



HAL
open science

Ontologies du Web : Histoire refoulée et perspectives paradoxales

Aurélien Bénel

► **To cite this version:**

Aurélien Bénel. Ontologies du Web : Histoire refoulée et perspectives paradoxales. *Intellectica - La revue de l'Association pour la Recherche sur les sciences de la Cognition (ARCo)*, 2014, 61 (1), pp.123-141. 10.3406/intel.2014.1041 . hal-02362390

HAL Id: hal-02362390

<https://utt.hal.science/hal-02362390>

Submitted on 21 Oct 2021

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

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

Ontologies du Web : Histoire refoulée et perspectives paradoxales

Aurélien Bénel

ICD/Tech-CICO
Université de technologie de Troyes
aurelien.benel@utt.fr

Résumé : En analysant les formats des « ontologies du Web » au crible de l'Organon, nous montrons que, malgré leur vocation informatique apparemment banale, ces ontologies posent de réelles questions philosophiques sur l'être et la vérité. Établissant ensuite par quels relais philosophiques elles se sont retrouvées dans le discours des informaticiens, nous déterminons quelles réponses, incorporées dans les ontologies, sont apportées à ces questions. Nous étudions enfin les conséquences de ces options philosophiques sur des terrains réels, afin de construire notre critique des ontologies de l'intérieur même du système de pensée de leurs promoteurs.

Mots-clés : Histoire de la logique, Épistémologie, Ingénierie des connaissances, Sémiotique des textes.

Title: Web ontologies: Paradoxical origins and perspectives

Abstract: In the light of the Organon, we show how Web ontologies, despite their mundane use by computers, raise real philosophical issues about *being* and *truth*. 'Embedded' answers to these issues are then identified by drawing a lineage of the term 'ontology' from philosophers to computer scientists. To finish with, we expose the consequences of these philosophical options on real cases, so that our criticism of ontologies is made from the inside of their proponents' system of thought.

Keywords: History of Logic, Epistemology, Knowledge engineering, Texts semiotics.

En philosophie, une ontologie est une théorie sur la nature de l'existence, sur les types de choses qui existent ; en tant que discipline, elle étudie ce genre de théories. L'intelligence artificielle et les chercheurs du Web ont détourné ce terme pour leur propre jargon, et pour eux une ontologie est un document ou un fichier qui définit formellement les relations entre termes. Le cas le plus typique d'une ontologie pour le Web étant une taxinomie et un ensemble de règles d'inférences.

(Berners-Lee *et al.*, 2001).

Introduction

Un volume sur la philosophie du Web ne peut faire l'impasse sur les « ontologies » du Web sémantique. La référence à la philosophie est en effet explicite, même si le rapport à la *théorie de l'être* l'est beaucoup moins pour ces « collections d'information [permettant de] découvrir des significations communes quelles que soient les bases de données » (Berners-Lee *et al.*, 2001).

Ces « ontologies » avaient déjà gagné depuis plusieurs années le domaine de la représentation des connaissances. Certains les voyaient comme la partie « portable » (Gruber, 1993) d'une base de connaissances à une autre. Bien avant que les sciences de l'information ne s'en inquiètent (Shirky C., 2005), notre communauté francophone d'ingénierie des connaissances appelait à la prudence :

La tâche conditionne la construction de l'ontologie qui dès lors ne peut être portable et encore moins universelle. Cela plaide de plus en faveur d'une vision non logique mais plutôt constructiviste de la connaissance (...). Comme toute connaissance, les ontologies sont interprétées par un expert humain en fonction de l'idée qu'il a de la tâche attribuée au système.

(Charlet, Bachimont *et al.*, 1996)

Dans un premier temps, nous essaierons de nous approprier cet appel à la prudence. En analysant leurs formats, leur structure, nous tenterons de comprendre ce que sont ces ontologies et pourquoi, malgré leur vocation

informatique apparemment banale, elles posent réellement des questions philosophiques sur l'*être* et la *vérité*.

Ensuite, pour déterminer quelles sont les réponses (à ces questions) incorporées dans les ontologies, nous tenterons d'établir d'où viennent ces ontologies et par quels relais philosophiques elles se sont retrouvées dans le discours des informaticiens.

Enfin, ces options philosophiques étant probablement très différentes de celles que nous défendons (Bénel et al., 2004), nous adopterons, par méthode, une certaine « empathie » afin d'analyser, sur des exemples de terrains réels, les conséquences de ces options.

I. Un sujet philosophique

1. Être

Cet article aurait pu se clore dès la citation mise en exergue : les ontologies du Web n'auraient tout simplement rien à voir avec l'Ontologie des philosophes, rien à voir par exemple avec les réflexions métaphysiques d'Aristote pour faire du langage défini dans les *Catégories* un outil (*Organon*) pour exprimer la connaissance de la substance du monde. Simples « fichiers », ces ontologies du Web ne mériteraient pas l'attention et les débats qu'elles suscitent.

Pour un lecteur de l'*Organon*, la structure de ces ontologies (cf. Fig. 1) présente pourtant une certaine familiarité, on y retrouve notamment :

- des termes universels (*Class*), comme « homme », « juste »,
- des termes singuliers (*Individual*), comme « Callias »,
- des propositions universelles affirmatives (*x subclassOf y*), comme « Tout homme est juste »,
- des propositions universelles négatives (*x disjointWith y*), comme « Aucun homme n'est juste »,
- des propositions singulières affirmatives (*x type y*), comme « Callias est juste ».

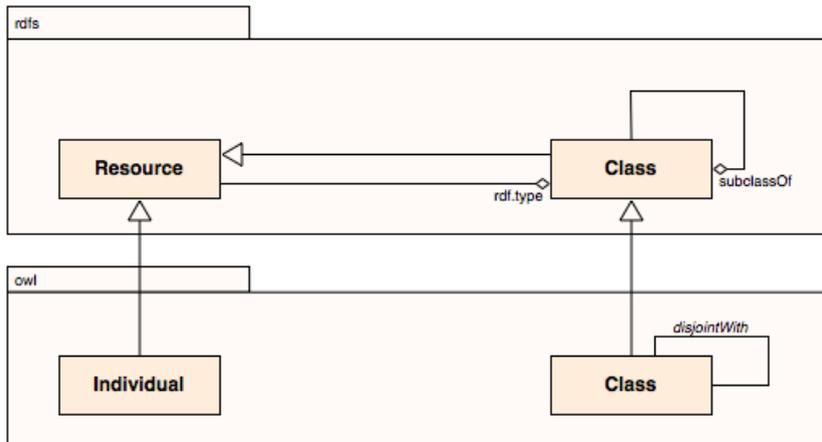


Fig. 1 – Le verbe « être » dans les ontologies du Web : Métamodèle (diagramme de classe UML)

Pour autant, une logique formelle – même du verbe « être » – est *logique* avant d’être *onto-logique*. À l’image de ce qu’Aristote écrit dans les *Seconds Analytiques* ou dans la *Métaphysique*, il faudrait tenter d’exhumer de cette logique les principes métaphysiques qui la sous-tendent. La tâche nous est grandement facilitée par la cohabitation de différents métamodèles : RDF, RDFS et OWL. Cette cohabitation, loin d’être aussi harmonieuse que ne le laissent penser les célèbres schémas les présentant comme des couches d’une même « pièce montée », nous révèle au contraire l’existence de controverses graves (Hayes, 2003) autour des principes (formels et métaphysiques) de la logique du Web.

Le point de divergence le plus visible (cf. Fig. 1) concerne ce qu’est une classe. OWL adopte une définition beaucoup plus contrainte que RDFS pour qui toute classe est également une ressource (Hayes, 2003). Autrement dit, exprimer à la fois « Callias est un *homme* » et « *homme* est une classe » est valide en RDFS mais ne l’est pas en OWL¹. La notion d’individu est ajoutée pour matérialiser cette différence.

Une autre différence notable repose sur le fait que les classes d’OWL doivent trouver leur place dans la hiérarchie des subsomptions entre deux bornes : *thing* et *nothing*. Ainsi « Tout rien est une chose ». Une fois passé le moment d’étonnement amusé, on ne peut que mesurer l’écart avec la métaphysique d’Aristote selon laquelle la *substance* des espèces (classes) provenait de celle des individus qui y étaient contenus. Autrement dit, on se refusait de parler de classes vides (Léonhardt, 2008, p. 28).

Enfin, il est remarquable que les différentes manières d’être dans OWL et RDFS (*type*, *subclassOf*, *disjointWith*) soient définies en RDF comme des prédicats. Autrement dit, le verbe « être » est ici un verbe comme les

¹ Tout au moins dans son premier niveau (OWL-DL).

autres. Ainsi est enjambée la ligne de démarcation entre la logique des ensembles et la logique des relations, ligne qui parcourt l'histoire de la logique (Blanché, 1970, p. 125, p. 196-201, p. 293-295) de la philosophie grecque jusqu'à certains formalismes contemporains comme les graphes conceptuels ou UML.

À cette étape de l'article, nous espérons avoir convaincu le lecteur que les ontologies du Web, présentées par certains comme de simples fichiers sans rapport avec l'« Ontologie », posent au contraire de réelles questions philosophiques relatives à la « substance ». Pour autant, le mystère reste entier sur ce que seraient ces « choses ». Appartiendraient-elles au réel ? au langage ? à la pensée ?

2. Vérité

Comme nous l'avons dit plus haut, la science, pour Aristote, a pour objet la connaissance de la substance du monde. Il distingue les substances *premières* identifiables par un nom propre² (ex. : *Callias*) et les substances *secondes* identifiables par un nom commun³ (ex. : *homme*), les premières étant, en quelque sorte, garantes des secondes. Dans un élan de prudence, Aristote considère que la sensation ne concerne que les premières alors que la science ne concerne que les secondes (Léohardt, 2008, p. 43-44). De là, découle une épistémologie assez alambiquée (cf. Fig. 2) dans laquelle :

- la vérité est transmise par déduction de l'universel (*Tout homme est juste*) au particulier (*Quelques hommes sont justes*),
- l'induction est considérée comme une faute logique,
- mais la substance des classes et la vérité des propositions universelles est révélée au philosophe de manière assez mystérieuse à partir du singulier (*Quand je vois Callias, c'est l'homme que je vois*⁴).

² Appelé plus haut « terme singulier ».

³ Appelé plus haut « terme universel ».

⁴ « Certes la sensation est distribuée individuellement, mais la sensation est de l'universel : ainsi celle d'un homme, mais non de l'homme Callias. » in Aristote, *Seconds Analytiques II*, 19, 100a, 20 (trad. Jacques Cazeaux).

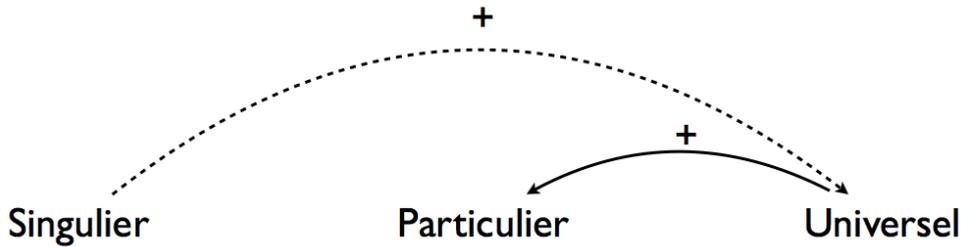


Fig. 2 – Parcours de la vérité chez Aristote (Léonhardt, 2008)

De manière plus « moderne », les ontologies du Web confondent le singulier et le particulier. L'objet n'est plus perçu comme singularité, mais comme régularité, comme instance d'une classe (Rousseaux, 2005). Dès lors, les mêmes substances seraient concernées par la sensation (qui devient « expérience ») et par le discours scientifique (en quête de vérité). Ce changement ontologique n'est pas sans conséquence sur l'épistémologie.

La première manière d'adapter l'épistémologie aristotélicienne serait celle du Cercle de Vienne (Positivisme Logique) dans laquelle la vérité des propositions universelles serait fournie par l'expérience (cf. Fig. 3). Comme l'a dénoncé Karl Popper (1934), une telle démarche est de nature inductive. Comment à la fois se réclamer de la logique et avoir de manière cachée une méthode basée sur un sophisme ?

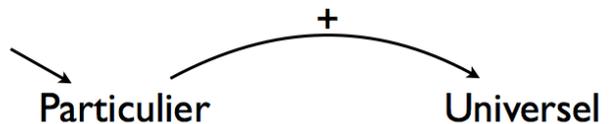


Fig. 3 – Parcours de la vérité chez les positivistes logiques (d'après Popper)

Au contraire, Karl Popper (ou avant lui Pierre Duhem) propose une méthode scientifique dans laquelle la vérité⁵ des propositions universelles n'est plus un prérequis ni même un objectif atteignable de la science. L'origine même de ces propositions devient extérieure à la méthode scientifique (« psychologique » pour Popper, « sans fondements » pour Feyerabend). Leur caractère hypothétique est transmis aux propositions particulières qui, elles, sont testables par l'expérience. Si le test est négatif, la fausseté des propositions particulières est transmise aux propositions universelles (cf. Fig. 4). On parle alors de réfutation.

⁵ Au sens de correspondance au monde.

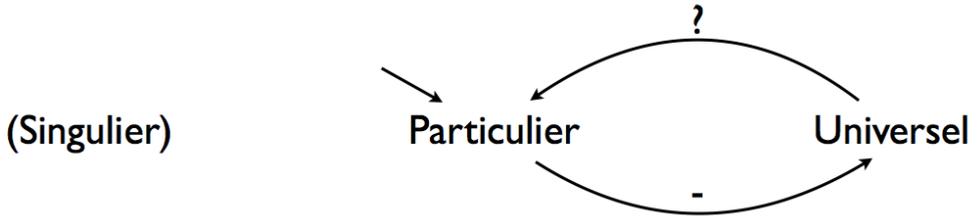


Fig. 4 – Parcours de la vérité chez Popper

Que dire alors des ontologies du Web (ex. Fig.5) à l'aune des sciences modernes ? Tout d'abord que leur « besoin de substance » (Farouki, 1996, p. 59-60) – au sens « d'expliquer le monde » – ne dépend plus de l'existence des classes ou des individus mais de la capacité de ces ontologies à pouvoir sans cesse être projetées sur le réel, à montrer les limites de ce qu'elles affirment et à être corrigées. Dès lors, on ne peut qu'être troublé par le fait que le protocole permettant de mettre à jour les ontologies (SPARUL) ait été normalisé 5 années après celui permettant de les interroger (SPARQL), que ces ontologies soient le plus souvent cachées de l'utilisateur, que leur complexité formelle rende quasiment impossible leur édition par les experts eux-mêmes au fil des réfutations successives.

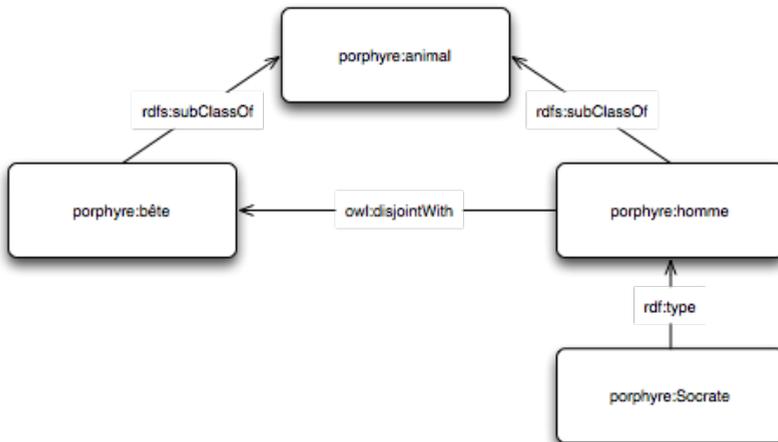


Fig. 5 – Exemple succinct d'ontologie pour le Web d'après l'*Isagoge* de Porphyre (graphe orienté étiqueté)

Ainsi, puisque les ontologies du Web ne semblent pas être accompagnées des outils et des méthodes requis par les théories de la connaissance modernes, dans quel courant s'inscrivent-elles et quelles en sont les conséquences sur leur statut et leur devenir ?

II. Une histoire refoulée

En première hypothèse, l'absence de méthodes et d'outils permettant de réfuter les ontologies du Web pourrait être expliquée par un certain positivisme « résiduel », mais cela ne suffirait pas à expliquer l'emprunt du terme « ontologie ». Dans la communauté anglophone de modélisation des connaissances, un des rares auteurs à faire référence à des lectures philosophiques est John Sowa (Sowa, 1992). Pour autant, quand il utilise le mot « ontologie » c'est pour attirer l'attention sur les problèmes philosophiques posés par l'analyse conceptuelle (Sowa, 1984) et pour attaquer vivement une intelligence artificielle qui ne prendrait pas acte de l'échec du Positivisme Logique (Sowa, 2001). En absence de filiation directe évidente du Cercle de Vienne aux ontologies du Web, nous étudierons, deux hypothèses de trajectoire de l'idée et du terme d'« ontologie ».

Une première piste nous est fournie par les travaux de Monique Slodzian sur un certain Eugen Wüster, à Vienne, qui, au moment de la dissolution du Cercle, s'en croit l'héritier. Contrairement aux membres du Cercle, Wüster n'est ni physicien, ni mathématicien, ni logicien, ni philosophe. Il s'agit d'un chef d'entreprise qui a vu dans le programme scientifique du Cercle l'opportunité de régler le problème de la communication entre ingénieurs de langues différentes. Pour cela, il définit la « Théorie Générale de la Terminologie ». De la langue, il ne garde ni les verbes, ni la syntaxe : seuls les termes ont grâce à ses yeux. Ces termes sont ensuite structurés dans ce qu'il appelle lui-même une « ontologie » (Slodzian, 2006). Il est important de noter que sa thèse est intitulée « Standardisation Internationale du Langage en Ingénierie » (1931) et que, de fait, Wüster sera à l'origine de l'une des commissions de l'ISO. Ainsi, entre le programme du Cercle de Vienne et celui de Wüster, l'objectif a sensiblement changé : là où le premier faisait de la science, le second fait de l'ingénierie ; là où le premier espérait décrire la Nature, le second décrit l'Artificiel. Finalement, l'Ontologie devient nomenclature, et le terme devient un symbole purement conventionnel (Slodzian, 2006). De fait, nous ne sommes pas très loin de ce que François Rastier reproche à la « Web Science » (Rastier, 2008), pas très loin non plus de ce que Tim Berners-Lee, lui-même, déclare dans son entretien avec Alexandre Monnin et Harry Halpin :

Quand on conçoit un protocole, en réalité on crée... On a la chance, en réalité, de définir comment fonctionne un nouveau monde. (...) Quand vous créez un protocole, vous obtenez le droit

*de jouer à Dieu, de définir ce que les mots veulent dire. (...) Les gens doivent coopérer (...) en étant d'accord.*⁶

Tim Berners-Lee (Halpin et Monnin, 2010)

En partant des articles d'aujourd'hui et en remontant les références, on pourrait espérer trouver d'autres pistes. Cependant la tâche se révèle plus difficile que prévue. Il semble en effet que toutes ces références nous mènent à un article court de Thomas Gruber (1991). Cet article lui-même ne cite que six articles, dont les plus anciens sont de l'année précédente. Face à une telle impasse, le lecteur curieux n'a plus qu'à chercher le mot « ontology » dans une base bibliographique (cf. Tab. 1). Au-delà de la croissance géométrique de son usage⁷, on peut noter une évolution dans les courants les plus cités à son sujet.

Décennie	Nombre	Auteurs les plus cités (3)
1930-1939	$3,8*10^2$	Coomaraswamy, Ginzburg, Somerville
1940-1949	$6,0*10^2$	Cerf, Quine, Cellars
1950-1959	$1,5*10^3$	Quine, Cartwright, Wells
1960-1969	$3,1*10^3$	Sartre, Bazin, Sommers
1970-1979	$5,8*10^3$	Cavell, G'ivon, Hellman
1980-1989	$1,4*10^4$	Moens, Cracraft, Horgan
1990-1999	$6,6*10^4$	Gruber, Guarino, Ushold
2000-2009	$3,1*10^5$	Euzenat, Ashburner, McGuinness

Tab. 1 – Recherche dans Google Scholar des articles (avec résumé ou texte complet) contenant le mot « ontology » (ou ses variantes).

Dans les années 1930, à l'époque du Cercle de Vienne, le mot semble rester l'apanage de l'étude des religions et de la philosophie des sciences. Ce n'est que dans les années 1940 et 1950, que l'on perçoit les conséquences du Cercle de Vienne avec l'arrivée de Quine dans les auteurs les plus cités. La philosophie analytique semble passer au second plan dans les années 1960, derrière la phénoménologie et les réflexions sur l'art. Elle revient au premier plan en 1970, accompagnée d'une linguistique d'inspiration viennoise. Cette tendance semble se renforcer dans les années

⁶ « When we design a protocol, we're actually creating... we get the chance to actually define the way a new world works. (...) When you create a protocol, you get the right to play God, to define what words mean. (...) People (...) have to join in, (...) with agreeing. »

⁷ Croissance sans doute biaisée par une mauvaise représentativité des contenus numérisés, et surtout par l'explosion du nombre de publications scientifiques au XXe s.

1980 sous la forme d'une linguistique « computationnelle »⁸. Enfin dans les années 1990, on retrouve l'ontologie informatique de Gruber et Guarino ; et dans les années 2000 les ontologies du Web sémantique.

On ne peut évidemment se satisfaire complètement d'une étude quantitative : les résultats restent assez grossiers et leur valeur purement heuristique. On notera cependant que, même si les filiations sont difficiles à établir, on observe une réelle continuité chronologique de l'usage du mot « ontologie » du Positivisme Logique jusqu'à l'informatique. Par ailleurs, on notera que l'un des maillons essentiels de cette continuité semble être Quine. De plus amples recherches sur ceux qui en informatique se réclament de Quine et parlent d'ontologie nous permettent de découvrir qu'il s'agit de John McCarthy, le même qui en 1956, à la célèbre conférence de Dartmouth, présentait le concept d'« Intelligence Artificielle ». Celui-ci évoque, au hasard d'une discussion sur la gestion des exceptions aux règles, la nécessité d'inclure la formalisation de ces « *connaissances de sens commun* » dans « *notre ontologie (les choses qui existent)* » (McCarthy, 1980 ; Monnin, 2013, p. 45). On découvre ensuite que Patrick Hayes suit son exemple en 1985 et que son ontologie est celle d'une « physique naïve », naïve, non pas au sens d'une simulation du raisonnement des humains dans leur vie quotidienne, mais au sens d'un monde saisi de manière pré-théorique et sur lequel on raisonne de manière formelle (Smith et Casati, 1993). Aujourd'hui, comme nous l'avons évoqué plus haut, Patrick Hayes participe à l'élaboration de RDF-core, SPARQL et OWL.

Ainsi, se dessine bien une filiation directe entre le Positivisme Logique et les ontologies du Web, mais à travers Quine qui, nous devons l'admettre, développait des vues sur la vérité fort différentes de celle du Cercle de Vienne. Il serait téméraire de les résumer en une seule définition. L'œuvre de Quine, en refusant la distinction entre propositions analytiques et synthétiques et en s'opposant à toute épistémologie normative, reste assez inclassable.

Nous nous en tiendrons à son petit essai « *sur ce qui est* » (1948) et à l'intérêt que ce dernier pouvait représenter pour McCarthy et ses « systèmes intelligents » basés sur la logique. Dans cet article, Quine rejette l'idée selon laquelle l'existence pourrait être découverte ou inventée. Pour lui, la référence n'est pas une question de *noms* mais de *pronoms* (de « variables unifiées »⁹ en logique formelle). En d'autres termes, le référent est interne au langage. Aussi singulière que puisse être cette position sur l'ontologie, elle offrait des avantages considérables au fondateur de l'Intelligence Artificielle :

⁸ On notera aussi une percée en zoologie.

⁹ *Bound variables* en anglais.

- exprimée dans les termes de la logique formelle et avec des références incessantes à Frege et Russel, elle était en quelque sorte « prête à l'emploi » pour des systèmes basés sur la logique,

- en affirmant que les mathématiques étaient seulement un langage parmi d'autres, elle laissait penser qu'elle pouvait s'appliquer à tous les champs couverts par le langage humain. McCarthy s'intéressait d'ailleurs également à d'autres philosophes du langage qui partageaient cette idée de la référence interne (Grice, Searle, etc.).

Il est significatif qu'une telle ontologie soit décrite par Quine comme un *mythe* que nous nous « engageons » à croire ou à ne pas croire, avec lequel nous pouvons être en désaccord, que nous pouvons comparer à d'autres, et dont la vérité (ou la fausseté) ne dépend pas de l'expérience. Appliquée à l'Intelligence Artificielle, une définition de « l'ontologie » aussi hétérodoxe, n'est pas sans conséquences : elle lui confère un statut spéculatif et difficilement compatible avec des applications au monde réel si ce n'est au prix de dangereux compromis. Cela pourrait expliquer pourquoi en 1982, quand Newell proposa d'ajouter la « couche des connaissances » aux systèmes basés sur la logique, l'Intelligence Artificielle était encore aussi éloignée des applications réelles (Rousseaux, 2005). Cela pourrait expliquer pourquoi « l'invention des connaissances » apporta à l'Intelligence Artificielle, dans un même mouvement, des domaines d'application insoupçonnés et de nouvelles ambiguïtés sur le statut de ses productions.

III. Des perspectives paradoxales

Comme nous venons de le voir, les ontologies du Web prennent leur source dans les œuvres de Wüster, de Quine et de McCarthy, œuvres dans lesquelles la vérité sort du cadre de la théorie de la connaissance de Popper, mais également de celle du Cercle de Vienne ou d'Aristote. L'épistémologie des travaux étudiés¹⁰ est davantage celle des mathématiques que celles des sciences : comme chez Hilbert redéfinissant de manière formelle la géométrie ou Lewis Carroll énonçant ses énigmes burlesques, ce qui est vrai l'est par *convention*, dans *l'univers du discours*. Pour autant, les raisons de cette démarche diffèrent d'un auteur à l'autre :

- Wüster s'imagine en ingénieur-démiurge – il *dit* et les artefacts *sont*,
- McCarthy, en pédagogue, a recours à des énigmes pour apprendre aux machines à raisonner,

¹⁰ Nous rappelons que nous ne traitons pour Quine que de son essai « On what there is » et non de ses œuvres complètes.

- Quine, qui fait ici figure d'observateur des comportements linguistiques, prend au sérieux ce qui est dit, indépendamment de toute réalité extra-linguistique.

Essayons maintenant d'explorer vers quel type d'ingénierie des connaissances nous mène chacune de ces trois pistes.

1. Connaissances et ingénierie

Un projet de gestion des connaissances dans l'industrie¹¹ (Zhou, 2009, p. 66-72) nous a permis de nous rapprocher du monde de Wüster. Chez notre partenaire industriel, la nomenclature occupe effectivement une place centrale dans son système d'information, la composition des produits fournissant la structure d'intégration des données de l'entreprise (on parle alors de « PLM » – *Product lifecycle management*).

Après un premier examen, la correspondance un peu trop parfaite entre la structure méréonomique¹² des données et celle des produits concrets fait place à une complexité et un raffinement rares, notamment autour de la notion de « pièce » (*part*) :

- la *pièce* n'est pas l'objet concret, mais une définition,
- par ailleurs la pièce est aussi une abstraction de ses différentes versions,
- la pièce peut être définie à partir d'autres pièces, permettant ainsi la réutilisation des définitions (et des processus de fabrication) au sein du même produit ou même entre plusieurs produits,
- dans une définition donnée, la référence à une autre (appelée *instance*) possède ses propres attributs (notamment de nom, de position, de couleur),
- cette « instance » étant elle-même abstraite (ex : « moteur de (tout) préhenseur FlexGrip »), il est possible de préciser les attributs spécifiques d'une *occurrence* individualisée dans le produit (ex : « moteur du premier préhenseur FlexGrip du module Flexicell d'assemblage de carrosserie dans la chaîne de fabrication de la Clio »).

En somme, les référents de la *pièce* (et par conséquent du *produit*) ne sont pas les objets manufacturés mais les documents de conception. Autrement dit, le caractère d'ontologie conventionnelle des nomenclatures industrielles ne provient pas de la correspondance immédiate à un monde artéfactuel (comme on aurait pu le croire avec Wüster), mais de leur caractère éminemment documentaire.

¹¹ *i-Semantec*, projet financé de 2007 à 2009 par le programme « Technologies logicielles » de l'Agence nationale de la recherche, et réunissant les entreprises ABB, CADeSIS et Topica tools ainsi que notre laboratoire.

¹² Méreonomie : terme construit à partir de «méréo-», membre et «-nomie», loi. Structure hiérarchique basée sur la relation entre le tout et ses parties.

Autre conséquence : cette place éminente des documents de *conception* se fait souvent au détriment de la spécificité des autres métiers concernés. Dès lors, ces métiers établissent leurs propres bases de données et de documents qui échappent alors à la « capitalisation ». Par ailleurs, dans un contexte d'instabilité du marché, il semble nécessaire, pour répondre aux besoins des clients, d'adapter le système d'information aux reconfigurations permanentes (postérieures à la conception) des processus et des réseaux de partenaires.

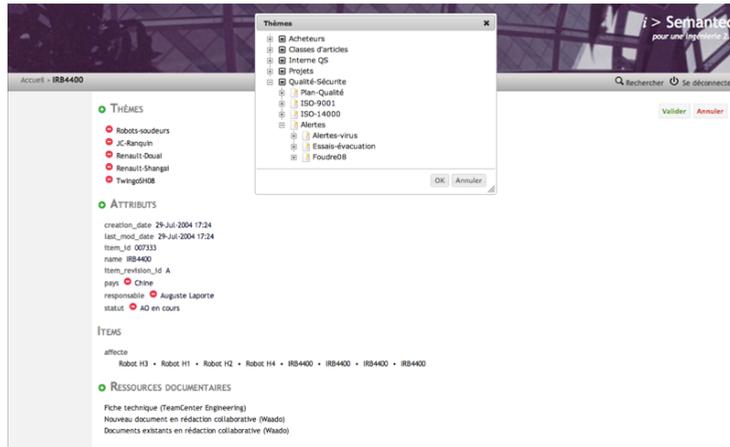


Fig. 6 – Accès unifié aux attributs stabilisés, aux attributs émergents, aux ressources documentaires ainsi qu'à la thématization participative (copie d'écran d'*Agorae*)

La solution proposée (cf. Fig. 6) consiste à extraire à la volée les *produits* et *pièces* du PLM sous forme d'*items* « documentarisés » (Briet, 1951 ; Zacklad, 2005), de sorte que les autres métiers (service des achats, service qualité...) puisse les thématiser avec leurs propres *points de vue*, les enrichir avec des attributs et des documents émergents, et enfin les relier à de nouveaux items non pris en compte au moment de la conception.

2. Connaissances et jeux éducatifs

Dans le cadre d'un partenariat avec une entreprise développant un jeu pour l'éducation au développement durable¹³, nous avons pu aborder la question des micro-mondes si chers à l'Intelligence artificielle et aux pédagogues. Comme chez McCarthy ou chez Carroll, l'invention d'un micro-monde donne la liberté de définir conventionnellement les règles qui le régissent. Cependant, dans notre cas, ce qui est enseigné n'est pas le

¹³ Société SOS-21, <http://www.sos-21.com/>

raisonnement indépendamment des contenus manipulés, mais la compréhension de l'impact environnemental des comportements quotidiens. Dès lors, il ne peut plus être question de décider de ces règles indépendamment de ce que l'on en sait dans le monde réel.

Comme toute connaissance scientifique, ces règles quand elles sont réfutées doivent pouvoir être remplacées. De plus, le développement durable, au croisement de plusieurs disciplines scientifiques mais aussi de l'éthique et du politique, requiert constamment de débattre, y compris de ce qui pourrait passer pour des évidences. Enfin, contrairement à ce que défendent certains chercheurs, l'apprentissage par le jeu ne peut se résumer à un modèle béhavioriste dans lequel la répétition de bonnes pratiques au sein du jeu « persuaderait » les joueurs d'adopter les mêmes comportements dans le monde réel. Cet apprentissage, selon nous, doit passer par la compréhension fine des règles et de leurs enjeux, ce qui nécessite souvent, au contraire, de tester les effets des « mauvaises » pratiques. Il faut fournir au joueur toutes les ressources lui permettant de se construire son propre modèle des règles individuelles et de ce qui constituerait une stratégie globale gagnante.

La solution que nous préconisons (El Mawas et al., 2012) consiste d'une part à gérer les règles non plus comme des programmes mais comme des données¹⁴, et d'autre part à associer à chacune de ces règles un fil de discussion dans un forum (cf. Fig. 7). Autour de ces fils de discussion vont se retrouver tous ceux qui s'intéressent aux règles : les concepteurs du jeu, les experts du domaine, les joueurs « bloqués » ou en quête de meilleurs scores, les joueurs vétérans souhaitant partager leurs « trucs et astuces ». Autour de cet *item* partagé, les *points de vue*, parfois conflictuels, viennent s'enrichir les uns les autres. Dans chaque message, les participants peuvent, sous la forme d'un tableau, proposer la nouvelle définition de la règle. Si une proposition est validée (par la personne ou l'équipe en charge), la règle est automatiquement mise à jour dans le jeu. Pour autant, cette décision ne clôt la controverse que pour un temps, son historique reste visible et la discussion ouverte.

À l'image de ce que Bruno Latour dit de la science (Latour, 1987), on pourrait affirmer que, dans les jeux éducatifs, à condition de regarder les règles en train de se faire, la boîte noire devient boîte de Pandore.

¹⁴ Ce qui, d'ailleurs, renoue avec les principes de l'Intelligence artificielle de la fin des années 1950 (McCarthy, 1980).

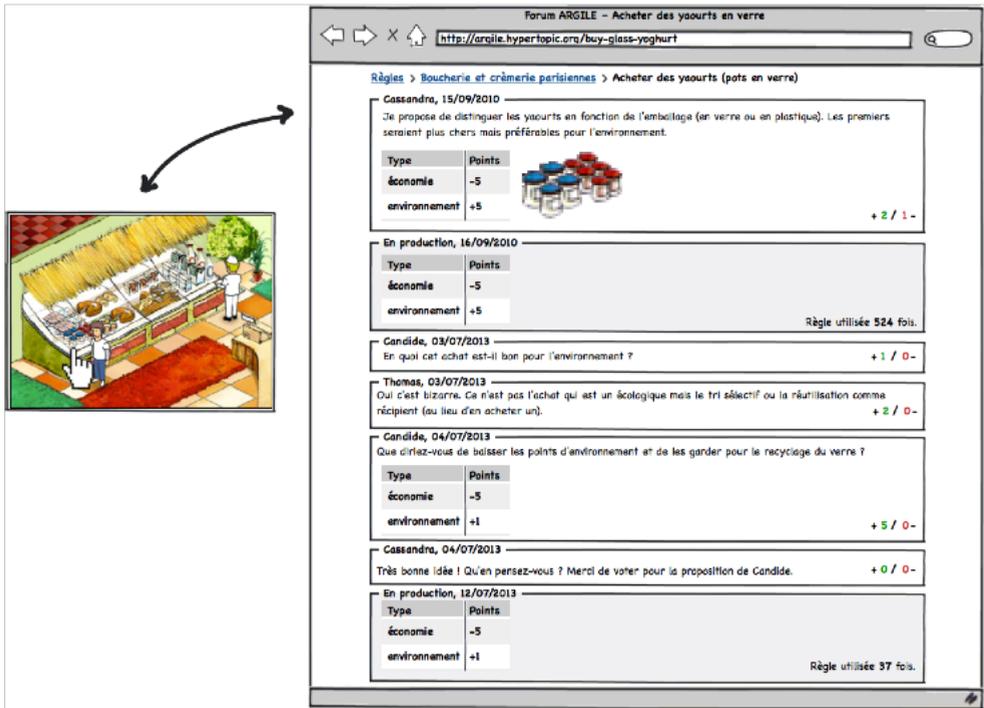


Fig. 7 – Quand apprendre par le jeu passe par la discussion des règles (maquette d'ARGILE, d'après une scène de SOS-21)

3. Connaissances et musées

Pour dénoncer les approches du langage considérant les mots comme de simples étiquettes interchangeable épinglées sur des significations ou des objets concrets, Quine parlait du « mythe du musée ». L'étude, à l'aide de nos outils, de vases conservés au Musée du Céramique (à Athènes) par des étudiants en histoire de l'art¹⁵ nous offre donc une occasion d'illustrer et de questionner les vues de Quine.

La manière la plus naïve de traiter ces vases de manière quinienne consisterait à formaliser l'étiquette (indépendamment de l'objet concret) :

Lécythe
trouvé dans une tombe du Céramique
daté de 1200 à 1100 av. J.C.
Inv. 450

deviendrait :

¹⁵ Sous la direction de Jean-Marc Luce, Équipe PLH-CRATA, Université Toulouse II.

type : Lécythe
lieu : Céramique
datation min. : -1200
datation max. : -1100
ID : Kerameikos#450

La valeur ontologique (au sens de Quine) de ces attributs et de ces valeurs viendrait du fait qu'ils soient également présents :

- dans la notice générique des céramiques grecques antiques,
- dans les autres notices basées sur ce modèle,
- ou, en ce qui concerne le nom propre « Céramique », dans les cartes géographiques.

À condition d'utiliser des identifiants uniques à l'échelle du Web et une syntaxe standard, on obtiendrait ce que certains appellent les « données reliées » (*Linked data*). Pour autant, le résultat, même mondialisé, demeure une simple base de données. Proposer un tel système aux étudiants reviendrait à leur présenter l'historien d'art comme un administratif qui applique une procédure ou, pour reprendre le personnage de l'archéologue dans *l'Automne à Pékin* de Boris Vian, comme un maniaque qui casse des pots pour les ranger dans des boîtes de format standard (Bruneau, 1976).

La prise en compte de l'*expertise* de l'historien d'art nécessite un travail d'ingénierie des connaissances et, pour commencer, de déterminer ce qui a du sens pour lui et pourquoi. Ainsi, dire que ce vase est un « lécythe » consiste à reconnaître dans sa forme un usage possible (vase à huile ou à parfum) et à le différencier d'autres formes existant pour le même usage (« alabastre », « aryballe »...). Pour autant, cet usage primaire n'empêche pas un détournement ou une réutilisation pour un autre usage (funéraire notamment). Pour le lieu de la découverte, « Céramique », le toponyme connu de Google Maps, ne suffirait pas, pas plus que la longitude et la latitude. Dans l'espace tridimensionnel de la nécropole, ce qui a du sens est la sépulture (« Céramique Tombe 42 »). En ce qui concerne la datation, celle du physicien présente relativement peu d'intérêt pour l'historien d'art. L'objet, en tant qu'artefact, doit être avant tout rattaché à une culture. Parce que sa forme a été reconnue comme celle d'un « lécythe », on va pouvoir, en fonction de son style, le positionner dans des séries de lécythes, et le rattacher à la période « HR III C récent », c'est-à-dire la partie la plus récente de la plus récente de la plus récente de la plus récente de l'Helladique, juste avant le sub-mycénien, autrement dit, à l'aube d'un changement de culture majeur.

Dans cette deuxième version du musée, les étiquettes deviennent beaucoup plus intéressantes. Le caractère ontologique (au sens de Quine) de « HR III C récent », « Céramique Tombe 42 », « Lécythe » peut être compris comme l'effet de « résonance » des étiquettes entre elles (qui reprennent certaines de ces valeurs), mais aussi des étiquettes avec une frise chronologique et un plan de la nécropole. Le visiteur du musée peut jouer à l'historien d'art, comprendre les raisonnements qui ont permis d'aboutir à certaines notices et essayer de deviner ce que l'expert a inscrit sur les autres.

Si la première version correspondait aux « données reliées », celle-ci correspondrait au « Web sémantique », tout au moins tel qu'il devrait être¹⁶.

Cette approche, très instructive, pose cependant un problème dès lors que l'on s'en sert comme instrument pour les chercheurs, dès lors que l'ingénieur des connaissances se retire et que demeure son système. Un chercheur n'est pas seulement un praticien. Il doit pouvoir critiquer le savoir de référence et construire peu à peu, par essai-erreur, sa propre théorie.

Dans une troisième et dernière version du musée, les chercheurs viennent, à n'importe quelle heure du jour ou de la nuit, épingler leurs propres étiquettes, frises et plans. Chacun de ces éléments participe d'un discours, porté, à un moment donné, par le chercheur ou son équipe. Se déroule, sous les yeux du visiteur, l'histoire de l'art en train de s'écrire, avec ses controverses et ses révolutions conceptuelles.

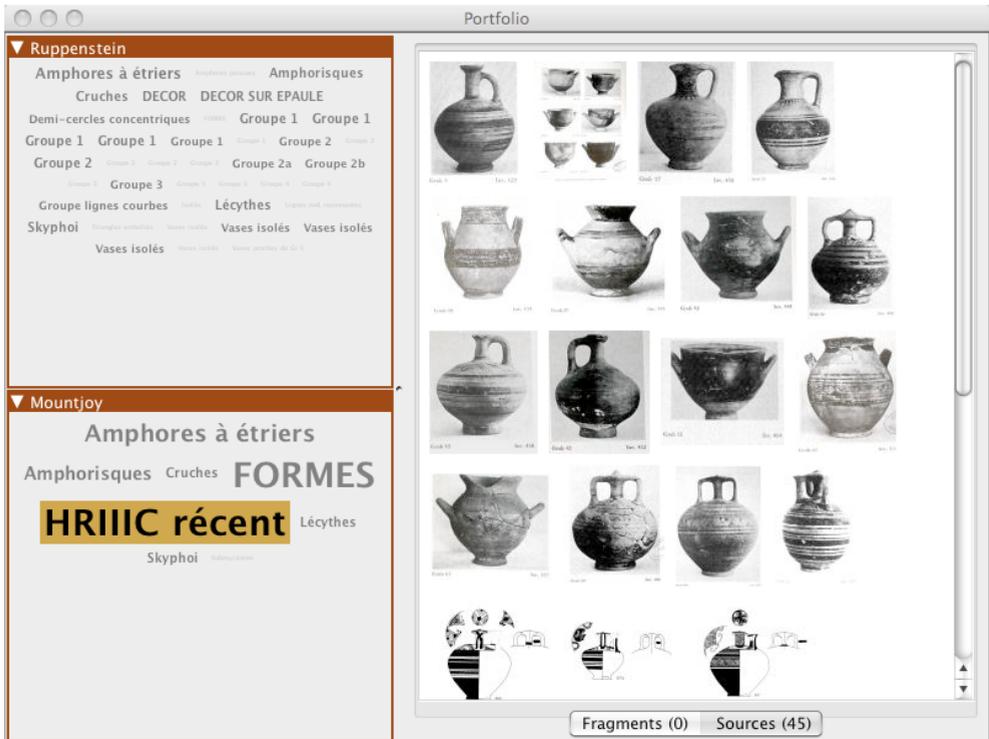
Si un tel dispositif est difficilement réalisable dans un espace physique, nous avons pu montrer qu'il l'était dans un espace numérique, préfigurant ce que nous appelons le « Web socio-sémantique ». Dans notre exemple (cf. Fig. 8), on peut voir deux études en concurrence sur le même corpus. Ces *points de vue*, beaucoup moins formalisés que les « ontologies », sont modélisés par les historiens d'art eux-mêmes¹⁷ et non plus par des ingénieurs de la connaissance. Dans les points de vue, les *thèmes* peuvent être construits comme des *classes*, déterminées dès le départ (« HRIIC récent », « submycénien »), ou comme des *collections* (Rousseaux, 2005), constituées de manière plus émergente, au fur et à mesure de la découverte de singularités dans le corpus (« groupe 1 », « groupe 2 » ...).

Plus encore que dans l'approche précédente, le caractère ontologique (au sens de Quine) du système vient des effets de résonance provenant des réemplois successifs des mêmes éléments. Par conséquent, à la différence des outils du Web sémantique qui fournissent des inférences déductives, ceux du Web socio-sémantique se focalisent sur des cooccurrences. Pour ne pas être interprétées comme des inductions amplifiantes, ces cooccurrences ne sont pas exploitées dans des « réponses automatiques à des questions », mais comme des supports à l'exploration, l'exploration restant pilotée par l'utilisateur. Celui-ci va par exemple interpréter la figure 8 comme la remise en question de l'étude qui faisait autorité par une nouvelle étude. Là où la première étude rattachait les vases à une seule époque (« HRIIC récent »), la nouvelle distingue, pour chaque type de vases, deux groupes stylistiques. En faisant cela, elle introduit un style transitoire (« groupe 2 ») qui

¹⁶ On comprendra que, selon notre typologie, les projets qui se revendiquent du Web sémantique sans faire l'objet d'une réelle ingénierie des connaissances appartiennent de fait aux « données reliées ».

¹⁷ Les points de vue peuvent être modélisés par les auteurs des études eux-mêmes, ou comme c'est le cas ici par un autre chercheur du domaine.

permettra de mieux comprendre cette époque de changement culturel (Bénel



et al., 2010).

Fig. 8 – Quand le savoir de référence est remis en question par une nouvelle étude (copie d'écran de Porphyry)

Ce regard quinien sur le Web socio-sémantique ne serait pas complet si on ne relevait pas la substitution de l'identifiant du vase par celui de sa photographie numérique – l'identifiant du vase ne devenant qu'un simple attribut de la photographie. Les publications en histoire de l'art et archéologie font déjà un recours massif aux photographies, plans, coupes. Même si ces documents visuels sont eux-mêmes une « vision du monde », ce sont les seuls à permettre la reconnaissance de l'objet analysé malgré les changements d'inventaire, inévitables, vu l'ancienneté de ces disciplines (Iacovella et al., 2005). Plus encore, le remplacement du « document » muséal (Briet, 1951) par un document photographique permet en quelque sorte d'*internaliser* le référent au discours des chercheurs, permet à ce discours de remplir lui-même les conditions de conservation et d'accès aux sources, nécessaires à de nouvelles interprétations.

Conclusion

À l'issue de cet itinéraire, au-delà de la seule question des « ontologies du Web », se dessine celle de ce qu'est l'informatique et surtout ce qu'elle pourrait être.

Dans un premier temps, nous avons montré qu'une réflexion philosophique sur les « ontologies du Web » ne pouvait se borner à critiquer un choix terminologique malheureux, et que, de fait, ces artefacts posaient de manière aigüe la question de l'articulation du langage avec le réel, la raison, la connaissance et sa genèse.

Contrairement à nos travaux précédents où nous avons mené une attaque plus « frontale » (Bénel et al., 2004), nous avons adopté cette fois-ci une méthode « compréhensive » visant à construire notre critique des ontologies de l'intérieur même du système de pensée de leurs promoteurs. Pour ce faire, afin de mieux cerner leur « métaphysique », nous avons tout d'abord essayé de retrouver l'histoire refoulée de l'emprunt du terme « ontologie » en informatique. À notre grande surprise, les trois maillons principaux de cette transmission (Wüster, Quine et McCarthy) défendent, pour des raisons et avec des modalités différentes, une ontologie *interne* à l'univers du discours.

Illustrant les univers de chaque auteur à l'aide de nos terrains d'expérimentation, nous avons tenté de montrer comment leur prise en considération nécessitait un dépassement des *ontologies* via le recours à des *points de vue* et des *documents*. En somme, nous avons essayé d'étayer l'idée selon laquelle les défenseurs des ontologies du Web, s'ils assumaient pleinement les soubassements philosophiques de leur approche, en viendraient, de manière tout à fait paradoxale, à adopter nos « alternatives à l'ontologie ».

Pour autant, notre approche (inspirée de l'herméneutique et de la sémiotique des textes) ne peut se réduire à ce qui peut en être dit à partir du cadre philosophique des ontologies. Pour pleinement comprendre ce qui est à l'œuvre dans nos terrains, il serait nécessaire par exemple de prolonger la réflexion de Quine sur les pronoms à ce qui se joue réellement dans un texte (et entre textes) dans le cas de *reprises* d'une expression ou d'une figure. Contrairement à une variable mathématique qui est « unifiée » une fois pour toutes, le sens de l'expression ou de la figure reste en suspens. De reprise en reprise, au contact de ses différents contextes, ce sens glisse, vole en éclat ou prend forme. Collections de singularités plutôt que classes à instancier (Rousseaux, 2005), ces nouveaux objets viennent ébranler les bases mêmes de l'informatique. Les ouvriers seront-ils assez nombreux pour les reconstruire ?

Bibliographie

Bénel A., Calabretto S. (2004), Ontologies... déontologie : réflexion sur le statut des modèles informatiques, *Linguistica Computazionale*, XX-XXI, 31-47.

- Bénel A., Zhou C., Cahier J.-P. (2010). Beyond Web 2.0... And Beyond the Semantic Web. In D. Randall, P. Salembier (Eds.) *From CSCW to Web 2.0: European Developments in Collaborative Design* (pp. 155-171). Springer Verlag.
- Berners-Lee T., Hendler J., Lassila O. (2001). The Semantic Web, In: *Scientific American*, May 2001, 34-43.
- Blanché R. (1970). *La logique et son histoire*. Armand Collin, 2002. Première édition en 1970, revue et complétée en 1996.
- Briet S. (1951). *Qu'est-ce que la documentation ?* Editions documentaires, industrielles et techniques.
- Bruneau, Ph. (1976). Quatre propos sur l'archéologie nouvelle. *Bulletin de Correspondance Hellénique*, 100, 1, 103-135.
- Charlet J., Bachimont N., Bouaud J., Zweigenbaum P. (1996). Ontologie et réutilisabilité : expérience et discussion. In N. Aussenac-Gilles, P. Laublet, C. Reynaud (Eds.) *Acquisition et ingénierie des connaissances* (pp. 69-87). Cepaduès Editions.
- El Mawas N., Cahier J.-P., Bénel A. (2012). ARGILE : Apprendre par le jeu des connaissances en construction. *Revue des Interactions Humaines Médiatisées (RIHM)*, XIII, 1, 93-120. Europaia.
- Farouki N. (1996). Les motivations de la connaissance. *La foi et la raison : Histoire d'un malentendu* (pp. 58-67). Paris : Flammarion.
- Gruber T. R. (1991). The Role of Common Ontology in Achieving Sharable, Reusable Knowledge Bases. In J. A. Allen, R. Fikes, E. Sandewall (Eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference* (pp. 601-602). Cambridge MA: Morgan Kaufmann.
- Gruber T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, V, 2, 199-220.
- Halpin H., Monnin A. (2010). *Tim Berners-Lee on PhiloWeb and Philosophical engineering*, Interview. Disponible sur Dailymotion.
- Hayes P. (2003). Response to "owl:Class vs rdfs:Class causing pain. Is owl:Class really needed?", In : *www-rdf-comments mailing list*, W3C, 17 sept. 2003.
- Iacovella A., Bénel A., Calabretto S., Helly B. (2005). Assistance à l'interprétation dans les bibliothèques numériques pour les sciences historiques. *Actes du colloque de bilan du programme interdisciplinaire "Société de l'information"* (pp. 167-179).
- Latour B. (1987), *La science en action* (Édition originale en anglais en 1987 et en français en 1989), Paris : Gallimard .
- Léonhardt J.-L. (2008). *Le rationalisme est-il rationnel ? L'Homme de science et sa raison*, Parangon.
- Lifschitz V. (1990). *Formalizing Common Sense: Papers by John Mc Carthy*. Norwood (New Jersey): Ablex Publishing.
- McCarthy J. (1980). Circumscription a Form of Nonmonotonic Reasoning. *Artificial Intelligence*, XIII, 27-39.
- Monnin A. (2013). *Vers une Philosophie du Web : Le Web comme devenir-artefact de la philosophie (entre URIs, Tags, Ontologie(s) et Ressources)*. Thèse de doctorat. Université Paris I.
- Popper K. R. (1934). *The Logic of Scientific Discovery* (translated from German in 1959). New York: Basic Books.
- Quine W. V. O. (1948). On What There Is, *Review of Metaphysics*.

- Rastier F. (2008). Sémantique du web vs semantic web ? – Le problème de la pertinence, *Syntaxe & Sémantique*, IX, 15–36.
- Ray K. (2010). *Web 3.0: A story about the Semantic Web*. Documentaire vidéo. Disponible sur Vimeo.
- Rousseaux F. (2005). Par delà les Connaissances inventées par les informaticiens : les Collections ? *Intellectica*, XLI-XLII, 2-3, 225-252.
- Shirky C. (2005). Ontology is Overrated: Categories, Links, and Tags. *O'Reilly Emerging Technology Conference (ETech'2005)*, San Diego, March 14-17, 2005.
- Slodzian M. (2006), La terminologie : historique & orientations, Conférence invitée, *Semaine de la Connaissance*, Nantes, 26-30 juin 2006.
- Smith B., Welty C. (2001). Ontology: Towards a new synthesis, *Proceedings of the International Conference on Formal Ontology in Information Systems* (pp. 3-9).
- Smith B., Casati R. (1993). La physique naïve : un essai d'ontologie, *Intellectica*, XVII, 2, 173-197.
- Sowa J. F. (1984). *Conceptual Structures: Information Processing in Mind and Machine*. The Systems Programming Series. Reading: Addison Wesley.
- Sowa J. F. (1992). Semantic Networks, In: S. C. Shapiro (Ed.), *Encyclopedia of Artificial Intelligence*, 2nd edition, vol. 2 (pp. 1493-1511). New York: Wiley.
- Sowa J. F. (2001). Signs, processes, and language games: Foundations for ontology, *Ninth International Conference on Conceptual Structures*, « Broadening the Base », California, July 30-August 3, 2001. Invited talk.
- Zacklad M. (2005). Processus de documentarisation dans les Documents pour l'Action (DopA) : statut des annotations et technologies de la coopération associées. *Actes du colloque "Le numérique : Impact sur le cycle de vie du document pour une analyse interdisciplinaire"* (Montréal, 13-15 octobre 2004). Éditions de l'ENSSIB.
- Zhou C. (2009). *Scaling up the Socio-semantic Web: Application protocol and user interfaces*. Thèse de doctorat. Université de technologie de Troyes.