

Chapitre 1

Les ontologies

Chapitre 1: Les ontologies

1. Introduction

WordNet étant une ontologie générique pour la langue, dans ce chapitre nous allons donc aborder la notion d'ontologie et décrire les démarches connues pour leur construction.

2. Définitions

Ontologie

Une ontologie est une spécification explicite d'une conceptualisation. Le terme « conceptualisation » fait référence à un système de concepts. L'expression « spécification explicite » signifie que la conceptualisation est représentée dans un langage (langue naturelle ou langage formel).

Le terme «ontologie» est un emprunt à la philosophie. Il désigne (Petit ROBERT, 1979) : la partie de la métaphysique qui s'applique à l'être en tant qu'être, indépendamment de ses déterminations (voir figure 1.1).

3. Concept

Un concept est une entité structurée. Il peut se définir comme une entité composée de trois éléments distincts :

- ✓ Le(s) terme(s) exprimant le concept en langue.
- ✓ La signification du concept, appelée également « notion » ou « intension » du concept.
- ✓ Le(s) objet(s) dénotés par le concept, appelé(s) également « réalisation » ou « extension » du concept.

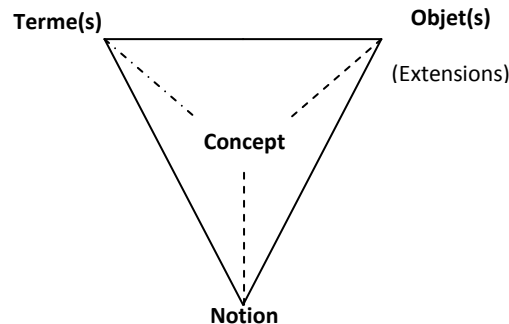


Figure 1.1 : Le triangle sémantique (Ogden et Richards, 1923)

4. Les relations entre les concepts

Il existe plusieurs types de relations entre les concepts :

Généralisation Spécialisation

Le concept CONCEPT1 généralise (ou subsume du terme subsumption) le concept CONCEPT2 (respectivement le concept CONCEPT2 spécialise le concept CONCEPT1) si et seulement si l'extension du concept CONCEPT2 est incluse dans l'extension du concept CONCEPT1.

Les liens de subsumption constituent le principal mode de structuration d'une ontologie.

Autres types de relations

Deux termes sont dits synonymes s'ils ont même sens.

Deux termes sont homographes (homonymes) s'ils s'écrivent de la même façon mais ont des sens différents, c'est-à-dire qu'ils expriment des notions différentes.

5. Les objectifs de l'ontologie

L'ontologie peut servir les objectifs de :

- ✓ La communication (humains et organisations) : Dans l'ontologie, il n'y a jamais deux termes ayant la même sémantique. Cette situation se produit souvent si l'on utilise un langage nature pour la communication.
- ✓ L'interopérabilité (machine et systèmes) : l'ontologie sert à définir le format d'échange entre les systèmes.
- ✓ L'ingénierie des systèmes : l'ontologie peut servir divers aspects du développement des systèmes d'information. Elle assiste le processus de construction de la spécification des systèmes .Elle soutient aussi l'automatisation du processus de vérification de la fiabilité des systèmes.

6. Les types d'ontologies

Il existe plusieurs types d'ontologie. Nous allons citer les catégories suivantes (voir figure 1.2) :

Les ontologies du domaine : elles sont appelées de la sorte parce qu'elles expriment des conceptualisations spécifiques à un domaine. Elles rendent compte du vocabulaire d'un domaine spécifique au travers de concepts et de relations qui modélisent les principales activités, les théories et les principes de base du domaine en question. La plupart des ontologies existantes sont des ontologies du domaine, elles sont réutilisables pour plusieurs applications concernant le domaine pour lequel elles ont été créées car elles ont été conçues de façon aussi indépendante que possible du type de manipulations qui vont être opérées sur ces connaissances.

Les ontologies applicatives (ou ontologies d'application) : Sont les ontologies les plus spécifiques, elles contiennent les connaissances requises pour une application particulière et ne sont pas réutilisables. Elles peuvent en outre inclure une ontologie de domaine.

Les ontologies génériques ou ontologies de haut niveau (upper ontology) : Elles 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.

Les ontologies de représentation : 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.

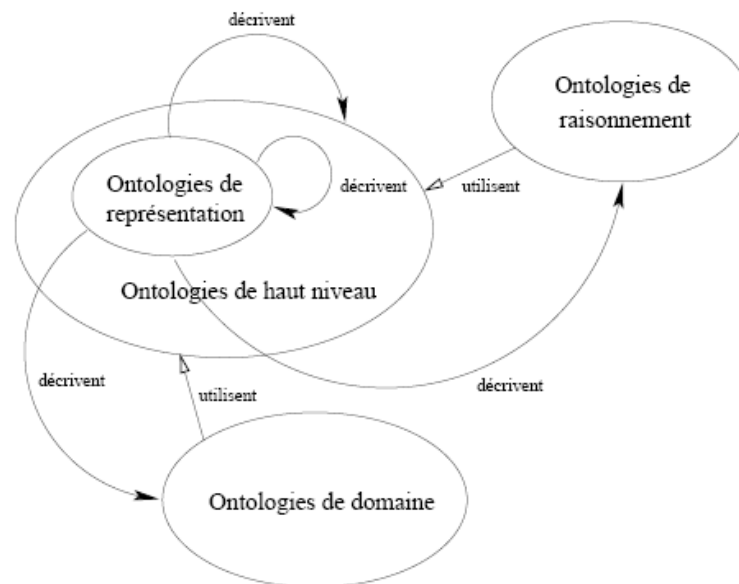


Figure 1.2 : Différents types d'ontologies [MIZOGUCHI, 1997]

7. Le cycle de vie des ontologies

Le cycle de vie des ontologies est inspiré du génie logiciel, Il comprend une étape initiale d'évaluation des besoins, une étape de construction, une étape de diffusion, et une étape d'utilisation. Après chaque utilisation significative, l'ontologie et les besoins sont réévalués et l'ontologie peut être étendue et, si nécessaire, en partie reconstruite.

La phase de construction peut être décomposée en 3 étapes : conceptualisation, ontologisation, et opérationnalisation (voir figure 1.3).

L'étape d'ontologisation peut être complétée d'une étape d'intégration au cours de laquelle une ou plusieurs ontologies vont être importées dans l'ontologie à construire.

La conceptualisation : identification des connaissances contenues dans un corpus représentatif du domaine .ce travail doit être mène par un expert du domaine, assisté par un ingénieur de la connaissance.

L'ontologisation : formalisation, autant que possible, du modèle conceptuel obtenu à l'étape précédente .ce travail doit être mène par l'ingénieur de la connaissance assisté de l'expert du domaine.

L'opérationnalisation : transcription de l'ontologie dans un langage formel et opérationnel de représentation de connaissances. Ce travail doit être mène par l'ingénieur de la connaissance.

Il est à noter que ce processus n'est pas linéaire et que de nombreux allers -retours sont nécessaires pour bâtir une ontologie opérationnelle adaptée aux besoins. Ce modèle de construction d'ontologie est ascendant, c'est -à-dire que l'on part des connaissances à représenter, pour obtenir à une représentation formelle .Une construction descendante est possible, qui consiste à choisir un modèle opérationnel de représentation, en fonction de l'objectif d'utilisation de l'ontologie, puis à instancier ce modèle avec les connaissances du domaine.

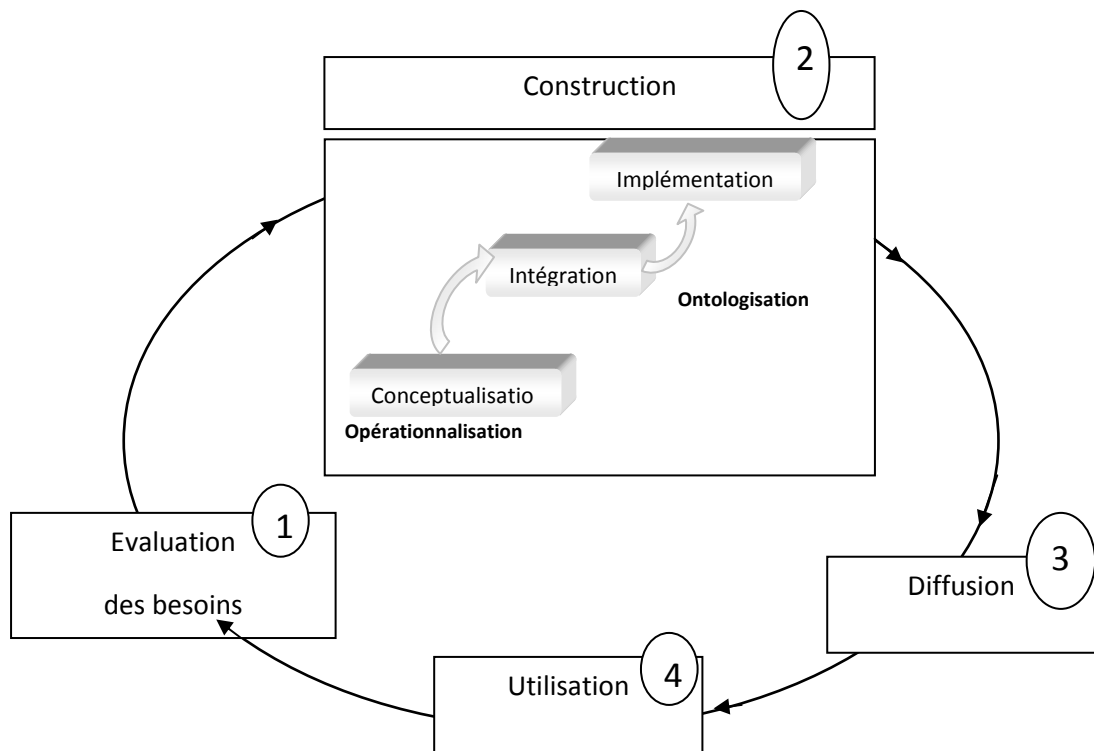


Figure 1.3 : Le cycle de vie d'une ontologie [Furst, 2002]

La construction d'une ontologie passe donc par plusieurs étapes, la figure 1.4 suivante schématise ces étapes. [Furst, 2002]

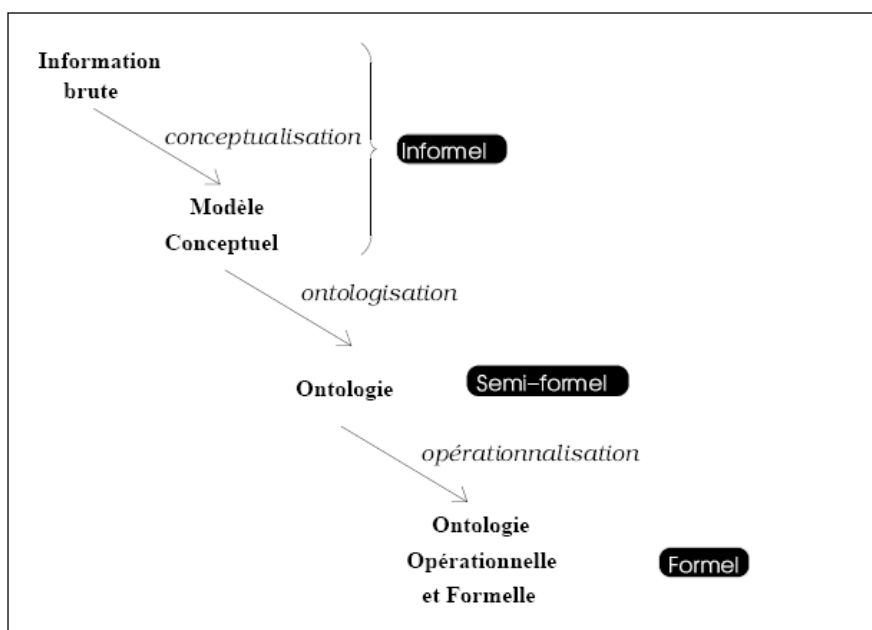


Figure 1.4 : Construction d'une ontologie opérationnelle. [Furst, 2002]

8. Méthodologie et outils de construction d'ontologies

Il existe des méthodologies et des outils pour créer des ontologies. Nous allons décrire les plus importants [Benaïssa, 2010].

Les méthodologies

- ✓ 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.
- ✓ La méthodologie On-To-Knowledge¹ (OTK) : La méthodologie On-To- Knowledge propose de construire une ontologie en tenant compte de comment l'ontologie va être utilisée par l'application plus tard (cette méthodologie est très dépendantes de l'application).
- ✓ METHONTOLOGY : Cette méthodologie a été développée au sein du groupe d'ontologie à l'université polytechnique de Madrid.

Les outils

- ✓ TERMINAE, intègre des outils linguistiques, elle permet la visualisation des résultats des extracteurs de candidats-termes Lexter et/ou Syntex. Ces concepts doivent ensuite être triés par un expert et organisés hiérarchiquement, puis la sémantique du domaine est précisée à travers des axiomes.
 - ✓ DOE (DIFFERENTIAL ONTOLOGY EDITOR) est un éditeur d'ontologie.
-

- ✓ ODE (ONTOLOGY DESIGN ENVIRONMENT), construction des ontologies au niveau connaissances (Méthodologie METHONTOLOGY).
- ✓ ONTOEDIT (ONTOLOGY EDITOR), est indépendant de tout formalisme. Il permet l'édition des hiérarchies de concepts et de relations et l'expression d'axiomes algébriques.
- ✓ PROTEGE2000 Il permet l'édition, le contrôle, la visualisation et l'extraction d'ontologie à partir des textes.
- ✓ ONTOLINGUA qui constitue une extension du langage KIF (KNOWLEDGE INTERCHANGE FORMAT). Ce serveur d'édition permet la fusion d'ontologie. L'ontologie est immédiatement représentée dans un formalisme.
- ✓ OILED (OIL EDITOR) est un éditeur d'ontologie s'inspirant du formalisme OIL. Un éditeur de petites ontologies avec un moteur d'inférence pour tester la cohérence de l'ontologie construite.

9. Conclusion

Dans ce chapitre, nous avons essayé d'éclaircir la notion d'ontologie. Pour cela nous sommes partis des origines philosophiques du terme, ensuite définir son sens en ingénierie des connaissances. Ensuite, nous avons étudié la manière de concevoir et de réaliser une ontologie en énumérant ses composants et en proposant les méthodes et les outils pour leur construction.

Le chapitre suivant sera consacré à la description de la base lexicale WordNet, qui est une ontologie lexicale.