

Ministère de la Culture et de la Communication  
Délégation au développement et aux affaires internationales  
Mission de la recherche et de la technologie

## « Muséologie, Muséographie et nouvelles formes d'adresse au public »

En collaboration avec  
L'Institut de Recherche et d'Innovation du Centre Georges Pompidou  
L'Ecole du Louvre  
le Musée des civilisations de l'Europe et de la Méditerranée  
et le LEDEN/Paris 8

21 mars 2007

### Ingénierie des connaissances

Bruno Bachimont, Université de Technologie de Compiègne  
Nathalie Leleu, MNAM/Centre Pompidou  
Jean Delahousse, Mondeca

#### Contexte de l'ingénierie des connaissances

« La quantité d'informations disponible mondialement double chaque année. Cette prolifération des données, disponibles sur différents média publics ou privés (TV, Internet, bases de données textuelles, enregistrements vidéo, enregistrements sonores...), combinée à leur hétérogénéité (texte, voix, image et vidéo, données numériques) rend la compréhension et l'analyse de l'information de plus en plus difficiles et pourtant indispensables. La croissance exponentielle des informations, loin de rendre l'environnement plus intelligible, en éloigne le sens. Après la révolution des PC, mettant le traitement des données à la portée de tous, la révolution des télécommunications mobiles et l'Internet, permettant aux individus de communiquer librement entre eux et d'avoir accès à des données nombreuses, les techniques de recherche et d'analyse de l'information nous acheminent vers une troisième révolution, celle de la connaissance »<sup>1</sup>.

---

<sup>1</sup> Source : Agence régionale de développement Paris Ile-de-France  
[www.paris-region.com/upload/document/D169.pdf](http://www.paris-region.com/upload/document/D169.pdf)

## Définitions de l'ingénierie des connaissances

1. La recherche que je mène s'inscrit sous l'égide de l'ingénierie des connaissances, en adoptant une double approche,

- une approche technique et scientifique,
- une approche philosophique et sémiotique.

### Qu'est-ce que l'ingénierie des connaissances ?

L'ingénierie des connaissances (IC) ne porte pas directement sur les connaissances, car ces dernières ne sont pas des objets matériels sur lesquels effectuer des manipulations et transformations. L'IC porte sur l'inscription matérielle des connaissances, en se fondant sur le fait que :

- toute connaissance est l'interprétation d'une inscription qui en est l'expression,
- toute inscription est matérielle et soumise à ce titre à une physique et peut faire l'objet d'une ingénierie.

L'IC ne porte pas sur toutes les inscriptions de connaissances, mais celles qui adoptent le support numérique comme substrat d'inscription. Le numérique confère une cohérence et une unité à l'ingénierie des connaissances :

- le support numérique est universel, et tout contenu peut s'inscrire sur un tel support,
- le support numérique est homogène au sens où les contenus inscrits peuvent être soumis à un même système technique.

L'ingénierie des connaissances sera donc l'ingénierie des inscriptions numériques de connaissances. Elle élabore des outils, méthodes et dispositifs mobilisant les inscriptions numériques pour assister le travail de la pensée et l'exercice de l'esprit.

### Les principales déclinaisons de l'ingénierie des connaissances

L'ingénierie des connaissances comprend deux modalités essentielles :

- l'ingénierie des représentations formelles de connaissances,
- l'ingénierie des inscriptions documentaires de connaissances.

Les inscriptions formelles correspondent aux représentations logiques formalisées modélisant le sens et la signification, que ce soit la signification de contenus, documents, pensées, expressions linguistiques, etc. Le principe est que la syntaxe formelle à la base de ces inscriptions contrôle et détermine la signification associée. Cette syntaxe possède donc une sémantique formelle.

Les inscriptions documentaires correspondent aux contenus mis en forme, produits, consultés et transmis à l'aide des documents. Les documents numériques reposent sur une numérisation du contenu où les unités numériques manipulées techniquement sont dans un rapport arbitraire avec les unités de sens interprétées par un lecteur :

- une image est faite de pixels manipulés techniquement,
- une image est vue et comprise en termes de formes visuelles ou graphiques,
- les formes visuelles interprétées n'ont pas d'existence technique dans le document : on ne sait pas manipuler tous les pixels composant un présentateur de télévision, et rien que ceux-là.

L'intérêt des inscriptions formelles est de pouvoir déléguer à la machine des traitements reposant sur la sémantique des contenus manipulés ; l'inconvénient est de contraindre cette sémantique et de la réduire aux capacités expressives et calculatoires des formalismes. L'intérêt des inscriptions documentaires est de pouvoir exprimer n'importe quel type de contenu et de signification associée, car la forme technique reflète l'apparence et non le sens du document. L'inconvénient est de ne pouvoir introduire de l'« intelligence » dans les traitements effectués<sup>2</sup>.

<sup>2</sup> Source : <http://www.utc.fr/~bachimon/Recherche.html>

2. L'ingénierie des connaissances propose des concepts, méthodes et techniques permettant de modéliser, de formaliser, d'acquérir des connaissances dans les organisations dans un but d'opérationnalisation, de structuration ou de gestion au sens large.

L'essor et l'utilisation croissante des Sciences et Technologies de l'Information et de la Communication (STIC) dans des environnements professionnels divers modifient parfois profondément, les conditions de la représentation et de l'échange des informations et des connaissances entre acteurs au sein d'organisations. L'ingénierie des connaissances appréhende les changements induits par l'utilisation des STIC dans la mesure où elle «réfléchit» sur l'instrumentation technique des contenus pour leur exploitation dans un cadre où ils sont mobilisés pour leur signification. La discipline contribue aux nouvelles « technologies de la connaissance » en développant une ingénierie permettant de diversifier et d'exploiter les modes d'inscription de la connaissance, les modalités d'organisation et de diffusion des savoirs et de démultiplier les interactions entre les utilisateurs.

Du fait de son intérêt pour la connaissance en tant qu'objet à construire, à exprimer, à transmettre, à acquérir ou à exploiter, l'ingénierie des connaissances s'associe à de nombreuses disciplines :

- d'une part, dans sa démarche d'ingénierie, l'IC mobilise les concepts et techniques de la représentation des connaissances, les méthodes d'analyse et de conception à objets, le raisonnement à base de cas, l'ingénierie documentaire ou l'ingénierie éducative, la conception de systèmes d'information...
- d'autre part, dans sa démarche de modélisation des connaissances, l'IC doit se rapprocher de disciplines permettant de caractériser et décrire les connaissances d'un domaine et d'évaluer leur mise en œuvre dans les SBC. La sociologie, la gestion ou l'ergonomie peuvent ainsi concourir à une démarche d'explicitation de ce que sont les connaissances dans un contexte humain et organisationnel<sup>3</sup>.

### ***Qu'est-ce qu'une ontologie ?***

Entretien avec Bruno Bachimont, directeur scientifique à l'INA et enseignant-chercheur contractuel (HdR) à l'Université de Compiègne.

Propos recueillis par Marie-Noëlle Rohart pour le laboratoire CRIS - Université Paris X, Printemps 2004.

La définition des ontologies est héritée d'une tradition philosophique qui s'intéresse à la science de l'Être. Aujourd'hui, elle signifie la « science des étants » c'est-à-dire l'ensemble des objets reconnus comme existants dans un domaine.

L'ontologie est utilisée, depuis plusieurs années, dans l'Ingénierie des Connaissances (IC) et l'Intelligence Artificielle (IA) pour structurer les concepts d'un domaine. Les concepts sont rassemblés et ces derniers sont considérés comme des briques élémentaires permettant d'exprimer les connaissances du domaine qu'il recouvre.

Les ontologies sont utiles pour partager des connaissances, créer un consensus, construire des systèmes à base de connaissances. De nombreux projets d'ontologies sont en œuvre comme celle du Web sémantique. Le problème fondamental est de respecter la diversité des langages et des représentations du monde, tout en permettant les échanges d'informations.

---

<sup>3</sup> Source : Actes des 16èmes journées francophones d'Ingénierie des Connaissances, 2005.  
<http://www-sop.inria.fr/acacia/afia2005/actes/ActesIC2005.pdf>

### **Comment définiriez-vous une ontologie ?**

On peut caractériser une ontologie comme une structuration des concepts d'un domaine. Ces concepts sont rassemblés pour fournir les briques élémentaires et exprimer les connaissances dont on dispose dans ce domaine.

### **Pouvez-vous faire la distinction entre une ontologie, un thésaurus et un index ?**

Un thésaurus est l'ensemble des mots empruntés à la langue d'un domaine. On les structure par des liens d'hyponymie, d'hyperonymie et par des liens d'association, qui sont en général assez vagues puisque ce sont des liens du type « voir aussi » ou « voir également ».

Une ontologie va reprendre la structuration hyperonymique et hyponymique du thésaurus mais va considérablement la contraindre. La signification de la relation d'hyponymie, le long de l'arborescence de l'ontologie, devra être toujours la même, sans changement du point de vue de l'interprétation de cette relation. Par ailleurs, on définira beaucoup plus précisément les relations d'association qu'il y a entre les différents concepts de l'ontologie. Dernière distinction, c'est qu'une ontologie est une relation de concepts (ou notions) alors que dans un thésaurus ce sont davantage des mots plutôt que des concepts. Cela signifie par exemple qu'on peut avoir des concepts qui ne possèdent pas d'appellation habituelle dans la langue usuelle, alors on est obligé d'avoir des périphrases ou des locutions un peu élaborées pour pouvoir désigner un concept dans le cadre d'une ontologie.

On peut prendre un index au sens bibliographique, au sens de l'énumération de terme ou, plus généralement, au sens d'une liste alphabétique, avec simplement le lieu des occurrences dans un contenu donné. On peut considérer que l'ontologie peut servir de répertoire ou de ressources pour constituer un index. Cela donne les concepts que l'on va reprendre pour constituer l'index et voir quels sont les contextes d'apparition de ces concepts dans un contenu donné.

### **Quelle différence faites-vous entre une ontologie de l'ingénierie des connaissances et une ontologie philosophique ?**

On dit souvent que cela n'a pas de rapport. Il faut bien distinguer une ontologie avec un petit « o » et l'ontologie avec un grand « O », qui est celle étudiée par la philosophie. L'ontologie philosophique signifie la « Science de l'être en tant qu'être » tandis que les ontologies, avec un petit « o » et la marque du pluriel, sont les ontologies mobilisées dans les technologies de l'information et de la communication. Malgré tout on peut établir un lien dans la mesure où une ontologie, dans le contexte des ingénieries des connaissances, est l'énumération des concepts renvoyant à ce que l'on considère exister dans un domaine.

Vous avez à ce niveau deux manières de concevoir les ontologies avec « o » minuscule et pluriel : soit l'ontologie est un répertoire des objets du domaine - on peut dire que c'est une vision réaliste, qui est souvent celle adoptée dans les travaux anglo-saxons sur les ontologies, soit, autre interprétation, l'ontologie ne désigne pas les objets du domaine mais renvoie à notre manière de penser le domaine. À ce niveau-là, il n'y a donc pas de prétention ontologique au sens philosophique du terme mais une prétention à fournir le terme nécessaire à la pensée dans le domaine.

La première approche est de déduire quels sont les objets qui existent dans ce domaine et donc à lister toutes les marques qui vous servent à désigner les objets d'un domaine. L'autre approche prend en compte la manière dont travaillent les spécialistes du domaine et fait le répertoire des notions dont ils se servent pour travailler dans ce domaine. Mais on ne s'occupe pas de savoir si ces notions renvoient à des objets ou pas et quel est le niveau de véracité de ces notions. C'est le travail des spécialistes de savoir si ce qu'ils disent est vrai ou pas. Dans l'ingénierie des connaissances, l'ontologie a simplement pour mission de fournir un répertoire ou un référentiel de concepts qui servira à soutenir et à soutenir les expressions de connaissance du domaine.

La première approche renvoie à une vision métaphysique de l'ontologie : on voit ce qui existe dans le domaine en s'appuyant sur la métaphysique au sens strict du terme. Nicolas Guarino notamment fait référence à des ouvrages de métaphysique pour asseoir les ontologies qu'il élabore.

### **Qui conçoit les ontologies ? Pour quel public ?**

C'est en général quelqu'un qui est formé aux techniques de l'ingénierie des connaissances et qui est versé dans les techniques relevant du web sémantique. Cet ontologue peut également être un terminologue, puisque le lien entre terminologie et ontologie est très fort. La terminologie recense les termes exprimant des concepts et désignant des objets.

À qui ça s'adresse ? L'ontologie s'adresse à tous ceux qui ont besoin de systématiser l'expression des connaissances qu'ils manipulent. Cela concerne donc à la fois les personnes qui vont faire de la capitalisation des connaissances, des référentiels métiers ou des systèmes d'information.

L'étendue des applications est relativement vaste. En revanche, il est encore relativement difficile de savoir ce que l'on peut faire de plus avec une ontologie, par rapport à ce que l'on est déjà capable de faire avec un thésaurus. Les applications ne savent pas encore bien exploiter l'apport qu'apportent les ontologies, du point de vue de la rigueur de leur structuration. Le thésaurus a évidemment les défauts et les qualités inverses : il est très souple, facile à utiliser ; en revanche l'utilisation est vite bloquée par le manque de rigueur dans la structuration des termes et dans leur organisation.

Le thésaurus est davantage adapté à l'utilisation par un humain, comme typiquement un documentaliste ou un ingénieur, tandis que l'ontologie sera davantage utilisée dans des systèmes dans lesquels on sera capable de spécifier une exploitation automatique.

### **Comment évolue une ontologie ?**

Une ontologie évolue mal parce que c'est une organisation relativement rigoureuse des significations entre différents concepts. Lorsqu'un concept évolue, il modifie tout le système et avec lui l'ontologie. L'ontologie doit être réélaborée à chaque fois qu'il y a une évolution. Cela reste un élément relativement bloquant. De nombreux travaux sont engagés pour faire évoluer les ontologies. Mais à chaque fois elles reposent sur des hypothèses formalisatrices assez fortes, qui ne rendent pas forcément compte de la manière dont les choses évoluent pratiquement dans la réalité.

Ma position personnelle est de penser que l'ontologie doit plutôt être un répertoire ou un référentiel pour pouvoir associer et annoter les usages. Les annotations évoluent mais non la colonne vertébrale de l'organisation. L'ontologie, à terme, aura cette fonction-là : une structuration au sein des connaissances, mais non la connaissance elle-même. Cela permet aux connaissances d'évoluer sans forcément remettre en cause l'organisation de ces connaissances. Il s'agit donc d'aller plutôt vers une interprétation nominaliste : les concepts sont des appellations pour des manières de penser, sans prétention à désigner des objets réels ni à dire le vrai ; c'est du ressort du spécialiste que de s'emparer du système de concepts pour exprimer des connaissances dont il prend en charge la signification et la vérité. Avec cette interprétation, ce que dit l'ontologie n'est pas forcément vrai. Car son problème est d'être capable d'organiser les connaissances. Ce sont ceux qui vont faire évoluer et qui vont utiliser les connaissances, qui se chargeront de leur vérité et de leur organisation.

### **Comment construit-on les ontologies ?**

On va partir de corpus de textes qui expriment les connaissances. Lorsqu'on appréhende une connaissance, on a à sa disposition une verbalisation, une explicitation linguistique. À ce niveau là, le meilleur marchepied pour la modélisation des ontologies, c'est de recourir à des textes issus de la pratique du domaine. Cela peut être des ouvrages didactiques, des notes techniques, des interviews, des dictionnaires ou tout ce qui permet d'avoir des explicitations d'usage des concepts du domaine.

A partir de là, on va en général mettre en œuvre une démarche du type « extraction terminologique ». Des outils de linguistique de corpus vont analyser, lemmatiser et extraire des candidats termes soit à partir de syntagmes nominaux, soit à partir de syntagmes verbaux. Ces candidats termes seront proposés à un expert qui pourra les sélectionner ou les ignorer pour constituer une ontologie.

Il y a une autre étape qui consiste, une fois qu'on a sélectionné les termes, à les structurer et à leurs associer un sens de référence. Parce que sélectionner un terme, c'est une chose, savoir ce qu'il signifie en est une autre. A un moment donné, il faut être capable d'explicitier le sens sous lequel on va retenir un terme plutôt qu'un autre. Pour cela, il y a une méthodologie qui s'appuie sur la conception linguistique des ontologies et qui permet de spécifier ces termes.

Après cette « normalisation sémantique », une dernière étape consiste à rendre praticables les ontologies dans une exploitation informatique. En général on va utiliser des langages de représentation des connaissances ; aujourd'hui ce sont plutôt des langages de type « OWL : Ontology Web Language », qui sont issus des travaux du web sémantique. Ils fournissent un langage de représentation de la logique qui est restreinte pour pouvoir exprimer les connaissances ontologiques.

### **Quelles difficultés rencontrez-vous dans l'élaboration d'une ontologie ?**

Tout d'abord, il est toujours difficile d'avoir accès aux connaissances, à la constitution du corpus.

La deuxième difficulté est l'organisation conceptuelle d'une ontologie. Les principes d'organisation sont souvent abstraits et font appel à des notions philosophiques. Il faut donc avoir une théorie soit sur la connaissance, soit sur le monde. On se retrouve souvent à faire ce qu'on appelle l'upper ontology, c'est à dire les concepts les plus abstraits, ceux qui vont organiser le reste de l'ontologie. On se retrouve à faire de la philosophie de la connaissance. Sans formation, ce n'est pas évident. Récupérer une théorie faite par quelqu'un d'autre, la comprendre et la manipuler, ce n'est pas chose aisée non plus.

La troisième difficulté est de pouvoir modéliser des concepts, non pas fondateurs de l'ontologie et qui sont de niveau philosophique, mais des concepts du domaine. Là, il faut avoir l'intelligence du domaine. Il va falloir travailler avec des spécialistes qui soient capables de confirmer, infirmer, suggérer les concepts qui sont importants pour structurer la connaissance de leur domaine.

Une fois qu'on a les concepts qui organisent le domaine, les concepts philosophiques qui organisent l'ontologie, il suffit de remplir celle-ci. En médecine, par exemple, on va énumérer tous les os humains une fois qu'on a défini la notion de squelette et la notion d'anatomie. L'anatomie et le squelette sont les concepts structurants du domaine. Par contre on n'a pas besoin d'être un grand spécialiste pour lister les médicaments ou les os.

### **N'y a-t-il pas un danger d'uniformisation culturelle ?**

Je ne le crois pas. Une ontologie ne marchera que là où il y a suffisamment de consensus culturel autour. C'est plutôt à travers l'extension des communautés de pratiques qu'une ontologie pourra avoir un rôle de plus en plus vaste. L'ontologie est à prendre en compte par rapport à une thématique donnée et une tâche donnée. Elle est donc valable uniquement pour une communauté de pratiques. Si cette communauté de pratiques s'étend, parce que tout simplement il y a de plus en plus de gens qui travaillent ensemble, vous aurez une ontologie qui va être la conséquence de cette uniformisation mais elle n'en sera pas directement la cause.

### **Classer le patrimoine culturel ne pose-t-il pas de difficultés ?**

Le patrimoine culturel est déjà classé. Il y a des catalogues, une manière d'inscrire les œuvres au patrimoine. C'est toujours pareil : c'est un débat qui excède les ontologies. L'uniformisation vous

permet de faire des choses de manière plus partagée et plus facilement circulaire mais, à ce moment-là, vous atténuez la diversité. Cela dépend donc de ce que vous voulez faire.

En général, la bonne nouvelle c'est que toute uniformisation est également un moteur à la diversification, c'est à dire que toute convergence crée de la divergence, toute uniformisation crée de la diversification. Seulement ce n'est pas forcément de la manière dont vous le pensez. L'uniformisation linguistique crée de nouveaux dialectes, soit de l'argot, soit du langage de banlieue, etc., qui sont très différents des patois géographiques qu'on avait autrefois. Il y a toujours de la diversité linguistique mais elle n'est plus située au même endroit.

### **Pouvez-vous donner un exemple d'application concrète ?**

Un exemple que je connais bien, c'est dans le contexte médical : pouvoir associer, catégoriser plus ou moins automatiquement, des dossiers « patient » par rapport à la classification internationale des maladies. C'est donc utiliser une ontologie qui sert de pivot pour pouvoir associer un code de cette classification internationale des maladies à une catégorie d'actes médicaux qui ont été faits sur un patient.<sup>4</sup>

### **Bruno Bachimont : Pourquoi n'y a-t-il pas d'expérience en ingénierie des connaissances ?**

Les modèles de l'ingénierie des connaissances permettent d'explicitier des points de vue interprétatifs proposés à un utilisateur pour mobiliser des sources documentaires ou contenus. Les modèles ne modélisent pas le monde, mais proposent une instrumentation pour exploiter des sources que l'utilisateur mettra lui-même en rapport à sa situation d'usage. Contextuelle, imprédictible, l'interprétation des sources n'est pas modélisable. Exploitant des sources instrumentées, l'utilisateur inscrit son propre usage en annotant le contenu des différences qu'il fait avec le point de vue proposé. Le système est une mémoire d'usages, qui sédimente les expériences passées pour penser les usages à venir.

L'ingénierie des connaissances n'aura donc pas d'expérience scientifique, mais des expériences humaines d'utilisation dont il faut garder la mémoire. En s'intéressant à la connaissance, l'ingénierie des connaissances doit prendre acte qu'elle est critique, contextuelle et variable. Le formel n'est pas une réponse, mais une instrumentation.<sup>5</sup>

### Bibliographie :

▪ BACHIMONT Bruno, « Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en Ingénierie des connaissances ». In Charlet J., Zacklad M., Kassel G. & Bourigault D., *Ingénierie des connaissances, évolutions récentes et nouveaux défis*, 2000, Paris : Eyrolles, p.305-323.

<sup>4</sup> Source : [http://www.technolangue.net/article.php?id\\_article=280](http://www.technolangue.net/article.php?id_article=280)

<sup>5</sup> Source : [http://www.utc.fr/~bachimon/Publications\\_attachments/BachimontIC2004.pdf](http://www.utc.fr/~bachimon/Publications_attachments/BachimontIC2004.pdf)

▪ GIBOIN Alain, GANDON Fabien, GRONNIER Nicolas, GUIGARD Cécile, et CORBY Olivier, « Comment ne pas perdre de vue les usage(r)s dans la construction d'une application à base d'ontologies ? Retour d'expérience sur le projet KmP ». In *Les Actes de la Conférence IC 2005 – 16<sup>èmes</sup> journées francophones d'Ingénierie des Connaissances*, p.133-144.

Lien : <http://www-sop.inria.fr/acacia/afia2005/actes/ActesIC2005.pdf>

▪ DIENG-KUNTZ Rose, CORBY Olivier, GANDON Fabien, GIBOIN Alain, GOLEBIEWSKA Joanna, MATTA Nada, RIBIERE Myriam, *Knowledge management, Méthodes et outils pour la gestion des connaissances*, Octobre 2005 (3<sup>ème</sup> édition), Dunod, 464 pages.

▪ TISSEYRE René-Charles *Knowledge management, théorie et pratique de la gestion des connaissances*, Octobre 1999, Hermès, 190 pages.

### Webographie :

- [www.utc.fr/~bachimon](http://www.utc.fr/~bachimon) : site de Bruno Bachimont.
- [www.irit.fr/GRACQ](http://www.irit.fr/GRACQ) : site du Gracq, Communauté française d'acquisition et d'ingénierie des connaissances.
- [www-test.biomath.jussieu.fr/~jc/Cederilic/ic\\_tome1/ic\\_tome1.html](http://www-test.biomath.jussieu.fr/~jc/Cederilic/ic_tome1/ic_tome1.html) : fac-similés des chapitres du livre *Ingénierie des connaissances, évolutions récentes et nouveaux défis*, CHARLET J., ZACKLAD M., KASSEL G. & BOURIGAULT D. (éd.), Eyrolles, Paris, 2000.