Intelligibilité du numérique

N° 1 | 2020
Faire dialoguer les disciplines via l'indexation des connaissances : la recherche interdisciplinaire en débat

Les ontologies informatiques au service de la communication interdisciplinaire : l'interopérabilité sémantique

Charles BODON *Doctorant*
UFR 10
EXeCO (UMR 8103)
Pantheon-Sorbonne University, Paris
Jean CHARLET *Chargé de recherche*
LIMICS UMRS 1142
Assistance Publique - Hôpitaux de Paris

Édition électronique :
URL :
https://intelligibilite-numerique.numerev.com/articles/revue-1/2163-les-ontologies-informatiques-au-service-de-la-communication-interdisciplinaire-l-interoperabilite-semantique
DOI : numerev_1693
Date de publication : 13/05/2020

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : BODON, Ch., CHARLET, J. (2020) Les ontologies informatiques au service de la communication interdisciplinaire : l'interopérabilité sémantique. *Intelligibilité du numérique*, (1). https://doi.org/10.34745/numerev_1693

La communication interdisciplinaire rencontre divers problématiques : l'hétérogénéités des informations (sémantique, standardisation des formats), les confusions linguistiques (intentions dans les énoncés, polysémie des mots) ou encore épistémologiques (expertises divergentes, terminologies inadéquates). Nous proposons pour favoriser l'intercommunication entre différents domaines de spécialité, l'emploi des ontologies informatiques : des artefacts informatiques permettant une représentation des concepts au sein d'un domaine de spécialité. Autour de la notion « d'interopérabilité sémantique », la capacité pour une machine et des acteurs de communiquer ensemble, nous viserons à montrer en quoi les ontologies peuvent permettre à des acteurs de différentes disciplines de clarifier les notions propres à leurs domaines en fonction d'une analyse de leurs usages des mots. Nous tâchons ici de mettre en avant la capacité des ontologies de traiter le langage selon différents niveaux du sens d'un mot (du plus concret et appliqué, au plus abstrait et conceptuel), afin de mettre en perspective ce en quoi cette analyse basée sur les textes d'un corpus, permet de rendre compte de, si ce n'est un état du réalisme d'une discipline, du moins, servir d'outils pour aider les acteurs à normaliser leurs pratiques. En ce sens, les ontologies informatiques peuvent être considérées et utilisées comme autant de bases terminologiques servant comme outils de références pour faciliter la communication interdisciplinaire.

Interdisciplinary communication encounters various issues: the heterogeneity of information (semantics, standardization of formats), linguistic confusion (intentions in statements, polysemy of words) or even epistemological (divergent expertise, inadequate terminologies). To promote intercommunication between different fields of specialty, we propose the use of computer ontologies: computer artefacts allowing a representation of concepts within a field of specialty. Around the notion of "semantic interoperability", the ability for a machine and actors to communicate together, we will aim to show how ontologies can allow actors from different disciplines to clarify the concepts specific to their fields according to 'an analysis of their word usage. We are trying here to highlight the capacity of ontologies to process language according to different levels of the meaning of a word (from the most concrete and applied, to the most abstract and conceptual), in order to put into perspective what this analysis based on the texts of a corpus, makes it possible to account for, if not a state of the realism of a discipline, at least, to serve as tools to help the actors to standardize their practices. In this sense, computer ontologies can be considered and used as terminological bases serving as reference tools to facilitate interdisciplinary communication.

Mots-clés :
Interopérabilité sémantique, Ontologie informatique, Langage, Concept, Contexte, Interdisciplinarité, Terminologie, Interdisciplinarity, Language, Context, Semantic interoperability,, Computer ontology, Concept, Terminology

1. Introduction

Nous proposons un axe épistémologique pour la recherche interdisciplinaire dans la présentation et l'usage d'outils informationnels que sont les ontologies informatiques. Pour permettre à plusieurs domaines de spécialité de communiquer ensemble, il s'agit ici de déterminer les conditions pour établir une compréhension entre ces domaines, mais aussi, une base théorique permettant cette compréhension.

En ce sens, nous introduisons les ontologies informatiques en tant que modèles de représentation des connaissance d'un domaine de spécialité, permettant de clarifier la communication des acteurs à propos de ce domaine.

Les ontologies informatiques pouvant faire l'inventaire des connaissances d'une discipline, l'objectif de cette présentation est de questionner la notion « d'interopérabilité sémantique » entre ontologies informatiques. L'interopérabilité sémantique est définie comme : la disposition pour plusieurs agents ou systèmes de pouvoir échanger et communiquer tout en préservant le contenu et la qualité de l'information.

L'idée principale sous-jacente à cette mise en perspective entre communications entre systèmes informatiques et communications interdisciplinaires est celle-ci : montrer qu'une fois l'interopérabilité entre les outils informatiques obtenue, ces ontologies peuvent servir, si ce n'est à désambiguïser totalement le langage naturel de ses polysémies, tout du moins permettre d'obtenir une base terminologique de référence pour les acteurs afin de communiquer à propos de leurs disciplines respectives.

Nous présentons les différents types d'ontologies, ainsi que leur usage spécifique traitant des objets et termes les plus concrets, aux plus abstraits. Cette typologie consiste à mettre en avant la capacité des ontologies informatiques de situer et représenter plusieurs niveaux du discours. A cela, nous présentons également les types d'hétérogénéités sémantiques que les ontologies peuvent rencontrer, et qui coïncident avec les confusions linguistiques communes à toutes tentatives d'alignements du discours de plusieurs disciplines différentes.

Nous visons également à problématiser le rapport à la connaissance telle qu'elle peut se manifester sous un aspect contextuel, tout en proposant une approche théorique de cet aspect fondée sur l'analyse de la signification des mots selon leur usage. En ce sens, nous rapportons des travaux sur les terminologies fondées sur une approche « sémasiologique », c'est-à-dire élaborant ses corpus à partir de l'occurrence d'un mot dans un contexte précis.

Nous chercherons également à proposer des réponses à deux circularités que nous identifions concernant la communication interdisciplinaire.

Une circularité « définitionnelle » au niveau des primitives du langage, on verra qu'on ne peut établir de terminologie sans partir d'une base lexicale. Or comment la définir si, comme nous allons le voir, pour définir un terme on doit dans une régression faire toujours appel à un autre ? Pour répondre à cela, nous proposerons une approche focalisée sur la récurrence des mots pour la constitution de corpus, laquelle permet de figer le contexte linguistico-pratique d'une discipline, ceci afin d'en déterminer les degrés d'usage et le genre du corpus à partir de son support textuel numérique.
Une circularité « praxéologique » concernant la nécessité d'employer l'interdisciplinarité pour définir l'interdisciplinarité. Cette circularité remet en cause le modèle référentialiste traditionnel pour la normalisation des pratiques disciplinaires : sur quelle base définir les pratiques si celles-ci sont interdépendantes et qu'il n'en existe pas une suffisamment monolithique pour toutes les englober ?
Nous proposerons pour répondre à ce deuxième problème, une conception « instrumentale » des ontologies qui s'accorde avec la démarche sémasiologique : en ce qu'il est nécessaire de trouver dans les ontologies le moyen de les considérer non pas comme bases fondationnelles pour la communication interdisciplinaire, mais davantage comme des outils d'accompagnement cognitif.
Les ontologies peuvent être considérées comme des outils qui permettent une accommodation de la science, non pas comme des moyens de « conceptualiser » ou « théoriser » celle-ci ou la réalité des entités traitées dans un domaine.

C'est-à-dire, qu'en tant que bases terminologiques disposant d'outils informatiques permettant d'affiner avec une certaine précision les modalités sémantiques d'un mot et ses liens avec d'autres (comme sa fréquence d'apparition dans un contexte, ses synonymes, les termes qui lui sont associés, celui qui est le plus souvent employé, le second plus employé, etc.), les ontologies informatiques peuvent servir à cartographier les connaissances d'une discipline pour permettre à ses acteurs une représentation conventionnelle de son contexte langagier.

1.2. Même langage ou même connaissance ? Une problématique linguistique et épistémologique de la communication interdisciplinaire

Le langage peut être défini comme le moyen ou le véhicule qui permet la transmission d'informations à propos d'objets du monde grâce au partage de concepts et de définitions entre deux interlocuteurs. En conservant une définition du langage comme une correspondance entre une proposition et un objet auquel elle réfère, nous obtenons alors un usage du langage comme étant un témoignage du rapport au monde de l'interlocuteur.

L. Wittgenstein (1921) dans le *Tractatus* indique le rôle du langage et son rapport à la connaissance en ces termes :

> *« 5.6 : Les frontières de mon langage sont les frontières de mon monde. ».*

La connaissance que nous avons de notre propre connaissance, de ses limites et conditions, est en ce sens dérivée du langage. Dans le cas de la communication interdisciplinaire particulièrement, c'est son usage, sa pragmatique, qui va permettre l'échange de connaissances.

En jugeant de la qualité du discours d'un individu, un sujet est ainsi en mesure de définir ou d'émettre des hypothèses sur les objets auxquels l'interlocuteur réfère en fonction de celui-ci. Le discours qui devient alors le témoin de la représentation du monde d'un individu et permettant à un autre d'évaluer la connaissance dont dispose son interlocuteur.

Mais en ce cas, rechercher la communication interdisciplinaire est-ce chercher à employer un discours univoque, une même langue ? Ou est-ce plutôt viser à établir les connaissances que nous avons en commun ?

Lorsque nous visons à définir un terme, nous pouvons rechercher la signification qu'il délivre et ce en quoi cette signification capture pertinemment, l'idée, l'activité ou l'objet à laquelle elle réfère. Cependant, lorsque plusieurs individus de champs disciplinaires différents emploient le même terme pour désigner une chose, peut-on dire qu'ils font référence à la même chose ?

Il s'agit ici de déterminer sur quoi repose la communication interdisciplinaire. Pour cela, il nous faut distinguer la connaissance commune d'une notion, de l'emploi d'un langage commun pour la désigner. C'est-à-dire interroger si l'échange d'informations entre agents repose sur une représentation mentale uniformément partagée ou sur un accord conventionnel vis-à-vis de la définition d'une connaissance. Nous faisons donc ici une première distinction entre l'épistémologie interdisciplinaire et la communication interdisciplinaire.

Il existe en effet une différence entre connaitre la même chose et employer différents mots pour la désigner.

Il est possible de parler d'une même chose, de plusieurs façons, et avec différents mots. Pour illustrer cela on peut repenser à l'exemple de [Frege (1892)] et de la référence à

Vénus et des termes « Phosphoros », soit l'étoile du matin, et de « Hèsperos », l'étoile du soir. Or, il s'avère que l'étoile du matin et l'étoile du soir sont toutes les deux des manières de nommer « Vénus ». Cependant, on peut remarquer que c'est une manière de nommer deux contextes de Vénus en fonction de sa position orbitale. Frege dit ensuite :

> *« La pensée contenue dans la proposition : « l'étoile du matin est un corps illuminé par le soleil » est différente de la pensée contenue dans « l'étoile du soir est un corps illuminé par le soleil ». Si quelqu'un ignorait que l'étoile du soir est l'étoile du matin, il pourrait tenir l'une de ces pensées pour vraie et l'autre pour fausse. La pensée ne peut donc pas être la dénotation de la proposition ; bien plutôt faut-il y voir le sens de la proposition. »*

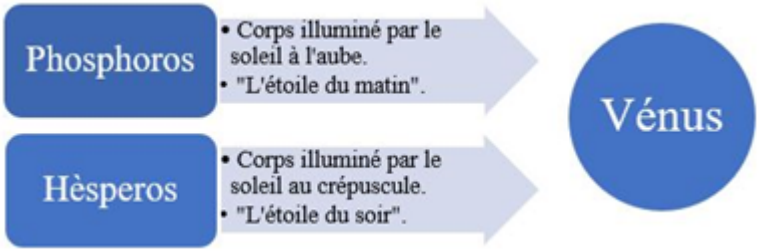


Figure 1. Frege, Sens et Dénotation *(1892)*

Ici, « l'étoile du matin » et « l'étoile du soir » sont donc deux expressions distinctes, disposant de leur sens propre, mais ont la même référence, car elles dénotent du même objet céleste à savoir la planète Vénus.

Cette distinction entre « sens » et « dénotation/référence » permet d'introduire deux manières de parler d'un concept : l'intension et l'extension.

Intension : C'est la définition, le sens d'un concept, ou sa compréhension. Par exemple, l'intension de « Vénus » pourrait être : « est la deuxième planète du système solaire par ordre de proximité avec le Soleil ».

Extension : C'est la dénotation ou référence d'un terme vers un concept. Il s'agit de l'ensemble des entités auxquelles l'intension s'applique. Ex : Vénus = {Phosphoros, Hèsperos, Etoile du Berger}.
On peut maintenant concevoir que lors d'une communication entre deux personnes de deux disciplines différentes : même si la définition – l'intension – d'une notion comme « Vénus » est commune, pour peu que la représentation contextuelle diffère entre eux, la communication peut alors être confuse ou simplement impossible.

Nous nous retrouvons dans une situation paradoxale, en ceci que la communication interdisciplinaire peut parler des mêmes choses, en ayant les mêmes connaissances parfois, et cependant employer des termes différents pour communiquer.

Pour reprendre notre distinction : on peut dire que les agents peuvent parler d'un même concept en utilisant différentes extensions de celui-ci (Phosphoros ou Hèsperos), sans savoir au préalable que celles-ci disposent de la même référence (Phosphoros et Hèsperos → Vénus).

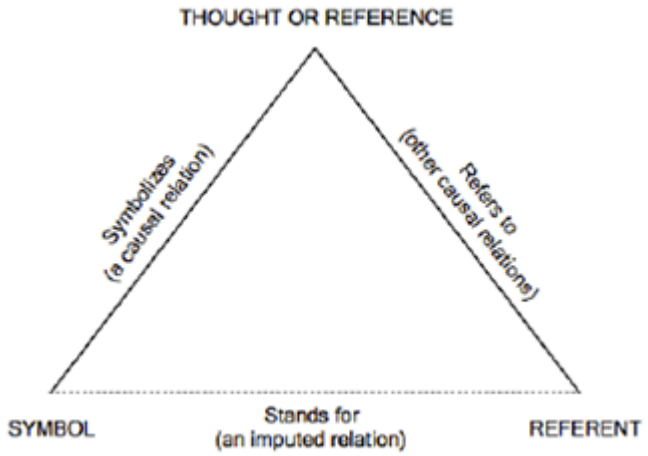


Figure 2. Triangle sémiotique d'Ogden-Richards, The Meaning of Meaning *(1923)*

On peut représenter cette relation triadique pensée-référence-symbole que nous évoquons par un processus de codification qu'est le triangle sémiotique ci-dessus. Ici, on observe plus concrètement que la pensée contenue dans une phrase (la pensée, ou référence : ex. Phosphoros est l'étoile du matin) désigne, réfère à un même objet (le référent, Vénus), qui lui-même peut être représenté et servir à établir une correspondance avec la phrase qui le désigne par un symbole (ici le mot pris comme signe « Phosphoros »).

Du point de vue de notre problématique, la communication interdisciplinaire rencontre « *l'inscrutabilité de la référence* » (Quine, 1969) : n'ayant accès qu'aux énoncés langagiers d'un interlocuteur et non à la connaissance contextuelle ou pratique personnelle qu'il en a, nous ne pouvons jamais savoir exactement à quel objet le terme qu'il emploie fait référence en son esprit. Autrement dit, on ne peut jamais savoir en quel sens un individu emploie un terme, ni quelle pensée est contenue dans ce qu'il dit. De fait, deux individus peuvent employer des termes différents, des pensées différentes mais ignorer qu'en réalité ces termes et pensées désignent le même référent.

La communication entre plusieurs champs disciplinaires n'est ainsi possible que si tous les acteurs ont la même connaissance de la signification des termes qu'ils utilisent et de leur domaine d'intension et d'extension. C'est-à-dire, si tous les acteurs disposent des mêmes référentiels sémantiques.

Or, il est difficile de définir sémantiquement un concept indépendamment du contexte dans lequel celui-ci apparaît ou de sa compréhension commune, rendant par conséquent sa définition ou celle d'un terme possiblement vague lors du passage d'un

domaine de spécialité à un autre.

La communication interdisciplinaire peut être alors définie comme : *la recherche du fait que tous les acteurs à partir d'un mot disposent de la même référence vis-à-vis d'un concept ou d'un objet.*
Plus précisément : le fait que les acteurs à partir de la même intension d'un concept, puissent communiquer à son sujet en ayant connaissance de l'ensemble de son extension.

Que les énoncés « l'étoile du matin » ou « l'étoile du soir » n'aient pas la même signification ne serait plus un problème épistémologique pour deux interlocuteurs, si ces derniers savent tous deux que ces énoncés ont la même référence et qu'ils sont tous deux une manière de parler de la même chose dans différents contextes.

Par conséquent, l'échange de connaissances lors de la communication interdisciplinaire devient problématique lorsque pour communiquer doivent se rencontrer :

- La subjectivité, la représentation individuelle et contextuelle d'une connaissance ;
- L'intersubjectivité, la représentation commune et partagée de cette connaissance entre les acteurs qui génèrent, alimentent, et organisent celle-ci.

Si l'on souhaite s'interroger sur comment faire communiquer différents domaines ensemble, il semble nécessaire de prendre en compte la différence entre les connaissances des acteurs des domaines ainsi que leur rapport à celles-ci. Il faut en ce cas considérer également le niveau linguistique sur lequel repose le terme, et la dimension extra-linguistique, mentale ou pratique, sur laquelle repose l'idée ou le concept auquel il fait référence.

C'est ici que l'ingénierie des connaissances emploie les ontologies informatiques, nous permettant de représenter les connaissances d'un domaine de spécialité en explicitant et spécifiant leur contexte. Les ontologies vont nous servir à faire l'inventaire de l'ensemble des termes qui se rapportent à une notion, et de clarifier les degrés de proximité des termes selon leurs occurrences. Elles font usage pour y parvenir à une formalisation du langage naturel, c'est-à-dire une réduction du langage et des énoncés à un aspect syntaxique.

2. Principes et typologies des ontologies

L'Ontologie est un terme philosophique qui signifie *être,* venant du grec ancien *ôn, onton*, et *discours, étude, science*, venant de *logos*. L'Ontologie correspond à une partie de la métaphysique, métaphysique qui peut se définir comme la science qui étudie les premiers principes et les premières causes, ou « l'être en tant qu'être », c'est-à-dire le *pourquoi* ou le *sens* de l'être. L'Ontologie quant à elle, étudie le *comment* de l'être, en

tant qu'il est organisé dans l'existence, c'est-à-dire son *étant*. En généralisant, elle vise à décrire dans la démarche d'un inventaire le « ce qu' il-y-a » dans un domaine d'objet.

L'ontologie est une discipline qui va traiter de questions à propos de la relation entre les objets d'un domaine, de leurs propriétés : sur quoi peut-on fonder cette propriété ? cette propriété elle-même peut-elle être instanciée dans plusieurs entités ? Lui est-elle propre ? Comment représenter ces relations ?

L'apport de l'ontologie pour l'intelligence artificielle et les ontologies informatiques, va être de fournir des outils conceptuels permettant de représenter les relations logiques et formelles, abstraction faite de leur contenu matériel. On va par exemple pouvoir trouver le meilleur moyen de décrire une collection d'entités en se focalisant avant tout sur ce qui les met en relation, c'est-à-dire d'être une collection. Puis l'on va pouvoir spécifier cette relation en cherchant les principes hiérarchiques qui organisent, qui ordonnent les entités qui composent cette collection.

Les ontologies informatiques héritent de cette recherche inventoriale car elles peuvent s'appliquer à des domaine de spécialité en établissant les relations entre les concepts qui lui sont propres. Elles peuvent faire l'inventaire des entités qui existent dans un domaine de spécialité en se limitant à représenter les entités d'un contexte précis. Cependant, la subjectivité et l'intersubjectivité des acteurs limitent la représentation logique de ces connaissances, et l'intelligibilité des concepts et notions est conditionnée par *l'interopérabilité sémantique* entre machines et acteurs.

L'interopérabilité sémantique est la capacité pour plusieurs systèmes d'échanger des informations entre eux, tout en conservant la signification et la qualité de l'information lors de l'échange. Cette interopérabilité pour être efficiente doit répondre à certains principes et est définie par l'Initiative Council on SDO Global Health Informatics Standardization comme :

> *« La capacité pour plusieurs systèmes d'échanger de l'information entre eux de telle sorte que chacun de ces systèmes puisse interpréter la signification de l'information reçue et utiliser cette information en articulation avec ses données locales. »*

Apparues au début des années 1990 dans les recherches sur l'acquisition des connaissances pour les systèmes experts, les ontologies s'inscrivent dans le développement du Web sémantique par les travaux de Berners-Lee (2001). Le Web sémantique vise à donner au contenu des ressources du web un accès utilisable par des programmes et agents logiciels, grâce à une compréhension logique des données. C'est-à-dire que la gestion de l'information se déploie en des réseaux sémantiques dans lesquels les données sont reliées entre elles, permettant ainsi à un ordinateur de les comprendre, les hiérarchiser et les parcourir rapidement.

La définition d'une ontologie informatique peut alors être proposée initialement comme étant un outil, artefact informatique qui représente sémantiquement la connaissance

d’un domaine de spécialité. Cependant, au fil des années de recherches, plusieurs définitions se sont proposées :

> Uschold et Gruninger (1996) « *Il s’agit du terme utilisé se référant à la compréhension partagée d’un domaine d’intérêt qui peut être utilisé comme cadre unificateur pour résoudre les problèmes de communication entre les gens et d’interopérabilité entre les systèmes.* »
> *Sowa (1999) « Une ontologie est un catalogue des types de choses supposées exister dans un domaine, du point de vue d’une personne utilisant un langage pour parler du domaine. »*
> *Arp, Smith, Spear (2015) « Une ontologie est un artefact de représentation, compris comme une taxinomie dont le but est de modéliser les relations et combinaisons entre classes d’objets et universaux. [...] Elles représentent (ou cherchent à représenter) la réalité, de manière à ce que différentes personnes puissent comprendre les termes qu’elles contiennent.*»

Bien que ces définitions reflètent différentes applications possibles des ontologies, celle de Gruber (1993) : « Une ontologie est une spécification partagée d’une conceptualisation. », semble être communément admise pour caractériser ce qu’est une ontologie. Une ontologie représente alors un consensus à propos d’un domaine de spécialité dont les connaissances et objets sont préalablement définis conceptuellement.

Le problème de l’interopérabilité qui nous occupe va donc être de parvenir à distinguer ce qui constitue la connaissance d’un objet ou d’un concept, du mot auquel il est raccroché, par l’analyse logique que permettent les ontologies. Ceci, afin de voir s’il est possible, en fonction des problématiques contextuelles que nous allons introduire, de modéliser cette relation entre mot et concept dans une ontologie informatique.

On pourrait ainsi désambiguïser les types de relations que les mots entretiennent entre eux en fonction d’un domaine, et proposer une méthodologie qui permettrait à plusieurs domaines de spécialité de communiquer ensemble sur la base de cette clarification du langage.

En ce sens, les ontologies informatiques ont pour but de produire un objet informatique respectant les règles syntaxiques de la logique formelle, tout en visant à expliciter et capturer les nuances de proximité et distance entre les mots du langage naturel.

Il faut néanmoins distinguer plusieurs méthodes de classification des connaissances pour voir l’intérêt particulier des ontologies.

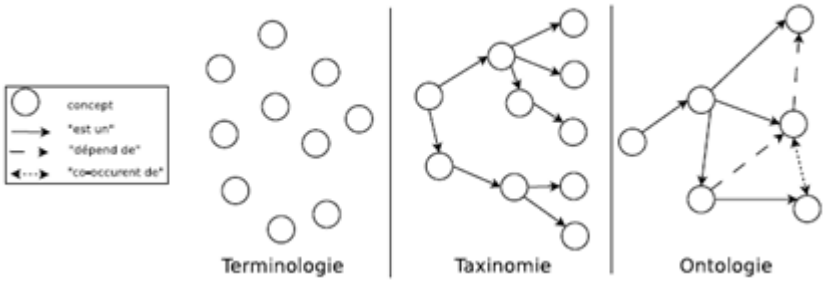


Figure 3. Schéma adapté de D. Zarebski (2018)

Par exemple, ici la terminologie est une *collection* de termes, tandis que la taxinomie est une *classification* des termes scientifiques. L'ontologie a ceci de spécifique qu'elle réunit ces deux méthodologies et introduit des principes logiques permettant de spécifier et structurer les relations entre les entités et de donner une signification à ces relations.

Par exemple, les principes différentiels (Bachimont, 2001) permettent de définir les termes par les relations d'identité qui les unissent ou les distinguent. Toutes les entités prennent racine à partir d'une entité « *Père* » comme concept fondamental à laquelle elles appartiennent. La signification relationnelle d'une entité dépend alors de sa position dans l'arbre de l'ontologie.

Organisées ainsi en réseaux, les ontologies informatiques proposent une représentation des connaissances selon une relation d'inclusion où un concept particulier est englobé par un plus général suivant une logique ensembliste.

Il s'agit ici de la *conceptualisation* de l'ontologie d'un domaine (Genesereth and Nilsson, 1987), soit : la représentation abstraite des entités d'un domaine par l'usage d'un corps formel.

Ces principes s'appuient sur ceux de l'inférence de la logique formelle et de l'engagement ontologique : le fait que l'existence d'une entité soit impliquée par sa quantification logique (Quine, 1948).

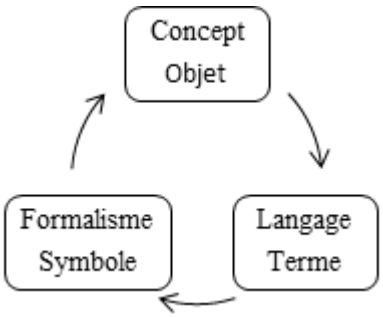


Figure 4. Correspondance logico-sémiotique dans une ontologie

Les ontologies informatiques permettent d'assurer cette correspondance logico-sémiotique par l'emploi de la relation de subsomption grâce à la formule « *est_un* ». Cette formule a pour fonction d'attribuer une caractéristique d'appartenance ou d'identité entre plusieurs entités. On peut dès lors organiser une hiérarchie d'entités et décrire dans le détail leurs relations par ce formalisme.

L'apport de la logique ensembliste va nous permettre d'expliciter les relations représentées dans nos arbres ontologiques. Si l'on suit le formalisme de [N. Guarino, Daniel Oberle, Steffen Staab (2009)] pour décrire les relations d'un contexte, on prend une structure O = telle que :
§ *D* un domaine de discours qui va servir d'ensemble des entités que nous allons représenter.
§ *W* un ensemble de mondes possibles pour représenter les différentes acceptions d'un terme en fonction de plusieurs contextes.
§ *R* un ensemble qui va représenter les relations conceptuelles de l'espace du domaine *<D, W>*.
On associe également un vocabulaire V qui coïncide avec les relations de *R*. C'est ici une recherche de l'engagement ontologique : on cherche à mettre en correspondance le symbole formel (ou textuel) que nous nommons par exemple « Vénus 1 » avec l'entité conceptuelle de « Vénus », et procéder ainsi pour tous nos symboles. La définition logique des entités est cependant une démarche plus épistémologique qu'ontologique (au sens philosophique), en ceci que c'est l'aspect définitionnel qui va importer pour l'ontologie informatique plus que la question de l'existence propre des objets représentés.

Une fois la structure en place, on peut ensuite définir le détail du contenu de nos principes différentiels exposés plus haut afin de rendre explicites les relations entre les entités.

Une ontologie est généralement une structure composée d'entités où O = < C, I, R, T, V> telle que : « O » est l'ontologie, « C » un ensemble de classe, « I » un ensemble d'individus, « R » un ensemble de relations, « T » un ensemble de données, « V » un ensemble de valeurs associées aux ensembles précédents [J. Euzenat and P. Shvaiko (2013)].

On définit alors trois types généraux d'ontologies [G. Declerck, A. Baneyx, X. Aimé, Jean Charlet (2014)], mais qui n'interagissent pas avec les mêmes entités. On parle ici de « niveaux » de généralité des entités, ou des concepts traités par les différentes ontologies, allant du degré d'abstraction du concept le plus généralisé comme le Temps ou l'Espace, au degré de concrétisation le plus spécifié, sur des objets comme « Cœur » ou « Organe ».

Ontologie de domaine (*Domain Ontology*) : Il s'agit du premier niveau d'ontologie et du plus concret. Il traite des entités propres à une discipline spécifique comme la pneumologie ou la chirurgie par exemple, et s'attache à établir une relation entre les

entités du domaine.

Ontologie de noyau (*Core Ontology*) : L'ontologie de noyau traite de concepts plus généraux qui disposent d'une interdisciplinarité conséquente pour plusieurs ontologies. Elle a la charge de permettre la communication de plusieurs ontologies à propos de concepts partagés par plusieurs sous-disciplines. Par exemple, les concepts tels que « *malade* », « *démence* » et « *myocarde* » couvrent plusieurs champs médicaux comme la cardiologie et la psychiatrie. Une ontologie de noyau assure donc l'interopérabilité entre deux ontologies de domaines différents qui partagent certains termes.

Ontologies fondationnelles (*Top Ontology*) : L'ontologie fondationnelle traite quant à elle des concepts méta-disciplinaires, c'est-à-dire des concepts suffisamment généraux pour être appliqués à tous les domaines d'une ou plusieurs sciences. Les concepts qu'elles traitent relèvent essentiellement de la métaphysique et se concentrent sur des concepts tels que l'évènement, le temps, le processus, la qualité, etc. Ainsi, elle permet la communication entre plusieurs ontologies en précisant le contexte des objets particuliers d'un domaine, à partir de concepts plus généraux.

3. Hétérogénéité des données et alignement des ontologies

Les principes des ontologies permettent ensuite d'identifier plusieurs caractéristiques de la polysémie interdisciplinaire et plusieurs méthodes d'interopérabilité sémantique.
L'hétérogénéité de l'information lors de la communication interdisciplinaire, ou son caractère polysémique, peut prendre plusieurs formes dans le cas des ontologies informatiques. Le but est alors de la réduire en alignant les ontologies entre elles afin d'établir des correspondances entre les entités de différents domaines de spécialité. On peut lister quelques types d'hétérogénéité au niveau même des ontologies qu'il s'agira de résoudre (J. Euzenat et P. Shvaiko, 2013) :
Hétérogénéité syntaxique : Cela correspond au fait que deux ontologies n'emploient pas le langage logique de la même manière, c'est-à-dire ne formalisent pas les entités représentées de la même manière. Quand bien même elles pourraient parler des mêmes objets avec les mêmes règles logiques, les façons de les représenter étant différentes, elles ne pourraient communiquer ensemble.
Hétérogénéité terminologique ou référentielle : Cela correspond au fait de l'emploi de différents termes pour référer à un même objet sans que ces termes ne se retrouvent réciproquement dans les deux ontologies ni ne visent à définir le même aspect de leur référence. Il s'agit généralement d'un problème de synonymie. Par exemple : différence synonymique avec une sub-différence temporelle entre : o1 = « Quotidien » et o2 = « Périodique » pour désigner « Journal ».
Hétérogénéité conceptuelle : Niveau extra-linguistique. Apparait lorsque deux ontologies sont fondées sur différentes théories pour représenter les entités primitives. Par exemple : o1 = essentialisme : les entités disposent d'une essence qui préexiste à

leur vécu ; et o2 = existentialisme : les entités forment leur essence à partir de leur vécu.
Hétérogénéité sémiotique : La granularité de la communication interdisciplinaire est également composée par l'intention, le « vouloir-dire » des différents agents entre eux. Cette hétérogénéité correspond au caractère interprétatif des entités en fonction de l'individu qui la comprend : la manière personnelle et privée dont un individu va interpréter un concept, mot, ou symbole. Cette hétérogénéité est hors de portée des ontologies car elle relève spécifiquement du caractère intentionnel propre à l'esprit humain.

Pour répondre à ces problématiques, comme nous l'avons vu avec les ontologies de noyaux, il est possible de faire correspondre plusieurs ontologies de domaines, et donc plusieurs contextes, grâce à différentes méthodes d'alignement (Declerck, Baneyx, Aimé, et Charlet, 2014) :
La fusion : Il s'agit ici de créer une base terminologique « T » unique, en fusionnant deux terminologies des ontologies « X » et « Y ». Ainsi, « X » et « Y » utilisent la ressource « T » pour communiquer.
L'alignement direct : L'alignement consiste en une sorte de contrepoint sémantique. C'est-à-dire que pour chaque terme d'une ontologie « X », on associe son équivalent dans l'ontologie « Y » à l'aide d'un langage informatique (SKOS, *Simple Knowledwge Organization System*) qui permet la mise en relation des termes synonymes, antonymes et conserver l'autonomie des terminologies de chaque ontologie sans les confondre en une seule.
La projection sur une terminologie de référence : La projection permet d'associer chaque terme d'une ontologie « X » et « Y » à un champ terminologique plus large nommé « Z » mais sans les confondre. « Z » sert alors de référence pour la définition des termes de « X » et « Y ». L'intérêt de cette méthode est de permettre à plusieurs domaines relatifs aux ontologies « X » et « Y » de se référer à une unique source « T ».

Ces méthodes d'alignement permettent de développer des bases terminologiques associant les différents sens qu'un terme peut porter selon le domaine où il se trouve employé.

C'est donc avant tout le domaine de spécialité, son contexte, sa problématique et son niveau de généralité qui vont restreindre les connaissances à représenter, ainsi que les normes et consensus qui vont organiser la place des objets du domaine étudié.

Il s'agit ensuite de modéliser des primitives dans le langage du domaine qui serviront à la base de la construction des connaissances. C'est-à-dire définir un terme qui soit suffisamment élémentaire pour être utilisé comme base d'un champ lexical, et dont l'identité soit suffisamment univoque pour ne pas se confondre avec les notions qu'il définit.

Cependant, il n'existe *a priori* pas de primitives dans le langage naturel répondant à ces critères [Bachimont (2000)]). Pour définir l'identité d'une chose il est nécessaire d'entrer

dans une « division de la référence » [Quine (1977)] ; c'est-à-dire une régression qui nécessite pour définir un terme d'en interroger toujours un autre.

4. Figer le sens ou figer le contexte ?

Nous prenons donc ici le parti que la communication interdisciplinaire repose sur une recherche terminologique. Charles S. Peirce (1978) rapporte sa « morale terminologique » à suivre pour les sciences :

> *« (2.222) Quant à l'idéal à atteindre, il est souhaitable en premier lieu, que, pour chaque branche de la science, il y ait un vocabulaire fournissant une famille de mots apparentés pour chaque conception scientifique et que chaque mot ait une seule signification exacte, à moins que ses significations différentes ne s'appliquent à des objets de catégories différentes qu'on ne peut jamais prendre l'une pour l'autre. »*

Notre démarche terminologique se caractérise alors au départ par la recherche d'une clarification d'un sens unique pour chaque terme d'une discipline et la désambiguïsation du sens qu'ils pourraient partager en apparence.

Pour tenter de répondre à cette problématique, on peut proposer une analyse de la structure logique du langage afin de désambiguïser et contextualiser la communication interdisciplinaire, consistant en :

> 1- La recherche d'une démarche épistémologique pour établir un inventaire des termes employés par plusieurs disciplines.
> 2- Établir clairement leur sens propre et celui ou ceux qu'ils peuvent partager.
> 3- Un moyen permettant d'aligner et mettre en relation ces termes entre eux afin de rendre explicite la référence qu'ils partagent.
> 4- Tenter de normaliser les définitions des propriétés au sein des disciplines, et viser à distinguer les potentielles conceptions entre experts au sein d'un même domaine.

Car si différents domaines de spécialités peuvent employer différents vocabulaires, les *règles* grammaticales, ainsi que celles de la syntaxe logique du langage, restent invariables en dépit du contexte. Nous allons pouvoir nous appuyer sur celles-ci pour proposer une méthode de clarification des échanges interdisciplinaires.

Présentée ainsi, cette recherche terminologique dispose d'une nature que l'on peut qualifier de « *fixiste* » : de viser le fait qu'un mot ne dispose que d'une seule et unique signification dont il est le seul porteur. Cette recherche se peut être pertinente quand appliquée à un seul domaine de spécialité.

Mais dans le cas de l'interdisciplinarité, la démarche terminologique et particulièrement

son caractère fixiste peuvent rencontrer le problème de la *polysémie* [F. Récanati (1997)] : lorsque des termes sont communs à plusieurs champs disciplinaires, mais dont la signification et la référence varient en fonction de leur domaine d'usage.

Comment en ce cas pouvons-nous rendre compte du caractère polysémique inhérent à la communication interdisciplinaire par une démarche terminologique ?

Pour répondre à cette problématique, les ontologies permettent de formaliser les énoncés et les concepts pour en donner une représentation commune qui va alors servir de base primitive pour l'élaboration de thésaurus et de terminologie. La connaissance interdisciplinaire rencontre le cadre épistémologique du contextualisme : les énoncés ne disposent d'une signification que dans un contexte d'énonciation.

Pour répondre à cela l'usage des ontologies informatiques comme outils informationnels peuvent nous permettre de tenter de trouver un « point fixe » à partir de l'usage des mots.

Les ontologies informatiques peuvent servir en ce sens à :

a) Contextualiser un domaine : Analyser les énoncés et leur domaine d'usage comme donnant leur signification aux mots. C'est là l'usage des ontologies en tant qu'outils permettant de figer un contexte de domaine par l'analyse de la composition des textes. Les ontologies vont nous permettre de représenter le genre d'un corpus et le domaine de provenance des ressources à décrire, et donc de situer un corpus vis-à-vis d'un autre comme nous le verrons dans la gestion des ressources termino-ontologiques (RTO).
Les ontologies permettent ainsi de rendre compte du contexte d'un domaine de spécialité grâce à l'interprétation textuelle et la récurrence de l'usage d'un mot dans un contexte donné.
b) Désambiguïser les relations entre les termes : Les ontologies peuvent préciser la signification des mots qui composent un énoncé ainsi que leurs fréquences d'utilisations. C'est-à-dire établir les notions d'un ou plusieurs champs disciplinaires, en considérant les relations logiques qu'elles entretiennent et en distinguant les multiples ou uniques références qu'elles peuvent avoir en les alignant les unes vis-à-vis des autres.

En précisant des caractères comme la proximité, l'association, la synonymie, la préférence, les ontologies permettent de clarifier quel usage est fait d'un mot, ainsi que son réseau de relations : c'est-à-dire l'écosystème terminologique qui évolue autour de lui.

Pour un terme relatif à une pathologie particulière, on peut par exemple définir la préférence de l'usage de tel ou tel terme, ceux qui sont employés les moins souvent, ceux qui s'y rapportent de manières dérivés, ceux qui y font référence, ou encore leur correspondance extra-disciplinaire : c'est-à-dire des notions hors du domaine médical

en question, mais qui désignent la même pathologie. La logique descriptive des ontologies permet de rendre compte de l'agencement des composants élémentaires, les mots, de nos énoncés complexes, nos propositions. Cette approche s'accorde avec la démarche sémasiologique qui sera introduite plus loin.

Pour résumer l'emploi des ontologies :
a) Pour répondre au contextualisme les ontologies permettent à la communication interdisciplinaire de les utiliser comme des terminologies permettant de préciser l'usage contextuel des mots, ainsi que les réseaux d'utilisations et des termes proches qui leur sont associés. En représentant ces réseaux, les ontologies permettent de figer les contextes disciplinaires en « cartographiant » les connaissances de ces disciplines et en précisant la nature exacte des relations qu'entretiennent ces connaissances entre elles.
b) Méthodologiquement, les ontologies utilisent une syntaxe logique basée sur l'usage d'éléments simples (symboles, mots) pour en former des plus complexes (formules, énoncés), permettant de définir les termes et entités. Elles peuvent ainsi représenter de façon logico-mathématique le rapport que les mots d'une ou plusieurs disciplines ont entre eux, et inférer d'elles-même les relations qu'ils entretiennent, puis les représenter sous la forme de graphes.

Par exemple [A. Barton (2016)], dans une ontologie biomédicale les composants élémentaires *« Organe, Cœur, Ventricule »* sont organisés sur une syntaxe logique fondée par des axiomes plus complexes tels que « *Cœur est_un Organe »*. Le raisonnement de l'ontologie est ensuite en mesure de déduire des théorèmes à partir des axiomes formalisés préalablement tout en suivant différentes propriétés formelles des relations. Par conséquent, le moteur d'inférence de l'ontologie à partir d'axiomes tels que : si « *a partie_de b » et « b partie_de c »* est en mesure d'en inférer qu'alors « *a partie_de c »* par transitivité.

Une ontologie biomédicale peut également spécifier le degré de proximité d'un concept vis-à-vis d'un autre selon leur usage : le concept « myocarde » peut être labellisé comme étant dans un certain contexte, le mot favori employé pour désigner le « cœur ». Tout comme elle peut spécifier la classe de l'entité du concept « myocarde » en indiquant qu'il fait partie des « organes » et par déduction que le mot « cœur » rentre dans la classe des « organes ».

Les ontologies peuvent de cette manière spécifier la sémantique des termes, et mettre en rapport les notions et concepts fondamentaux avec les mots et leurs modalités (antonymes, synonymes, hypéronymie, hyponymie, etc.). L'explicitation que proposent les ontologies peut également aller jusqu'à établir des métadonnées sur l'élaboration des concepts qu'elles prennent en charge, par l'ajout de commentaires ou de labellisation sémantique des notions.

Elles sont ainsi capables de traiter les termes et concepts des disciplines comme des symboles logiques manipulables. En modélisant les relations qu'entretiennent les notions de spécialités entre elles, elles permettent ainsi de rendre compte et de figer le

contexte de ces notions tout en prenant en compte le degré d'utilisation et de proximité d'un terme vis-à-vis d'un autre.

Nous insistons donc sur l'aspect représentationnel des ontologies, en tant qu'outils permettant de spécifier des contextes d'une notion et d'un terme à partir d'une structure logique.

On peut de cette manière appliquer une méthode et suivre une épistémologie nous permettant de mettre en relation symboles et significations, mots et contextes d'usage.

Cette relation peut se définir ainsi *: pour un terme disposant d'une référence, si un symbole capture la même référence, alors le symbole à la même signification que le terme*.

Les ontologies informatiques permettent ainsi de proposer une démarche terminologique visant à conserver le facteur contextuel de l'usage d'un terme. En ce sens, les ontologies telles que les définit Bachimont (2001) :

> *« [...] permettent de représenter de manière systématique les notions utiles pour déterminer et expliciter les connaissances qu'il faut mettre en œuvre pour traiter une tâche dans un domaine de spécialité. Les ontologies ne sont donc pas des bases de connaissances, ni n'expriment de connaissances sur un domaine ou dans une application. Elles explicitent de manière systématique les notions utiles à la formulation des connaissances. Elles ont davantage affaire avec le sens, ou le cadre sémantique dans lequel les assertions sur le domaine peuvent prendre une signification, qu'avec la vérité de ces assertions. »*

Pour les ontologies informatiques définies comme « une spécification d'une conceptualisation », « ce qui existe est ce qui est représenté » (Gruber, 1993) et elles visent une définition formelle des concepts et de leurs relations. L'ontologie informatique n'entretient alors pas de rapport direct avec les objets qu'elle décrit : elle vise à figer un contexte disciplinaire par une approche syntaxique de la communication et de l'échange d'information.

Nous insistons donc sur le fait que les ontologies définissent les relations entre nos énoncés, et ne s'occupent pas de l'existence des objets. L'usage des termes que nous faisons est avant tout théorique, en ceci que les termes dans nos ontologies sont des objets conceptuels visant à rendre compte de ce qu'une théorie dit qu'« il-y-a » dans un domaine d'objet, et non pas de l'existence matérielle ni des propriétés à proprement parlé de l'objet.

5. L'aspect terminologique et interdisciplinaire des

Cependant, pour que cette relation soit efficiente il est nécessaire de passer par la normalisation sémantique nécessaire de l'ontologie, par la constitution d'un corpus pour un domaine. C'est après la formation de ce corpus que l'on va pouvoir appliquer cette normalisation à l'usage et aux interprétations multiples du contexte interdisciplinaire.

La difficulté de cette entreprise réside dans le fait qu'un corpus vise à établir des notions non-contextuelles comme références : des notions dont on fixe et standardise la signification de manière à ce que toute interprétation ait à se conformer à celle-ci. Cependant, plus il y a de domaines à représenter, plus les notions augmentent en conséquence. La recherche du sens d'une notion est alors conditionnée par un référent physique en la qualité de l'expert du domaine qui va établir quelles notions peuvent être considérées comme étant à la base d'une construction terminologique.

Le texte étant le support sur lequel va se fonder la représentation graphique de l'ontologie, il convient alors également d'assurer l'homogénéité des informations qu'il va fournir. On passe ici au traitement du corpus

Le thésaurus d'une terminologie - l'ensemble de termes sélectionnés dans le champ lexical d'un domaine de spécialité - est essentiellement lié au support qui a permis de la construire, ainsi qu'à l'influence de son environnement (Poibeau, 2005) :

> *[...] La terminologie reflète l'organisation lexicale des éléments propres à la langue de spécialité en question. [...] Comme il n'existe pas deux textes (corpus) semblables, les termes et les relations entre les termes que l'on peut extraire à partir d'un corpus sont propres à ces textes (ou à ce corpus). Une terminologie est donc un objet intimement lié à sa source, au corpus et à – l'extrême – à l'application qui a permis de le constituer.*

Par conséquent, si l'on emploie les ontologies informatiques comme des bases terminologiques il convient donc de sélectionner et restreindre le corpus à appliquer aux ontologies d'un même champ scientifique pour préciser leur référence. C'est-à-dire qu'il est nécessaire que le corpus textuel d'une ontologie relève du contexte scientifique auquel il fait référence. Un corpus destiné à la physique ne peut pas décrire un contexte psychiatrique, quand bien même ils pourraient partager des termes en commun.

Une connaissance médicale par exemple n'a pas de signification médicale par et pour elle-même (Charlet, Cordonnier, et Gibaud, 2002). C'est-à-dire qu'un énoncé n'a de signification qu'en vertu d'un contexte où il est appliqué. Une assertion comme « *Un seul patient est passé au bloc opératoire* », nécessite tout un réseau d'informations et de compléments sémantiques pour dispenser une signification.

Pour en clarifier la signification, il est alors nécessaire d'aligner les termes avec leurs définitions afin de préciser l'objet auquel ils réfèrent. Pour permettre l'interopérabilité

sémantique et l'échange que favorise l'interface humains-ontologies, il faut s'assurer que la signification contextuelle d'un terme s'accorde avec celle du concept qui est représenté textuellement dans l'ontologie. Il est alors nécessaire que la représentation du terme « Patient » par un agent humain corresponde bien à sa représentation en un concept codé appelé « Patient(x) » dans l'ontologie informatique.

Les ontologies informatiques peuvent ainsi fonctionner sur des documents et du texte qu'elles interprètent en données. C'est lors de l'interprétation et de choix de ces textes que les ontologies informatiques vont être au service de l'interdisciplinarité.

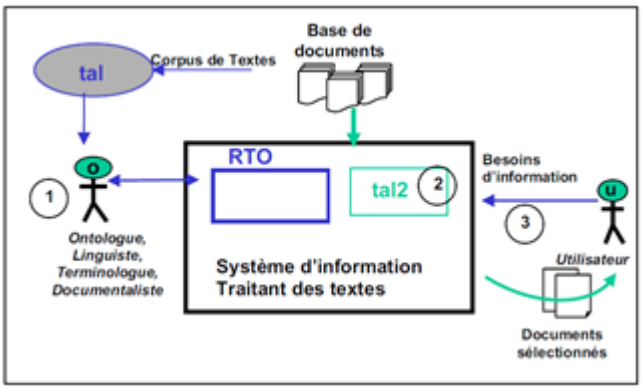


Figure 5. Schéma adapté de Nathalie Aussenac-Gilles, Anne Condamines (2004)

N. Aussenac-Gilles, A. Condamines (2004) rapportent comment les ressources termino-ontologiques (RTO) permettent de systématiser ces échanges interdisciplinaires au service de l'ingénierie des connaissances.
Ici, le Traitement Automatique des Langues (TAL) et les sciences de l'information se joignent à l'ingénierie des connaissances pour « fouiller » les textes et les analyser. La numérisation des données pousse ainsi vers une interdisciplinarité motivée par une recherche sur la pertinence des termes à modéliser dans une ontologie. En ce sens, l'ontologue, le linguiste, le terminologue et le documentaliste associent leur rôle dans le cadre d'élaboration des terminologies et des ontologies.

Finalement les documents sont modélisés sous une représentation relationnelle des connaissances entre elles. Cette synergie entre les différents acteurs et étapes pour l'élaboration des ressources permet alors de générer un circuit pour le traitement de l'information.

La base de documents permet à l'utilisateur **(1)** via le TAL d'interagir directement avec les corpus de texte et de modéliser des représentations relationnelles entre les connaissances **(2)**. Les RTO sont quant à elles traitées par le TAL, qui peut alors retransmettre l'information nécessaire à l'utilisateur **(3)** selon les documents sélectionnés.

Les ontologies pour être opérationnelles et assurer l'interopérabilité sémantique doivent

suivre plusieurs protocoles qui appliquent les mêmes normes (Aubin, 2016). Ces normes s'appliquent à des formats, où il s'agit d'utiliser pour la gestion des données les mêmes standards comme par exemple XML et RDF. Ceux-ci vont nous servir à structurer les « Terminologie de référence », c'est-à-dire les bases lexicales qui fournissent des définitions univoques des concepts codés pour un domaine de spécialité.

Il est donc nécessaire de coder les termes et les mots d'usages avec la même source informatique afin de permettre à différents systèmes de se référer au même terme lorsqu'ils rencontrent la même ressource numérique. Pour cela on va faire usage de « Jeux de valeurs », un processus qui représente les concepts et les associe à des éléments de données correspondantes**.** Ainsi qu'à des « alignements sémantiques » où l'on va établir une équivalence entre les concepts de plusieurs terminologies de références différentes.

Pour résumer, les terminologies sont alignées entre elles, puis renvoyées à des jeux de valeurs qui structurent les termes en données. Ensuite lorsque le concept est codé, une correspondance est alors établie entre le code, le terme, et la base terminologique dont il provient, créant ainsi une correspondance entre eux pour permettre à l'agent et à la machine d'assurer une même lecture d'un terme.

Cependant, d'un expert à l'autre du même domaine une notion peut être plus importante qu'une autre. Par conséquent, dans le cas d'une recherche interdisciplinaire la tâche se complexifie grandement de par les multiples interprétations nécessaires pour faire consensus.

On peut néanmoins mettre cette problématique en perspective de la notion d'*usage* et de *genre* qu'indiquent N. Aussenac-Gilles, A. Condamines (2004) pour classer et interpréter les RTO.

Pour la communication interdisciplinaire, on ne peut *a priori* pas concevoir une unique terminologie. Cependant, on peut faire le constat empirique qu'il existe autant de RTO qu'il y a de pratiques et de domaines de spécialité. C'est-à-dire qu'il existe une pluralité d'acteurs et de paramètres qui peuvent rentrer en compte dans l'élaboration d'une terminologie : les textes, linguistes, experts, utilisations, utilisateurs, but visé, logiciels adéquats.

Pour rendre compte de cet aspect collectif de l'organisation des RTO, on peut considérer qu'il n'y a pas autant d'interprétations et d'usage que d'acteurs, ce qui serait indénombrable, mais plus simplement des *genres* d'*usage*.

La notion de genre : renvoie à une hypothèse d'appartenance à un même groupe pour la catégorisation des textes. La dimension extra-linguistique intervient en ce que l'on peut considérer le texte comme un médium entre une pratique et une conception. Le genre d'un texte vise à aligner les occurrences langagières textuelles avec la dimension psychologique et extra-linguistique de la communication interdisciplinaire.

La notion d'usage : renvoie quant à elle au caractère *situé* d'une pratique ou d'un texte. Elle permet également de spécifier la notion de genre en considérant la composante multi-pragmatique qui rentre en compte lorsque plusieurs disciplines envisagent de communiquer.

La double considération des notions de genre et d'usage permettrait de modéliser la dimension collective de la communication interdisciplinaire. Le genre permettant d'inscrire l'usage dans le cadre d'une analyse du discours pour établir les textes ; et l'usage d'intégrer et classifier le genre des textes en inscrivant la recherche de normalisation sémantique à partir des occurrences et de la pratique langagière.

On peut ensuite considérer deux démarches pour la normalisation d'une terminologie (Aussenac-Gilles, 2005). La première, appelée *onomasiologique*, soit la signification à partir d'une idée, concept ou notion ; est normée à partir d'une conception monosémique d'une notion, c'est-à-dire une relation idéalement univoque entre le terme et le concept. L'onomasiologie est une partie de la lexicologie (étude des mots), mais prenant pour primitive le concept, l'idée pour étudier ensuite les mots qui s'y raccrochent. Une telle idéalité n'existant *a priori* pas, il n'y a pas d'unique concept qui puisse subsumer l'ensemble d'un système interdisciplinaire -si tel était le cas il n'y aurait qu'une discipline-, l'approche que nous préconisons à partir de la récurrence des mots selon leurs usages dans un texte, semble capturer plus pertinemment le caractère contextuel d'une connaissance et sa définition.

Cette approche terminologique est nommée *sémasiologique*, soit la signification à partir de l'usage mot pour en définir le concept. Elle s'adapte mieux au caractère polysémique de l'interdisciplinarité, en ce qu'elle considère d'abord les phénomènes linguistiques, et ensuite établit des normes à partir de l'interprétation de l'occurrence d'un terme dans un contexte. Cette approche part des signes et de leurs agencements et variabilité dans un corpus, pour ensuite tenter de délimiter et définir le concept vers lequel ils tendent. Cette démarche va se fonder à partir du format textuel pour fournir une représentation des connaissances et des données à l'ontologie.

En somme, en plus d'une recherche épistémologique pour obtenir une homogénéité entre les termes et significations pour favoriser le dialogue, il est nécessaire également d'orienter la méthode vers une normalisation et opérationnalisation des pratiques. Sur la base d'une informatisation de la récurrence de l'usage des mots, on peut espérer ainsi figer un contexte en ce que, la cartographie informatique des discours que va fournir l'ontologie, va permettre une certaine représentation des pratiques langagières d'une discipline qui pourra par suite être conventionnelle.

On peut observer concrètement avec le besoin de cette standardisation informatique, la nécessité d'une pragmatique vis-à-vis de l'épistémologie contextualiste que nous avons présentée : en ceci qu'il semble qu'on ne puisse rendre compte d'un phénomène multi-contextuel qu'en le faisant converger vers une pratique unique.

Résiste cependant une problématique de taille : celle des primitives qui semblent ne pouvoir être définie qu'arbitrairement. Autrement dit, tout le processus de normalisation et d'opérationnalisation des ontologies informatiques est une pétition de principe en ce que, pour accorder conventionnellement les pratiques entre elles, nous faut-il à l'origine de l'entreprise, conventionnellement définir ces pratiques.

6. Conclusion : remarques, instrumentalisme et opérationnalisation des ontologies

Cette recherche de standardisation des pratiques, vise ici à représenter les connaissances et la sémantique humaine de manière formelle. Cette méthode nous permet ainsi d'expliciter les contextes d'usages des notions ainsi que la structure syntaxique des énoncés de différents domaines, facilitant l'alignement des notions entre elles avec leurs termes correspondant et favorisant les consensus.

Mais il nous faut remarquer une certaine circularité dans la recherche de la communication interdisciplinaire : comme nous l'avons vu pour la normalisation sémantique, pour faire communiquer différentes disciplines entre elles, l'interdisciplinarité est nécessaire.

Cette pétition de principe remet alors en cause le paradigme référentiel. Etant donné la multiplicité des domaines et qu'aucun ne prime ou n'a d'autorité *a priori* sur l'ensemble des autres, on ne peut alors considérer *un* modèle sémantique ou pratique sur lequel baser la représentation et l'alignement des connaissances.

Selon l'approche instrumentaliste de G. Declerck et J. Charlet (2014) c'est ici l'aspect spécifiquement technique, utilisable et performant des ontologies qui doit être au service d'un *accommodement* de la science et des phénomènes : servir à faciliter la tâche des acteurs pour laquelle ils utilisent les ontologies.

Les ontologies informatiques servent avant tout d'outils fonctionnels et non de modèles exacts de la réalité et de la sémantique. C'est-à-dire que les ontologies informatiques sont d'abord des outils permettant de rendre opérationnelle et effective la normalisation des échanges et pratiques, et non de représenter ou servir de témoin du sens véritable ou du statut épistémologique d'une notion ou d'un domaine de spécialité : « La carte n'est pas le territoire qu'elle représente ». Les ontologies ne peuvent rendre compte des nuances et implicites du langage naturel. Par conséquent, leur emploi n'a pour but que de nous orienter dans le caractère complexe (toujours au sens de E. Morin) de l'organisation de la ou des sciences entre elles et de leurs dialogues.

Cette conception des ontologies s'inscrit dans l'optique d'augmenter l'exercice de la pensée humaine pour dépasser ses limites cognitives (mémorielles, calculatoires, etc.), en mettant en relation les faits observables et en intégrant les données

correspondantes grâce à l'ingénierie des connaissances. Les ontologies informatiques sont alors à considérer comme des *supports cognitifs* qui accompagnent l'activité intellectuelle des acteurs.

De cette façon, l'interopérabilité sémantique peut être réalisée par les ontologies informatiques si l'on met en œuvre leurs capacités propres à l'ingénierie :

1- Employer ces dernières comme *outils d'interface sémantique* entre acteurs et terminologies, en résorbant l'hétérogénéité sémantique et en alignant termes et données. L'ontologie fait office ici de « pivot » entre le langage naturel et les données informatiques.
2- Normaliser les systèmes terminologiques entre eux en les faisant communiquer. On peut employer une ontologie comme *outil d'interopérabilité* visant à connecter les terminologies entre elles pour, à court/long terme, les subsumer en une unique terminologie homogène.

Ainsi, au modèle référentiel sémantique des terminologies où celles-ci servent à mettre en place les consensus entre les termes et les choses du domaine auxquels ils sont associés, les ontologies informatiques se proposent de substituer un « référentialisme instrumental ».

C'est-à-dire que les ontologies permettent de servir de « plateformes » à partir desquelles les théories et activités scientifiques peuvent se référer pour normaliser leurs pratiques, sans pour autant interférer dans les principes qui seraient à suivre idéalement ou inférer sur le réalisme présupposé d'un domaine.

Cette position instrumentale des ontologies permet ainsi de considérer ces artefacts informatiques comme autant d'outils dont l'opérationnalisation et l'efficience servent de témoin et de référent pour la construction des connaissances d'un ou plusieurs domaines. Les ontologies permettent de faire communiquer sinon les sciences ensemble, du moins les acteurs entre eux à propos des possibles confusions sémantiques, et ainsi clarifier le débat scientifique. On peut dès lors espérer des ontologies informatiques que, plus celles-ci sont interopérables plus on peut alors se fier aux méthodes et données homogénéisées qu'elles retransmettent.

7. Remerciements

Nous remercions vivement Jacques Hilbey, ingénieur d'études au LIMICS/INSERM pour ses remarques et relectures attentives.

8. Bibliographie

Arp,R., Smith, B., & Spear, A.D. (2015). Building Ontologies with Basic Formal Ontology, *The MIT Press*, Cambridge, Massachussetts.

Aubain, D. (2016). *Les référentiels sémantiques dans l'interopérabilité des systèmes d'information de santé*, avril 2016

Aussenac-Gilles, N., & Condamines, A. (2004). Documents électroniques et constitution de ressources terminologiques ou ontologies, *Information-Interaction-Intelligence*, Vol.4, n.1.

Barton, A. (2016). Ontologies appliquées biomédicales et ontologie philosophique : un développement complémentaire. « Une solution au problème de Babel : les ontologies appliquées. » *Revue de la société de philosophie des Sciences*, Vol.3, n.1.

Bachimont, B. (2001). Modélisation linguistique et modélisation logique des ontologies : l'apport de l'ontologie formelle. In *Actes de la conférence « IC'2001 »*, Grenoble.

Bachimont, B.(2000). Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en Ingénierie des connaissances, *Revue Ingénierie des connaissances : évolutions récentes et nouveaux défis*, 2000/4

Berners-Lee, T. Hendler, J., & Lassila, O. (2001). The Semantic Web, *Scientific American*. Livre digital. [En Ligne] http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21

Charlet, J., Cordonnier, E., & Gibaud, B. (2002). Interopérabilité en médecine : quand le contenu interroge le contenant et l'organisation. *Information-Interaction-Intelligenc*e, Vol.2 n.2.

Declerck, G., Baneyx, A., Aimé, X., & Charlet, J. (2014). Les ontologies fondationnelles peuvent-elles débabeliser le Web ? *Revue d'intelligence artificielle* – n° 2-3/2014, 191-216.

Declerck, G., & Charlet, J. (2014). Pourquoi notre sémantique naïve n'est pas formalisable et pourquoi c'est (presque) sans conséquence sur l'ingénierie ontologique, *Intellectica*, 2014/1, 61, pp.

Euzenat, J., & Shvaiko, P. (2013). *Ontology matching*, Springer-Verlag, 978-3-642-38720-3

Frege, G. (1971). *Sens et Dénotation*, 1892 in *Ecrits logiques et philosophiques*, trad. C. Imbert, Seuil, Paris, 1971

Genesereth M. R., & Nilsson, N. J. (1987). *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann, Los Altos, CA.

Guarino, N., Staab, S., & Oberle, D. (2009). What Is an Ontology?, *Handbook on Ontologies*, pp.1-17.

Ogden, C.K., & Richards, I.A. (1923). *The Meaning of Meaning*, A Harvest Book, Harcort, Brace&World, Inc. New York.

Pierce, C.S. (1978). *Ecrits sur le signe*, « Editions du Seuil ».

Poibeau, T. (2005). *Parcours interprétatifs et terminologie*. pp.12. ffhal-00005508f

Quine, W.V.O. (1969). *Relativité de l'ontologie*, Paris, Aubier-Montaigne.

Quine, W.V.O. (1977). *Le Mot et la Chose*, éditions Flammarion.

Quine, W.V.O. (1948).*On What There Is*. Review of Metaphysics Harvard University

Press.
Récanati, F. (1948). *La polysémie contre le fixisme*. In: Langue française, n°113, 1997. Aux sources de la polysémie nominale.
Sowa, J. F. (1999). *Knowledge representation: Logical, philosophical and computational foundations.* Brooks Cole Publishing Co.: Pacific Grove, CA USA.
Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, Methods and Applications, *Knowledge Engineering Review*, vol.11, n°2.
Wittgenstein, L. (2006). *Tractatus logico-philosophicus.* Gallimard.

Thèses :

David R.L. Zarebski, (2018). *Ontologie Naturalisée et Ingénierie des Connaissances*, 27 septembre 2018

HDR :

Nathalie Aussenac-Gilles, (2005). *Méthodes ascendantes pour l'ingénierie des connaissances*, HDR, Université Paul Sabatier – Toulouse III, 1e décembre 2005