

## Une expérience de co-construction de « carte de thèmes » dans le domaine des logiciels libres

Jean-Pierre Cahier<sup>1</sup>, L'Hédi Zaher<sup>1</sup>, Jean-Philippe Leboeuf<sup>1, 2</sup>, Xavier Pétard<sup>1, 2</sup>,  
Claude Guittard<sup>3</sup>

<sup>1</sup>*Tech-CICO, Institut des Sciences et Technologies de l'Information de Troyes,*

<sup>2</sup>*LIMSI, Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur,*

{[jean-pierre.cahier](mailto:jean-pierre.cahier@utt.fr), [mohamed\\_hedi.zaher](mailto:mohamed_hedi.zaher@utt.fr), [jean-philippe.leboeuf](mailto:jean-philippe.leboeuf@utt.fr), [xavier.petard](mailto:xavier.petard@utt.fr)}@utt.fr

<sup>3</sup>*BETA ULP Strasbourg*

[guittard@cournot.u-strasbg.fr](mailto:guittard@cournot.u-strasbg.fr)

### Résumé

Dans cet article, nous présentons un dispositif de « cartes de thèmes » dans le domaine du logiciel libre. Dans une approche volontairement ouverte, basée sur l'approche de « Place de marché à base de connaissances » et sur le modèle HyperTopic proposés par le laboratoire Tech-CICO, cette expérimentation collective vise la construction d'un espace d'information partagée et repérable pour et par les concepteurs et utilisateurs de logiciels libres. Grâce à ce dispositif, les membres de la communauté du logiciel libre peuvent décrire et rechercher des logiciels, par navigation dans une « carte de thèmes » multi-points de vue. Chacun peut déclarer les caractéristiques d'un logiciel selon une structure d'index comportant plusieurs arborescences, permettant ainsi, idéalement, une construction sémantique collective « au fil de l'eau » par les acteurs.

Nous présentons ce projet comme un exemple de notre approche du « Web socio sémantique ». Nous détaillons le modèle HyperTopic, qui supporte cette approche, et l'outil collecticiel AGORÆ qui permet sa mise en oeuvre informatique. Inspiré des réseaux sémantiques simples de type « Topic Map », HyperTopic les complète en visant à la fois la représentation des objets métiers (ici les logiciels) et celle des acteurs sociaux et de leur activité, et en offrant un soutien à la manipulation de la sémantique des notions relevant de ces deux familles d'objets.

**Mots-clés :** Web socio sémantique, gestion des connaissances, applications, logiciel libre, co-construction au fil de l'eau.

### Introduction

De nombreuses opinions convergent autour de l'idée que la construction d'une « Société de la Connaissance » doit s'appuyer sur un perfectionnement des standards du Web, et beaucoup de recherches se focalisent dans le sens d'un « Web sémantique » appelé à fournir de meilleurs services grâce à une nouvelle vague de standards.

Mais quel type de buts du « Web sémantique » - et au fond, quelle Société de la Connaissance - devons nous contribuer à favoriser? Nous faisons ici l'hypothèse que la sémantique du Web, celle qui sera la plus utile et la plus efficace, sera celle qui sera co-créée au sein de groupes tels que les communautés. Le « Web sémantique » signifie donc pour nous qu'il faut continuer à inventer de nouvelles façons de participer à l'univers informationnel et de se repérer consciemment dans une information foisonnante, en inventant aussi les nouveaux modèles de coopération nécessaires pour créer cette conscience collective.

Pour cela, la priorité de la recherche-développement ne doit pas être de concentrer tous les efforts sur des systèmes automatisés, reléguant en arrière-plan l'apport cognitif humain et la maîtrise sociale de « la sémantique ». Il faut utiliser l'ordinateur et le Web avant tout pour des systèmes d'aide à l'individu et aux groupes, augmentant leur capacité d'appréhender cognitivement et socialement l'information au sein de leurs activités, pour créer collectivement du sens et partager les sémantiques complexes qu'exigent le monde actuel.

Aussi la voie que nous proposons pour le « Web sémantique » est-elle celle du « Web socio sémantique ». Dans le présent article, nous souhaitons illustrer concrètement cette approche. Pour cela nous présentons une application à un projet de co-construction, permanente et à distance, d'une « carte de thèmes », s'adressant à une communauté ouverte, nombreuse et déjà existante sur le Web.

L'application, que le lecteur pourra facilement consulter sur le Web<sup>1</sup> afin de se faire directement un avis, vise à la fois à donner un exemple de la perspective du Web socio sémantique, et à expérimenter sur une plus longue période la co-construction concrète d'une carte de thèmes, en direction d'une communauté donnée, avec un nombre élargi d'acteurs. Son terrain est celui des logiciels libres (« Free Software »). La « carte sémantique » à construire visant la base « d'expertise » commune à cette communauté, l'élaboration de cette carte de thèmes suppose une réelle compréhension et une connaissance approfondie du domaine, en termes de savoirs et de savoir-faire de métier.

### ***Plan de l'article***

Nous présentons dans une 1<sup>ère</sup> partie le « Web socio sémantique » dans le cadre d'une problématique générale de gestion de connaissances.

Dans la 2<sup>ème</sup> partie nous détaillons les outils conceptuels, méthodologiques et pratiques constitutifs de notre approche et de l'application prise en exemple. Il s'agit principalement du modèle HyperTopic, du modèle de place de marché à base de connaissances (KBM) qui s'appuie sur le modèle HyperTopic, et de la plate-forme logicielle AGORÆ.

Le dispositif de carte de thèmes dans le domaine du logiciel libre est ensuite décrit dans la 3<sup>ème</sup> partie.

Nous mentionnons enfin dans la 4<sup>ème</sup> partie en conclusion quelques éléments de bilan et de perspectives, ainsi que quelques questions ouvertes à ce stade du projet.

## **1. Enjeux et problématique : Le « Web socio sémantique »**

### **1.1. Un enjeu pour la gestion collective de connaissances**

Dans des organisations et des communautés de toutes sortes, les progrès des NTIC doivent aider à l'émergence de nouveaux modèles de coopération, appuyés sur des outils de type « collectif » (groupware). S'ils parviennent à mieux prendre en compte la dimension sémantique collective, ces modèles seront riches de perspectives comme générateurs de nouveaux espaces, lieux du partage de connaissances vivantes et sources d'innovation.

---

<sup>1</sup> lien <http://www.yeposs.org/>, Yellow Pages for Open Source Software

Pour se repérer dans des « territoires » d'activités complexes dont les lexiques et la sémantique sont volumineux, particuliers et changeants, les acteurs ont besoin de *cartes*. Il faut les aider à la fois à utiliser, structurer et recréer en permanence la « carte de thèmes » décrivant leur domaine, leurs activités et leur positionnement d'acteur. Et comme il existe souvent plusieurs points de vue, la carte de thèmes doit prendre en compte cette pluralité.

La « carte de thèmes multi-points de vue » est donc un artefact qui joue un rôle clé sur la voie d'une maîtrise accrue des sémantiques partagées, avec l'aide des NTIC. Cette idée sert notamment de pivot à l'approche « place de marché à base de connaissances » (ou KBM – « Knowledge-Based Marketplace ») étudiée par le Laboratoire Tech-CICO : une KBM peut être définie [CAHIER 02] comme un espace d'intercompréhension, langagière et cognitive, basé sur une carte de thèmes multi-points de vue, reliant les acteurs à des ressources du Web ; son modèle de coopération comporte une variété assez simple de rôles d'acteurs (consulter, contribuer, éditer la carte de thèmes...), équivalents aux rôles clients et fournisseurs d'une « place de marché » virtuelle.

Mais il existe bien d'autres cas d'utilisation de cartes de thèmes, et plus généralement il apparaît que le principe de la « carte multi-points de vue » est potentiellement utile à une variété de nouveaux modèles de coopération, avec des rôles adaptés à chaque situation, dans des univers informationnels, décisionnels, éducatifs, d'écriture ou d'annotation collective, etc. C'est pourquoi nous avons élaboré un cadre conceptuel générique, le cadre du « Web socio sémantique » facilitant la mise en œuvre de cartes de thèmes multi-points de vue, en s'appuyant sur les standards du Web, au service des formes de coopération de demain. Ce cadre conceptuel peut aider à multiplier les expériences pour imaginer, évaluer et perfectionner ces nouvelles formes, afin que le « Web sémantique » qui se profile facilite la création de sens partagé et évolutif, et non l'inverse.

### **1.2. La « Toile socio sémantique » pré-existe au Web**

Indépendamment du Web et de toutes les formes informatisées de systèmes d'information et de coopération existantes, nous pouvons déjà noter qu'au sein des innombrables activités humaines collectives, il en existe certaines, fort nombreuses, où l'inscription d'aspects sémantiques dans des artefacts partagés prend de l'importance. Dans certains de ces cas, des groupes ont besoin de rendre explicites certains éléments de « structure sémantique » partagée, prenant la forme de thésaurus, de cartes, de diagrammes, de guides, de dictionnaires, d'annuaires de ressources ou de compétences, de catalogues, etc.

Les recherches récentes ont permis de mettre en évidence et d'approfondir théoriquement l'existence de plusieurs catégories de communautés – que l'on peut caractériser comme communautés de pratique, communautés épistémiques, communautés « intensives en connaissances », ou encore communautés d'actions. Sans approfondir ici ces caractérisations, il apparaît qu'une communauté peut recourir à un artefact sémantique non seulement pour aider ses membres à se repérer dans leurs tâches et dans l'information, mais aussi, dans bien des cas, d'une façon que l'on pourrait qualifier de « stratégique », pour se construire en tant que communauté à travers l'activité des membres.

Pour créer leurs artefacts sémantiques, informatisés ou non, et ensuite pour assurer leur « maintenance » une fois institués, les groupes tels que les collectifs de métier inventent en permanence : les catégories, les connaissances et le sens des mots changent, des néologismes apparaissent dans les lexiques de métier, et ces groupes déploient une variété de solutions pour construire et réactualiser en permanence les constituants de ces artefacts,

suivant diverses formes de division du travail (dont le métier de documentaliste). Les structures d'index multi-points de vue sur lesquelles s'appuie le système que nous proposons n'échappent pas à cette règle : comment naissent les « thèmes » et les relations d'une structure sémantique, comment évoluent-ils dans l'activité des acteurs, dans leurs interactions ou dans les activités focalisant plus spécifiquement cette structure sémantique ?

Si la gestion de cartes de thèmes et de structure d'index pose une problématique particulière de gestion collective de connaissances quand elle est « assistée pas ordinateur », et donc doit être régulée explicitement, la problématique d'ensemble de la régulation explicite est loin d'être une nouveauté. Toute activité collective primaire comporte indissociablement une part d'activité collective secondaire « socio sémantique », plus ou moins implicite, au plan du langage et de la sémantique. Clairement le « socio sémantique » possède sa sphère propre, souvent implicite et fortement intégrée dans une « toile socio sémantique » humaine, interactionnelle et langagière, antérieure à « la toile » du Web informatisé et à ses standards technologiques.

### **1.3. La sémantique se construit dans l'activité**

Dans ce Web socio sémantique, qu'ils tissent souvent sans en avoir conscience, les acteurs sont en situation d'utiliser une sémantique existante, mais aussi, en même temps qu'ils interagissent dans leurs activités, ils mettent en œuvre certaines interactions plus particulières à travers lesquelles ils modifient, complètent, discutent de nouvelles catégories, bref « co-construisent » la sémantique partagée dont ils ont besoin. En suivant les approches constructivistes, il existe une boucle entre langage et activité [VYGOTSKY 97] de même que, non sans liens, il existe une boucle entre langage et connaissance [ALQUIER 93] [PRINCE 96].

Pour simplifier, on peut retenir que la « sémantique » se construit dans l'activité, et que les acteurs de ce type de communauté sont donc à la fois utilisateurs et constructeurs de cette sémantique.

Ces rapports des groupes à leur sémantique peuvent être immédiats ou au contraire utiliser des médiations plus ou moins développées, jusqu'à prendre la forme de dispositifs, très variés, dont certains peuvent être considérés comme des *artefacts organisationnels* [TEULIER 03]. Du double point de vue des sciences de gestion et de l'ingénierie des connaissances, de tels artefacts modélisent des interactions entre des normes et des connaissances mises en pratique dans l'organisation, et en retour ils nécessitent un apprentissage pour être intégrés dans le système d'action des acteurs. L'accent mis par certains courants des Sciences de Gestion, notamment [HATCHUEL 00] pour la théorie axiomatique de l'action collective, fondée de façon non séparable sur les opérateurs « savoir » et « relation », montre que la problématique de la gestion collective de connaissances pour les artefacts sémantiques devrait pouvoir se rapprocher des approches en termes d'outil de gestion (notamment pour l'évaluation). Pour [DUPOUET 02] « à l'intérieur d'une communauté intensive en connaissances, le comportement des membres se caractérise comme dans toute communauté, par le respect des normes sociales qui sont définies, mais aussi dans ce contexte particulier par l'engagement volontaire dans la construction, l'échange, et le partage d'un répertoire de ressources cognitives communes ». Par ailleurs, d'autres approches, bien que très différentes les unes des autres en gestion [KAPLAN 01] ou en sociologie [BOLTANSKI 91], [CHATEAURAYNAUD 03] peuvent utilement nourrir la réflexion sur les usages possibles de sémantiques basées sur de multiples points de vue en lien avec les enjeux importants pour les groupes humains.

En particulier, dans un système de « carte de thèmes » tel que celui que nous considérerons dans la suite, un objectif de gestion sera de dégager des points de vue et d'y nommer et catégoriser les thèmes, caractérisant des aspects importants, non dans une optique de connaissance universelle (par exemple encyclopédique), mais *pour des activités* déterminantes pour les acteurs à un moment donné. Les approches de gestion basées sur l'activité, l'analyse fonctionnelle ou systémique des objets considérés, l'analyse des facteurs de décisions, etc., pourront servir de repère pour dégager les « points de vue » qui vont structurer de façon importante la carte de thèmes. Dans notre exemple, s'agissant d'une classification multi-points de vue de logiciels libres, les thèmes ne seront pas seulement des facettes ou attributs simples des logiciels, mais souvent des propriétés « heuristiques » importantes d'un certain point de vue dans le jugement des experts, comme le fait que le logiciel obéisse à un standard, un langage ou un label (enjeu d'interopérabilité dans la sphère industrielle), qu'il relève de différentes formes de gratuité ou suive un certain business model (enjeu économique), soit placés sous tel régime de propriété industrielle (enjeu juridique), etc. Les thèmes sont des expressions linguistiques qui peuvent d'ailleurs nécessiter jusqu'à une dizaine de mots pour exprimer leur sujet. Ces « attributs heuristiques », qui condensent une réelle expertise et ne vont pas sans controverses, vont peser suivant de multiples dimensions dans l'évaluation du logiciel, dans les critères de choix entre outils concurrents, dans la faisabilité technique d'un assemblage de composants, etc. Un enjeu est donc l'élaboration, à la fois de la culture partagée par la communauté et du cadre de normes, issus d'explicitation de connaissances souvent implicites au départ dans le collectif.

#### **1.4. La voie « socio sémantique » pour le « Web sémantique »**

Les artefacts sémantiques peuvent bien entendu être basés sur l'informatique, ce qui devient souvent nécessaire lorsque croissent la taille de la communauté ou la complexité de son organisation et de son domaine d'activité. Nombreuses sont désormais les situations, comme la conception ou le choix de logiciel, où les connaissances sont particulièrement mouvantes, obsolètes, peu stabilisées. C'est aussi le cas pour les classifications de produits dans des places de marchés associant des communautés de fournisseurs concurrents, aussi bien que dans le système de Pages Jaunes de compétences dans certaines grandes entreprises [CAHIER 04a]. L'utilisation des technologies du Web pour répondre aux besoins de ce type de communautés doit selon nous être approchée en considérant un objectif socio-technique d'ensemble, d'assistance à cette activité de co-construction « naturelle » des structures sémantiques dans les groupes concernés. Le « Web socio sémantique » ne remplace pas, mais au contraire accompagne toujours une « Toile socio sémantique » humaine extrêmement vivante, laquelle rappelons-le n'est pas forcément ni explicite, ni encadrée, ni unifiée par des normes contraignantes.

Nous définissons le Web socio sémantique (W2S) comme un Web social qui participe de la construction d'une représentation structurée du domaine et du collectif. S'adressant à des communautés d'utilisateurs poursuivant des objectifs similaires, il contribue à une structuration progressive des réseaux cognitifs et sociaux, ce que nous souhaitons démontrer par l'exemple dans le présent article (pour une définition théorique plus complète du W2S, on pourra se reporter à [ZACKLAD 03] [CAHIER 04a]). Notons bien que cette approche n'est pas antinomique, mais au contraire tout à fait complémentaire (cf. Figure 1) avec le courant que nous avons caractérisé [CAUSSANEL 02] comme « Web computationnellement sémantique », privilégiant les inférences automatisées et plus proche de la définition du Web Sémantique apportée par Tim Berners-Lee [BERNERS-LEE 01]. Des

standards nous semblent aussi nécessaires pour le Web socio sémantique. Ils peuvent prendre appui sur les niveaux les plus « syntaxiques » proposés par le W3C (ce point sera développé au § 2.4.)

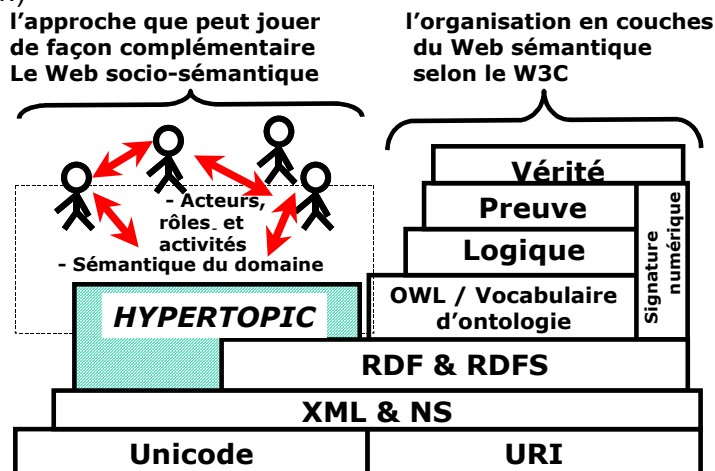


Fig.1- Le Web socio sémantique s'appuie sur les couches basses des standards « Semantic Web » du W3C

### 1.5. Une triple sémantique : objets de domaine, acteurs et activités

Notre approche du Web socio sémantique suppose le principe de « modélisation engagée ». Ce principe affirme la possibilité de visibilité, pour les modélisateurs, de la représentation de leurs actions en tant qu'acteurs du monde et en tant que modélisateurs. Concrètement, cela signifie la possibilité et la maîtrise de la construction de la structure sémantique par les acteurs eux-mêmes. La représentation n'est pas dissociée de la pratique directe qui suppose pour les acteurs d'être tout à la fois existants sur le « territoire », représentés sur la « carte », acteurs du territoire et acteurs de la carte. Parmi les nombreuses questions qui se posent : la disponibilité et la maîtrise de la carte, les critères de qualité d'une bonne carte, les moyens organisationnels et techniques nécessaires pour réaliser de cette qualité.

Jusqu'ici l'ingénierie a produit un nombre considérable de systèmes dans lesquels elle est parvenue à fournir un degré souvent élevé de maniement des structures sémantiques très volumineuses liées au domaine : dans les applications récentes, on dispose ainsi d'artefacts très performants (structure de navigation, cartes de thèmes, systèmes de requêtes...) traitant l'accès à des contenus Web selon de volumineux lexiques de domaine.

Dans la perspective de gestion collective de connaissances dans laquelle nous nous inscrivons, d'autres réponses doivent être apportées par l'ingénierie. La modélisation engagée suppose que les acteurs sont considérés en tant que tels, et l'accent est mis sur le groupe d'acteurs, utilisateur et producteur de ces connaissances, susceptible de mettre éventuellement en œuvre une certaine organisation et des rôles pour des activités socio sémantiques. L'ingénierie est plutôt démunie face à toutes les difficultés liées à la complexité de cette dimension des acteurs. Les systèmes de recherche d'information et de gestion de contenus que produit l'ingénierie butent sur des problèmes tels que :

- les droits d'accès ou les niveaux de responsabilité par rapport à l'information sont de plus en plus différenciés en fonction de rôles ou à cause de la segmentation des organisations ;

- les moyens de coopération restent faibles entre des acteurs impliqués, selon des rôles complémentaires ou des procédures de coordination (workflows) dans la co-construction ou la co-maintenance des structures sémantiques ;

- sur le terrain, il existe naturellement une grande *diversité de types d'interactions* et de rôles entre acteurs ; ces derniers ne sont pas une fois pour toutes « experts » ou « utilisateurs » mais sont dans la réalité susceptibles de jouer alternativement une variété de rôles contributeurs, éditeurs ou responsables de la structure sémantique pour certains sous-domaines, etc..

Comme le montre la figure 2, les questions qui se posent – et la nécessité de « modélisation engagée » pour tenter de répondre à leur complexité - concernent à la fois le plan du *domaine*, mais aussi celle des *acteurs, des activités et des rôles* intervenant dans la co-construction.

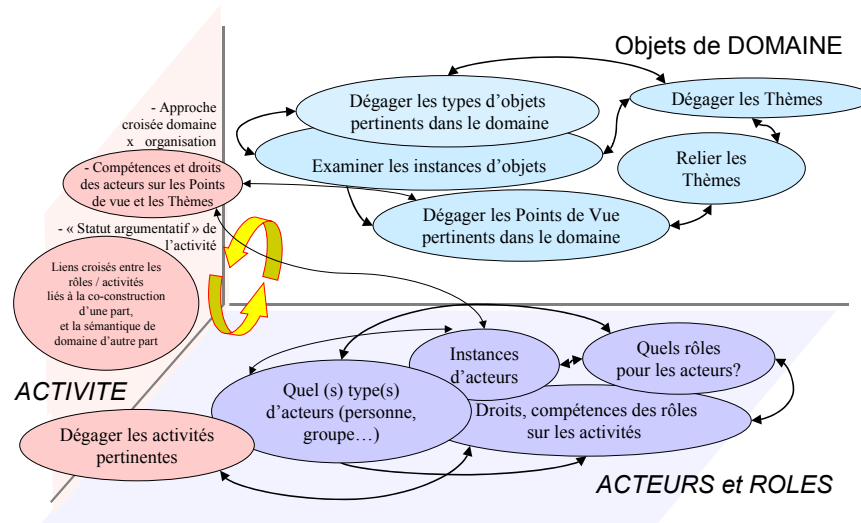


Fig.2 Les trois dimensions d'analyse et de représentation

### 1.6. Quelques défis à l'approche de Web socio sémantique

Les enjeux du Web socio sémantique sont donc importants pour les organisations :

- Dans beaucoup de cas, la construction de structures sémantiques, pour des raisons d'étendue du domaine, d'évolutivité, de volumétrie, etc., ne peut ni reposer sur un seul acteur ni être effectuée a priori. L'introduction de règles de gestion régissant une *co-construction* entre plusieurs acteurs est alors inévitable. Ces règles de gestion peuvent aussi inclure la possibilité pour la co-construction de s'effectuer « au fil de l'eau », en tant que processus permanent relevant de la gestion de connaissances dans l'organisation.

- Les acteurs ont besoin de s'organiser pour tirer parti des différences et des complémentarités, et d'explicitier d'éventuels conflits entre métiers ou entre points de vues. Le modèle doit permettre de décrire un *cadre d'interaction*, pouvant être plus ou moins décentralisé, aidant à prendre en compte la diversité des vocabulaires et des systèmes conceptuels dont les acteurs sont porteurs dans leurs différents contextes.

- Les acteurs doivent aussi pouvoir exprimer des connaissances et participer à la structuration sémantique, tout en se limitant aux sous-domaines pour lesquels ils estiment

disposer effectivement de compétence (phénomène de responsabilité et de « *relevance* » personnelle et sociale).

- Enfin il est nécessaire de prendre en compte le « cadre argumentatif » ce qui plaide aussi pour un niveau de représentation fin des acteurs impliqués, en rapport avec un répertoire d'activités riche : un participant à la construction ou la gestion de la structure sémantique doit pouvoir exprimer un niveau de responsabilité et un spectre étendu d'argumentations et de « nuances » dénotant sa relation d'un acteur donné à la connaissance sur laquelle il est amené à intervenir, allant de la prise de position ferme à l'hypothèse de travail soumise au groupe, en fonction par exemple des étapes d'un processus de validation collectif.

### **1.7. Une sémiotique de l'outil**

L'approche du Web socio sémantique esquissée précédemment appelle une parenthèse épistémologique, du point de vue du positionnement au sein de l'Ingénierie des Connaissances, concernée de façon majeure par ce sujet. Il est clair que nous faisons ici, dans le sillage de certaines réflexions récentes de cette discipline [CHARLET 03] [ZACKLAD 03b] [CAUSSANEL 02] [BACHIMONT 03] le choix d'une approche éloignée du paradigme cognitiviste traditionnel (qui, pour simplifier, viserait à modéliser un système formel de connaissance autorisant des inférences automatiques).

L'ingénierie des connaissances s'intéresse heureusement aussi, de façon croissante, aux approches sémiotiques basée sur l'outil. Ce sont ces approches qui fondent notre démarche. L'outil laisse en effet une part majeure à l'activité humaine et sociale, il est ouvert à une multiplicité d'acteurs et à une variété d'interprétations de régimes d'usages, et suppose des structures plus faiblement contraintes. Dans cette approche on ne situe donc plus le problème au niveau d'un modèle « d'agent cognitif », mais au niveau de l'outillage nécessaire pour accéder à des ressources documentaires et donc des fonctions nécessaires à l'activité cognitive humaine pour se repérer, annoter, naviguer, indexer des documents semi-structurés, comme dans l'approche des « documents pour l'action » [ZACKLAD 04] ou de « l'artefacture » [BACHIMONT 03].

Cela permet un déplacement majeur des difficultés et des enjeux, par rapport à la modélisation formelle : la modélisation de connaissances, telle que nous la proposons dans la suite, comme « cartographie » prenant appui sur le modèle HyperTopic (cf. §2.1 et Fig.3), n'est pas tributaire des contraintes d'une modélisation exhaustive supposant l'explicitation d'ensembles complets de concepts. Nous essayons aussi de cette façon d'intégrer les acquis du Travail Coopératif assisté par ordinateur (TCOA / CSCW), qui se réfère davantage à la visualisation qu'à la modélisation des connaissances. Comme on le verra au §2 (Fig.3), HyperTopic est un modèle de représentation des connaissances qui l'on peut situer au niveau « épistémologique » de la signification des réseaux sémantiques défini par [BRACHMAN 79] comme « *structure formelle des unités conceptuelles et de leurs interrelations (...) permettant de donner une définition formelle des primitives structurant la connaissance, plutôt que des primitives de connaissance particulière (comme dans les réseaux de Schank)* ». Ce modèle s'applique à l'information non-structurée ou semi-structurée, qui repose sur l'aspect pragmatique de l'usage d'un outil et sur la façon dont un groupe s'organise pour réguler les significations. Avec cette approche, nous ne sommes pas en présence de *concepts* comme dans les modélisations formelles de domaine, mais de simples inscriptions d'expressions linguistiques - les *thèmes* (en anglais, *topic*) - requérant la présence des acteurs pour réguler la représentation. Dans les cartes de thèmes telles que nous les envisageons, il n'existe pas, ou très peu, de régulation interne par le modèle : « la

sémantique » est recrée de façon externe, en permanence, dans l'activité et l'interaction des acteurs. Cela n'exclut pas bien sûr que l'on puisse faire appel ponctuellement, pour certains sous-ensembles où cette complémentarité serait utile, à des techniques plus formelles (d'ontologie formelle, d'inférence automatique, de mesure automatique de distances sémantiques, de traitement statistique, de traitement automatique du langage, ou de systèmes multi-agents).

## **2. Concepts, méthodologie et outils**

### **2.1. Modèle générique et choix d'une architecture de rôles**

Cette partie introduit brièvement les éléments généraux d'ingénierie des connaissances (modèle, méthodologie, outils) auxquels nous avons fait appel pour le dispositif de carte de thèmes pour le logiciel libre, sachant que ces éléments peuvent être transposables à d'autres situations et applications du Web socio sémantique.

Le modèle est le cadre formel sur lequel vont reposer les deux facettes très imbriquées de l'ensemble d'activités considérées :

- l'expression par des cartes de thèmes d'une sémantique explicite d'objets métiers dans un domaine, à l'intention de divers rôles des acteurs,
- et la co-construction de cette sémantique explicite, y compris des sémantiques des activités et des rôles, par les acteurs.

Le modèle générique, proposé pour faire face à ce double problème, avec une visée de réutilisabilité, est le modèle HyperTopic. Ce modèle sert de langage de représentation de connaissances et de noyau pour la construction, autour de lui, d'un ensemble d'éléments semi-formels, sous forme de « cartes de thèmes », construites par les utilisateurs du système.

La méthodologie suivie est donc fortement structurée par les composants du modèle HyperTopic. HyperTopic intervient donc à un niveau très générique, sans définir un choix figé de rôles, ni un modèle unique de coopération. Il offre en revanche une base à l'expression, pour des applications particulières, des architectures de rôles adaptées.

Dans l'application de cartes de thèmes considérée pour les logiciels libres, l'architecture de rôles retenue est celle d'une « place de marché à base de connaissances » (« Knowledge-Based Marketplace » ou « KBM »), modèle de coopération qui a déjà donné lieu à plusieurs mises à l'épreuve sur le terrain [CAHIER 02b], [CAHIER 04a].

Comme le montre la Figure 3, l'architecture générale utilisée requiert donc deux niveaux de généralité. Le niveau le plus général, HyperTopic, sert d'appui pour des modèles plus spécialisés pour s'adapter à des genres d'activité particuliers utilisant des cartes de thèmes, comme l'écriture collective, l'annotation, la négociation, la gestion de contenus Web. Dans le cas présent le modèle retenu (KBM) apporte à ce niveau à la fois un ensemble prédéfinis de rôles et un modèle de coopération. Aux niveaux supérieurs, nous introduisons une distinction entre un niveau de carte de connaissances (expression de que nous préférons à « modèle de connaissances », compte tenu des précautions épistémologiques énoncées au §1.7) et le niveau informatique proprement dit, incluant les données mémorisées par le système.

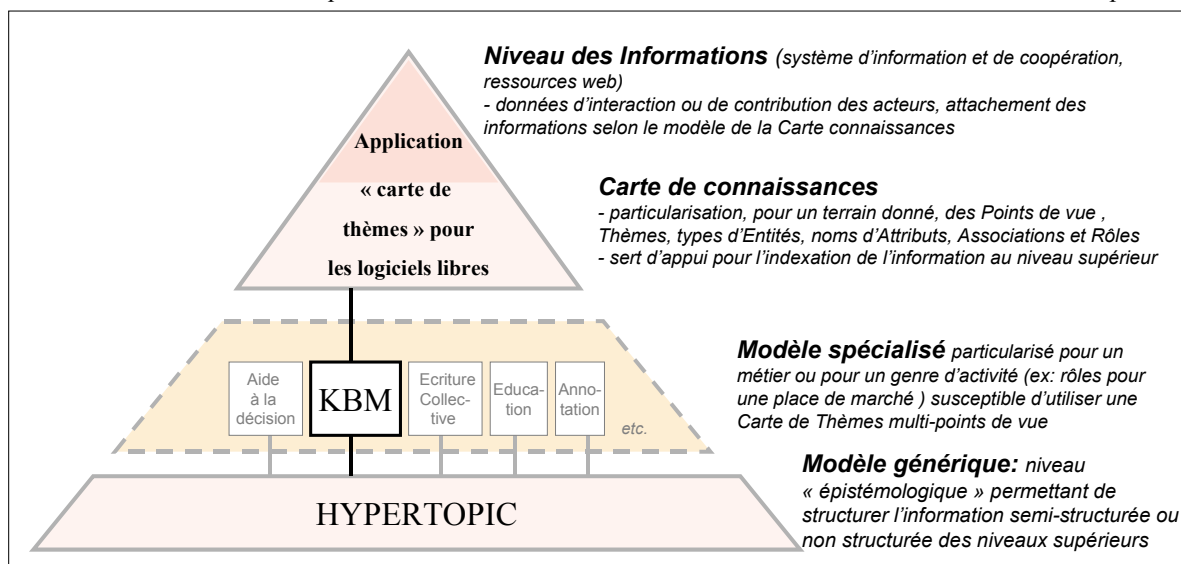


Fig.3 – architecture d'une application HyperTopic

Le modèle de Place de Marché à base de Connaissances (KBM) est le modèle spécialisé le plus développé aujourd'hui pour HyperTopic. Ce modèle KBM est commode pour illustrer, dans l'une de ses voies possibles, le concept de Web socio sémantique, pour trois raisons

- Le modèle KBM est simple. Il comporte un petit nombre de rôles (dérivés des « place de marché » virtuelles) faciles à comprendre par des utilisateurs non-spécialistes ;
- Il s'applique bien à des collections importantes d'objets métier d'un seul type (ici, les « logiciels »)
- Il peut s'appuyer sur d'un outil logiciel, AGORÆ, conforme au modèle HyperTopic, ayant permis de développer déjà plusieurs applications KBM sur d'autres terrains, et donc de disposer de repères méthodologiques accélérant la mise en place d'applications.

Dans la suite de cette partie, nous décrivons les concepts du modèle HyperTopic (§2.2) avant d'évoquer certains repères méthodologiques associés au modèle KBM (§2.3). Nous donnons enfin quelques éléments sur la plate-forme d'outils AGORÆ qui a été utilisée pour la réalisation de la carte de thèmes (§2.4).

## 2.2. Le modèle HyperTopic

Les figures 4a et 4b résument les composants du modèle HyperTopic, respectivement dans la version actuelle de ce modèle (HyperTopic V1, implémentée dans l'outil AGORÆ) et dans la version future (V2) qui étend les cartes de connaissances aux acteurs et aux activités.

L'objectif général est, comme dans le standard Topic Map [XTM 01] de fournir les éléments nécessaires pour décrire une carte de thèmes (Topics), auxquels sont rattachées des ressources Web liées aux objets du monde. Les *thèmes* ne sont pas des concepts (cf §1.7) mais des expressions linguistiques simples ou complexes exprimant « un sujet dont on parle ». Mais dans HyperTopic, les concepts de base pour structurer une carte ont été enrichis par rapport aux Topic Maps, notamment pour faciliter leur manipulation. Pour cela, en plus des *thèmes*, *associations* et *ressources* qui reprennent les concepts normalisés des Topic Maps, HyperTopic utilise les notions d'*entité* et de *point de vue*.

C'est l'*entité* (et non la ressource documentaire) qui est reliée aux thèmes. Nous introduisons cette *entité*, car dans beaucoup d'applications la recherche d'information s'applique d'abord à des « objets » possédant une structure générique qui agrège un certain nombre de descripteurs permettant leur caractérisation « primaire », et auxquels sont associés des attributs standards et une ou plusieurs occurrences de ressources matérielles porteuses de l'information cible. Par exemple, dans l'application logiciel libre, l'entité est un logiciel, et la ressource associée peut être une fiche de description ou bien directement l'URL du producteur de ce logiciel libre, permettant éventuellement de télécharger le logiciel.

Le *point de vue* est un descripteur de mise en situation d'une entité correspondant à une vue de certains acteurs. Il correspond à une famille de caractéristiques de l'entité, regroupées et hiérarchisées en plusieurs niveaux, en fonction d'un angle de vision « faisant sens » pour un acteur ou un ensemble d'acteurs donnés dans certaines situations (par exemple, un point de vue correspondant à « un métier »). Cette définition du point de vue distingue HyperTopic des approches de type FacetMap (XFML, 2002) dans lesquelles les « facettes » traduisent au départ essentiellement des couples attribut-valeur attachés aux objets considérés, sans poser au niveau du modèle la possibilité explicite de regrouper les facettes dans des points de vue signifiants pour les acteurs, comme le propose HyperTopic.

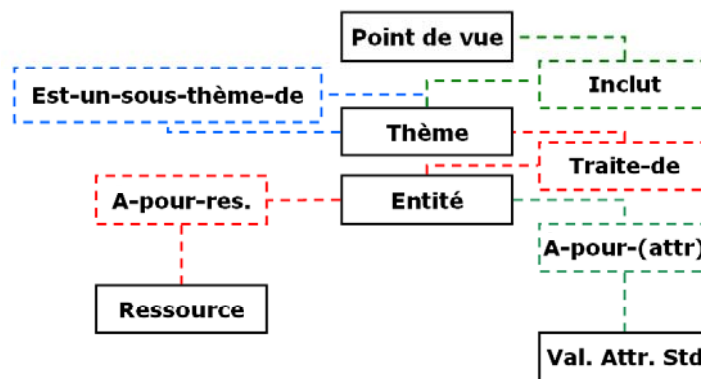


Figure 4a – Base du modèle HyperTopic (V1) (présentation simplifiée avec, en pointillés, les relations)

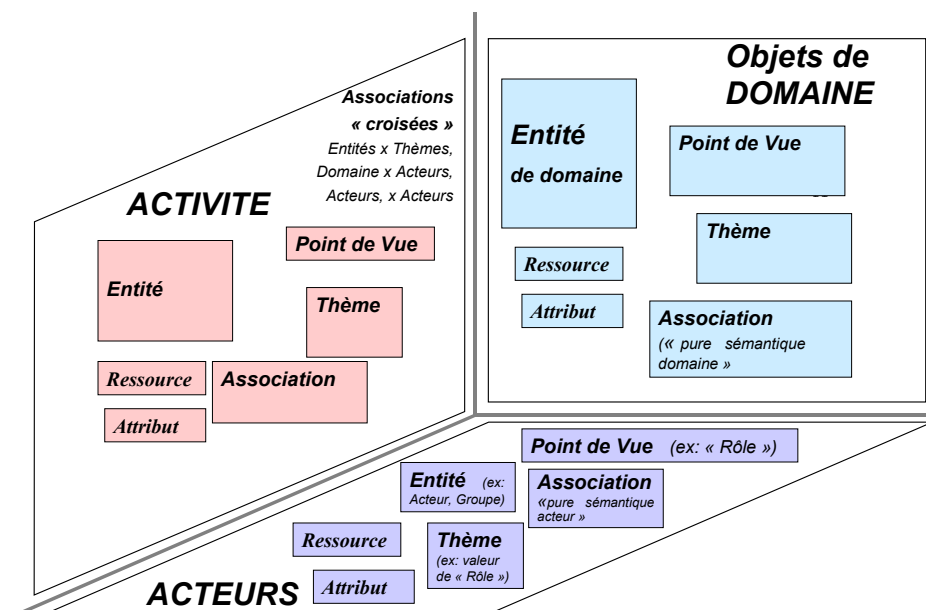


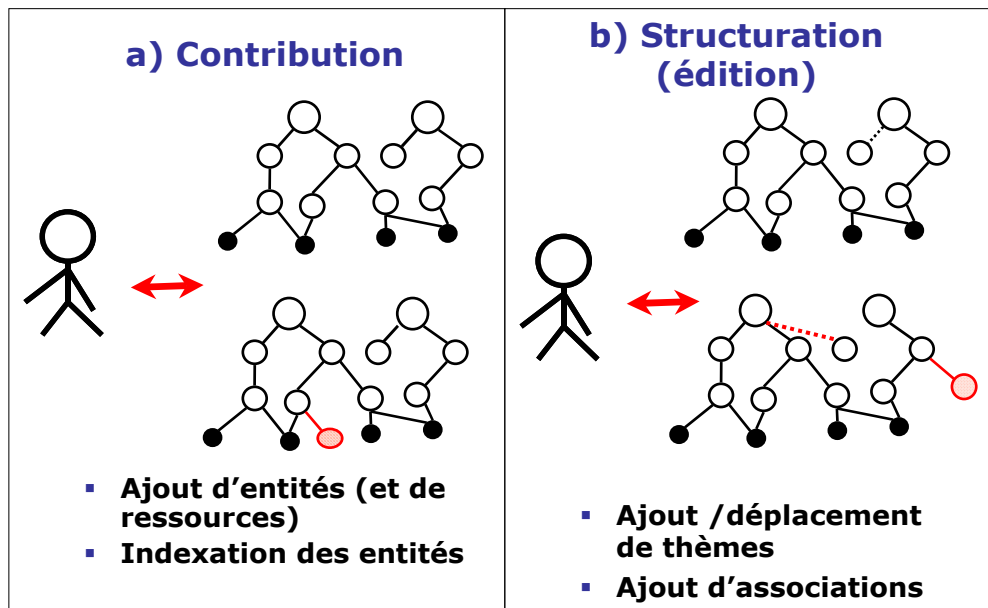
Figure 4b – Le Modèle HyperTopic (V2) type l'entité pour réaliser des cartes de thèmes sur trois plans

Notons bien que dans HyperTopic, un point de vue est un point de vue *sur* une entité. Dans l'application logiciels libres considérée, il existe une seule entité «logiciel», sur laquelle portent cinq points de vue (qui seront détaillés au §3.2).

La relation utilisée entre les thèmes dans un point de vue est l'association standard « type / sous-type ». En dehors de cette relation hiérarchique, le modèle prévoit la possibilité d'associations transverses de thème à thème. Ce sont des relations nommées, par exemple dans l'application logiciel libre, on utilise une seule de ces relations transverses de thème à thème « voir aussi » (pour des raisons de simplification, cette relation n'a pas été représentée sur la figure 4a).

Dans sa version 2, HyperTopic permet de spécifier et de retrouver *ce que sont* les objets métiers, et *ce qui change* dans ces objets au fil du temps, en termes de définition externe des objets par des attributs heuristiques (les thèmes associés à une instance d'entité). Dans sa partie acteurs / activités, le modèle permet aussi de définir *par qui* sont modifiées la collection d'entités et la carte de thèmes, *avec qui, quand, comment*, avec quel degré de *certitude* pour l'acteur, avec quel degré de *validation* pour l'organisation, etc. Les acteurs ne disposent pas forcément des mêmes droits ou compétences pour intervenir aux diverses étapes, d'où l'importance de disposer aussi d'une marge de malléabilité dans la définition de rôles d'acteurs.

On peut ainsi décrire la façon dont plusieurs acteurs contribuent à une séquence ou à une action en complémentarité. Par exemple considérons le scénario de *création* d'une nouvelle entité. Dans l'exemple la carte de thèmes de logiciels libres, une séquence serait initialisée par un développeur souhaitant décrire un nouveau logiciel et l'associer à certains thèmes, selon différents points de vue (Figure 5a), et créer un nouveau thème ou déplacer un thème existant si la carte existante est incomplète (Figure 5b). Déplacer (ou supprimer) un thème peut impliquer de retrouver les acteurs ayant créé ce thème et ayant attaché des entités à ce thème, pour les prévenir ou engager une discussion. En prenant en compte la dimension des acteurs et des activités, le modèle HyperTopic (V2) crée les conditions permettant de gérer ces activités collaboratives au niveau de la carte de connaissances. Sans cela, la maintenance d'une carte de thèmes dans des conditions opérationnelles, avec des acteurs participants et des entités rattachées en grand nombre, est très difficile.



**Fig.5 - Contribution (a) et structuration sémantique (b) avec HyperTopic** : les pastilles pleines représentent les entités, les pastilles creuses les points de vue et les thèmes. En gris l'entité (Fig 5a) ou le thème (Fig 5b) ajoutés.

De plus, ce n'est pas forcément le même acteur qui va être impliqué dans tous les événements que comporte ce scénario. Par exemple, dans l'activité de contribution, un développeur peut s'estimer compétent pour indexer le logiciel dans les points de vue « Fonctionnalités » et « Génie logiciel » (cf.3.2), mais moins dans les points de vue « Juridique » et « Business Model », donc souhaiter s'adresser à un autre acteur pour passer le relais ou lui demander de valider ses choix ou ses créations de thèmes. Le modèle HyperTopic V2 est conçu pour permettre de mieux prendre en compte de ces étapes de coopération entre les acteurs. D'une façon générale HyperTopic (V2) est conçu pour permettre une définition facile de rôles selon une granularité fine, par des non-informaticiens utilisant des cartes de thèmes (cf. Fig.8 pour l'administration des autorisations).

### 2.3. Les rôles et les étapes de mise en place du système

#### 2.3.1. L'amorçage et la conception initiale du système

En phase initiale, il faut s'accorder sur la ou les entités considérées et construire, sur chaque entité, un ensemble de points de vue ayant un sens dans une organisation (cf. Figure 2). Dans un but « d'amorçage » du système, il peut également être important de préparer dans cette phase une version initiale, même très incomplète ou imparfaite, de la carte des thèmes. Version d'amorçage qu'un cercle plus nombreux d'acteurs perfectionnera ensuite, dans la phase de co-construction sémantique élargie. Dans le cas de l'application logiciels libres, le choix s'est par exemple arrêté rapidement sur une seule entité « logiciel ».

#### 2.3.2. La construction sémantique par un cercle élargi d'acteurs

Une fois le système initialisé, le dispositif doit être capable de supporter les formes d'utilisation et de co-construction sémantique du système en « vitesse de croisière », autour des différents types de rôles nécessaires à la communauté s'organisant autour du système. Par exemple – c'est l'une des différences entre la phase d'initialisation et la phase de

croisière – il est recommandé que les éditeurs n'aient plus alors le droit de modifier directement les points de vue. Les points de vue, qui jouent un rôle très structurant, ne pourront être modifiés qu'après un débat sur le forum et une décision collective.

La co-construction à grande échelle est plus facile s'il existe une culture de confiance au sein de la communauté, et si l'on peut connaître l'auteur d'un choix et entamer le dialogue avec lui, aussi bien dans les situations d'accord que de désaccord. Le logiciel AGORÆ possède une fonction de mémorisation du nom d'utilisateur ayant effectué la dernière mise à jour sur un thème ou sur une entité, ce qui représente aussi un premier niveau de responsabilisation, en attendant le développement de fonctionnalités de trace et de collaboration plus poussées.

#### 2.4. La plate-forme logicielle AGORÆ

Conforme au modèle HyperTopic (Version 1), la plate-forme d'outils AGORÆ actuellement utilisée pour l'application logiciel libres est développée par le Laboratoire Tech-CICO avec des logiciels libres (PHP, MySQL), selon une architecture et un code source modulaire, dans une optique de généricité et de portabilité de l'outil.

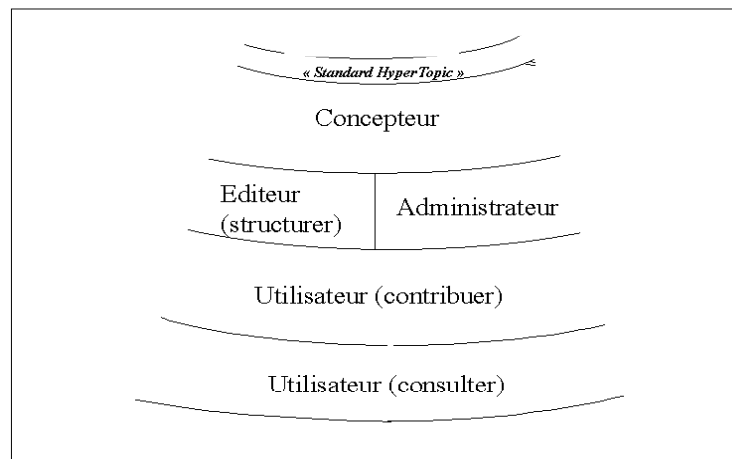


Fig.6 – Architecture des rôles KBM repris par l'outil AGORÆ

Bien qu'il s'agisse d'un outil de recherche, destiné surtout à développer et valider les concepts du Web socio sémantique, l'outil AGORÆ V1 présente aujourd'hui la plupart des fonctionnalités nécessaires aux deux phases évoquées, de conception initiale et d'utilisation normale en régime « de croisière », pour des applications recourant à des rôles KBM standard : la création et la gestion d'une structure par points de vue et thèmes sous forme d'un arbre, la création et la gestion des entités et des ressources associées (documents, URL). L'outil implémente aussi certains mécanismes de contribution à plusieurs sur certaines fonctionnalités (structure de l'arbre, ajout de ressources...) et de trace des actions.

L'outil AGORÆ propose également des éléments d'interface adaptés pour le rôle « concepteur », en phase de conception initiale d'une nouvelle application basée sur HyperTopic. Les membres du groupe concepteur ont accès à toutes les fonctions des autres rôles, notamment d'éditeur pour la création *ex nihilo* des points de vue et des arborescences de thèmes. Ils disposent aussi de fonctions d'*importation* d'éléments d'arborescences thématiques d'origine diverses (thésaurus, ontologies...) depuis d'autres bases AGORÆ ou depuis XML ou depuis des formats « bureautiques » (Excel, MindManager...). Pour le

groupe concepteur, il peut en effet être commode d'utiliser de telles fonctions pour une construction suivant diverses méthodes (incrémentale, en mode « brain storming », « descendante », « guidée par les instances », etc.) de la carte initiale.

Notons enfin qu'une base créée avec AGORÆ peut d'ores et déjà être transcrite automatiquement au format XML selon un schéma HyperTopic. Une carte de thèmes, un point de vue ou un sous-ensemble d'arborescence ou de composants HyperTopic, par exemple l'application de cartes de thèmes pour les logiciels libres, peut être représentée et exportée à partir des outils AGORÆ dans un format d'extension XML, pour lequel nous proposons l'appellation XHT (XML HyperTopic) (pour plus de détails, voir [www.sociosemanticWeb.org](http://www.sociosemanticWeb.org)). Cela participe de notre effort pour proposer à terme rapproché HyperTopic comme un standard pour le Web socio sémantique. Une telle normalisation, complétant XML et éventuellement RDF (donc à un niveau de syntaxe, cf. Fig.1), peut être très utile y compris pour les besoins très pratiques de récupérer et fusionner facilement des cartes de thèmes dans un format partagé, facilitant l'échange et accélérant les manipulations manuelles. Des passerelles et des conversions sont également envisagées à partir de XHT vers les formats XTM et XFML [XTM 01] [XFML 02]).

### **3. Le dispositif et l'application « logiciels libres »**

#### **3.1. Objectifs, approche sur le terrain et motivation des acteurs**

Cette partie vise à préciser les objectifs et la méthodologie particulière, à partir des éléments méthodologiques généraux esquissés précédemment (§2.3), qui ont été suivis pour la carte de thèmes du logiciel libre. Outre le fait de réaliser un système informatique opérationnel sur le Web, c'est bien en effet un système *socio-technique* singulier dont il s'agit de réussir la construction, sachant que cette construction est surtout le fait des acteurs. Les concepteurs d'une application de collecticiel socio sémantique, telle que cette carte de thèmes, s'ils peuvent amorcer le processus et continuer ensuite à y participer, ne peuvent ni amener toutes les connaissances, ni coopérer à la place des acteurs. C'est pourquoi nous précisons rapidement quelques objectifs et choix d'approche (de management, et de management de connaissances) opérés relativement à la communauté concernée.

La carte de thèmes « logiciels libres » vise potentiellement l'ensemble des offreurs et des utilisateurs (francophones pour l'instant) de ces logiciels : chacun peut dans ce système proposer, répertorier et retrouver facilement en ligne des logiciels entiers ou des composants : un acteur peut y déclarer le logiciel qu'il a conçu, mais aussi un logiciel qu'il connaît, par exemple pour l'avoir utilisé ou testé, selon une variété thématique très détaillée. Quelques semaines après son lancement, la carte comporte déjà plusieurs centaines de thèmes, et une centaine d'outils de logiciel libre y sont rattachés.

À ce premier niveau, l'objectif est de fournir un outil simple et facilement utilisable pour chercher et décrire des logiciels libres : une sorte de Place de Marché, rejoignant d'une certaine façon l'inspiration de départ du modèle KBM. Cet outil doit aussi condenser dans sa carte de thèmes un socle d'expertise commun à la communauté, dans un souci d'ouverture, d'indépendance par rapport aux pressions commerciales, et de qualité.

Par rapport à ces objectifs idéaux, le travail de structuration sémantique de la carte présente une vraie difficulté, notamment au début pour l'équipe d'amorçage. Pour limiter cette difficulté, cette équipe réunit des compétences à la fois en Ingénierie des Connaissances et dans le domaine des logiciels libres, en associant des personnes ayant déjà au départ des

expériences approfondies et différentes sur ces logiciels. Le groupe impliqué dans la conception initiale de la carte de thèmes est composé au départ d'une dizaine de personnes – incluant notamment les auteurs du présent article - notamment de trois laboratoires universitaires (BETA-ULP à Strasbourg, Tech-CICO à Troyes, LIMSI à Orsay).

Cette condition de double compétence semble importante pour réussir la phase d'amorçage des points de vue. Car la majorité des spécialistes en logiciels libres disposent rarement d'un second savoir-faire en Ingénierie des Connaissances. Comme le notait déjà P.Y. Raccach, « *si l'expert est, par définition, compétent dans son domaine d'expertise, il n'est pas, en tant qu'expert du domaine, compétent sur son expertise* » [RACCAH92].

Quant aux motivations pour construire le système, celles-ci doivent être fortes en phase d'amorçage, mais aussi dans la phase d'élargissement, pour réussir à amener la participation voire la « fidélisation » de nombreux acteurs comme éditeurs de la carte. Comment motiver ces acteurs, au-delà de la simple curiosité pour un concept innovant de co-construction de carte de connaissances ? L'intervention pour structurer, même très partiellement une carte de thèmes, est une opération de prise de recul qui reste difficile, inhabituelle, présentant même pour certains acteurs des formes d'inhibition, dans un contexte où les experts dont on recherche la participation sont trop occupés. Elle représente aussi un engagement dans une communication de groupe (même si l'outil doit la rendre plus agréable et en réduire la difficulté et la durée). Il semble exclu de demander à des acteurs réels en grand nombre de « construire la carte pour construire la carte », sans contrepartie de l'effort pour réfléchir, classifier, communiquer, débattre, négocier que cela suppose.

Pendant, ces contreparties existent. Au départ, l'appartenance ou le désir d'appartenance communautaire est important et même indispensable ; dans ce cadre, un acteur peut considérer comme gratifiant d'être « auteur » d'éléments de connaissances, s'il sait que la trace de son apport créatif subsiste d'une façon ou d'une autre. Il y a un but commun et une stratégie « gagnant-gagnant » qui veut que si l'acteur fournit des connaissances dans le système, il en retire(-ra) aussi de celles que les autres membres y (aur-)ont fournies.

Ces premiers aspects de la motivation existent, mais ne suffisent pas toujours à créer une forte motivation. Ils sont selon nous renforcés par le fait que le modèle HyperTopic prévoit la notion d'entité. C'est un enjeu concret qui conduit à ne pas (seulement) construire « la connaissance pour la connaissance ». Le fait de décrire une entité (un logiciel donné, pas n'importe lequel) en rapport avec un enjeu métier (pour le faire connaître, le valoriser, le diffuser, vendre des services associés, faire connaître sa compétence de critique, etc. ) est aussi un puissant ressort d'incitation. Placer ou discuter le déplacement de thèmes sur la carte demande un effort cognitif et social important. Mais l'acteur y consentira plus volontiers, parce qu'il met en rapport l'effort nécessaire avec le retour attendu d'une bonne description du produit, donc avec son but de métier, qui est également important.

Ce « principe de réalité » amène naturellement l'acteur non seulement à consulter et contribuer, mais à structurer et à entrer potentiellement en communication avec d'autres éditeurs de la carte, via le dispositif proposé, dans des discussions et des négociations (convergences ou micro-conflits de points de vue, de lexique, de sens d'un thème, de description d'une entité, etc.). Ce faisant, il participe alors à un second niveau de communauté de « co-construction sémantique », qui doit se doter de tous les moyens de communication et d'arbitrage. Cette seconde communauté, que nous avons analysée plus précisément comme une communauté d'action [ZACKLAD 03b] [CAHIER 04b], se fixe comme objectif externe (but de prestation) de développer la carte et de la maintenir à jour, compte tenu des évolutions des entités indexées (ici, les logiciels libres), des connaissances, de

l'environnement, et de ses rapports des forces internes. La communauté du logiciel libre ne reste d'ailleurs pas inactive<sup>2</sup> pour élaborer des systèmes de suivi de ces changements, dans un contexte où le nombre de projets open source devient astronomique, et nous souhaitons bien entendu nous situer comme une source d'idées expérimentales en complémentarité éventuelle avec ces initiatives.

Dans notre cas, la particularité des logiciels libres a permis d'avancer l'hypothèse d'un recouvrement sensible entre ces deux niveaux de communautés : dans le monde du logiciel libre, les utilisateurs sont bien souvent en même temps développeurs ou ré-utilisateurs, renforçant le fait que ce sont souvent les mêmes personnes qui sont susceptibles de consulter, de contribuer et d'éditer la carte de thèmes, et qui sont compétents pour cela. Au moins dans un premier temps, il a donc semblé inutile de dissocier ces niveaux de communautés. Cela représente un avantage pour une mise en place facilitée du dispositif – inutile notamment de mettre en place un système de droits et d'autorisations sophistiqué et coûteux. La tradition de responsabilité et de confiance dominante dans la communauté du logiciel libre a aussi plaidé pour le choix du logiciel libre comme terrain d'application. Toute personne souhaitant contribuer ou éditer reçoit systématiquement un mot de passe pour ces rôles (Nul n'étant cependant à l'abri du risque de manipulation erronée ou mal intentionnée, il a été toutefois décidé de sauvegarder très régulièrement le site).

Au-delà de la phase d'amorçage, si le dispositif suscite l'adhésion des intéressés, la tradition de débat et d'inventivité collective dont sait faire preuve la communauté du logiciel libre devrait permettre que l'élaboration de solutions, par exemple pour complexifier le système de droits ou améliorer le collecticiel, soit reprise au niveau de la communauté elle-même.

### **3.2. Quels points de vue pour classier les logiciels libres ?**

Cherchant à refléter une sémantique à la fois usuelle dans la communauté et efficace pour la classification des logiciels, la carte de thèmes peut donc être considérée comme la structure d'index d'un annuaire « pages jaunes » des logiciels libres, servant de relais pour accéder via le Web à des ressources descriptives de ces outils. Les usagers en consultation du système vont chercher à évaluer ou comparer les outils en fonction de buts métiers (par exemple intégrer un logiciel dans une application, faire un choix pérenne, savoir quel est l'état d'achèvement du logiciel, sur quel service complémentaire ils pourront compter, etc.). Pour prendre en compte toutes ces dimensions d'activité, la carte de thème a pour le moment été organisée, à l'issue de la phase d'amorçage, selon les points de vue suivants :

- « *Thématiques / Fonctionnalités* » : un point de vue classificatoire des thématiques des logiciels et des fonctionnalités qu'ils proposent : développement logiciel, utilitaires système, multimédia, jeux, etc.;
- « *Génie logiciel* » : un point de vue des techniques et méthodes de développement et d'intégration, décrivant la nature des logiciels, les doctrines et les modèles de leurs développements, etc. ;
- « *Business models* » : un point de vue économique, s'inspirant au départ de la classification proposée par [GUITTARD 02] : modèle non-business, modèle à façon, modèle défensif, modèle des distributions Linux, modèle service et hardware ;

---

<sup>2</sup> Des projets, extensions XML modernes de RPM et métadonnées eMerge sont actuellement en développement, notamment par Edd Dumbill (chairman XML Europe) pour intégrer des descriptifs projets aux sources des développements open source.

- « *Juridiction et droit* » : un point de vue juridique, incluant les licences, les droits des tiers, les brevets mis en jeu, etc. ;
- « *Acteurs* » : un point de vue plus organisationnel traitant les communautés des logiciels, certaines sociétés, organisations, organismes ou projets de recherches jouant un rôle par exemple pour des labels, etc. (« Sun », « Inria », « Sourceforge ») ;

Ces points de vue non réductibles les uns aux autres reflètent le fait que le logiciel libre est un domaine complexe, en évolution rapide. Ils sont loin d'être donnés au départ ni de faire l'unanimité, même dans le groupe restreint d'amorçage du système. Le fait de les dégager, même à titre d'hypothèses, revient à produire des connaissances décisives sur le domaine.

Ces points de vue remplissent de façon imbriquée plusieurs fonctions. Ils servent de passerelles en assurant des traductions entre des langages d'offre et des langages de demande, très spécialisés et moins spécialisés, des dialectes de métiers voisins ou complémentaires dans l'organisation du travail. Dans des contextes où les acteurs appartiennent simultanément à plusieurs communautés (entreprises, communauté des informaticiens, communauté du logiciel libre, communautés épistémiques plus particulières), les lexiques de métiers posent des problèmes d'ambiguïté, d'évolution et de variation tant synchronique que diachronique, qui ne peuvent être levés que par la contextualisation effectuée par les experts. Les points de vue (ainsi que les chemins complets depuis le point de vue jusqu'à un thème en passant par tous les sous-thèmes du point de vue) constituent alors des moyens pour optimiser la couverture sémantique et réduire les conflits de sens dus par exemple à la coexistence de dialectes, un problème bien étudié par [PRINCE 96] notamment pour les dialectes informaticiens qui nous intéressent ici. Enfin, les points de vue correspondent aux langages particuliers des rôles d'acteurs concernés, notamment ceux qui *proposent* des logiciels (développeurs, groupes de développeurs, éditeurs de logiciels) et ceux qui souhaitent les *acquérir* pour les utiliser, les intégrer avec d'autres programmes, les « reprendre » pour les développer dans de nouvelles directions (bifurcations), etc. Les termes de lexique de métier, contribuant dans la carte de thèmes à définir l'identité ou les propriétés des outils de logiciel libre, notamment le coût réel, sont importants pour des activités-clés de ces acteurs, comme la promotion, la comparaison, le choix, etc.

La « carte sémantique » à construire pour cela prend pour toile de fond, en première analyse, la base « d'expertise » commune à la communauté du logiciel libre, soit une part de connaissances globalement partagées, sans pour autant que cette connaissance puisse y être considérée ni comme unanime, ni comme bien explicitée, ni comme comportant le niveau de cohérence que l'on attendrait d'une approche scientifique.

### **3.3. Les possibilités d'interaction suivant les rôles**

a) Page d'accueil de consultation : l'utilisateur dispose d'une approche globale sur les points de vue, regroupant plusieurs centaines de thèmes sur 5 niveaux. Il peut ensuite naviguer (en mode « Yahoo ») dans les thèmes.

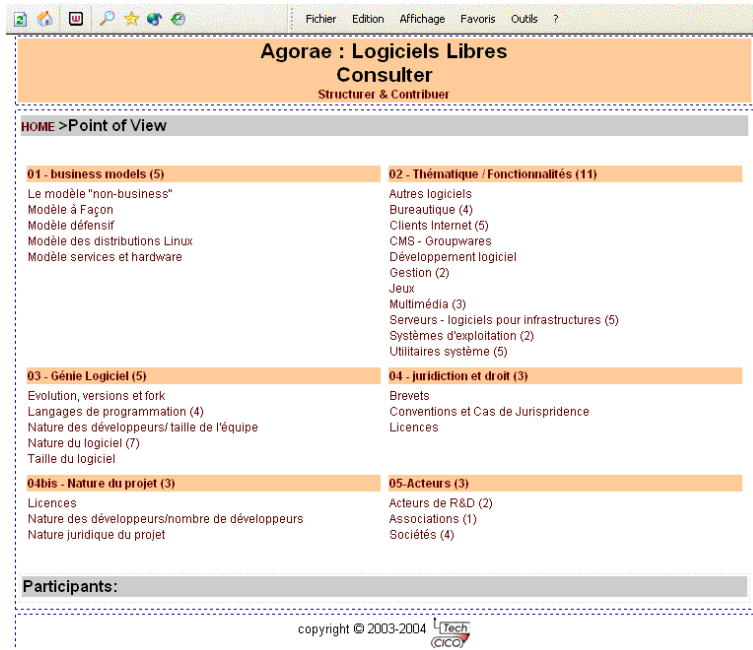


Fig. 7a - L'application « logiciels libres », page d'accueil de consultation

b) Pour un thème donné (dont le chemin d'accès est fourni en haut de l'écran) de niveau quelconque dans l'arborescence, l'utilisateur peut connaître les produits correspondants sur ce thème (« logiciels ») ainsi que les thèmes associés (« voir aussi »).

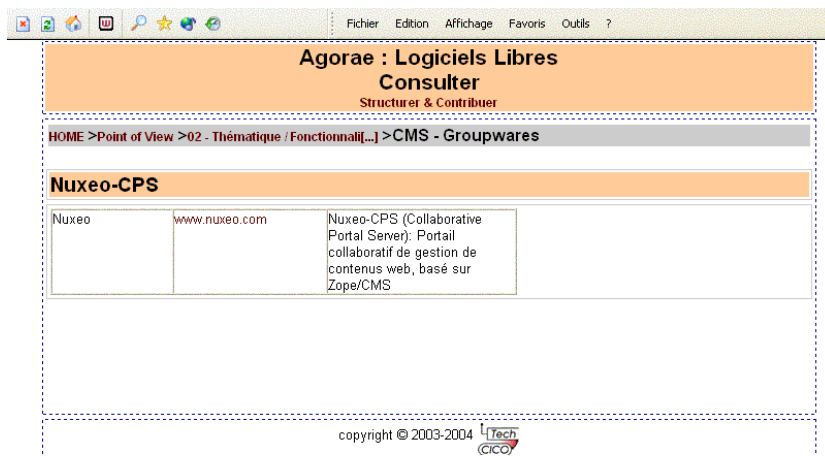


Fig. 7b- L'application « logiciels libres » : fiche d'une entité reliée au thème « CMS - Groupware »

c) Création d'une instance d'entité : le contributeur définit le contenu de la fiche du logiciel (ici par copier-coller depuis un tableur) et peut associer à chaque entité ajoutée tous les thèmes qu'il souhaite, par exemple le thème « CMS-Groupware » sous le point de vue *Thématiques / Fonctionnalités*. Le contributeur dispose pour cela sur la droite de l'arbre de l'ensemble des points de vue et thèmes qu'il peut cocher ou décocher à sa guise. Contrairement aux écrans de consultation, l'interface permettant de structurer et contribuer est réservée aux membres identifiés (tout membre concerné par le logiciel libre souhaitant participer en contribuant ou éditant est a priori admis, et reçoit un identifiant)

Fig. 7c – L'application « logiciels libres », en mode contribution

d) Création ou modification d'un thème : le rôle « éditeur » permet de créer ou modifier le nom du thème, les commentaires associés (ex : définition, traduction), sa position dans l'arbre. L'éditeur coche le thème parent (dans l'arborescence de gauche) et éventuellement tous les thèmes associés (dans l'arborescence de droite).

**Fig. 7d - L'application « logiciels libres », en mode édition**

## **4. Perspectives**

### **4.1. Les conditions réunies d'un « passage à l'échelle »**

Du point de vue des objectifs généraux de l'expérimentation, l'objectif, actuellement en cours de réalisation, est d'observer le comportement du système dans des conditions ouvertes et non idéales, nécessitant de faire face notamment à des désaccords de classification ou de signification des termes dans la communauté. Pour cela, nous souhaitons atteindre un rythme de croisière de plusieurs dizaines de membres « fidélisés » et un élargissement à de nombreux membres plus ou moins occasionnels éditant les thèmes sur le site. Ce niveau d'utilisation permettra, comme nous l'avons expliqué, d'étudier la faisabilité d'une telle co-construction sémantique « au fil de l'eau » à cette échelle plus importante.

Pour mieux assurer cette co-construction, comprendre comment elle peut s'établir et quelles conditions peuvent la favoriser ou permettre de dépasser les difficultés, nous nous appuierons en particulier sur des fonctions de collecticiel complémentaires (forum, annotations partagées...) et sur l'enregistrement des versions successives de la carte de thèmes et de certains échanges verbaux entre membres de la communauté, que permettra prochainement l'outil AGORÆ. Il apparaît en effet nécessaire de fournir des espaces de discussion structurés dans la forme (fils de discussion relatifs à un thème, une entité) mais libres dans le contenu.

Cet article est donc à considérer comme une première étape dans un travail en cours. L'étape actuelle ouvre déjà un certain nombre de pistes intéressantes, notamment en ce qu'elle a permis de préciser un cadre d'objectifs pour la suite, et de commencer à mieux décrire les formes d'outillage « collecticiel » nécessaires pour établir plus solidement la praticabilité à l'échelle de l'ensemble d'une communauté et une réelle appropriation sociale par les acteurs.

### **4.2. Perspectives d'amélioration de l'outil**

Pour les outils génériques AGORÆ, un certain nombre d'extensions sont envisagées ou en cours, contribuant à rendre la perspective du « Web socio sémantique » davantage opérationnelle et accessible.

*Interface concepteur.* Nous travaillons notamment à compléter les outils existants avec une Interface complète pour un rôle « Concepteur », de type « assistant », pour guider le groupe de concepteurs non-informaticiens dans la création d'une nouvelle application, définir ses préférences et spécifier la partie applicative (points de vue, rôles, activités, associations et éventuelles contraintes sur les associations), qui ne pourront plus être modifiées ensuite par les autres rôles. Ce paramétrage va constituer le squelette de la structure sémantique, qui sera co-construite ensuite par un cercle plus large d'éditeurs respectant ce cadre de départ.

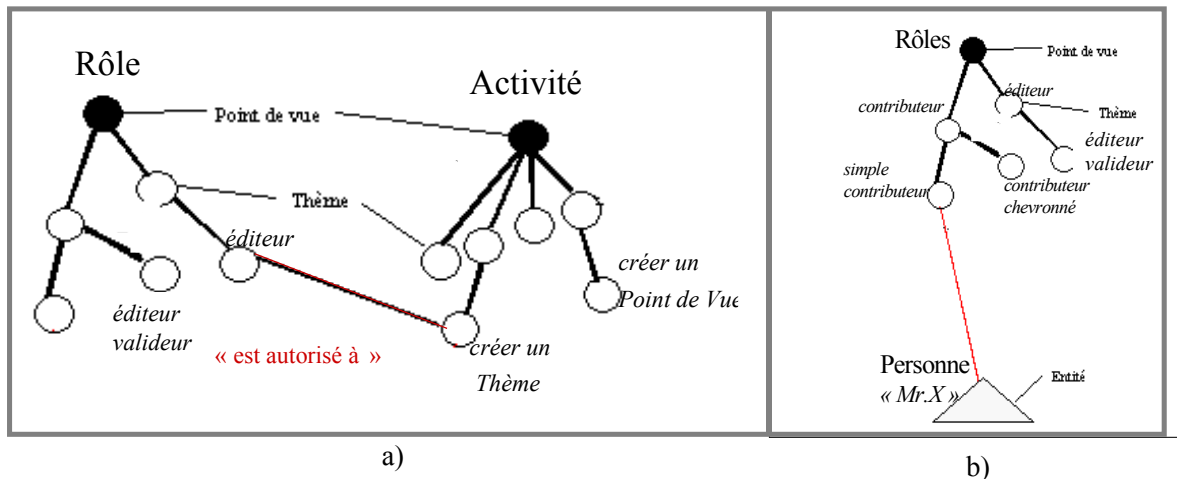


Fig. 8 –Edition malléable, basée sur HyperTopic, des rôles (a) et des autorisations (b)

Vers une spécification « sémantique » des rôles et des autorisations des acteurs. L'association des acteurs aux rôles est gérée de façon classique (matrice d'autorisation). Les rôles devenant plus nombreux et différenciés pour leur composition en activités, nous étudions la possibilité de permettre à des non-informaticiens d'éditer ces rôles à la façon des cartes de thèmes dans le concept général HyperTopic (cf. Figures 4b et 5), ce qui permettra de réutiliser les outils AGORÆ pour des fonctions d'administration des rôles et des acteurs.

*Rôles KBM.* Dans le cadre des rôles dédiés du modèle KBM, il peut être intéressant, grâce au socle HyperTopic v2, de pouvoir particulariser ces rôles pour une application donnée. Par exemple, le logiciel libre est de plus en plus structuré au niveau mondial, avec un rôle important de pôles communautaires, créateurs d'éléments de thématique en anglais ou d'autres langues, qui peuvent être régulièrement importés dans la carte de thèmes. Dans une hypothèse où l'application de carte de thèmes existerait en plusieurs langues, les thèmes devraient être traduits alors que tous les participants en rôle d'« éditeurs sémantiques » n'en ont pas la compétence. Cela conduirait à spécifier des rôles plus fins, par exemple « d'éditeur-traducteur », autorisé par exemple à traduire des thèmes créés par d'autres, ou à rectifier leurs traductions erronées. De même, pour certains thèmes faisant l'objet de définition ou de renvoi vers des sigles ou des URL d'organisations, ces activités plus spécifiques de terminologie ou de relevance organisationnelle relèveraient alors de rôles d'« éditeurs » plus finement particularisés. Là encore des fonctionnalités de définition malléable de rôles (Fig.8a) par cartes de thèmes basées sur HyperTopic V2 seront particulièrement appréciables.

## Conclusion

Ce type de système pose à notre avis des défis importants de gestion de connaissances. Ces défis concernent d'abord l'Ingénierie des Connaissances, le Travail Coopératif Assisté par Ordinateur (TCAO ou CSCW) et les Sciences de Gestion. Ils incitent à la recherche et à l'évaluation de nouveaux outils - vus autant comme outils de gestion que comme outils informatiques - pour la gestion collective de connaissances. De nouvelles modalités doivent être proposées pour répondre au caractère très collectif des formes de gestion de connaissances que nécessitent de telles « cartes de thèmes », construites collectivement au

fil de l'eau, et pour créer les conditions de leur réelle appropriation sociale par tous les acteurs concernés.

Au stade actuel, nous ne pouvons que modestement jeter de premières bases dans ces directions. A travers le projet évoqué de carte de thèmes de logiciels libres, le système apparaît opérationnel au plan informatique, et sémantiquement représentatif d'une convergence au niveau d'un premier petit groupe de co-constructeurs. L'objectif de cartographie de connaissances des outils de logiciel libre n'est pas encore complètement rempli, mais apparaît à portée de main et grandement facilité par le procédé de multi-points de vue. La possibilité de coopérer à distance de façon médiatisée pour réaliser une carte de thèmes, même à cette petite échelle, a pu être testée. Si l'expérience ne permet pas encore de généraliser des conclusions définitives, elle nous encourage toutefois dans notre hypothèse qu'il est envisageable de co-créeer des îlots de sémantique vivante sur le Web à partir de communautés.

Mais nous sommes loin de pouvoir extrapoler à une communauté numériquement plus importante, représentative de la réalité. Il est donc nécessaire de poursuivre et d'étendre cette expérience de terrain en invitant toutes les personnes intéressées par le logiciel libre à la rejoindre. Cela crée l'opportunité d'une évaluation plus rigoureuse, dans une approche de recherche-intervention s'appuyant sur les acquis de l'Ingénierie des Connaissances, du Travail Coopératif et des Sciences de Gestion. Complétant d'autres applications sur le terrain, basés sur le modèle HyperTopic, ces résultats contribueront selon nous à faire du « Web socio sémantique » une perspective opérationnelle et accessible à une grande variété d'applications.

## **Références**

[ALQUIER 93] "Modélisation des systèmes d'information : modèle coopératif", Thèse HDR, Université de Toulouse I, 1993

[BACHIMONT 03] ,"Intelligence économique et écriture dynamique: de la raison graphique à la raison computationnelle", Bachimont B., Intervention au Colloque "Connaissances, activité, organisation", Cerisy, septembre 2003

[BERNERS-LEE01], "The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities", Berners Lee T., Scientific American, May 2001

[BOLTANSKI91], "De la justification - Les économies de la grandeur", Boltanski L., Thévenot L., Editions Gallimard, coll. NRF Essais, Paris 1991

[BRACHMAN 79], « On the Epistemological status of semantic networks », Brachman, R.J., in N.V. Findler, Associative Networks, pp. 3-50, Academic Press, New York 1979

[CAHIER 02a], « Approche cognitive de la co-construction d'un catalogue multi-points de vue et multi-experts », Cahier J.-P., Zacklad M., Actes de la Conférence Extraction et Gestion des Connaissances EGC 2002, Montpellier du 21 au 23 janvier 2002, p.221-226, Hermes

[CAHIER 02b] ,"Towards a Knowledge-Based Marketplace model (KBM) for cooperation between agents", Cahier J.-P., Zacklad M, Actes conference COOP'2002, St Raphael, 4 - 7 June 2002, IOS Press

- [CAHIER 04a], « Une application du Web socio sémantique à la définition d'un annuaire métier en ingénierie », Cahier J.-P., Zacklad M., Monceaux A. ,in Actes de la Conférence Ingénierie des Connaissances IC 2004, Lyon Mai 2004
- [CAHIER 04b], "Socio-Semantic Web applications: towards a methodology based on the Theory of the Communities of Action", Cahier J.-P., Zacklad M., COOP'04 Workshop on Knowledge Interaction and Knowledge Management
- [CAUSSANEL 02] ,« Les Topic Maps sont-ils un bon candidat pour l'ingénierie du Web Sémantique ? »,Caussanel J, Cahier J.-P., Zacklad M., Charlet J., ,Conférence Ingénierie des Connaissances IC2002, Rouen Mai 2002
- [CHARLET 00], "Ingénierie des connaissances", Charlet J., Zacklad M., Kassel G. , Bourigault D., Eyrolles, 2000
- [CHARLET 03], "Ingénierie des connaissances pour les systèmes d'information", Charlet J., Reynaud C., Teulier R., Chapitre 10
- [CHATEURAYNAUD 03], "Prospero : une technologie littéraire pour les sciences humaines ", Chateauraynaud F., CNRS éd. , 2003, Collection : CNRS communication , ISSN 1290-6824 ISBN 2-2710-6133-4
- [DUPOUET 02] ,« Morphogenèse de communautés de pratique », Dupouët O., Yildizoglu M., Cohendet P., Doct. de travail E3I n°2002-4, nov.2002, <http://147.210.86.202/ifrede/e3i/publications/2002/2002-4.pdf>
- [GUITTARD 02], Les Business Models du Logiciel Libre, Guittard C., Working Paper, Laboratoire BETA – ULP Strasbourg, 2002, <http://claudeguittard.free.fr>
- [GUITTARD 04], « Le Forum virtuel, un CyberBA? » , Guittard C., Working Paper, Laboratoire BETA – ULP Strasbourg, 2004
- [HATCHUEL 00], "Quel Horizon pour les sciences de gestion? Vers une théorie de l'action collective", Hatchuel A., in "les Nouvelles Fondations des Sciences de Gestion, éléments d'épistémologie de la recherche en management", David A., Hatchuel A., Laufer R., Collection FNEGE, Vuibert, 2000
- [KAPLAN 01], The Strategy Focused Organization, Harvard Business School Press. Traduction française : Comment utiliser le tableau de bord prospectif pour créer une organisation orientée stratégie, Les Editions d'Organisation, 2001
- [LEBOEUF 04] "PHP5 et My-SQL", Leboeuf J.P., Editions Eyrolles 2004
- [PRINCE 96], « Vers une informatique cognitive dans les organisations, Le rôle central du langage », Prince, V. , Ed. Masson, Paris, 1996
- [RACCAH92], « Expertise et sciences cognitives », Raccach P.-Y., Revue Internationale de Systémique, vol6, n°1-2, pp.159-166, 1992
- [TEULIER 03], "Modélisation des interactions entre normes et savoir-faire: le concept d'artefact organisationnel", Teulier R., Bouche R., Actes IC 2003  
14es journées francophones d'Ingénierie des connaissances, Laval, 1-4 juillet 2003, pp.1-17
- [VYGOTSKY 97], "Pensée et langage", Vygotski, L., Paris: La Dispute, SNEDIT, 1997
- [XFML 02], eXchangeable Faceted Metadata Language, XFML Core Specification, 8 oct.2002 <http://xfml.org>,
- [XTM 01], XML Topic Maps (XTM) 1.0, TopicMaps.Org Specification, 2001. <http://www.topicmaps.org/xtm/1.0>],,

[ZACKLAD 01], "Management des Connaissances, Modèles d'entreprise et applications", Zacklad M., Grundstein M. & al., Editions Hermes Science, 2001

[ZACKLAD 03a], Du Web Cognitivement Sémantique au Web socio sémantique, Zacklad, M., Cahier, J.P., Pétard, X., Journée « Web Sémantique et SHS » du 7 mai 2003, <http://www.lalic.paris4.sorbonne.fr/stic/as5.html>

[ZACKLAD 03b] (2003b) *Communities of Action: a Cognitive and Social Approach to the Design of CSCW Systems*, Zacklad, M. , in Proceedings of GROUP'2003, pp. 190-197, Sanibel Island, Florida, USA.

[ZACKLAD 04] Processus de documentation dans les documents pour l'action (DopA) : statut des annotations et technologies de la coopération associées, in Actes du colloque « le numérique : impact sur le cycle de vie du document pour une analyse interdisciplinaire, 13-15 octobre 2004, Montréal (Québec)