

Vers le Web Socio Sémantique : introduction aux ontologies sémiotiques

Manuel Zacklad

► **To cite this version:**

Manuel Zacklad. Vers le Web Socio Sémantique : introduction aux ontologies sémiotiques. Feb 2005, 2005. <sic_00001347>

HAL Id: sic_00001347

https://archivesic.ccsd.cnrs.fr/sic_00001347

Submitted on 20 Feb 2005

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.

Vers le Web Socio Sémantique : introduction aux ontologies sémiotiques

Manuel Zacklad

Manuel.Zacklad@utt.fr
Université de Technologie de Troyes
Equipe Tech-CICO / ISTIT FRE CNRS 2732

Résumé :

Dans cette communication, nous présentons le positionnement du Web Socio Sémantique (W2S) par rapport au Web Sémantique (WS) et nous discutons en particulier la problématique de l'ingénierie des ontologies sémiotiques. Préalablement nous introduisons un cadre théorique pour l'analyse des activités collectives distribuées, la théorie des *transactions communicationnelles*. Après avoir défini la notion de *production sémiotique* nous introduisons les stratégies permettant une *distribution spatio-temporelle* des transactions. Parmi ces stratégies, nous étudions plus en détail la stratégie de *formalisation de l'expression* qui permet de construire des ontologies sémiotiques dans les communautés distribuées qui utilisent des *Document pour l'Action* (DopA) au sein du W2S. Nous présentons différents critères pour l'établissement des *accords définitionnels* dans la construction des ontologies issus des approches *logique*, *contextuelle* et *situationnelle* (ou pragmatique). En conclusion, nous proposons une hypothèse expliquant l'effectivité partielle du WS autrement que par le recours à une approche logique de la signification. Selon nous, si les systèmes automatisés sont efficaces dans les transactions fermées c'est parce qu'ils recourent à une *formalité machinale* basée sur des conventions sociales qui ferment le potentiel sémiotique des signes par opposition à la *formalité de la signification* qui l'amplifie.

Mots-clefs : Web Socio Sémantique, Web Sémantique, Ontologies, Ontologies Sémiotiques, Coopération, Transactions Communicationnelles, Document

Introduction

Les infrastructures du Web dont l'importance ne cesse de croître répondent à la fois à des besoins de communication et de partage d'information. Relevant du premier besoin, on trouve des applications variées allant de la messagerie à la visio-conférence en passant par le chat et la téléphonie. Relevant du second besoin on trouve des applications elles aussi très diverses allant des sites Web statiques aux bibliothèques numériques de grandes tailles en passant par les weblogs et les sites interactifs de toutes natures offrant des accès à de nombreuses applications administratives (au sens des « legacy system » anglo-saxons). Dans cette typologie, les forums constituent certainement l'intermédiaire le plus fameux entre communication à distance et mise à disposition publique du corpus constitué par des interactions scripturales.

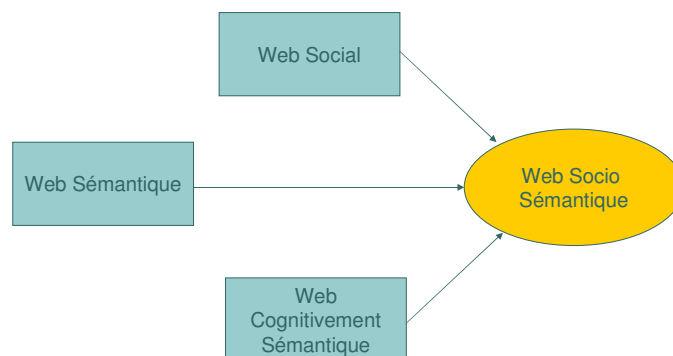
Portées par des ambitions d'ingénierie, différentes visions du Web ont émergé avec un succès plus ou moins grand ces dernières années. La plus célèbre est celle du Web Sémantique bien connue en ingénierie des connaissances et résolument tournée vers les besoins informationnels. L'objectif est de proposer une formalisation logique d'un certain nombre de contenus ou de systèmes d'index de ces contenus permettant à des agents logiciels de répondre automatiquement aux requêtes complexes de leurs utilisateurs. Moins structuré, mais correspondant à un nombre très important d'applications, le Web social, parfois associé au

« social computing » s'adresse aux besoins de communication en visant des usages associés à la conduite d'interactions éphémères entre utilisateurs distants tout en offrant des représentations, souvent de nature graphique, des réseaux sociaux ainsi constitués.

La troisième vision du Web, dans laquelle nous nous inscrivons, est celle du Web Socio Sémantique (W2S) (Zacklad 2003c, Cahier et al. 2003) une approche qui combine les deux précédentes en postulant l'existence d'une co-détermination des usages informationnels et communicationnels du Web. Cette vision est principalement promue actuellement par le l'équipe Tech-CICO à l'Université de Technologie de Troyes, une équipe spécialisée dans l'étude des usages coopératifs des TIC, en relation avec d'autres communautés comme le groupe « Pratiques Collectives Distribuées » dont l'animateur, Bill Turner, est au LIMSI (CNRS). Le Web Socio Sémantique prolonge pour partie les première critiques que nous avons formulées à l'égard de certaines attitudes hégémoniques du courant du Web Sémantique, en proposant l'approche alternative du Web Cognitivement Sémantique (Caussanel et al. 2002).

Alors que le *Web Sémantique* est surtout destiné à l'inférence automatique d'agent logiciels portant sur des contenus stables et assez simples pour être décrits en logique formelle, le *Web Cognitivement Sémantique*, vise à soutenir les activités de recherche d'utilisateurs humains dans des corpus complexes et évolutifs. Tout en prolongeant cette approche, le *Web Socio Sémantique* (W2S) inscrit les pratiques de recherche et d'élaboration informationnelle des usagers du Web dans des activités de « coopération structurellement ouvertes » (Zacklad 2003a) qui reposent également sur des pratiques communicationnelles intensives. Les collectifs visés sont de taille très variée comme l'est leur degré de structuration qui va de communautés de pratiques informelles à des projets sous pression temporelle en passant par des communautés d'action. Leur caractéristique commune principale est qu'une partie de leur activité est « distribuée », spatialement, temporellement voire socialement ce qui justifie leur usage d'un système d'information et de communication (cf. infra).

Dans cet article nous présenterons un certain nombre de directions de recherche associées à l'ingénierie du W2S. Nous commencerons par exposer rapidement un cadre théorique permettant l'analyse de l'activité collective distribuée qui donne une place essentielle aux *productions sémiotiques* des acteurs dans la médiation de l'activité. Ce cadre théorique nous permettra de préciser le rôle des ontologies sémiotiques qui sont une des composantes principale de l'ingénierie du W2S. Nous présentons différents critères pour l'établissement des *accords définitionnels* dans la construction des ontologies issus des approches *logique*, *contextuelle* et *situationnelle* (ou pragmatique). Enfin, nous suggérerons en conclusion une hypothèse expliquant l'effectivité partielle du WS autrement que par le recours à une approche logique de la signification.



Un cadre théorique pour l'analyse des activités collectives distribuées

Les transactions communicationnelles symboliques

Dans la théorie psycho-socio-économique d'orientation pragmatique que nous sommes entrain de développer, l'activité est toujours envisagée de manière transactionnelle, que les transactions aient lieu entre des « personnes » distinctes ou entre la même personne engagée avec elle-même dans un dialogue intérieur prolongé, le cas échéant, par des gestes. Cette approche s'inscrit dans la filiation de plusieurs courants théoriques en psychologie et psychologie sociale comme l'interactionnisme symbolique (dont le principal fondateur est G.H Mead, 1934) ou la théorie de l'activité, une dénomination qui regroupe les travaux de psychologues comme Vigostky (1934) ou Leontiev (1981) mais qui engloberait également dans notre cas des linguistes comme Bakhtine (1977). Contrairement à une acception du terme fréquente en économie, les transactions ne relèvent pas exclusivement des activités commerciales, mais sont avant tout des activités de production de formes nouvelles énergétiques et/ou sémiotiques véhiculées par des médias aussi divers que l'environnement matériel, des objets malléables, ou des supports de transcription¹.

Les transactions sont de nature diverse, leur analyse dépendant à la fois du degré d'institutionnalisation et du niveau d'interprétation de l'activité considéré (cf. infra). Quand le niveau d'interprétation est celui de l'action, dans l'acception sociale du terme, les transactions sont souvent « communicationnelles symboliques », mais dans certains contextes elles peuvent être considérées comme contractuelles au sens formel (juridique) ou monétaires (un sous-ensemble des transactions contractuelles). Quand l'activité est interprétée au niveau des opérations, les transactions ont toujours une dimension énergétique associée à leur matérialité. La plupart des actions impliquent une diversité de transactions sur des plans divers. Par exemple, la prestation rendue par un kinésithérapeute implique à la fois des transactions énergétiques, communicationnelles symboliques, contractuelles et monétaires.

Quand le niveau d'interprétation est celui de l'action, les transactions communicationnelles symboliques « *sont des interactions entre selfs² cognitivement interdépendants médiatisées par des productions sémiotiques leur permettant de créer de nouvelles significations visant à réduire leur incertitude mutuelle dans la poursuite ultérieure de leurs projets. La transaction se réalise par un échange réciproque de connaissances et d'engagements permettant de partager des représentations, attitudes ou affects communs facilitant la poursuite de l'action collective (...)* » (Zacklad 2003b). Les transactions communicationnelles se réalisent dans des *situations transactionnelles* qui comprennent (Zacklad 2004b) :

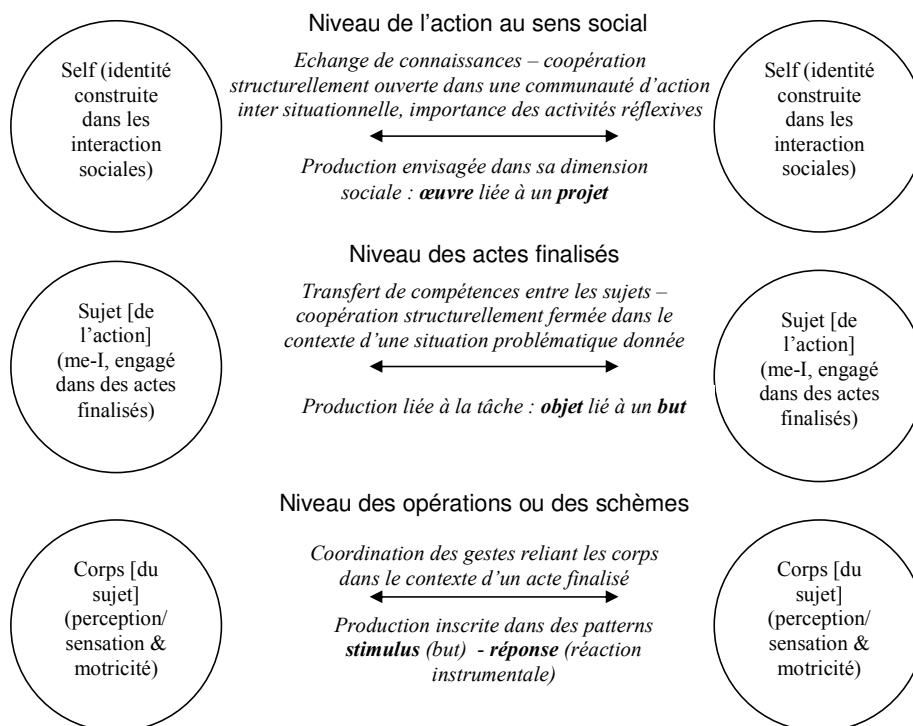
- un ou plusieurs *réalisateurs* ;
- un ou plusieurs *bénéficiaires* ;
- des *paramètres de la situation transactionnelle* au sens où ce terme a pu être utilisé dans le domaine de l'analyse pragmatique de l'énonciation, qui relèvent des domaines suivants :

¹ Les transactions peuvent se référer à l'échange, mais à nouveau le terme d'échange n'est pas limité à l'échange commercial, les échanges entre personnes peuvent être de nature énergétique ou sémiotique et sont une composante indispensable des processus de production comme peuvent l'être les échanges énergétiques et informationnels dans le fonctionnement d'une cellule.

² Dans les transactions communicationnelles, au niveau de l'action, nous utilisons le terme de self (Mead 1934) au lieu de celui d'acteur. Les selfs, qui peuvent être individuels ou collectifs sont transformés par les transactions dont ils sont parties prenantes.

- un *projet commun* sur la base d'intérêts ou d'objectifs pour partie partagés qui justifient la participation à cette situation collective ;
- des *relations sociales* entre le réalisateur et le bénéficiaire dont découlent pour partie les intérêts ou objectifs communs poursuivis et des *compétences* liées à la situation et à la transaction ;
- un *cadre spatio-temporel* et des *conditions environnementales* de nature non strictement sociales susceptibles d'influencer tant l'atteinte des objectifs que le processus de production sémiotique et faisant notamment peser des contraintes portant sur le choix du *média* associé
- un *terrain représentationnel commun* correspondant à l'ensemble des représentations communes évoquées à partir des connaissances associées à la situation transactionnelle et de l'histoire de la transaction, terrain commun indispensable à l'intercompréhension ;
- une *production sémiotique* véhiculant un *contenu sémiotique* via un *média* élaborée par les réalisateurs à l'intention des bénéficiaires dont nous détaillerons les caractéristiques plus bas.

Niveaux d'interprétation de l'activité



En suivant la plupart des traditions d'analyse de l'activité (théorie de l'activité version Leontiev 1981, échelle double de Rasmussen, 1986), nous considérons que celle-ci peut-être décrite à différents niveaux d'interprétation qui renvoient à son organisation hiérarchique (l'interprétation pouvant être le fait des sujets eux-mêmes ou de leurs observateurs). Conformément à l'approche de Leontiev mais en ne suivant pas strictement sa terminologie³, on considérera que *l'action sociale* repose sur la mise en œuvre d'*actes finalisés* qui reposent

³ En particulier, conformément aux usages de la sociologie, le niveau « supérieur » est nommé niveau de *l'action* au lieu de niveau de *l'activité*.

eux-mêmes sur des *opérations*. Les acteurs engagés dans les transactions sont considérés au niveau de l'action comme des *selves*, des (auto)-constructions sociales représentant la personne, au niveau des actes finalisés comme des *sujets* et au niveau des opérations comme des *corps* porteurs des sensations et de la motricité. Les productions associées aux niveaux d'interprétation de l'activité sont également sujettes à différentes descriptions, *œuvres symboliques* au niveau de l'action, *objets* au niveau des actes finalisés, simple *stimulus* dans des chaînes comportementales automatisées au niveau des opérations.

Décomposition de la production sémiotique dans les transactions communicationnelles

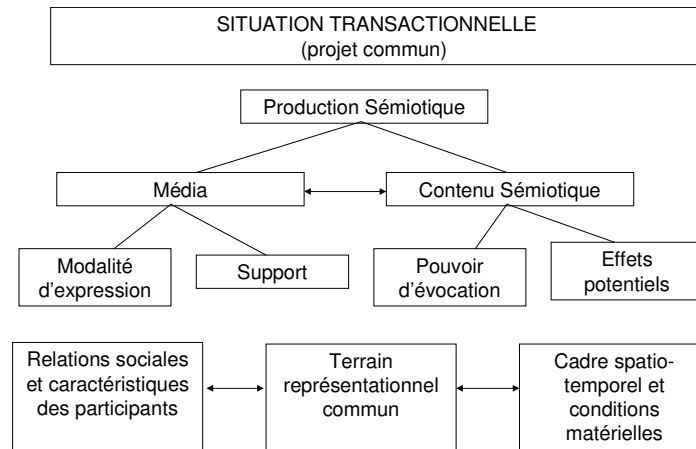
Dans le cadre des transactions communicationnelles, la production de la transaction est d'abord et avant tout une production sémiotique. Notons que ce n'est pas le cas de toutes les productions. La dimension sémiotique peut être présente sans constituer pour autant la facette dominante. Par exemple, la réalisation d'un déplacement, la préparation d'un aliment ordinaire, la confection d'un vêtement d'hiver ou la programmation d'un logiciel, ont certes une dimension sémiotique mais celle-ci est secondaire par rapport à leur valeur énergétique ou instrumentale. Cependant, dans la plupart des activités intellectuelles ou relationnelles qui nous intéressent d'avantage ici, la production de la transaction est avant tout sémiotique.

Dans l'analyse de la production sémiotique (Zacklad 2004b), on pourra établir une distinction entre d'une part la forme et le support de cette production, ou *média de la production sémiotique* et d'autre part le *contenu sémiotique* « véhiculé » par ce média. La production sémiotique se compose donc :

- du *média* qui peut lui même être analysé selon deux paramètres interdépendants :
 - le choix d'une *modalité d'expression* (langagière orale, scripturale, gestuelle, filmique, graphique, etc...), au sein de laquelle est conventionnellement déterminé un système de signes permettant des agencements syntagmatiques et paradigmatiques ;
 - le choix d'un *support matériel* associé à cette modalité d'expression (type de vocalisation, amplitude des gestes, nature du support pour un média scriptural, etc...) adapté aux caractéristiques de la modalité d'expression. Le support doit posséder une certaine malléabilité lui permettant de recevoir la forme et dans certain cas de la pérenniser ;
- du *contenu sémiotique* (ou signification) qui peut lui-même être appréhendé de deux manières en fonction de :
 - son *pouvoir d'évocation* qui est sa capacité à évoquer des *représentations communes*⁴ en fonction (i) de l'agencement des signes conforme aux possibilités offertes par le média et (ii) des paramètres de la situation transactionnelle (le *pouvoir d'évocation* correspond à la facette traditionnellement étudiée par la sémantique). Dans certaines situations extrêmement standardisées les signes n'évoquent pas de représentation mais déclenchent directement des « automatismes » et l'on considérera que les dimensions matérielles et communicationnelles de la transaction se confondent (cf. infra) ;
 - ses *effets potentiels* biologiques, psychiques et sociaux simultanés ou différés qui correspondent aux *conséquences possibles* de l'évocation de certaines représentations qui permettent d'attester de l'effectivité de la communication (et indirectement de celles des représentations communes évoquées). Ces effets sont

⁴ La notion de représentation commune correspond pour nous à la notion de « terrain commun » au sens de Clark 1996.

plus ou moins prévisibles selon le degré de standardisation de la situation transactionnelle et portent notamment sur l'actualisation du terrain commun et l'élargissement du « contenu sémiotique établi » entre les participants, considéré comme ayant publiquement ou officiellement fait progressé la transaction⁵ (les *effets potentiels* correspondent à une des facettes traditionnellement étudiée par la pragmatique linguistique).



Distribution des transactions dans un cadre spatio-socio-temporel hétérogène

Du fait de la distribution des activités collectives, réalisateurs et bénéficiaires des transactions communicationnelles peuvent ne pas être présents dans le même cadre spatio-temporel. Cela implique la mise en place de stratégies permettant de prolonger ces transactions de manière à ce qu'elles puissent être initialisées, interrompues, réactualisées, répétées, dans toutes les configurations de présence absence du bénéficiaire et du réalisateur. Par ailleurs, dans certains contextes, des réalisateurs distants peuvent se substituer aux réalisateurs ayant initié la transaction en « jouant leur rôle ». Dans ces contextes fréquents dans les organisations complexes et le monde économique nous parlerons de *distribution spatio-socio-temporelle des transactions*.

Dans le contexte des transactions dont la production, principalement sémiotique, repose sur des médias symboliques (transactions communicationnelles symboliques), il existe différents moyens permettant de distribuer la transaction dans un contexte spatio-socio-temporel hétérogène (Zacklad 2004b). Nous en présenterons ici six : la *normalisation de la situation transactionnelle*, la *formalisation de l'expression*, la *ritualisation mnémotechnique*, *l'abstraction*, la *médiation substitutive* et la *documentarisation*. Ces différents moyens ne sont pas indépendants les uns des autres et se renforcent souvent. Ces stratégies de distribution exerceront des effets permettant de stabiliser le cadre contractuel, le terrain représentationnel commun, la forme des productions sémiotiques ou encore leur transmission à distance.

1. *La normalisation de la situation transactionnelle* consiste à jouer sur les paramètres de la situation transactionnelle, homogénéisation des caractéristiques des acteurs en présence,

⁵ Correspondant à la notion de « Discourse Record » chez Clark.

reproduction des objectifs, des relations sociales, des compétences, des conditions environnementales, de manière à encourager la génération de productions sémiotiques similaires dans un cadre spatio-socio-temporel hétérogène.

2. *La formalisation de l'expression* consiste à travailler sur cette composante du média de manière à stabiliser le contenu sémiotique. Elle s'exerce dans les dimensions syntagmatique et paradigmatique. Dans le premier cas, elle vise à fixer précisément les conditions d'agencement des signes composant la production sémiotique par la détermination d'un ensemble de contraintes d'enchaînement. Dans le second cas, elle vise à positionner dans un réseau sémantique l'ensemble des signes équivalents ou proches et la nature de leurs associations en fonction de leurs ressemblances, de leur pouvoir d'évocation et de leurs effets potentiels dans une situation transactionnelle donnée. Le développement de langage de spécialité dans différents domaines d'activité professionnels relève directement de cette stratégie. Elle nous intéressera particulièrement dans le cadre de l'élaboration « d'ontologies sémiotiques ».

3. *La ritualisation mnémotechnique* exploite directement la normalisation de la situation transactionnelle et la formalisation de l'expression pour travailler sur les méthodes utilisées par les acteurs pour réaliser et recevoir la production sémiotique et en particulier pour la réactiver dans un contexte spatio-socio-temporel différent (cf. notamment Bachimont 2004, pour le rôle du corps dans les pratiques mnésiques). Si la ritualisation mnémotechnique s'installe pour une large part à l'insu des sujets elle relève dans d'autres cas de la mise en place de sessions de formation, de l'apprentissage volontaire d'une terminologie et de ses règles d'emploi, de la diffusion répétée de slogans visant à « marquer les esprits »...

4. *L'abstraction* consiste à travailler sur le contenu sémiotique lui-même pour systématiser les points de vue qui s'y expriment. Elle consiste à développer un niveau de discours théorique ou doctrinaire plus « universalisant » facilitant la transposition des transactions communicationnelles dans des situations dont la « structure de surface » (Zacklad 2004a) semble éloignée des situations d'origines et résistant donc à la normalisation de la situation transactionnelle. Le développement de principes méthodologiques abstraits, par opposition aux procédures relevant d'une méthode appliquée, est un exemple du recours à l'abstraction.

5. *La médiation substitutive* agit sur le média de la production sémiotique. Mais au lieu d'influencer la modalité d'expression et sa formalisation, elle consiste à travailler le *support matériel* (ce qui en retour ne sera pas sans conséquences sur la modalité d'expression). Elle substitue à la perception directe de la production sémiotique à partir du corps du sujet réalisateur une modalité de perception indirecte par le biais d'une *transcription sur un support d'écriture* ou *du codage automatique d'une des sources physiques* qui véhicule les signaux constituant la production sémiotique sur un support différent.

Le *support matériel* peut-être classé selon qu'il est de nature éphémère ou pérenne. Les *supports éphémères* sont caractérisés par le fait que les stimuli qu'ils véhiculent ne peuvent exercer qu'un effet transitoire sur les organes des sens des récepteurs notamment du fait que la forme donnée au support ne marque pas celui-ci de manière pérenne (c'est le cas de l'air qui est le support éphémère de la parole ou des flux lumineux qui renseignent l'interlocuteur sur les gestes de son partenaire dans une interaction de face à face).

Par contraste, les *supports pérennes* conservent la forme donnée au support ce qui permet au destinataire de réitérer l'effet exercé par la forme sur ses organes des sens. Quand il contrôle

ces supports, le destinataire peut exercer un contrôle sur les conditions dans lesquels les stimuli sont rendus accessibles à ses organes sensoriels dans des intervalles de temps courts (re-lecture de quelques lignes ou re-vision d'un extrait de film sur un support numérique personnel) ou longs (reprise d'un document après plusieurs jours, semaines, années...).

Il existe trois procédés utilisés dans la médiation substitutive : la *transcription*, le *codage automatique d'une source physique* et *l'enregistrement* des productions sémiotiques codées. Ces procédés permettent de transmettre à distance les productions sémiotiques et dans certains cas de les conserver de manière pérenne. La transcription relève des pratiques de l'écriture et repose sur un code graphique qui doit être maîtrisé par le rédacteur comme par le lecteur. Le codage automatique d'une source physique convertit une forme physique associée à une modalité d'expression pour la transmettre plus aisément sur un autre support avant de la décoder pour la restituer à son destinataire. L'enregistrement permet de conserver les signaux physiques codés sur des supports pérennes. Le codage automatique d'une source physique peut transmettre à distance une production sémiotique sans que le support de la transmission ne soit pérenne ni du côté du réalisateur ni du côté du bénéficiaire comme dans le cas d'une conversation téléphonique. Dans certains cas, retransmission radiophonique différée d'un événement enregistré, le support est pérenne pour le réalisateur mais pas pour le bénéficiaire.

En sélectionnant une source physique associée à une modalité d'expression (voix, données, image, kinesthésie...), la médiation substitutive appauvrit souvent la production sémiotique par rapport au potentiel disponible dans les situations de face à face (redondance de la multimodalité). Dans certains cas elle offre l'avantage de mettre en relief certains attributs supposer posséder une pertinence particulière. Enfin, elle peut, notamment dans le cas de la transcription, venir doubler les modalités d'expression utilisées dans le face à face, comme c'est le cas lors du dessin de schémas dans une réunion.

La documentarisation vient prolonger la médiation substitutive quand les supports utilisés sont pérennes. Elle consiste 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.

La stratégie de documentarisation nous permet de donner une définition de la notion de document. En effet, l'utilisation d'un support pérenne par le biais de la transcription ou l'enregistrement est la condition nécessaire mais non suffisante de la documentarisation. Le support doit faire l'objet d'un *investissement documentaire* permettant de le ré-exploiter dans des situations transactionnelles différées et distribuées. Résultat d'un investissement documentaire, le document désigne une *production sémiotique transcrite ou enregistrée sur un support pérenne qui est équipée d'attributs spécifiques visant à faciliter les pratiques liées à son exploitation ultérieure dans le cadre de la préservation de transactions communicationnelles distribuées*.

Soutien à la distribution des transactions communicationnelles dans le Web Socio Sémantique

Ce détour un peu long par la théorie des transactions communicationnelles nous semble cependant indispensable pour aborder certaines caractéristiques du Web Socio Sémantique et de son ingénierie. On l'aura compris, celui-ci contribue à faciliter la prolongation et la préservation des transactions communicationnelles à l'intérieur de communautés de pratiques ou d'action spatio-socio-temporellement distribuées. Alors que l'approche du Web Social privilégie souvent la médiation substitutive pour faciliter les interactions synchrones à distance sans recourir à des supports pérennes, l'activité « coopérative » des usagers du Web Socio Sémantique les amène le plus souvent à privilégier la stratégie de documentarisation.

Ce faisant, ils s'appuient sur de nouveaux types de média leur permettant de concilier l'interactivité et la pérennité offerte par les supports électroniques que nous avons appelés ces des « Documents pour l'Action » (Zacklad 2004b). Les forums, les usages publics de la messagerie, les systèmes d'annotation disponibles à partir des éditeurs de texte ou de sites Web (blogs, wikis, mais également annuaires évoluant sur la base des contributions de participants...) en constituent les exemples les plus connus. Les usagers des applications relevant du W2S, dont les transactions sont médiatisées par des Documents pour l'Action, sont ainsi confrontés à la gestion d'un nombre très important de documents évolutifs, souvent composés de fragments multiples non nécessairement intégrés, pris en charges par des auteurs différents dont les droits et les prérogatives à l'égard des fragments sont eux-mêmes sujets à négociation.

Face à cette diversité et à la complexité des relations entre les documents et souvent entre les fragments eux-mêmes (les annotations), le Web Socio Sémantique doit proposer des outils facilitant l'investissement documentaire de ses usagers et donc la qualité des transactions ultérieures. Si les moteurs de recherche offrent indéniablement une technologie de premier plan, celle-ci n'est généralement pas suffisante dans le contexte du W2S. En effet, la proximité sémantique des fragments, des documents et des dossiers, est potentiellement élevée et l'objectif est avant tout de les discriminer finement sur la base d'attributs aussi divers que les auteurs impliqués, le contexte de production et les thématiques abordées dans tel ou tel fragment. Qui plus est, la dimension coopérative de l'activité des usagers, qui signifie que les transactions impliquent des individus distincts dont les compétences sont parfois hétérogènes, nécessite de recourir à des systèmes d'indexation partagés par le collectif. Enfin, l'utilisation d'une ontologie sémiotique ne s'oppose pas à celle d'un moteur de recherche. Au contraire, elle est en mesure d'accroître la précision des réponses fournies.

L'élaboration de systèmes d'index partagés, relevant de Ressources Terminologiques et Ontologiques (Aussenac & Condamines, 2004) que nous nommons « ontologies sémiotiques » est donc une composante essentielle de l'ingénierie du W2S. Elle relève directement de la stratégie de formalisation de l'expression. En approchant les ontologies sémiotiques sous cet angle, nous serons en mesure de les différencier des ontologies formelles.

Critères pour l'établissement des accords définitionnels dans la construction des ontologies

Les ontologies sémiotiques sont des productions sémiotiques cohérentes qui regroupent des expressions stéréotypées extraites des transactions communicationnelles et organisées selon des axes paradigmatique et syntagmatique. Les expressions sélectionnées sont considérées comme des *concepts sémiotiques* à l'issue d'un processus *d'investissement définitionnel*. Cet investissement définitionnel, qui peut se limiter à un agencement différentiel de concepts sémiotiques dans des hiérarchies de proximités sémantiques basées sur leur fonction opératoire dans la situation transactionnelle, a pour objectif de faciliter la résolution des problèmes rencontrés dans cette situation. L'aide apportée par une ontologie sémiotique dans la recherche et la sélection d'informations comme dans la navigation sémantique est ainsi toujours relative à une classe de situations problèmes dans un cadre transactionnel donné qui justifie la cohérence d'ensemble de l'ontologie⁶.

La construction d'un concept sémiotique peut être plus ou moins délibérée, mais son intégration dans l'ontologie fait toujours l'objet d'un *accord définitionnel* explicite, contrôlé par un individu ou une communauté responsable de sa signification opératoire. L'accord définitionnel consiste à associer à un ensemble de signes un contenu sémiotique, c'est-à-dire à s'accorder sur une représentation commune et sur les effets que celle-ci est susceptible de déclencher. Il existe différentes modalités de l'accord qui renvoient plus ou moins aux grandes tendances utilisées par la sémantique pour aborder le signifié (par exemple Mounin 1968, p. 159). Celles-ci peuvent être caractérisées comme relevant de théories *logiques*, *contextuelles*, *situationnelles* (ou pragmatiques) du signifié même si ces approches visent souvent à caractériser des mots alors que nous visons principalement le signifié d'expressions.

L'approche logique qui est privilégiée dans les ontologies formelles veut asseoir la signification sur des critères qui soient à la fois indépendant du contexte sémiotique de l'expression (les autres signes composants la production sémiotique) et de la situation transactionnelle. L'accord définitionnel s'appuie sur la sélection d'un petit nombre de sèmes, pour partie hérités du graphe dans lequel se trouve le concept, qui confère à l'expression (au mot dans le cas de Web sémantique) une sorte « d'autonomie logique » lui permettant ainsi de se combiner librement avec d'autres expressions dans des arrangements syntaxiques inédits. L'accord sur les sèmes considérés comme pertinents peut-être sensé refléter une raison universelle, comme dans le cas du Web sémantique originel, ou peut être le fait d'une communauté plus restreinte. Ce qui caractérise l'approche logique c'est l'indépendance du sens vis-à-vis du contexte de la production sémiotique et vis-à-vis de la situation transactionnelle.

L'approche contextuelle ancre largement la signification de l'expression dans celle de la production sémiotique envisagée dans sa globalité, ou tout du moins, dans les parties de la production sémiotique, texte ou discours par exemple, qui « font sens ». En effet, comme l'expliquent les traducteurs professionnels⁷, non seulement toute traduction mot à mot est impossible mais le plus souvent toute traduction phrase à phrase l'est également. C'est en considérant des fragments signifiant de texte eux-mêmes délimités par les structures

⁶ On reproche parfois aux thésaurus leur incohérence... cela ne n'est pas le cas pour les ontologies sémiotiques qui revendent des critères d'organisation explicite justifiés par des situations problème génériques.

⁷ Monique Slodzian, communication personnelle.

« formelles » du document plus ou moins flexibles (saut de paragraphe, chapitre...) qu'une délimitation des unités signifiantes peut être effectuée.

L'accord définitionnel s'inscrit ici dans le contexte d'une production sémiotique donnée ou d'un « genre » de production sémiotique et s'appuie sur les relations de l'expression avec d'autres expressions voisines ou éloignées (selon leur proximité sémantique) sans qu'une analyse sémique ait nécessairement à être conduite. Prise au sérieux, l'approche contextuelle vient également relativiser l'importance de la syntaxe. La pertinence des combinaisons d'expressions dépend du genre mais dépend également de la proximité d'autres expressions dans une aire de signification donnée. La syntaxe apparaît d'avantage comme un soutien heuristique à l'interprétation que comme une condition préalable de toute signification.

De la même façon, l'approche contextuelle tendrait à remettre en cause l'indépendance de la signification de l'expression vis-à-vis de la matérialité de son support ou du moins de la manière dont le support serait appréhendé perceptivement dans le cas du recours à une médiation substitutive. En effet, les indices permettant d'interpréter la proximité sémantique sont souvent eux-mêmes reliés à des caractéristiques de proximité spatiale sur le support, voire de facilité manipulatoire associée aux activités de navigation sémantique dans celui-ci.

L'approche situationnelle (ou pragmatique) ancre la signification dans la situation transactionnelle elle-même et dans ses différentes composantes. Les expressions ne sont pas seulement rapportées à l'environnement interne de la production sémiotique considérée, à leur proximité sémantique, mais également aux paramètres de la situation transactionnelle dans leur diversité : projet commun, nature des relations sociales entre les participants et caractéristiques de ces derniers, cadre spatio-temporel et conditions environnementales, terrain représentationnel commun... Selon cette approche, la signification se construit sur la base d'une « pratique sociale continue, au cours de milliers et de milliers d'expériences d'usage des termes en situation, corrigées successivement par la méthode des essais et erreurs » (Mounin 1968, p. 163).

Les accords définitionnels s'ancrent dans une approche fonctionnelle de la signification dépendant des caractéristiques de la situation. Il ne s'agit d'ailleurs pas d'une approche *référentielle* qui rabattrait le sens sur l'existence « choses autonomes dans un monde objectif » servant de référent indépendamment des activités coopératives des acteurs et de leurs pratiques linguistiques de dénomination dans le cadre des situations transactionnelles problématiques (cf. Dewey et Bentley 1949). L'existence des « choses » en tant qu'entités appréhendables par la pensée et la langue dépend des projets poursuivis par les acteurs, projets dont le déroulement entraîne une dénomination des ingrédients des situations problématiques à l'aide d'expressions, expressions qui structureront la pensée et qui influenceront en retour le déroulement de l'activité.

Dans l'approche situationnelle, l'accord définitionnel, la signification de l'expression, dépend donc de la compréhension de la finalité de l'action dans laquelle elle est employée, de son réalisateur (énonciateur dans le cas du discours), du contexte matériel de la production, des relations sociales entre les réalisateurs et les bénéficiaires, etc... Selon que l'expression sera utilisée dans une transaction visant à produire des soins, assembler un dispositif technique, juger un prévenu, former un élève, piloter un navire, recruter un collaborateur, convaincre des électeurs potentiel, vendre un produit... sa définition ne sera pas la même et il n'y aura pas de compréhension possible entre des partenaires sans une expérience, même indirecte, de ces situations de référence.

Selon cette approche, dans une situation donnée, l'expression n'aura pas le même sens selon qu'elle émanera de l'un ou l'autre des participants, voire même n'aura tout simplement aucun sens. Ainsi, l'expression « Vous êtes embauché⁸ ! » émanant du candidat à l'emploi, est une expression a priori sans signification dans les transactions associées au recrutement. Son interprétation nécessiterait la référence à une autre situation transactionnelle où les rôles seraient inversés.

Dans l'approche situationnelle, l'investissement ontologique implique une entente entre les participants quant à la fonction des expressions dans une situation transactionnelle donnée, c'est-à-dire quant à leur contenu sémiotique effectif, leur pouvoir évocation et leurs effets potentiels. C'est non seulement l'autonomie combinatoire des expressions vis-à-vis de la syntaxe et du support de la production sémiotique, c'est-à-dire du média, qui est mise à mal, mais encore leur autonomie par rapport à l'ensemble des paramètres des situations transactionnelles dont elles sont issues et auxquelles elles sont destinées.

Conclusion : sémiotique fermée vs sémiotique ouverte, formalité machinale vs formalité de la signification

En conclusion, nous présentons une hypothèse provocatrice selon laquelle l'effectivité de certaines applications relevant du Web Sémantique ne tient pas à l'approche « logique » de la signification, ancrée dans une syntaxe formelle et une sémantique référentielle, pourtant souvent mise en avant, mais à la fermeture sémiotique opérée dans les applications, fermeture sémiotique qui ne fonctionne qu'en regard de la fermeture corollaire des situations transactionnelles de référence. Cette fermeture sémiotique est également celle qui permet à la plupart des applications automatisées de fonctionner efficacement sans qu'aucune référence ne soit faite à la « logique formelle ». Or, paradoxalement, cette caractéristique semble largement ignorée par les promoteurs du Web Sémantique, qui en s'identifiant au courant logiciste du programme de l'Intelligence Artificielle imputent leur effectivité à la remonté vers les couches hautes du célèbre cake, couches de la preuve et de la vérité (sic !).

Quelle est la condition de l'effectivité des calculs réalisés par les programmes informatiques dans les situations sociales ou techniques dans lesquelles ils sont employés dès lors que l'utilisation visée n'est pas un soutien heuristique à la réflexion mais une prise de décision automatisée ? Elle dépend entièrement de la fermeture de la situation transactionnelle dans laquelle le programme informatique transmettra des informations et appliquera des règles décisionnelles. Comme le savent tous les programmeurs, c'est uniquement à condition d'avoir listé toutes les alternatives possibles, que celles-ci résultent d'un ensemble de règles stockées dans une base de données ou d'un calcul basé sur des données issues de capteurs, qu'un programme ne « plantera » pas, dans un sens étroitement technique ou lié aux implications pragmatiques associées aux dégâts consécutifs pouvant être causés dans la situation d'usage.

Du point de vue de la situation transactionnelle, ce fonctionnement renvoie à des situations interactionnelles dites « fermée » dans lesquelles la dimension sémiotique de la communication s'efface devant le recours à des automatismes comportementaux. Le signe est alors réduit alors à un signal qui déclenche une réponse conditionnée sans recours à

⁸ On aurait pu utiliser l'expression « Je vous fais un rabais ! » émanant du client de l'agence de voyage pour rester dans l'univers bien connu des ontologistes.

l'interprétation. C'est uniquement à cette condition que sera possible à la fois, la conception des algorithmes et l'utilisation sûre du programme dans des situations sociotechniques de référence. Bien sur, l'utilisateur, dans sa phase de découverte de l'interface, pourra avoir une attitude anthropomorphique à l'égard du programme. Mais ses apprentissages seront vite orientés vers la compréhension des règles de fonctionnement déterministes du logiciel, afin qu'il puisse lui-même développer des schèmes sensori-moteurs de bas niveau économisant tout investissement sémiotique inutile.

Alors que dans une base de donnée automatisée la sémiotique des expressions utilisées est fermée et que le signe a vocation à devenir un signal, dans une base documentaire, les activités de recherche et de navigation doivent pouvoir s'appuyer sur des signes dont la sémiotique est ouverte en cohérence avec l'ouverture des situations transactionnelle associées. La recherche et la navigation tirera donc partie d'expressions du domaine à la fois pertinentes et contrôlées, des « concepts sémiotiques », organisés dans des ontologies du même nom et permettant des parcours guidés par des finalités, certes identifiées, mais non réductible à des procédures automatisées.

Ces concepts sémiotiques et leurs ontologies relèvent bien d'une formalisation. Mais cette formalisation est celle donnée à l'expression pour accroître son ouverture sémiotique vers le contexte et la situation au moyen de l'investissement définitionnel dans un cadre transactionnel donnée. La *formalité de la signification* vient ouvrir le sens en multipliant les ancrages possible du concept sémiotique à partir de différents points de vue tandis que la *formalité machinale* doit fermer le sens pour que le signe devienne autant que possible un signal univoque. Dans la formalité machinale, la syntaxe est rigide, les signes sont précisément dénombrés et infiniment combinables et leur organisation sur le support totalement indifférente.

Ainsi, ce n'est pas la formalité de la signification qui explique l'effectivité des « langages syntaxiquement formels », car ils ne possèdent en principe aucun potentiel sémiotique, mais la standardisation des situations transactionnelles de référence, celles dont ils sont issus et auxquelles ils se destinent (une vision proche des réflexions du deuxième Wittgenstein). L'automatisme de la signification renvoie en fait à l'automatisme comportemental qui fait que le signe est traité comme un stimulus attendant une réponse conditionnée. En d'autres termes, si les applications « touristiques » du Web Sémantique fonctionnent dans certaines transactions simples, c'est parce que l'industrie du tourisme se mondialise en proposant des prestations de plus en plus formatées à des consommateurs de plus en plus conditionnés !

La garantie de l'effectivité des traitements automatiques effectués lors des calculs dépend en fait de la force des sanctions sociales qui permettent de normaliser les comportements : l'énoncé $2+2=5$ vaut une sanction à l'élève et un licenciement au technicien de même que l'emploi inconsidéré d'une carte bleue dans les transactions commerciales envoie l'utilisateur inconsistant en prison. C'est dans l'automatisme des comportements conditionnés et des sanctions que réside le secret de la formalité non sémiotique, de la formalité machinale (littéralement des formules qu'il nous faut interpréter comme des machines, qui n'est pas celle des formules des mathématiques avancées ou le *style* de la démonstration importe). On ne contestera d'ailleurs pas l'intérêt de cette automatisme pour un fonctionnement organisationnel ordonné et le développement de pratiques d'ingénierie efficaces et fiables, mais on se souviendra que la force de la formalité machinale repose sur celle des conventions sociales qui restreignent le potentiel sémiotique des signes dans certaines situations

transactionnelles et pas sur une quelconque preuve ou vérité résultant des activités introspectives des « ontologies formels » accédant aux couches hautes du Web Sémantique.

Bibliographie

- AUSSENAC, N., CONDAMINES, A. (2004). Documents électroniques et constitution de ressources terminologiques ou ontologiques, *Information-Interaction-Intelligence*, Volume 4(1) : 75-94.
- BACHIMONT, B. (2004). *Arts et sciences du numérique : ingénierie des connaissances et critique de la raison computationnelle*. Habilitation à diriger des recherches de l'Université Technologique de Compiègne.
- BAKHTINE, M., VOLOCHINOV, V. N. (1997 trad.). *Le marxisme et la philosophie du langage*, Les Editions de Minuit, Paris.
- CAHIER ET AL. (2004). Cahier, J.-P., Zacklad, M., Monceaux, A., Une application du Web socio sémantique à la définition d'un annuaire métier en ingénierie, *Actes des journées Ingénierie des Connaissances*, mai 2004, Lyon.
- CAUSSANEL ET AL. (2002). Caussanel J, Cahier J.-P., Zacklad M., Charlet J., " Les Topic Maps sont-ils un bon candidat pour l'ingénierie du Web Sémantique ? ", Conférence Ingénierie des Connaissances IC2002, Rouen Mai 2002.
- CLARK, H. H. (1996). *Using language*, Cambridge University Press.
- DEWEY, J. & BENTLEY, A. F. (1949). *Knowing and the known*. In J. A. Boydston (1989), *John Dewey: The later works, 1925-1953* (Vol. 16, pp. 2-294). Carbondale: Southern Illinois University Press.
- LEONTIEV, A. A., (1981). *The Problem of Activity in Psychology*. In Wertsch, James V. *The Concept of Activity in Soviet Psychology*. Armonk, NY: Sharpe.
- MEAD G. H., (1934). *Mind, Self and Society from the Standpoint of a Social Behaviorist*, nouvelle édition (1967). The University of Chicago Press, Chicago.
- MOUNIN, G., (1968). *Clefs pour la linguistique*, Seghers, Paris
- RASMUSSEN J., (1986). *Information processing and human-machine interaction*, North-Holland, Amsterdam.
- VYGOTSKI L.S. (1934). *Pensée et Langage*. La Dispute. Paris.
- ZACKLAD M., (2003a). Un cadre théorique pour guider la conception des collecticiels dans les situations de coopération structurellement ouvertes, in Bonardi, C. GEORGET, P. ROLAND-LEVY, C. ROUSSIAU, *Psychologie Sociale Appliquée, Economie, Médias et Nouvelles Technologies*, In : Press, Coll Psycho, Paris, 135-164.
- ZACKLAD M., (2003b). Transactions communicationnelles symboliques et communauté d'action : réflexions préliminaires, présentation au colloque de Cerisy de septembre 2003 organisé par P. Lorino et R. Teulier, *Archive SIC*, <http://archivesic.ccsd.cnrs.fr/>
- ZACKLAD ET AL. (2003c). Zacklad, M., Cahier, J.P., Pétard, X. Du Web Cognitivement Sémantique au Web Socio Sémantique, *Journée « Web Sémantique et SHS » du 7 mai 2003*, <http://www.lalic.paris4.sorbonne.fr/stic/as5.html>
- ZACKLAD, M. (2004a). Transférabilité des connaissances : une reconceptualisation de la distinction tacite/explicite. *En route vers Lisbonne : 1er colloque luxembourgeois sur l'économie de la connaissance dans une perspective européenne*, 12-13 Octobre 2004, Luxembourg. 12 p.
- ZACKLAD, M. (2004b). Processus de documentarisation dans les Documents pour l'Action (DopA) : statut des annotations et technologies de la coopération associées. *Le numérique : Impact sur le cycle de vie du document pour une analyse interdisciplinaire* http://archivesic.ccsd.cnrs.fr/sic_00001072.html, 13-15 Octobre 2004, Montréal. Montréal: 19 p.