

Hypertopic : une métasémiotique et un protocole pour le Web socio-sémantique

Manuel Zacklad, Jean-Pierre Cahier, Aurélien Bénel, l'Hédi Zaher, Christophe Lejeune, Chao Zhou

► **To cite this version:**

Manuel Zacklad, Jean-Pierre Cahier, Aurélien Bénel, l'Hédi Zaher, Christophe Lejeune, et al.. Hypertopic : une métasémiotique et un protocole pour le Web socio-sémantique. 18eme journées francophones d'ingénierie des connaissances (IC2007) 02 au 06 juillet 2007, Jul 2007, Grenoble, France. pp.217-228. sic_00180793

HAL Id: sic_00180793

https://archivesic.ccsd.cnrs.fr/sic_00180793

Submitted on 21 Oct 2007

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

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

Hypertopic: une métasémiotique et un protocole pour le Web socio-sémantique. Manuel Zacklad, Aurélien Bénel, L'Hédi Zaher, Christophe Lejeune, Jean-Pierre Cahier, Chao Zhou. *Actes des 18eme journées francophones d'ingénierie des connaissances (IC2007)*. Francky Trichet (Eds.). pp. 217-228. Cépaduès. 2007. ISBN 978-2-85428-790-9.

Hypertopic : une métasémiotique et un protocole pour le Web socio-sémantique

Manuel Zacklad¹, Jean-Pierre Cahier¹, Aurélien Bénel¹, L'Hédi Zaher¹, Christophe Lejeune^{2,1}, Chao Zhou¹

Tech-CICO - Institut Charles Delaunay
Université de Technologie de Troyes / CNRS (FRE 2848)
12 rue Marie Curie, 10010 - Troyes Cedex, France
<http://www.tech-cico.fr>
{zacklad, cahier, benel, zaher, chao.zhou}@utt.fr
²CEMAD, 4000 Liège, Belgique
Christophe.Lejeune@ulg.ac.be

Résumé : Nous présentons les catégories de la métasémiotique Hypertopic et le protocole du même nom, que nous considérons être des outils particulièrement adaptés au Web socio-sémantique, en les illustrant par plusieurs exemples d'applications. Nous présenterons des éléments de description XML d'Hypertopic qui a vocation à faire l'objet d'un standard et nous le comparons à certains standards actuels du Web Sémantique.

Mots-clés : Web socio-sémantique, métasémiotique, Hypertopic, ontologies sémiotique, protocole.

1 Introduction

En suivant Bachimont (2004), nous considérerons que l'ingénierie des connaissances (IC) est surtout une ingénierie des supports numériques de connaissances. À ce titre, elle est une partie de l'ingénierie documentaire, si l'on adopte une vision élargie de la notion de document (Pédauque, 2006); ingénierie documentaire qui est elle-même une partie essentielle de l'ingénierie des systèmes d'information (Zacklad, 2007). Toujours en cohérence avec la vision de Bachimont, nous considérerons qu'au sein de l'ingénierie documentaire, la spécialité de l'IC est de concevoir les systèmes d'orientation et d'accès (Zacklad, 2007); ces systèmes contribuent alors à la documentarisation du support des productions sémiotiques pour permettre l'orientation des usagers à l'intérieur des documents numériques et des corpus de documents. De ce point de vue, l'IC peut aussi être considérée comme une part intégrante des sciences de l'information, notamment en tant qu'elles sont confrontées aux questions de catégorisation et de recherche d'information dans un contexte de numérisation (Ertzscheid, 2006). Il va donc de soi que les développements du Web constituent un domaine d'application central pour l'IC.

Dans ce contexte, nous développons une approche originale nommée Web socio-sémantique (W2S) au sein du laboratoire Tech-CICO. Selon cette vision, le web est

appréhendé avant tout comme un instrument de gestion documentaire facilitant des transactions coopératives interpersonnelles éventuellement très asynchrones et distribuées entre des acteurs individuels et collectifs engagés dans des échanges, débats, controverses, relevant de domaines très variés. Selon cette vision du Web, les modalités et les outils de gestion des documents doivent être pour partie conçus par les acteurs engagés dans une coopération active. Parmi ces outils, le Web socio-sémantique préconise le recours à des cartes de thèmes (Cahier, 2005) ou réseaux de description (Bénel, 2003) que l'on peut considérer comme relevant des ontologies sémiotiques (Zacklad, 2005), elles-mêmes proches du niveau des ontologies différentielles définies par B. Bachimont dans la méthodologie Archonte¹.

Pour décrire les ontologies sémiotiques, mais plus largement aussi pour permettre le développement de la recherche ouverte d'information qui combine exploitation d'attributs standardisés, de documents ressources et d'ontologies sémiotiques, le W2S préconise de recourir à une métasémiotique au sens de Hjelmslev, (Badir 2006) dénommée Hypertopic (Cahier *et al.* 2004, Zhou *et al.*, 2006). Une métasémiotique² peut être vue comme une métalangue. Mais celle-ci n'a rien à voir sur un plan épistémologique avec les métalangages de représentation des connaissances fondés sur la logique formelle et sa sémantique. La métasémiotique Hypertopic caractérise des expressions langagières selon leur fonction dans un processus d'organisation de l'information, lui-même lié aux processus transactionnels (on parle d'item, de thème, d'attribut, de document ressource, etc.). La métasémiotique Hypertopic permet de réaliser des *sémiotiques tâche-domaine*³, que l'on peut considérer comme des modèles caractéristiques d'un domaine et d'une tâche, si l'on comprend le terme de modèle dans le sens d'une schématisation langagière et non dans celui de la logique formelle.

La métasémiotique Hypertopic, est opérationnalisée à l'aide de XML (cf. infra) et peut être exploitée à l'aide d'un protocole informatique du même nom aujourd'hui au stade de V.1. Ce protocole permet de naviguer dans les ressources documentaires selon une logique hypertextuelle et facilite la recherche ouverte d'information au sein des corpus. En assistant la navigation, le protocole Hypertopic permet de réaliser des associations entre thèmes, ressources documentaires et attributs standardisés, dont on peut considérer qu'elles relèvent d'inférences pragmatiques correspondant à l'exploitation d'association heuristiques entre divers éléments. Mais, à nouveau, ces inférences n'ont rien de commun avec les inférences de la logique mathématique et la sémantique formelle sur laquelle elle s'appuie.

¹ Cependant, à la différence des préconisations de la méthodologie Archonte, le W2S ne considère pas les ontologies sémiotiques comme une étape préalable à la constitution systématique d'ontologies référentielles et computationnelles basées sur des langages formels et ne s'appuie pas sur la notion de concept dans la tradition des langages de représentation des connaissances de l'intelligence artificielle.

² Selon (Badir 2006) « *Les métasémiotiques prennent leur source dans des sémiotiques dotées, pour ainsi dire, d'un plan de contrôle. Grâce à ce plan, chaque élément du contenu trouve à s'adjoindre une expression sous un rapport de dénomination. C'est ce que l'on fait lorsqu'on dit, dans un exemple célèbre exploité par Roland Barthes, que dans telle publicité pour des pâtes françaises les teintes jaunes et vertes sur fond rouge (les couleurs du drapeau italien) signifient l'italianité (Barthes, 1982 : 26). Italianité est une expression métasémiotique pour désigner la signification d'éléments visuels (de couleurs).* ».

³ Chez Hjelmslev, les métasémiotiques connotent des sémiotiques dénotatives, les langues. La dénotation n'a pas de portée référentielle directe puisqu'elle vise d'abord un *contenu*, un signifié.

Dans cet article nous présenterons la métasémiotique et le protocole Hypertopic que nous considérons être des outils particulièrement adaptés au Web socio-sémantique. Nous présenterons les catégories de la métasémiotique et plusieurs exemples d'applications déjà réalisées par les membres de l'équipe. Nous présenterons ensuite des éléments de description en XML d'Hypertopic qui a vocation à faire l'objet d'un standard. Nous concluons par une discussion justifiant l'intérêt d'Hypertopic par rapport à certains standards actuels du Web Sémantique. Chacune des parties pourrait donner lieu à un ou plusieurs articles séparés qui sont, pour certains, en cours de rédaction : (1) métasémiotique Hypertopic, (2) applications sur un plan méthodologique, (3) fonctionnalités, choix d'implémentation et d'architecture du protocole, de la base de donnée et des clients, (4) discussion épistémologique. Notre choix a été volontairement de donner une vision synoptique de ces développements en cours qui permettra, nous l'espérons, de saisir la cohérence d'ensemble de notre projet. Certaines parties pourront sembler parfois un peu lacunaires, le lecteur intéressé par les approfondissements étant renvoyé à la bibliographie.

2 Catégories de la métasémiotique Hypertopic

Hypertopic est directement conçu pour faciliter des activités d'enquête (Dewey, 1938) mises en œuvre par les membres d'une communauté ou d'une organisation impliqués dans des tâches récurrentes et pour partie communes médiatisées par le Web socio-sémantique (Cahier, 2005). Ces activités d'enquête peuvent être catégorisées en référence aux tâches décrites par l'ingénierie des connaissances (Weilinga et al., 1992) : classification, diagnostic, conception, réparation, supervision, etc. Dès que ces tâches ont un certain niveau de complexité, de récurrence et de durabilité, il est important de mettre en œuvre des stratégies de gestion des connaissances (Dieng-Kuntz et al., 2001) (Zacklad & Grundstein, 2001) (Ermine, 2003), facilitant la transférabilité du connaissant et du connu vers d'autres acteurs (Zacklad, 2006).

Selon (Zacklad, 2006), différentes formes de tacitivité appellent différentes stratégies de gestion des connaissances. Quand la tacitivité est lié à la volatilité des connaissances (Zacklad, 2006), la stratégie de documentation, propre aux sciences de l'information et la stratégie de documentarisation qui la prolonge et qui relève plus spécifiquement de l'ingénierie des connaissances, sont indiquées. La documentarisation consiste à « *transcrire ou à enregistrer les contenus sémiotiques sur des supports pérennes puis à doter ces supports d'attributs spécifiques permettant (i) de faciliter leur gestion parmi d'autres supports, (ii) de faciliter leur manipulation physique, condition d'une navigation sémantique à l'intérieur du contenu sémiotique et enfin, (iii) de faciliter l'orientation des récepteurs, mais également de plus en plus des réalisateurs eux-mêmes, à l'intérieur du support en définissant une ou plusieurs cartographies des contenus sémiotiques susceptibles de guider la navigation sémantique.* » (Zacklad, 2007).

Un modèle Hypertopic résulte de l'application de la métasémiotique à un domaine et à une tâche pour documentariser des contenus sémiotiques afin de mettre en œuvre des raisonnements relevant de la classification heuristique (Clancey, 1985). La classification heuristique correspond pour nous à un raisonnement d'un très haut niveau de généralité applicable aux différents types de tâches précédemment citées. Vue à travers le prisme de la classification heuristique, l'enquête consiste à caractériser une situation ou un artefact qu'il s'agit d'analyser ou de spécifier en référence à d'autres situations ou artefacts analogues. Ces situations ou artefacts, dont la caractérisation est le résultat de l'enquête, correspondent à la catégorie *d'item*⁴. L'expression d'*item*⁵ vise notamment à souligner que la situation ou l'artefact étudié fait lui-même partie d'une collection et que l'analyse des similarités entre les éléments de cette collection est essentielle au raisonnement.

Table 1. Catégories utilisées dans la métasémiotique Hypertopic.

Catégorie principale	Catégorie associée
Item	
Point de Vue	Acteur porteur du point de vue (individuel ou collectif)
Attribut standard	Valeur (de l'attribut)
Ressource	Fragment (de la ressource)
Thème (ou attribut heuristique)	Thèmes parents (super et sous ordonnés)

Il existe deux grandes étapes itératives dans le cycle de vie d'un modèle Hypertopic. La première consiste à enrichir le modèle en caractérisant de nouveaux items qui ne figurent pas encore dans le modèle. La seconde consiste à rechercher des items déjà en partie caractérisés pour résoudre de nouveaux problèmes.

L'enrichissement du modèle implique trois activités. La première est la spécification référentielle de l'item qui est réalisée à l'aide de la catégorie *attribut standard*. Les attributs standards sont des descriptions conventionnelles et standardisées des situations et des artefacts qui s'appuient souvent sur leurs caractéristiques mesurables.

La seconde est la documentation de l'item, réalisée à l'aide de la catégorie *ressource*, qui consiste, quand cela est possible, à associer à l'item des documents ressources ou des fragments de ces documents qui en fournissent une description⁶ : fichier texte, photo, son, film, etc. La troisième activité est la thématization heuristique de l'item, réalisée à l'aide de la catégorie *thème* (ou topic), qui correspond à la caractérisation de l'artefact ou de la situation considérée en fonction de compétences métiers, d'hypothèses théoriques ou d'intérêts particuliers à l'aide d'expressions

⁴ Nommée entité dans les précédentes descriptions d'Hypertopic.

⁵ Nous sommes parfois tentés d'assimiler l'item à un « concept sémiotique » correspondant à la situation ou à l'artefact (concept sémiotique par opposition aux approches référentielles du concept). Cette discussion sort du cadre de cet article.

⁶ Dans le cas particulier où l'item considéré est lui-même un document numérique (par exemple, un livre électronique), la catégorie ressource permet d'accéder directement à un représentant de l'item (une version particulière de ce livre) ce qui n'est pas possible si l'item est un artefact matériel (par exemple, un voyage).

linguistiques si possible synthétiques et précises. Les thèmes sont organisés selon une arborescence qui traduit des relations de plus ou moins grande généralité (versus spécialisation) de l'expression dans le réseau sémantique ainsi constitué selon un point de vue donné (les thèmes n'étant pas des concepts ou des classes, la notion d'héritage ne s'applique pas).

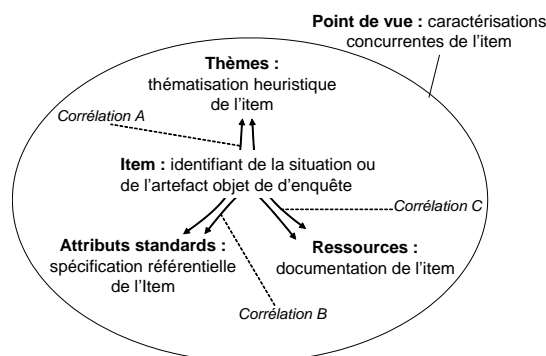


Fig. 1 – Catégories Hypertopic et types de corrélation

Ces trois activités et plus particulièrement celle de la thématisation heuristique, s'appuient sur une cinquième catégorie essentielle de la métasémiotique Hypertopic, celle de *point de vue*⁷. C'est la diversité des formes de caractérisation de l'item selon les communautés professionnelles, d'intérêt, de pratiques ou les hypothèses en cours d'examen qui justifie le recours à plusieurs thématisations potentiellement concurrentes de l'item. C'est également la multiplicité des points de vue qui structure le raisonnement heuristique en permettant selon les tâches de faire varier les caractérisations de l'item⁸. Dans les implémentations actuelles d'Hypertopic, les points de vue s'appliquent surtout à des thématisations différentes de l'item. Il pourrait être pertinent que les points de vue sélectionnent également certaines formes de documentation, voir des spécifications référentielles différentes⁹.

La Figure 1 schématise les différentes formes de caractérisation de l'item selon Hypertopic. Elle met également en évidence différentes formes de relations entre ces catégories qui correspondent à des grandes familles d'applications informatiques. Nous allons les passer brièvement en revue, en signalant que dans certaines de ces applications l'item lui-même peut ne pas être décrit à l'intérieur du système. Dans le cas de la corrélation A, les applications mettent en relation la thématisation heuristique d'une situation et sa spécification référentielle qui permet de l'identifier

⁷ Nous renvoyons à Bénel (2003) et Cahier (2005) pour une discussion approfondie de cette catégorie.

⁸ Dans les exemples canoniques du raisonnement diagnostique en médecine, un problème médical peut être décrit selon le point de vue des symptômes (point de vue des patients) ou selon celui des processus physiopathologiques en cause (point de vue des médecins).

⁹ Nous pensons, par exemple, à des guides de maintenance faisant intervenir différents métiers qui ont chacun un point de vue particulier sur un ouvrage d'art qui se traduit également par des fragments documentaires et des attributs standards spécifiques.

sans recourir à une documentation. Cette corrélation a été exploitée dans de nombreux systèmes experts qui souvent ne faisaient pas référence explicitement aux items (les cas) d'où les caractérisations heuristiques et référentielles étaient issues. Le raisonnement effectuait des corrélations entre les propriétés mesurables d'un nouveau cas en cours d'examen et sa caractérisation heuristique. Dans la corrélation B, les ressources documentaires et les attributs standards sont mis en relation sans recours à une thématization. Ce cas correspond à de nombreux systèmes de gestion documentaire dans lesquels la recherche ne fait pas appel à des descriptions heuristiques selon divers points de vue des ouvrages. Dans la corrélation C, la thématization heuristique est directement mise en relation avec une documentation sans relation avec la caractérisation référentielle des items évoqués par celle-ci. Cela correspond à des applications d'aide à la lecture par le biais d'annotation plus ou moins libres, voire à des applications de recherche d'information dans lesquelles les thèmes permettent d'accéder à des éléments de corpus sans autre référence aux situations visés par la documentation. Le principe de la recherche ouverte d'information (Zaher et al., 2006a), qui correspond aux phases d'exploitation des modèles Hypertopic, est de combiner l'ensemble des formes de caractérisation de l'item et des corrélations auxquelles elles donnent lieu.

3 Exemples d'applications Hypertopic

La Première application en date d'Hypertopic, « Agora-FT » a été développée en 2002 par Tech-CICO pour France Télécom sous la forme d'un prototype opérationnel centrée sur l'item « projet R&D » (Cahier et al., 2005). Le projet de R&D est associé à une documentation (*ressource*) et thématisé selon sept *points de vue* correspondant aux grands métiers de l'entreprise utilisant un vocabulaire distinct. L'objectif de l'application est de faciliter la promotion de ces projets et l'exploitation de leurs résultats¹⁰.

Également réalisée à partir du logiciel Agorae, la cartographie « DKN SeqXAM »¹¹ (Zaher et al., 2006b) a été construite en collaboration entre Tech-CICO et le LIMSI-CNRS, dans le cadre du projet DKN (Diaspora Knowledge Network) soutenu par l'UNESCO. La carte DKN SeqXAM a été construite selon une méthode de « co-conception controversée » (ou « co-construction conflictuelle »). La méthode permet la construction simultanée de plusieurs *points de vue* individuels et leur juxtaposition dans une même ontologie sémiotique, permettant alors la comparaison et la discussion. Les items sont des articles scientifiques caractérisés différemment par cinq biologistes. Les *thèmes* permettent d'accéder à la ressource documentaire en texte intégral et également aux fragments du corpus employant des termes du thème. Tirant parti d'autres travaux effectués au LIMSI-CNRS (Ferey, 2007) le système

¹⁰ La deuxième application industrielle a été une maquette de gestion des compétences réalisée en collaboration avec EADS et présentée dans (Cahier et al 2005).

¹¹ Cette cartographie et la méthode de « co-construction conflictuelle » font l'objet d'une démonstration et d'explications sur le site <http://www.dkn.tech-cico.fr/>

permet ainsi la comparaison de classifications réalisées par plusieurs experts sur un ensemble de termes identiques.

Également dans le domaine de la recherche scientifique, un travail d'analyse de vases est réalisé depuis 2003 par une équipe d'historiens d'art (laboratoire du CRATA) avec Porphyry¹², logiciel d'annotation désormais compatible avec Hypertopic. L'étude porte sur 300 vases de la région de Paestum (en Campanie) représentant des scènes de banquet ou liées à Dionysos. Chaque vase est documenté par des photographies présentant son profil ou les scènes figurées. Au printemps 2005, trois étudiantes ont contribué à l'étude iconographique de ces vases, dans le cadre de leur maîtrise en sciences de l'antiquité (Bénel, 2006).

Le serveur Hypertopic Cassandre accompagne les chercheurs en sciences humaines dans la dialectique interprétative de construction de catégories d'analyse (des thèmes semi-automatiques) congruentes avec leur matériau (les ressources, consistant en entretiens approfondis, en documents institutionnels ou en forums de discussion). Les études pilotes (actuellement en cours) portent sur la mobilisation autour des logiciels libres (Epicum & Lejeune, 2005), l'analyse de la diffusion institutionnelle de la notion de « société civile » (Pirotte, 2007) et l'analyse pragmatique des actes de langages sur le forum de discussion d'un quotidien français.

Table 2. Exemples d'applications Hypertopic

Agorae FT (Cahier 2005)	Porphyry Paestum (Bénel, 2006)	DKN-SEQXAM (Zaher, 2006b)	Cassandre (Epicum & Lejeune 2005)
Item Projet de R&D dans le domaine des télécoms	Item Vase /Scène figurée /Élément figuratif	Item Article scientifique	Item Entretien sociologique
Points de vue technologique; d'usage; de l'offre commerciale; du client ; etc.	Points de vue Projets de recherche en archéologie et en histoire de l'art	Points de vue Ceux des différents membres du projet	Points de vue Analyse sociologique des facteurs de mobilisation
Attribut standard Date de début; statut du projet; etc.	Attribut standard Musée ou collection privée, Numéro d'inventaire	Attribut standard Attributs bibliographiques standard	Attribut standard Nom de l'acteur, date et lieu de l'entretien
Thèmes IPV6; Services Web; voyages; santé; économies de bande passante, etc. (1500 thèmes)	Thèmes Analyse de la forme du vase, des scènes figurées, de la technique de fabrication...	Thèmes bacteria; bordetella spp; genome sequence; compartiment, etc. (400 thèmes)	Thèmes défi technique, distinction sociale, engagement politique, socialisation...
Ressources Sites Web ou Pièces Jointes descriptifs du détail des projets	Ressources Un dossier photo par vase, une photo par scène figurée, un fragment photo par élément figuratif	Ressources Textes des articles, et fragments de ces textes	Ressources Passages (textuel) de l'entretien

¹² Logiciel Porphyry (Réseau ARTCADHi) : <<http://www.porphyry.org/prototypes/expert/>>.

4 Implémentation XML d'Hypertopic, protocole HT V.1.0 et exemple d'application

La métasémiotique Hypertopic a donné lieu à une description en XML et au développement open source de deux serveurs (Argos et Cassandre dédié à l'analyse de textes), de clients et d'un protocole Hypertopic (HT v.1.0) permettant l'accès aux serveurs. Le serveur Hypertopic exploite l'approche REST de conception de services Web (Fielding, 2002). Conformément à cette approche, chaque nœud d'un graphe Hypertopic est identifié par un URI et matérialisé par un flux XML pouvant être consulté, créé, modifié et supprimé en adressant au serveur les méthodes HTTP correspondantes (*get*, *post*, *put* et *delete*).

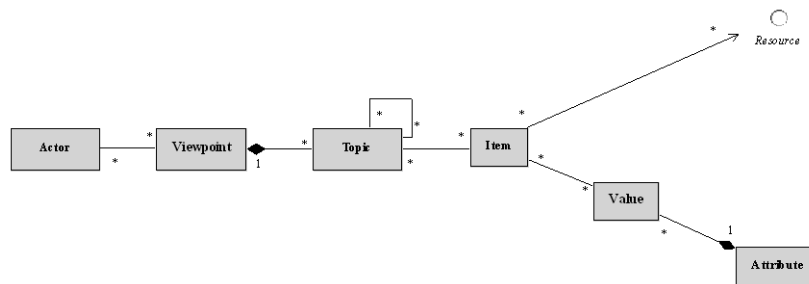


Fig. 2 – Les trois branches de description de l’item dans la métasémiotique Hypertopic (Diagramme de classe UML)

Dans l’architecture que nous utilisons aujourd’hui, les items, les points de vue, les thèmes et les attributs standards sont stockés par le serveur d’une manière transparente dans une base de données PostgreSQL structurée selon le modèle UML présenté dans la Figure 2¹³ et les documents dans un serveur WebDAV. Deux clients sont aujourd’hui disponibles pour l’enrichissement du modèle et son exploitation et font l’objet d’améliorations constantes : Agoræ développé en PHP et Porphyry en Java. Dans les figures qui suivent nous présentons la description XML sur laquelle s’appuie le protocole HT v.1.0 et une copie d’écran du client Agoræ.

```

element item {
  (element topic {attribute href {xsd:anyURI}, text}
  |element \attribute {attribute name {text},attribute value {text}, attribute href {xsd:anyURI}}
  |element resource {attribute name {text},attribute href {xsd:anyURI}}
  ) *
}

```

Fig. 3 – Spécification XML de l’item (Syntaxe : Relax NG compact [Clark, 2002]).

¹³La modélisation UML de la métasémiotique a bien comme c’est normal une portée référentielle puisqu’il s’agit de permettre de gérer (au sens administratif) des listes d’expressions. Mais quand les expressions sont des thèmes, elles n’ont pas de portée référentielle directe. « Emission de variété », renvoie de manière non univoque à un signifié (qui varie selon les thèmes voisins), signifié qui peut renvoyer lui-même de manière non univoque à des référents.

La figure 4 présente l'application CogDoc qui vise à partager un corpus de ressources pédagogiques élaborés par des enseignants en Sciences de l'Information et en Ingénierie des Connaissances. Les points de vue sont notamment ceux des cursus offerts par les différents établissements d'enseignement, ici celui de l'UTT.

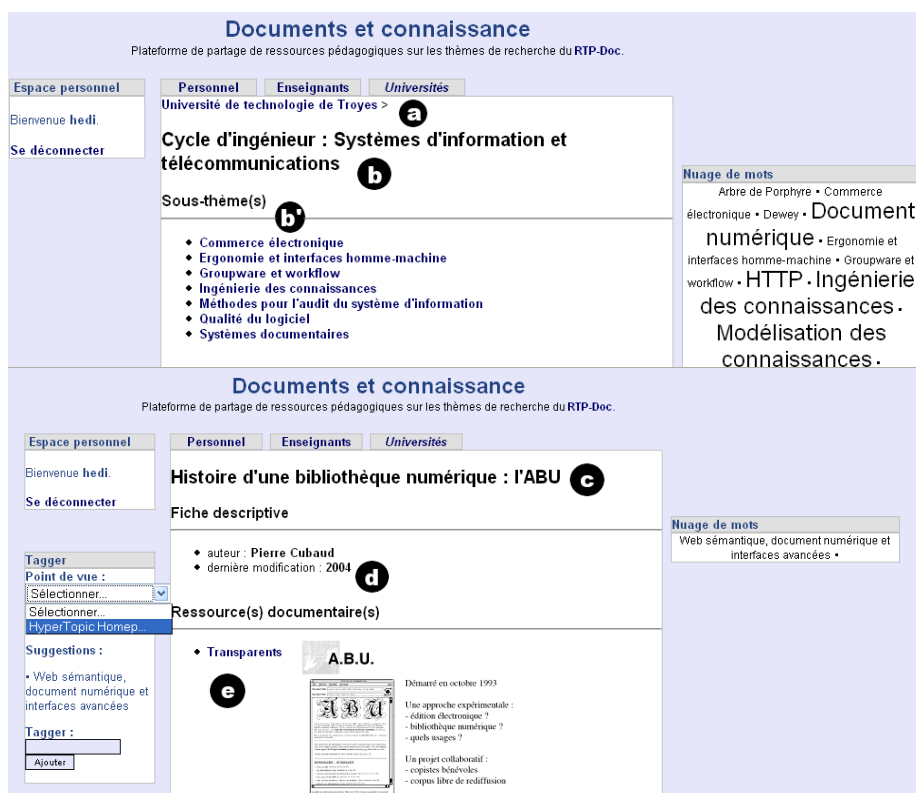


Fig. 4 – Le client HyperTopic Agorà : exploration selon le point de vue des cursus de l'UTT (a) *point de vue* (un établissement) (b) *thème* (un cursus) (b') *sous-thèmes* (des noms d'UV et leurs problématiques), (c) *item* (un cours), (d) *attribut standard/valeur* (auteur, dernière modification...) (e) *ressource documentaire* (ici des transparents). Les « Tags » permettent aux utilisateurs de proposer d'autres sous-thèmes selon une méthode « ascendante ». Ceux-ci peuvent ensuite être intégrés aux ontologies sémiotiques associées aux points de vue.

5 Discussion : Hypertopic versus le cake du Web Sémantique

Un des points les plus ouverts à la discussion est la relation entre la métasémiotique Hypertopic et les standards actuels du Web Sémantique. La proposition d'un nouveau standard semble toujours une entreprise périlleuse tant la raison d'être d'un standard est précisément sa généralisation à un nombre toujours plus grand d'utilisateurs. Pour

notre « défense », notre argumentation peut être organisée en deux points. Premièrement, soulignons le fait que nous nous inscrivons très largement dans les standards actuels du Web promus par le W3C, que cela soit par le choix des technologies XML ou par celui du recours au protocole HTTP (même si le W3C préfère SOAP à REST pour l'écriture des services Web).

Pourquoi ne remontons nous pas alors les couches du cake du Web Sémantique puisque notre objectif est précisément de faciliter l'orientation des usagers en explicitant certaines relations sémantiques qui caractérisent les items partagés sur le Web ? La raison, qui constitue notre deuxième argument, est simple : la remontée vers les couches hautes du cake de Tim Berners Lee s'appuie sur la sémantique formelle qui ne nous semble pouvoir exprimer qu'une petite partie des contenus sémiotiques véhiculés, notamment quand ceux-ci dépendent de facteurs culturels et/ou font l'objet de débats. Notons bien qu'il ne s'agit pas pour nous de nier la pertinence d'une représentation formelle et consensuelle (dans une aire culturelle donnée) de certaines dimensions conventionnelles ou mesurables des artefacts échangés ou coproduits dans les transactions médiatisées : références administratives des dates, des lieux, des dénominations de personnes, des prix, nomenclatures anatomiques, dimensions et références de pièces mécaniques ou d'éléments chimiques, etc.

Ces facettes des artefacts, qui visent précisément à faciliter les usages référentiels, ont tout à gagner à être représentées à l'aide des attributs standards qui pourraient être eux-mêmes représentés à l'aide d'ontologies justement nommées référentielles par (Bachimont, 2004), prolongées par les ontologies computationnelles qui utilisent les langages orientés objet et logiques préconisés dans les couches hautes du Web Sémantique. Mais ces dimensions formelles et consensuelles sont loin d'être suffisantes dès lors que les transactions ont une certaine complexité ; complexité qui correspond à celle des activités coopératives dans lesquelles il est nécessaire en contexte de proposer des modélisations ad hoc des situations transactionnelles qui sont souvent cognitivement et socialement conflictuelles (Zaher, 2006b). Dans ces situations, il est nécessaire d'utiliser d'autres approches que celles des langages référentiels et de recourir plutôt aux approches rhétorique/herméneutique (Rastier, 2003) qui nous semblent compatibles avec les logiques de l'enquête de la philosophie pragmatique d'inspiration juridique (Dewey, 1938).

Certain auteurs – (Bachimont, 2004), mais également (Gandon, 2006) dans la discussion du W2S¹⁴ – considèrent que les approches rhétorico-herméneutiques constituent une phase initiale de la modélisation qui doit être prolongée par une formalisation référentielle. Ils considèrent que les programmes informatiques doivent fonctionner sur la base de cette logique référentielle mais que leurs usagers interprètent les résultats de leurs calculs selon une perspective rhétorico-herméneutique. Nous ne souscrivons pas entièrement à cette vision. Pour nous, ces deux approches de la modélisation sont complémentaires mais distinctes et guident toutes deux des implémentations informatiques dont les principes de fonctionnement

¹⁴ Par manque de place nous ne pouvons poursuivre ici la discussion avec F. Gandon (2006) sur l'utilisation de RDF sans remontée dans les « couches hautes » du cake. Il est possible d'exprimer Hypertopic en RDF mais l'écriture des descriptions en est sensiblement alourdie. L'opportunité de cette traduction est une question conjoncturelle qui nous semble devoir se discuter au cas par cas.

diffèrent. Il est également possible d'utiliser la puissance des ordinateurs pour réaliser des associations heuristiques qui ne s'appuient pas sur une sémantique référentielle. Ces associations heuristiques, peuvent être basées sur des calculs statistiques, comme le font certains moteurs de recherche, mais peuvent également être basés sur une vision contextuelle, situationnelle et pragmatique du sens (Zacklad, 2005), proche de l'approche rhétorico-herméneutique, comme nous le préconisons dans les approches à base d'ontologies sémiotiques (annuaires, thésaurus, cartes de thème, réseau de description, etc.). Ces approches, qui sont également au cœur du paradigme hypertextuel, nous semblent devoir faire l'objet de recherches pluridisciplinaires actives au sein de l'IC.

6 Références

- BACHIMONT B.. Arts et sciences du numérique : Ingénierie des connaissances et critique de la raison computationnelle. Mémoire d'Habilitation à Diriger des Recherches, Université de Technologie de Compiègne, 12 janvier 2004.
- BADIR S. (2006), « La hiérarchie sémiotique », dans Louis Hébert (ed), *Signo*, Rimouski (Québec). Disponible sur : <http://www.signosemio.com>.
- BENEL, A. (2003). Consultation assistée par ordinateur de la documentation en Sciences Humaines : Considérations épistémologiques, solutions opératoires et applications à l'archéologie, Thèse de doctorat en informatique, INSA de Lyon.
- BENEL A. (2006). Porphyry au pays des paestans : usages d'un outil d'analyse qualitative de documents par des étudiantes de maîtrise en iconographie grecque, In: *Actes du colloque international "Corpus en Lettres et Sciences sociales : des documents numériques à l'interprétation", Albi, juillet 2006. Texto!, juin 2006, vol. XI, n°2*. p.182-189. Disponible sur : http://benel.tech-cico.fr/publi/benel_TEXTO_06.pdf
- CAHIER, J.P., ZACKLAD, M., MONCEAUX, A. (2004) Une application du Web socio-sémantique à la définition d'un annuaire métier en ingénierie, In N. Matta (ed), actes des journées francophone d'Ingénierie des Connaissances (IC 2004), PUG.
- CAHIER J.-P, (2005). Ontologies sémiotique pour le Web socio sémantique: étude de la gestion coopérative des connaissances avec des cartes hypertopiques. Thèse en informatique, Université de Technologie de Troyes.
- CLANCEY W., Heuristic Classification., *Artificial Intelligence Journal*, 27, pp. 289-350, 1985.
- CLARK J., RELAX NG Compact Syntax, Commitee specification, Organization for the Advancement of Structured Information Standards [OASIS], 2002. Disponible sur : <http://www.oasis-open.org/committees/relax-ng/compact-20021121.html>
- DEWEY J. "Logique, la théorie de l'enquête", trad. Gerard Deledalle, col. L'interrogation philosophique, PUF, 1938 (re-éd.1993).
- DIENG-KUNTZ, R., CORBY, O., GANDON, F., GIBOIN, A., GOLEBIOWSKA, J., MATTA, N., et al. (2001). Méthodes et outils pour la gestion des connaissances : Une approche pluridisciplinaire du Knowledge Management (2ème édition): Dunod.
- ERMINE, J. L. (2003). La gestion des connaissances: Hermès science publications.
- ERTZSCHEID, O. (2006). Indexation sociale et continents documentaires. Document numérique et société (DocSoc), Fribourg - Suisse.

- FEREY N., TOFFANO-NIOCHE C., MATTE-TAILLEZ O., GHERBI R., TURNER W., "REVIVOS: REAlité VIRTuelle pour la construction de VOcabulaires Spécialisés", 18ème journées francophones de l'ingénierie des connaissances IC 2007.
- FIELDING R. (2002). Architectural Styles and the Design of Network-based Software Architectures, PhD thesis, University of California, Irvine, 2002. Disponible sur : <<http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm>>
- GANDON F. (2006). Le Web sémantique n'est pas anti-social », actes des 17ème journées francophones de l'ingénierie des connaissances, IC 2006, In Mounira Harzallah, Jean Charlet, Nathalie Aussenac-Gilles (ed), actes de la semaine de la connaissance SdC'2006, vol. 1 , pp. 131-140..
- LEJEUNE C. & ERPICUM M. (2005). Résister sans s'engager : logiciels libres et politique, Journée d'étude "Engagement politique et social en Belgique. De nouvelles dynamiques à l'oeuvre?", Association Belge de Science Politique.
- PIROTTE G., (2007). La notion de société civile, Paris, La découverte (collection "Repères").
- PEDAQUE, R. T. (2006) Le document à la lumière du numérique, Caen : C&F Editions.
- RASTIER, F.. De la signification au sens. Pour une sémiotique sans ontologie. *Texto !* juin-sept. 2003 [en ligne]. Disponible sur : <http://www.revue-texto.net/Inedits/Rastier/Rastier_Semiotique-ontologie.html>.
- ROUSSEAU F. (2005). Par delà les Connaissances inventées par les informaticiens : les Collections ? *Intellectica* 2005/2-3, n° 41-42 pp. 225-252.
- WIELINGA, B., SCHREIBER, G. & BREUKER, J. (1992). KADS: A modeling approach to knowledge engineering. *Knowledge Acquisition 4, Special Issue on KADS*.
- ZACKLAD, M., GRUNDSTEIN, M., eds. (2001). Management des connaissances, Stanmore : Hermes Science Europe LTD.
- ZACKLAD, M. (2005a). Introduction aux ontologies sémiotiques dans le Web Socio Sémantique. In M.-C. JAULENT (ed) actes des 16èmes journées francophones d'Ingénierie des Connaissances, Grenoble: PUG.
- ZACKLAD, M. (2006). Gestion du connaissant et du connu dans la théorie transactionnelle de l'action (TTA), Document de travail (Working paper) disponible sur : <http://archivesic.ccsd.cnrs.fr/>
- ZACKLAD, M. (2007). Processus de documentarisation dans les Documents pour l'Action (DopA). *Babel - edit -, Le numérique : impact sur le cycle de vie du document*. ENSSIB - février 2007 [en ligne] <http://babel.enssib.fr/document.php?id=727>
- ZAHER L. H., CAHIER J.-P. & ZACKLAD M. (2006a). Information retrieval and e-service: Towards open information retrieval. In *Proceedings of International Conference on Service Systems and Service Management*, (IC SSSM'2006). pp. 41-46, IEEE, ISBN 1-4244-0451-7.
- ZAHER, L. H., CAHIER J.-P., TURNER, W. ZACKLAD M. (2006b). A conflictual co-building method with Agoræ, *Proceedings of COOP'06 Workshop on Knowledge Interaction and Knowledge Management*. 9-12 may 2006, Carry-le-Rouet, France, 2006.
- ZHOU C., LEJEUNE C. & BÉNEL A. (2006). Towards a standard protocol for community-driven organizations of knowledge, *Proceedings of the thirteenth international conference on Concurrent Engineering*, Frontiers in Artificial Intelligence and Applications, vol. 143. Amsterdam: IOS Press, 2006. p.438-449. Disponible sur : http://benel.techcico.fr/publi/zhou_CE_06.pdf