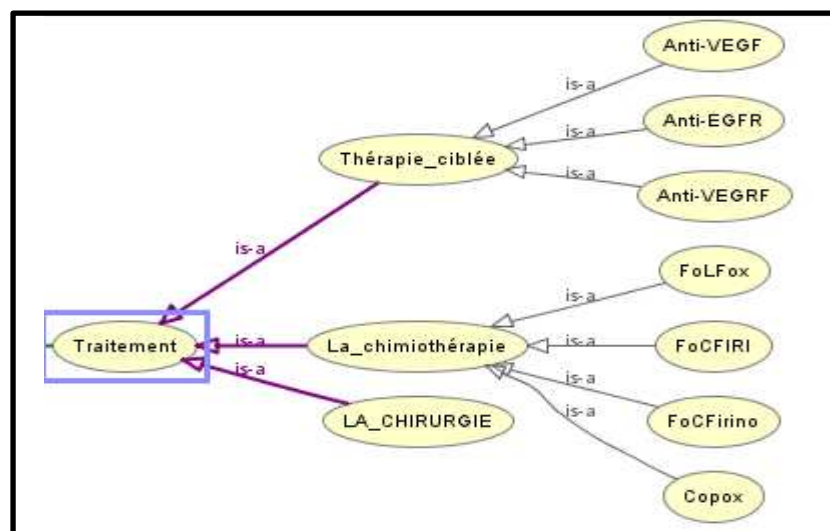


Figure 3.6 : la visualisation d'ontologie de traitement

3.2.6 Représentation hiérarchique des classes

L'ontologie choisit, couvre un bon nombre de termes, c'est pour cette raison que la hiérarchie de classe ci-dessous ne représente que les classes principales.

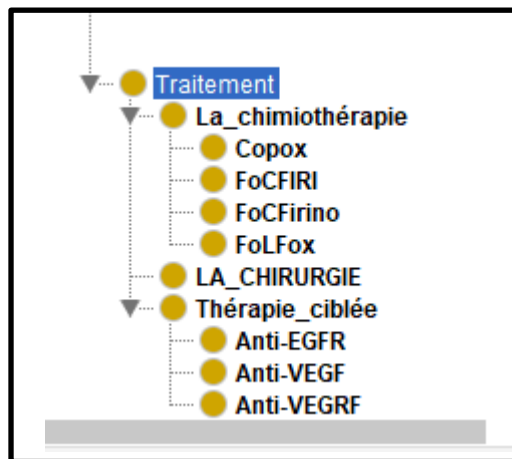


Figure 3.7 : représentation hiérarchique de classes principales de traitements.

3.2.7 Construction de la table des relations binaires

Le but de cette tâche consiste à construire une table de relations binaires décrites en détaille. Pour chaque relation utilisé dans le diagramme des relations binaires, nous définissons le nom de la relation, le nom des concepts sources et cibles et les cardinalités source et cible (voir Tableau 3.3).

Nom de la relation	Concept source	Concept ciblé	Card. Source	Card. source
Is-a	La chimiothérapie	Copox	0..n	0..n
Is-a	La chimiothérapie	FolFox	0..n	0..n
Is-a	Thérapie ciblée	Anti-EGFR	0..n	0..n
Is-a	Thérapie ciblée	Anti-VEGFR	0..n	0..n
(...)	(...)	(...)	(...)	(...)

Tab 3.3 : Table des relations binaires.

4. Définition de l'ontologie dédiée à la Traçabilité

Un malade cancéreux doit vivre le reste de sa vie sous surveillance et sous traitement, ce dernier dépend toujours du type de cancers ainsi que tous traitements pris. Un système de traçabilité des données du dossier « patient–maladies–traitements » prévoit l'intégration de multiples éléments générés par des différents protocoles incluant l'historique thérapeutiques du patient, les traitements oncologiques que ce soit chimio, thérapie ciblée ou hormonothérapie considèrent comme un processus d'identification du patient/médecin pour un suivi médical complet qui implique la connaissance de l'évolution du patient.

Dans le cadre de traitement oncologique de nombreux progrès médicaux ont été possibles grâce aux évolutions de la transmission des informations, aux évolutions technologiques et statistiques. Une bonne gestion du diagnostic ainsi que le choix convenable au type, nature et le stade de la tumeur ont eu comme effets de réduire les coûts, de contribuer au développement durable, d'optimiser le stockage de l'information et surtout a permis d'affiner le diagnostic via les post-traitements, la chose qui nous garantit une traçabilité de l'état du patient qui nous permette une meilleure cohérence du parcours de soins.

Ce type de système de traçabilité est mis en place pour réduire les erreurs médicales. Le suivi de l'évolution d'un patient sur ses interventions médicales est aujourd'hui possible en grand partie si on donne beaucoup d'importance à la télémédecine et à la collaboration entre médecins.

L'objectif d'une ontologie de traçabilité thérapeutique est de conserver l'ensemble des protocoles appliqués au patients depuis la détection de la tumeur, avec un principe d'optimisation de l'accessibilité aux données de grandes tailles qui sont manipulées et échangées. Car, comme nous l'avons montré dans l'état de l'art les ontologies sont le moyen le plus adapté pour l'accès aux traces des connaissances et d'une manière optimisée.

4.1. Traçabilité dans le domaine médicale

La sécurité, la traçabilité et l'efficacité dans le domaine médical apparaissent comme un avantage Compétitif pour la surveillance et la certification de l'information transmise. Dans les systèmes médicaux, les tâches les plus importantes sont d'identifier, de classer et de protéger l'information en utilisant des bases de données relationnelles. La traçabilité de l'information médicale a une fonction stratégique qui est le suivi des activités de procédures

et des traitements faits par les différents départements hospitaliers et les différents professionnels de santé.

Le développement d'une ontologie dédiée à la traçabilité thérapeutique et le suivi médical des patients cancéreux a pour principal objectif : l'amélioration du choix et de la qualité du protocole thérapeutique dans le domaine oncologique. L'ontologie de traçabilité avec l'ontologie du traitement va permettre un post-traitement et un diagnostic fiable lors des métastases sous la responsabilité des médecins oncologues.

4.2. Ontologie de traçabilité

L'objectif principal de cette ontologie est de permettre un échange d'information entre les professionnels de santé qui contribuent à l'amélioration de l'état de santé d'un malade cancéreux par le partage d'expériences et principalement l'aide à la prise de décision thérapeutique. Pour obtenir un diagnostic correct, il est nécessaire de pouvoir : (1) Récupérer les différents traitements liés au patient tout au long de la gestion de sa maladie ; (2) Afficher la chronologie de participation des médecins qui ont contribué aux diagnostics ainsi que à la décision thérapeutique (3) Afficher les commentaires, les observations ou les recommandations cliniques que les médecins ont proposé.

L'enjeu principal de cette ontologie constitue un facteur clé d'amélioration de la performance du traitement. Son utilisation dans le travail collaboratif est la réponse organisationnelle et l'unification entre (Ontologie de traitement – Ontologie de traçabilité). Les objectifs d'utilisation de cette ontologie sont :

- L'amélioration de l'accessibilité de tous les médecins travaillant de manière collaborative,
- L'impulsion d'une meilleure coordination entre les médecins (enregistrement de : actions de participation, contenus, informations du spécialiste, dates et heures de participations),
- La prise en compte des besoins et attentes du patient,
- L'amélioration en toute sécurité du partage de l'information entre les professionnels de santé,
- La favorisation d'une meilleure prise de décision thérapeutique.

La Figure (3.8) montre une représentation de l'information dans l'ontologie de traçabilité. Cette ontologie fournit l'information nécessaire pour la décision du choix thérapeutique et la modélisation du travail collaboratif entre les professionnels. L'ontologie de traçabilité est liée

directement à l'ontologie thérapeutique et a quatre fonctions principales : (1) Décrire les différents types de dysfonctionnement dans les maladies cancéreuse; (2) Donner un aperçu de tous les examens et traitements associés au patient ; (3) Montrer les institutions et les médecins qui ont collaboré pour améliorer le traitement et les observations individuelles pour la prise de décision thérapeutique ; et (4) Montrer la participation (en enregistrant la date) au conduite thérapeutique.

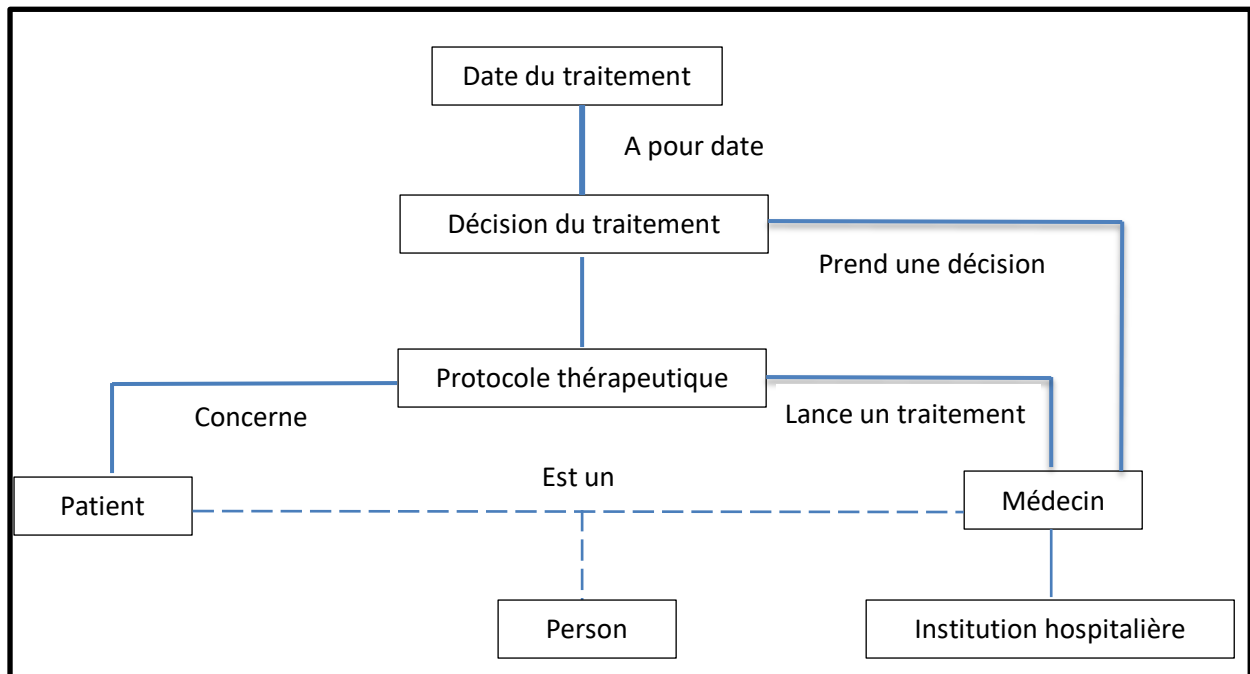


FIGURE 3.8– Représentation de l'information dans l'ontologie de traçabilité.

L'intégration de différentes sources d'information dans le cadre médical en utilisant des ontologies (Maladie–Traitement–Traçabilité) est rendue possible grâce à la communication entre les ontologies (Figure 3.9), c'est-à-dire l'information est récupérée et organisée grâce à la communication entre les ontologies permettant l'analyse de différents facteurs, vocabulaires et rôles. Cette information est présentée de manière à faciliter la compréhension de l'utilisateur (professionnels médicaux).

Dans le contexte de l'aide à la décision clinique, cette ontologie permet d'accéder à l'information et d'aider la communication entre les médecins. Elle fournit des connaissances pertinentes sur les protocoles thérapeutiques donnés aux patients. Elle a l'avantage de combiner les informations relatives à la conduite thérapeutique, de faciliter l'accès aux données du patient et d'identifier les indicateurs dans les données du patient pour améliorer la

détection probable des métastases. Principalement, cette ontologie fournit une terminologie standard pour les concepts médicaux qui peuvent faciliter l'intégration de données pour le travail collaboratif.

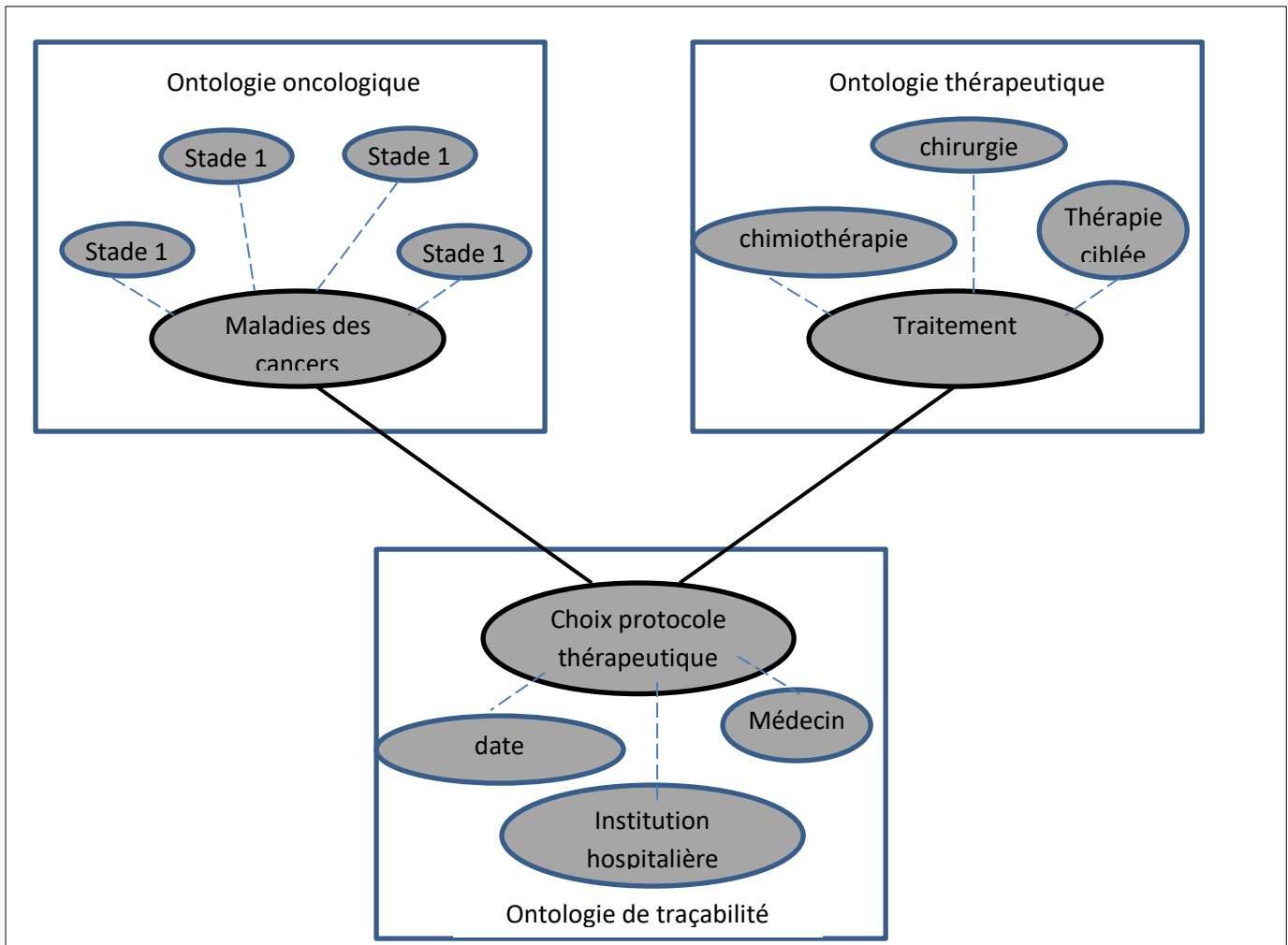


FIGURE 3.9– Diagramme conceptuel de l'ontologie de traçabilité

4.3 La visualisation d'ontologie de traçabilité

Dans ce graphe nous présentons tous les relations « is-a » de notre ontologie de traçabilité.



Figure 3.10 : Visualisation de traçabilité

Construction de diagramme de relations binaires

Cette tâche permet de représenter d'une manière graphique les différentes relations qui existent entre les divers concepts que ce soit de même ou de différentes hiérarchies (voir **figure 3.11**).

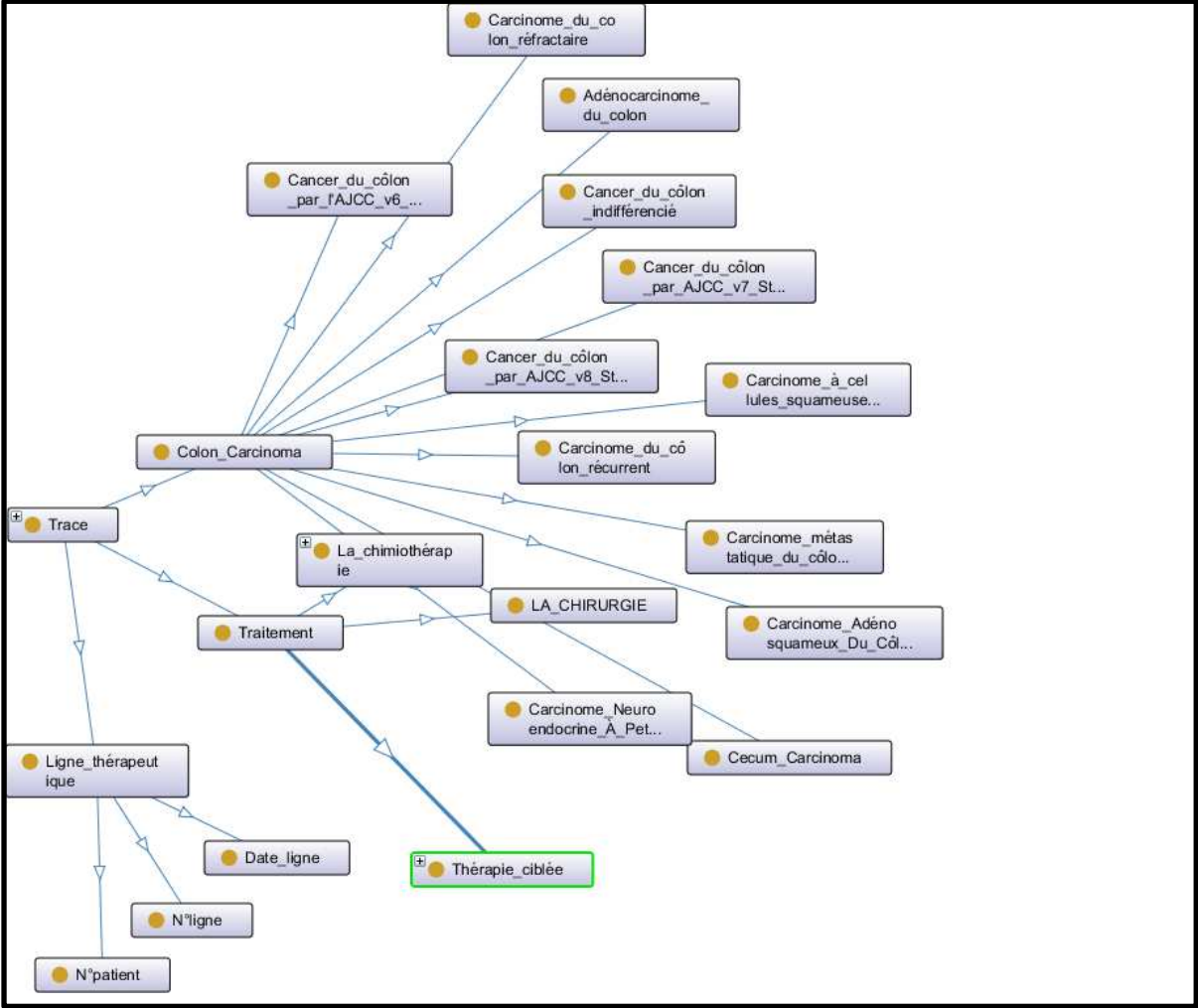


Figure 3.11 : diagramme de relations binaires

Conclusion

Dans ce chapitre nous essayons de définir tous les étapes de construction d'une ontologie de traçabilité à l'aide du langage protégé (owl), aussi bien les outils d'ontologie médicale.

Conclusion Générale

CONCLUSION GÉNÉRALE

L'objectif fixé au début de ce travail, était de proposer une nouvelle ontologie de traçabilité pour la conduite thérapeutique des patients cancéreux. Cette ontologie est définie à partir de protocoles médicamenteux appliqués sur les patients depuis la détection de la maladie. La traçabilité dans notre travail dépend de connaissances sur les maladies oncologiques ainsi que les différentes conduites thérapeutiques que ce soit chimiothérapie, hormonothérapie ...

L'étude bibliographique effectuée (cf. chapitre 2) a montré la quantité d'informations qu'on doit avoir pour que le médecin puisse prendre une décision envers le choix thérapeutique le plus adapté au cas de patient présenté.

Cette diversité nous a guidé à penser de construire une traçabilité thérapeutique à l'aide des ontologies afin de donner plus de facilité au médecin d'avoir l'historique thérapeutique des patients.

Au cours de ce travail, on a présenté une ontologie dans le domaine oncologique (domaine d'application « cancer de colon ») qui regroupe de nombreux concepts ainsi que les liens entre ces concepts (classes et sous classes), des propriétés et des instances.

L'ensemble du travail que nous avons réalisé au cours de ce mémoire est fondé sur deux phases. La première est consacrée à la description d'une ontologie dans le domaine oncologique (domaine d'application « cancer de colon ») qui regroupe de nombreux concepts ainsi que les liens entre ces concepts (classes et sous classes), des propriétés et des instances. Cette ontologie est définie en utilisant une base de connaissance à partir de l'IDO. La deuxième étape de la première phase vise à définir l'ontologie de traitements à partir de zéro en basant sur la collecte d'information et la description des concepts à partir des professionnels médicaux. La deuxième phase de notre travail considère le noyau de notre travail où on a essayé de définir une ontologie de traçabilité pour les conduites thérapeutiques à partir de deux ontologies définies précédemment.

Cette ontologie de traçabilité constitue une base pour la recherche d'information qui facilite et qui organise le travail du médecin oncologue lors de son suivi du patient cancéreux, cette traçabilité a comme but aussi de donner un historique détaillé sur les différentes conduites thérapeutiques appliquées sur le

patient pour n'importe quels oncologue traitant vue la durée de traitement importante et la probabilité d'avoir plus qu'un médecin traitant pour le même malade.

Perspectives :

Concernant les perspectives, ce travail peut être une source d'émergence des plusieurs idées, entre autres la création d'une platform pour favoriser les interactions entre médecins par un mécanisme solide (cloud computing basé) afin de faciliter l'accès à l'information médicale et thérapeutique comme elle peut nous permet d'avoir un moteur de recherche ontologique des connaissances cohérentes qui permet de créer de nouvelles connaissances pour le traitement des maladies oncologiques complexes.

Parmi les perspectives auxquelles nous aspirons le faire dans les futures travaux est l'intégration de la traçabilité dans toutes les spécialités liées à l'oncologie pour avoir une traçabilité complète sur tous les examens complémentaires (bilans biologique, anapath, imagerie), les différentes interventions chirurgicales ainsi que les différentes conduites thérapeutiques.

BIBLIOGRAPHIE

- [1] : Mémoire Pour l'obtention du diplôme de Magister en Informatique Thème : Construction d'une ontologie pour le domaine de la sécurité : Application aux agents mobiles Présenté par :Riad LEKHCHINE , Dirigé par : Pr. Zizette BOUFAÏDA
- [2] : Cours Ontologies et Web sémantique Présentée par : Dr Bourougaa-Tria Salima Email : bourougaasalima@gmail.com
- [3] : Chapitre 1 Les ontologies Pdf.
- [4]: Ontology Development 101: A Guide to Creating Your First Ontology , Natalya F. Noy and Deborah L. McGuinness Stanford University, Stanford, CA, 94305 noy@smi.stanford.edu and dlm@ksl.stanford.edu
- [5] : Modélisation et construction d'une ontologie pour le web sémantique , Ramahefy T.R1. – Rakotomiraho S.2– Rabeherimanana L. 3 Laboratoire de Recherche Systèmes Embarqués, Instrumentation et Modélisation des Systèmes et Dispositifs Electroniques (LR-SE-I-MSDE) Ecole Doctorale en Sciences et Techniques de l'Ingénierie et de l'Innovation(ED-STII) ,Université d'Antananarivo BP 1500 - Antananarivo 101 – Madagascar , 1ramahefy@yahoo.fr– 2rakmiraho@yahoo.fr – 3rabehlyliane@gmail.com
- [6] : Xavier Lacot - Introduction à OWL, un langage XML d'ontologies Web 2005.
- [7] : Conception et implémentation d'une ontologie médicale ;Cas : insuffisance cardiaque , Melle . BELABBES FATIMA. 15 Décembre 2015. Université Abou Bakr , Belkaid– Tlemcen
- [8] : Construction d'une ontologie pour le domaine de la sécurité : Application aux agents mobiles(mémoire de magister Riad LEKHCHINE.
- [9] : Gestion des Ontologies Médicales (G.O.M) Melle. HADJOUI Fatima Zohra 2012 , Université Abou Bakr Belkaid– Tlemcen
- [10] : Borst W. N. (1997). Construction of Engineering Ontologies. Center for Telematica and Information Technology, University of Tweenty, Enschede, NL.
- [11] : Bachimont, B., Engagement sémantique et engagement ontologique conception et réalisation d'ontologies en ingénierie des connaissances.
- [12] : Len at et Guha, 1990, Lenat, D. B. et Guha, R. V. (1990). Building large knowledge based

- systems. Reading, Massachusetts : Addison Wesley.
- [13] : GANGEMI A., PISANELLI DM., STEVE G. "An Overview of the ONIONS Project: Applying Ontologies to the Integration of Medical Terminologies. Data & Knowledge Engineering, 1999, vol. 31, No 2, pp 183-220.
- [14] : Construction d'une ontologie pour le domaine de la sécurité : Application aux agents mobiles(mémoire de magister Riad LEKHCHINE
- [15] : Ontologies et services aux patients : Application à la reformulation des requêtes
 « Radja Messai » Docteur de l'Université Joseph Fourier – Grenoble I
 (Spécialité : Ingénierie de la cognition, de l'interaction, de l'apprentissage et de la création)
- [16] :Bachimont, B., Engagement sémantique et engagement ontologique conception et réalisation d'ontologies en ingénierie des connaissances.
- [17] :Gestion des Ontologies Médicales (G.O.M) Melle. HADJOUI Fatima Zohra 2012. Université Abou Bakr Belkaid– Tlemcen
- [18] : Borst W. N. (1997). Construction of Engineering Ontologies. Center for Telematica and Information Technology, University of Twente, Enschede, NL.
- [19] : Lenat et Guha, 1990, Lenat, D. B. et Guha, R. V. (1990). Building large knowledge based systems. Reading, Massachusetts : Addison Wesley.
- [20]: Euzenat, 1995, Euzenat, J. (1995). Building consensual knowledge bases : context and architecture. Towards very large knowledge bases (actes 2nd international conference on building and sharing very large-scale knowledge bases (KBKS), Enschede (NL), (10-13 avril) 1995) IOS press, Amsterdam (NL), pages 143_155.
- [21] : Mémoire pour l'obtention du diplôme de doctorat Développement des ontologies multi-points de vue: une approche basée sur la logique de description Mounir HEMAM
- [22] : NOY N. et McGUINNESS D. L. "Ontology Development 101: A Guide to Creating Your First Ontology". Technical Report KSL-01-05Stanford: Knowledge Systems Laboratory, mars 2001.
- [23] : GANGEMI A., PISANELLI DM., STEVE G. "An Overview of the ONIONS Project: Applying Ontologies to the Integration of Medical Terminologies. Data & Knowledge Engineering, 1999, vol. 31, No 2, pp 183-220.
- [24] : The Unified Medical Language System (UMLS): Integrating Biomedical Terminology - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-various-subdomains-integrated-in-the-UMLS_fig1_8954995 [accessed 15 May, 2019]
- [25] : <https://sante-medecine.journaldesfemmes.fr/faq/13971-oncologie-definition>.
- [26] : Généralités en cancérologie, Formation Physiopathologie en ST2S,03/02/2010 Elisabeth MOERSCHEL – Section IMRT de Strasbourg.

- [27] : <https://www.elsan.care/fr/patients/oncologie-medicale>.
- [28] : Généralité en cancérologie ,Sensibilisation des étudiants IDE '1ère année à la prise en charge des patients atteints de cancer.
- [29] : Thèse pharmacie présentée et soutenue par Adama Samaké 2011-2012.
- [30] : Le cancer Johanne Marcotte et Renée Ouimet 2008 Bibliothèque nationale du Québec.
- [31] : Recherche Prévention et promotion du dépistage Actions pour les personnes malades, Société et politiques de santé (SPS) , LES CANCERS DU COLON ET DU RECTUM
- [32] : <http://www.disease-ontology.org/>
- [33]: BACHIMONT B., ISAAC A. & TRONCY R. (2002). Semantic commitment for designing ontologies : A proposal. In EKAW '02 : Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web, p. 114–121, London, UK : Springer-Verlag.
- [34]:Jiménez-Ruiz, E., Grau, B.C., Sattler, U., Schneider, T., Berlanga, R., 2008. Safe and economic reuse of ontologies: a logic-based methodology and tool support, in: Proceedings of the 5th European Semantic Web Conference on The Semantic Web: Research and Applications, ESWC'08. Springer-Verlag, Berlin, Heidelberg, pp. 185–199.
- [35]: Rastier, F., Cavazza, M., & Abeille, A. (1994). Semantique pour l'analyse- de la linguistique a l'informatique. Enseignement de Paris, France : Masson.
- [36]: Fernandez, M., Gomez-Pérez, A., Juristo, N. (1997) METHONTOLOGY: From Ontological Art Towards Ontological Engineering. Proceedings of the AAAI-97 Spring Symposium Series on Ontological Engineering, Stanford, CA, USA 1997, p 33-40.
- [37]: Extraction de connaissance à partir des données biomédicales guidée par une ontologie : Application au dépistage du cancer de sein. UNIVERSITE DE M'SILA
FACULTE DES MATHEMATIQUES ET DE L'INFORMATIQUE
Département de l'informatique- Mounira LAKAB

