

1. Introduction

L'exploitation de connaissances en informatique a pour objectif de ne plus faire manipuler en aveugle des informations à la machine mais de permettre un dialogue (une coopération) entre le système et les utilisateurs. Alors, le système doit avoir accès non seulement aux termes utilisés par l'être humain mais aussi à la sémantique qui leur est associée, afin qu'une communication efficace soit possible. Actuellement, la connaissance visée par ces ontologies est un sujet de recherche populaire dans diverses communautés (l'ingénierie des connaissances, la recherche d'information, le traitement du langage naturel, les systèmes d'information coopératifs, l'intégration intelligente d'information et la gestion des connaissances). Elles offrent une connaissance partagée sur un domaine qui peut être échangée entre des personnes et des systèmes hétérogènes. Elles ont été définies en intelligence artificielle afin de faciliter le partage des connaissances et leur réutilisation. La définition explicite du concept ontologie soulève un questionnement qui est tout à la fois d'ordre philosophique, épistémologique, cognitif et technique.

2. Définitions de l'ontologie

Ontologie est une branche de la métaphysique qui s'intéresse à l'existence, à l'être en tant qu'être et aux catégories fondamentales de l'existant. En effet, ce terme est construit à partir des racines grecques « ontos » qui veut dire ce qui existe, l'être, l'existant, et « logos » qui veut dire l'étude, le discours, d'où sa traduction par « l'étude de l'être » et par extension de l'existence. [21]

Dans la philosophie classique, l'ontologie correspond à ce qu'Aristote appelait la Philosophie première (protè philosopha), c'est-à-dire la science de l'être en tant qu'être, par opposition aux philosophies secondes qui s'intéressaient, elles, à l'étude des manifestations de l'être (les existants). [24].

Ontologie : partie de la métaphysique qui s'attache à l'étude ou à la théorie de l'être dans son essence, indépendamment des phénomènes de son existence. [7].

- Dans le cadre de l'intelligence artificielle, **Neeches** et ses collègues [23] furent les premiers à proposer une définition à savoir :
« *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* »
- En 1993, **Gruber** [14] propose la définition suivante : « *Spécification explicite d'une conceptualisation* ».
- Cette définition a été modifiée légèrement par **Borst** [4] comme « *spécification formelle d'une conceptualisation partagée* ».
- Ces deux dernières définitions sont regroupées dans celle de **Studer** [26] comme « *spécification formelle et explicite d'une conceptualisation partagée* ».
 - ✓ *Formelle* : l'ontologie doit être lisible par une machine, ce qui exclut le langage naturel.
 - ✓ *Explicite* : la définition explicite des concepts utilisés et des contraintes de leurs utilisations.
 - ✓ *Conceptualisation* : le modèle abstrait d'un phénomène du monde réel par identification des concepts clefs de ce phénomène.
 - ✓ *Partagée* : l'ontologie n'est pas la propriété d'un individu, mais elle représente un consensus accepté par une communauté d'utilisateurs.
- Pour **Guarino & Giaretta** [16] « *une ontologie est une spécification rendant partiellement compte d'une conceptualisation* ». **Swartout** et ses collègues [27] la définissent comme suit : « *une ontologie est un ensemble de termes structurés de façon hiérarchique, conçue afin de décrire un domaine et qui peut servir de charpente à une base de connaissances* ».
- La même notion est également développée par **Gomez** [13] comme :
« *une ontologie fournit les moyens de décrire de façon explicite la conceptualisation des connaissances représentées dans une base de connaissances* ».

- **GRUBER** [15]: « *In the context of computer and information sciences, an ontology defines a set of representational primitives with which to model a domain of knowledge or discourse. The representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members)*».

3. Cycle de vie

Puisque les ontologies sont 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 précisé. Dans ce contexte, les activités liées aux ontologies sont, d'une part, des activités de gestion de projet (planification, contrôle, assurance qualité), et d'autre part, des activités de développement (spécification, conceptualisation, formalisation) ; s'y ajoutent des activités transversales de support telles que l'évaluation, la documentation, la gestion de la configuration [2].

Un cycle de vie inspiré du génie logiciel est proposé dans [6] et [10]. Nous l'avons adapté à nos besoins et proposons notre vision du cycle de vie d'une ontologie (**figure I.1**). Il comprend une étape initiale de détection et de spécification des besoins qui permet notamment de circonscrire précisément le domaine de connaissances, une étape de conception qui se subdivise en trois phases, une étape de déploiement et de diffusion, une étape d'utilisation, une étape incontournable, d'évaluation, et enfin, une sixième étape consacrée à l'évolution et à la maintenance du modèle. Après chaque utilisation significative, l'ontologie et les besoins doivent être réévalués et 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.

haute abstraction tels que les entités, les évènements, les états, les processus, les actions, le temps, l'espace, les relations, les propriétés...

4.3. Ontologie Générique (Generic ontology)

Elle est appelée également noyau ontologique, modélise des connaissances moins abstraites que celles véhiculées par l'ontologie de haut niveau mais assez générales néanmoins pour être réutilisées à travers différents domaines. Cette ontologie inclut un vocabulaire relatif aux choses, évènements, temps, espace, causalité, comportement, fonction, etc.

4.4. Ontologie du domaine (Domain ontology)

Cette ontologie exprime des conceptualisations spécifiques à un domaine, elle est pour plusieurs applications de ce domaine. Elle fournit les concepts et les relations permettant de couvrir les vocabulaires, activités et théories de ces domaines. Selon **Mzoguchi** [22], l'ontologie du domaine caractérise la connaissance du domaine ou la tâche est réalisée. Par exemple, dans le contexte du e-Learning, le domaine peut être celui de formation.

C'est ce type qui nous sera utile pour définir les objectifs ciblés de notre projet.

4.5. Ontologie de Taches (Task ontology)

L'ontologie de tâches fournit un vocabulaire systématisé des termes employés pour résoudre des problèmes liés aux tâches qui peuvent être ou non du même domaine. Elle fournit un ensemble de termes au moyen desquelles nous pouvons décrire généralement comment résoudre un type de problèmes. Elle inclut des noms, des verbes et des adjectifs génériques dans les descriptions de tâches.

4.6. Ontologie d'application (Application ontology)

C'est l'ontologie la plus spécifique, elle contient des concepts dépendants d'un domaine et d'une tâche particuliers, elle est spécifique et non réutilisable. Ces concepts correspondent souvent aux rôles joués par les entités du domaine lors de l'exécution d'une certaine activité [17]. Il s'agit donc ici de mettre en relation les concepts liés à une tâche particulière de manière à en décrire l'exécution.

5. Composantes d'une ontologie

Comme nous l'avons abordé, les ontologies fournissent un vocabulaire commun d'un domaine et définissent la signification des termes et des relations entre elles. La connaissance dans les ontologies est principalement formalisée en utilisant les cinq types de composants [12] à savoir : **concepts** (ou classes), **relations** (ou propriétés), **fonctions**, **axiomes** (ou règles) et **instances** (ou individus).

- Les concepts, aussi appelés termes ou classe de l'ontologie, correspondent aux abstractions pertinentes d'un segment de la réalité (le domaine du problème) retenus en fonction des objectifs qu'on se donne et de l'application envisagée pour l'ontologie ;
- Les relations traduisent les associations (pertinentes) existant entre les concepts présents dans le segment analysé de la réalité. Ces relations incluent les associations suivantes :
 - Sous classes de (généralisation-spécialisation) ;
 - Partie de (agrégation ou composition) ;
 - Associe à ;
 - Instance de, etc.

Ces relations nous permettent d'apercevoir la structuration et l'interrelation des concepts, les uns par rapport aux autres ;

- Les fonctions constituent des cas particuliers de relations, dans laquelle un élément de la relation, (le nième) est défini en fonction des N-1 éléments précédents ;
- Les axiomes constituent des assertions, acceptées comme vraies, à propos des abstractions du domaine traduites par l'ontologie.
- Les instances constituent la définition extensionnelle de l'ontologie ; ces objets véhiculent les connaissances (statiques, factuelles) à propos du domaine du problème.

6. Les méthodes d'ingénierie ontologique

Il existe une multitude de méthodes d'ingénierie ontologique mais l'absence de directives structurées et communes ralentisse le développement d'ontologie à l'intérieur et entre les équipes, l'extension de n'importe quelle ontologie, la possibilité de réutilisation de l'ontologie. On entend par méthodologie, les procédures de travail, les étapes, qui décrivent le pourquoi et le comment de la conceptualisation puis de l'artefact construit. Dans la suite nous allons présenter que deux méthodes de l'état d'ontologies :

6.1. La méthode de Bachimont

Cette méthode propose de contraindre l'utilisateur à un engagement sémantique en introduisant une normalisation sémantique des termes manipulés dans l'ontologie. La méthode de normalisation suit trois étapes [1]

- **Normalisation sémantique** : l'utilisateur doit choisir les termes du domaine et les normaliser en explicitant leurs propriétés et en exprimant les identités et les différences dans leur voisinage proche. La place d'une *notion* dans l'ontologie doit être justifiée par rapport à la communauté et la différence avec le père et la fratrie.
- **Formalisation des connaissances** : Cette étape consiste à désambiguïser les notions de l'ontologie référentielle obtenue par l'étape précédente et choisir leurs sens pour un domaine spécifique. Cela peut nécessiter la création de nouveaux concepts, l'ajout de propriétés et d'axiomes.
- **Opérationnalisation des connaissances** : Le système utilise un langage opérationnel de représentation de connaissances qui possède les caractéristiques nécessaires pour répondre aux besoins exprimés lors de la spécification du système.

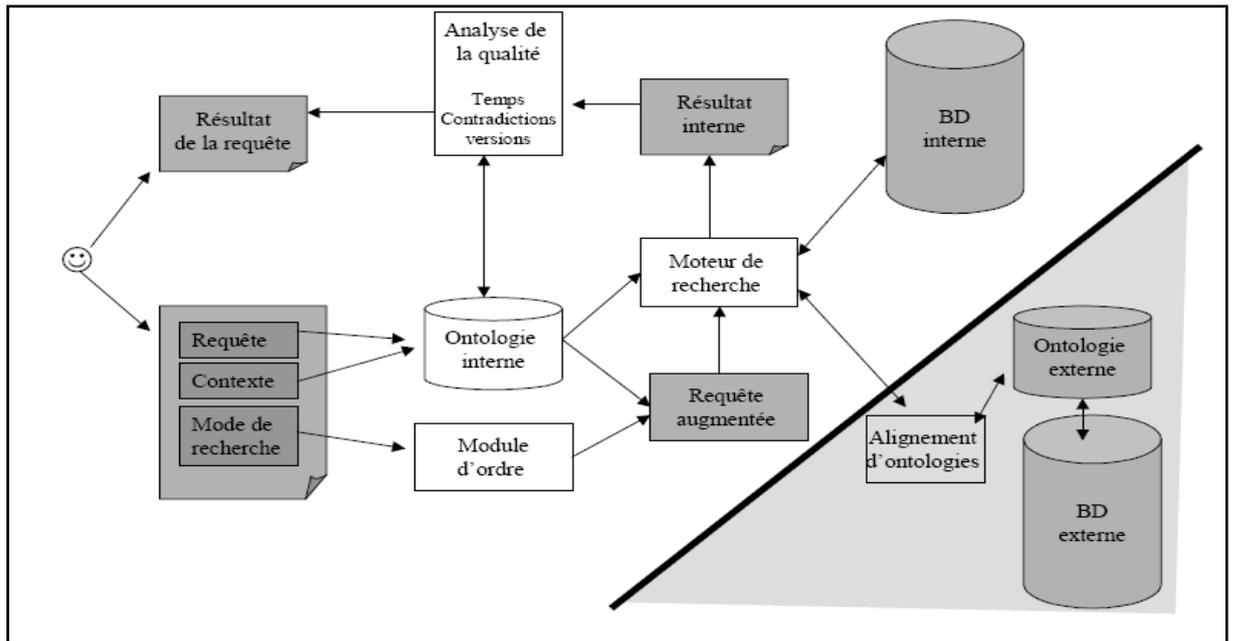


Figure (I. 2) : Processus de la méthode Bachimont [1]

6.2. La méthode METHONTOLOGY

C'est une méthodologie mise au point par l'équipe du laboratoire de l'intelligence artificielle de l'Université polytechnique de Madrid. Cette méthode inclut :

- L'identification du processus de développement de l'ontologie.
- Le cycle de vie basé sur l'évolution de prototypes.
- Les techniques de gestion de projet (planification, assurance qualité), de développement (spécification, conceptualisation, formalisation, implémentation, maintenance) et des activités de support (intégration, évaluation, documentation).

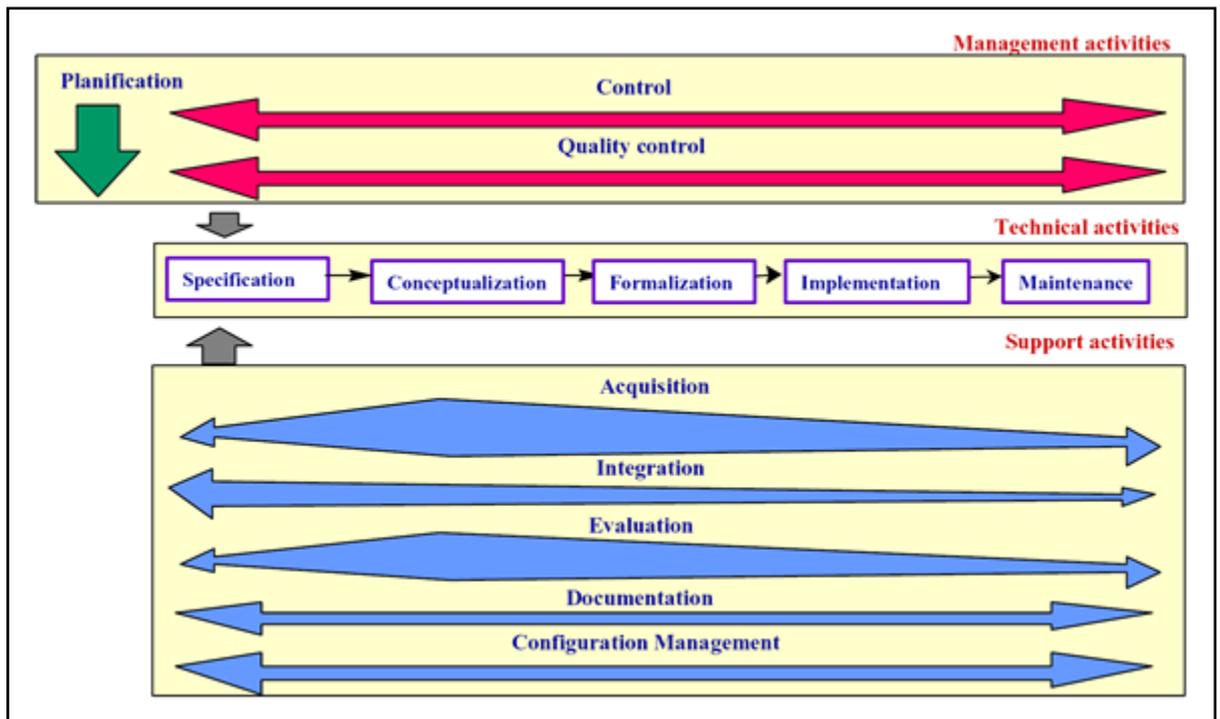


Figure (I. 3) : Processus de développement d'ontologie de METHONTOLOGY [5].

7. Les outils de construction d'ontologie

On distingue deux familles d'outils : les outils de construction d'ontologie dépendants de formalisme de représentation et les outils de construction d'ontologie indépendants de formalisme de représentation.

7.1. Les outils dépendants de formalisme de représentation

- **Ontolingua :** Ontolingua est un serveur d'édition d'ontologies. Il utilise des classes, des relations, des fonctions, des instances et des axiomes pour décrire une ontologie. Une relation peut contenir des propriétés nécessaires (contraintes) ou nécessaires et suffisantes qui définissent la relation. En plus le serveur Ontolingua offre la possibilité d'intégrer les ontologies Ontolingua, ce qui permet la construction modulaire des ontologies. [8]

- **OntoSaurus** : OntoSaurus de l'Information Science Institute de l'Université de Southern California est composé de deux modules : un serveur utilisant LOOM comme langage de représentation des connaissances, et 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.
- **WebOnto** : WebOnto du Knowledge Media Institute de l'Open University, est une application Web pour naviguer et développer collaborativement les ontologies. Il supporte la navigation collaborative, la création et l'édition d'ontologies sur le Web. Les ontologies WebOnto sont implémentées dans le langage OCML. WebOnto distingue quatre types d'ontologies : ontologie de domaine, ontologie de tâche, ontologie de méthode, et ontologie d'application.
- **OilEd** : OilEd (Oil Editor) est un éditeur d'ontologies utilisant le formalisme OIL. Il est essentiellement dédié à la construction de petites ontologies dont on peut ensuite tester la cohérence à l'aide de FACT, un moteur d'inférences bâti sur OIL.

7.2. Les outils indépendants de formalisme de représentation

- **Protégé 2000** : est une interface modulaire permettant l'édition, la visualisation, le contrôle d'ontologie, l'extraction d'ontologies à partir de sources textuelles, et la fusion semi-automatique d'ontologies. Le modèle de connaissances sous-jacent à protégé 2000 est issu du modèle des frames et contient des classes, des slots (propriétés) et des facets (valeurs des propriétés et contraintes), ainsi que des instances des classes et des propriétés. Il autorise la définition de méta-classes, dont les instances sont des classes, ce qui permet de créer son propre modèle de connaissances avant de bâtir une ontologie.

- **ODE et WebOde** : L'outil ODE (Ontology Design Environment) permet de construire des ontologies au niveau connaissance, comme le préconise la méthodologie METHONTOLOGY. L'utilisateur construit son ontologie dans un modèle de type frame, en spécifiant les concepts du domaine, les termes associés, les attributs et leurs valeurs, les relations de subsomption.
- **OntoEdit** : OntoEdit (Ontology Editor) est également un environnement de construction d'ontologies indépendant de tout formalisme. Il permet l'édition des hiérarchies de concepts et de relations et l'expression d'axiomes algébriques portant sur les relations, et de propriétés telles que la généralité d'un concept. 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.

8. Domaines d'applications des ontologies

8.1. Système d'information

L'intégration d'une ontologie dans un système d'information vise à réduire, voire éliminer, la confusion conceptuelle et terminologique à des points clés du système, et à tendre vers une compréhension partagée pour améliorer la communication, le partage, l'interopérabilité et le degré de réutilisation possible, ce qui permet de déclarer formellement un certain nombre de connaissances utilisées pour caractériser les informations gérées par le système, et de se baser sur ces caractérisations et la formalisation de leur signification pour automatiser des tâches de traitement de l'information.

L'ontologie retrouve maintenant dans une large famille de systèmes d'information. Elle est utilisée pour :

- Décrire et traiter des ressources multimédia ;
- Assurer l'interopérabilité d'applications en réseaux ;
- Piloter des traitements automatiques de la langue naturelle ;
- Construire des solutions multilingues et interculturelles ;
- Permettre l'intégration des ressources hétérogènes d'information ;
- Vérifier la cohérence de modèles ;
- Permettre les raisonnements temporel et spatial ;
- Faire des approximations logiques ; etc.

Ces utilisations des ontologies se retrouvent dans de nombreux domaines d'applications tel que :

- Intégration d'information géographique ;
- Gestion de ressource humaine ;
- Aide à l'analyse en biologie, suivi médicale informatisé ;
- Commerce électronique ;
- Enseignement assisté par ordinateur ;
- Bibliothèque numériques ;
- Recherche d'informations.

8.2. Web sémantique

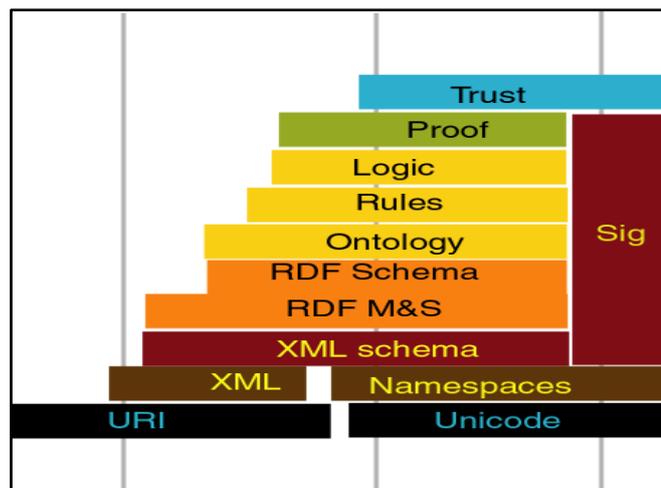


Figure (I. 4) : Pyramide du Web sémantique

Un courant particulièrement prometteur pour l'expansion des systèmes à base d'ontologies est celui du Web sémantique. Il s'agit d'une extension du Web actuel, dans laquelle l'information se voit associée à un sens bien défini, améliorant la capacité des logiciels à traiter l'information disponible sur le Web. L'annotation des ressources d'information du Web repose sur des ontologies, elles sont aussi disponibles et échangées sur le Web. Grâce au Web sémantique, l'ontologie a trouvé un jeu de formalismes standards à l'échelle mondiale, et s'intègre dans de plus en plus d'applications Web, sans même que les utilisateurs ne le sachent. Cela se fait au profit des logiciels qui à travers les ontologies et les descriptions qu'elles permettent, peuvent proposer de nouvelles fonctionnalités exploitant les effets d'échelles du Web pour en améliorer les effets.

9. Conclusion

Dans ce chapitre, nous avons rappelé la notion des ontologies, leurs différents types et composants, et quelques outils de manipulation.

A cet effet, nous allons présenter dans le chapitre suivant le langage OWL, le langage le plus utilisé pour la représentation des ontologies.