

Ontologie pour un système multi-agents dédié à une mémoire d'entreprise

Fabien Gandon, Rose Dieng

► **To cite this version:**

Fabien Gandon, Rose Dieng. Ontologie pour un système multi-agents dédié à une mémoire d'entreprise. IC'2001, Ingénierie des Connaissances, plateforme AFIA'2001, Jun 2001, Grenoble, France. <hal-01145808>

HAL Id: hal-01145808

<https://hal.inria.fr/hal-01145808>

Submitted on 27 Apr 2015

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.

Ontologie pour un système multi-agents dédié à une mémoire d'entreprise

Fabien Gandon, Rose Dieng-Kuntz

INRIA - Projet ACACIA, 2004, route des Lucioles
B.P. 93 - 06902 Sophia Antipolis Cedex
{Fabien.Gandon|Rose.Dieng}@sophia.inria.fr

Résumé

L'ontologie joue un rôle important dans la nouvelle génération de systèmes d'information ou à base de connaissances. Elle est aussi une clef de voûte des systèmes multi-agents utilisant une communication de haut niveau. Il n'est donc pas étonnant qu'elle soit centrale à un système d'information multi-agents destiné à la gestion d'une mémoire d'entreprise. Nous présentons ici les premiers éléments de notre retour d'expérience sur la conception d'une ontologie pour un tel système.

1 INTRODUCTION

L'explosion des technologies de l'information de cette dernière décennie a provoqué une modification des règles des marchés économiques et forcé les organisations à adapter leur structure et à améliorer leur temps de réaction. Les systèmes d'information sont devenus, pour elles, de véritables réseaux nerveux assistant la gestion orientée projet et les équipes virtuelles ainsi que le renseignement et la veille technologique pour détecter des mouvements aussitôt que possible et surtout essayer de les prévoir pour être novateur. L'intérêt industriel pour les méthodologies et les outils permettant la capitalisation et la gestion de la connaissance organisationnelle s'est donc considérablement accru. Les technologies du Web sémantique fournissent des options intéressantes pour matérialiser et structurer les mémoires afin de préparer leur exploitation et leur gestion. En parallèle, l'intelligence artificielle distribuée propose le paradigme des systèmes multi-agents, très adapté au déploiement d'une configuration logicielle au-dessus de ce paysage d'information distribué. Nous présentons dans un premier temps le projet CoMMA évaluant une approche multi-agents pour une mémoire d'entreprise. Nous justifierons l'importance de l'ontologie dans le système étudié, avant de nous concentrer sur l'approche adoptée pour sa conception. Nous décrirons en détail les phases de recueil, de réutilisation, d'analyse terminologique, de structuration et de formalisation.

2 CONTEXTE D'APPLICATION

2.1 Le projet CoMMA

La mémoire organisationnelle est une représentation persistante, explicite, désincarnée des connaissances et des informations dans une organisation, afin de faciliter leur accès, leur partage et leur réutilisation par les membres adéquats de l'organisation, dans le cadre de leurs tâches individuelles et collectives (Dieng *et al.*, 2000). L'enjeu de la construction d'un système de gestion de mémoire d'entreprise est donc de réussir l'intégration cohérente de la connaissance dispersée dans l'organisation afin d'en promouvoir la croissance, la communication et la préservation (Steels, 1993). La mémoire organisationnelle est un système distribué au sens large du terme. Elle n'est pas seulement distribuée entre des systèmes informatiques, sa distribution inclut d'autres artefacts, ainsi que les membres de l'organisation. Les solutions informatiques pour la gestion de la mémoire, doivent prendre en compte cette "super-distribution" pour s'intégrer aux organisations. Les modèles de la distribution doivent aller au-delà de la dimension technique et rendre compte de l'organisation, de son infrastructure et de son environnement.

Une mémoire distribuée et hétérogène. La mémoire organisationnelle a cela en commun avec le Web qu'elle est un paysage d'informations hétérogènes et distribuées. De ce fait, ils partagent tous les deux les problèmes de bruit et de précision lors de la recherche d'information. Pour reprendre la formule de John Naisbitt, "les utilisateurs du Web se noient dans l'information alors qu'ils ont soif de connaissance". Parmi les initiatives visant à résoudre ces problèmes le Web sémantique (Berners-Lee *et al.*, 2001) est une approche prometteuse, cherchant à rendre explicite la sémantique des documents au travers de méta-données ou d'annotations. Les intranets reposant sur la technologie internet, peuvent bénéficier des progrès du Web sémantique, et la mémoire d'entreprise peut alors être étudiée comme un "web sémantique d'entreprise".

Une population d'utilisateurs distribuée et hétérogène. La population des utilisateurs d'une mémoire organisationnelle est hétérogène car il existe plusieurs profils de membres concernés par la gestion et l'exploitation de la mémoire. Chacun a ses particularités, ses centres d'intérêts et son rôle dans l'organisation. Cette population est aussi distribuée car les utilisateurs sont dispersés dans l'infrastructure de l'organisation. Ils ne sont pas forcément conscients de l'existence des autres, de leur fonction ou de leurs capacités. Ils ignorent parfois l'existence, l'emplacement et la disponibilité d'artefacts et de gens avec lesquels ils partagent des centres d'intérêts. Pour s'adapter aux profils des utilisateurs et aux contextes d'utilisation de la mémoire, l'apprentissage symbolique propose des techniques permettant d'apprendre les particularités individuelles ou d'assister l'émergence de communautés d'intérêt et la diffusion proactive d'information par une exploitation collective des profils.

Une gestion distribuée et hétérogène. La mémoire organisationnelle est distribuée et hétérogène. La population de ses utilisateurs est distribuée et hétérogène. Il semble donc intéressant que l'interface entre ces deux mondes soit elle-même de nature distribuée et hétérogène pour d'une part pouvoir se calquer sur le paysage d'information, et d'autre part s'adapter aux utilisateurs. Comme le note (Wooldrige *et al.*, 1999), les progrès de la programmation se sont faits à travers le développement d'abstractions de plus en plus puissantes permettant de modéliser et de développer des systèmes de plus en plus complexes. Nous pensons, comme eux, que le paradigme des systèmes multi-agents (SMA) représente une nouvelle étape dans l'abstraction et qu'il peut être employé pour comprendre, modéliser, et développer une classe importante de systèmes distribués complexes. Le paradigme des SMA apparaît particulièrement adapté au déploiement d'une architecture logicielle au-dessus du paysage d'information distribué qu'est la mémoire d'entreprise. Par leur collaboration, les agents logiciels réalisent une intégration des connaissances de l'entreprise et en permettent la capitalisation globale tout en s'adaptant localement à la spécificité des ressources et des utilisateurs individuels.

Afin d'évaluer ces idées, le projet ACACIA est membre du consortium CoMMA - Corporate Memory Management through Agents (CoMMA, 2000). CoMMA est un projet IST (1999-12217), subventionné par la Commission Européenne, visant à développer et tester un environnement de gestion de la mémoire d'entreprise. Le domaine d'application est fixé par deux scénarios :

- L'aide à l'insertion d'un nouvel employé en assistant ce dernier dans la recherche d'informations ou en lui suggérant proactivement l'information dont il a besoin pour s'intégrer.
- Le support à la veille technologique en assistant l'identification et l'évaluation d'informations pertinentes pour l'activité de l'entreprise, et en les diffusant aux personnes concernées et compétentes.

La figure 1 illustre comment CoMMA cherche à intégrer un certain nombre de technologies émergentes : l'idée est de conjuguer au sein d'un système multi-agents des techniques et outils issus de l'ingénierie des connaissances, de la galaxie XML, de la recherche d'informations et de l'apprentissage symbolique.

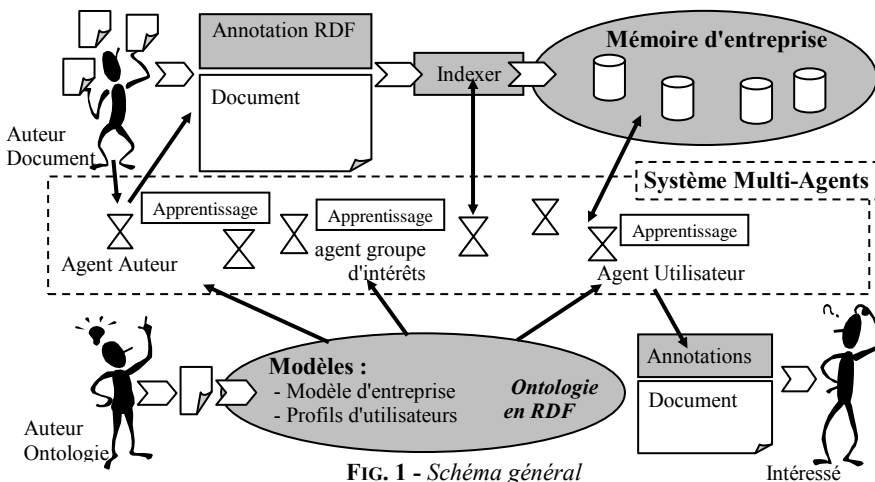


FIG. 1 - Schéma général

2.2 De la nécessité d'une ontologie

Dans le projet CoMMA, pour matérialiser la mémoire d'une entreprise, nous avons étudié les atouts du langage XML, afin de structurer les documents, et ceux du langage RDF utilisé pour le Web sémantique, afin d'annoter la mémoire. A la différence du Web, une mémoire organisationnelle a un contexte délimité et mieux défini : l'organisation. Nous pouvons envisager d'identifier les différents profils d'utilisateurs et comme la communauté de l'entreprise partage souvent un certain nombre de points de vue (politique d'entreprise, pratiques instituées...) on peut penser qu'un consensus ontologique est possible dans une certaine mesure. Partant de là, la figure 2 schématise notre architecture de modélisation : (1) l'ontologie est formalisée dans le langage RDFS (2) la situation actuelle de l'organisation et de ses membres est décrite en utilisant les primitives de l'ontologie, fournissant au système un modèle de son environnement qu'il peut exploiter dans ses interactions avec les utilisateurs ; cette situation est formalisée en RDF, en instanciant le schéma RDFS de l'ontologie (3) Les annotations des documents sont écrites pour structurer la mémoire ; elles sont aussi des instances du schéma RDF formalisant l'ontologie et réfèrent la situation actuelle si nécessaire ; par exemple: "le rapport de d'analyse de tendance "nouvelles normes de la téléphonie mobile" a été écrit par la cellule de veille technologique du département D12 et concerne les sujets: UMTS, WAP".

Nous pouvons distinguer plusieurs sources de besoin d'une ontologie dans

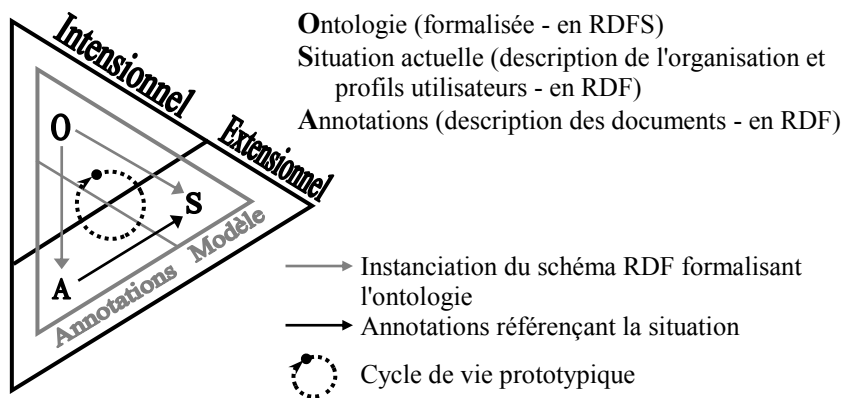


FIG. 2 - Schéma O.S.A.

CoMMA : une venant de la nature même des SMA où les agents ont besoin d'un vocabulaire conceptuel consensuel pour exprimer les messages qu'ils échangent ; une autre venant de l'aspect système d'information utilisant des requêtes et des annotations basées sur un vocabulaire conceptuel consensuel ; enfin l'ontologie est une composante de le mémoire d'entreprise qui capture une connaissance potentiellement intéressante en elle-même pour l'utilisateur. Dans CoMMA, nous ne développons qu'une seule ontologie mais les différents agents et les différents utilisateurs n'exploiteront pas les mêmes aspects de cette ontologie.

3 CONCEPTION D'UNE ONTOLOGIE

Notre travail sur l'ontologie a été particulièrement influencé par le rapport de retour d'expérience et d'analyse sur TOVE et l'Ontologie d'Entreprise écrit par Uschold et Gruninger (1996), la comparaison de méthodologies présentée dans (Fernandez *et al.*, 1997) et (Gómez-Pérez *et al.*, 1996) et les travaux effectués sur les bases théoriques des ontologies (Bachimont, 2000 ; Guarino, 1992 ; Guarino & Welty, 2000).

3.1 Position et Définitions

Ironiquement, le domaine de recherche de l'ontologie a beaucoup souffert de l'ambiguïté. Nous avons donc commencé par un état de l'art du domaine de l'ingénierie ontologique, du cycle de vie d'une ontologie et des différentes méthodologies proposées. Nous donnons dans cette partie les définitions et les points de vue que nous avons adoptés.

Le système exploite un modèle de la situation actuelle de l'organisation décrivant les aspects pertinents de l'environnement organisationnel (ex: organigramme, fonctions et spécialités d'un groupe...). Pour ce faire, nous avons besoin d'une ontologie définissant les primitives nécessaires à la représentation de cette description et la sémantique qui leur est associée. Guarino et Giaretta (1995) ont défini l'ontologie comme "une théorie logique qui rend compte partiellement mais explicitement d'une conceptualisation", la conceptualisation étant définie comme "une structure sémantique intensionnelle qui capture les règles implicites contraignant la structure d'un morceau de réalité". L'ontologie est une représentation explicite partielle parce qu'elle se concentre sur les aspects de la conceptualisation nécessaires au bon fonctionnement du système.

Un concept est un constituant de la pensée, né de l'esprit (un principe, une idée, une notion abstraite) sémantiquement évaluable et communicable. L'ensemble des attributs d'un concept s'appelle sa compréhension ou son intension et l'ensemble des êtres qu'il englobe, son extension. Il existe une dualité entre l'intension et l'extension: à des intensions incluses $I1 \supset I2$ correspondent des extensions incluses $E2 \supset E1$. L'intension est déterminée par l'identification de propriétés qualitatives ou fonctionnelles communes aux individus auxquels le concept s'applique, cet ensemble de caractères permettant de le définir. Pour exprimer et communiquer l'intension nous choisissons une représentation symbolique, presque toujours verbale, ex: les définitions des différentes notions associées à un terme et données dans les dictionnaires. Notons aussi que l'exemplification et les illustrations utilisées dans un dictionnaire montrent, respectivement, qu'il est parfois nécessaire pour clarifier une définition en langue naturelle d'exhiber un échantillon représentatif de son extension ou de faire appel à d'autres formes de représentations.

Les représentations des intensions peuvent être organisées, structurées et contraintes pour exprimer une théorie logique rendant compte des relations qui existent entre les concepts. L'ontologie est l'objet capturant les expressions des intensions et la théorie visant à rendre compte des aspects de la réalité choisis

pour leur pertinence dans les scénarios considérés pour le système envisagé. La représentation des intensions et de la structure ontologique peuvent faire appel à des langages plus ou moins formels selon l'opérationnalisation envisagée pour l'ontologie. La construction formelle de l'intension donne une représentation précise et non ambiguë de la manière dont on peut concevoir son sens, ce qui permet sa manipulation logicielle et son utilisation comme une primitive de représentations de connaissances pour les modèles et les annotations. L'expression de l'intension partant presque systématiquement d'une définition en langue naturelle, "définir une ontologie est une tâche de modélisation menée à partir de l'expression linguistique des connaissances" (Bachimont, 2000). Par raffinements successifs, nous augmentons l'ontologie en développant les pendants formels des aspects sémantiques pertinents pour le système dans les scénarios envisagés. Dans l'ontologie, les concepts en intension sont habituellement organisés en taxonomie, c'est à dire une classification basée sur leurs similitudes.

Quand des personnes s'accordent sur l'utilisation et la théorie spécifiée par l'ontologie, nous disons qu'ils ont pris un engagement ontologique. L'ingénierie ontologique s'occupe des aspects pratiques, essentiellement les méthodologies et les outils de l'application des résultats de la science ontologique à la construction d'ontologies, pour un cas précis et avec un but spécifique.

3.2 Analyse par scénarios et Recueil

Les scénarios sont des descriptions textuelles de l'activité organisationnelle et des interactions concernant le système. Nous pensons comme Caroll (1997) que les *scénarios* sont une façon naturelle et efficace de capturer les besoins des utilisateurs dans leur contexte, ce qui est essentiel pour une intégration symbiotique du système à l'environnement de travail. Les principaux avantages que nous avons identifiés en utilisant des scénarios pour CoMMA sont:

- Se focaliser sur les aspects de gestion des connaissances spécifiques au cas d'application considéré, capturant ainsi la portée du travail effectué.
- Capturer et présenter une image globale permettant de voir le système comme un composant d'une solution possible pour la gestion de connaissances dans une compagnie.
- Rassembler un ensemble concret d'interactions avec la mémoire d'entreprise, compréhensible et accessible pour tous les intéressés du projet et fournir un excellent point de départ pour construire des 'use case' et autres diagrammes d'interactions.
- Fournir un cadre pour évaluer et contrôler chaque nouvelle idée.

L'analyse par scénario nous a amenés à définir une grille suggérant les principaux aspects à considérer lors de la description d'un scénario. Durant la phase de recueil, cette grille rappelle les thèmes généraux des informations à rechercher et les points qui pourraient être affinés. Il aide aussi à définir le domaine de l'ontologie : le vocabulaire conceptuel doit fournir l'expressivité nécessaire aux interactions décrites dans les scénarios.

L'analyse par scénario produit aussi des traces : les rapports de scénario. Dans ces rapports les utilisateurs racontent leur histoire, leur quotidien et leurs vœux.

Ces rapports sont extrêmement riches et donc de très bons candidats à inclure dans le corpus d'une étude terminologique. Ils représentent notre première étape de recueil (figure 4a).

Plusieurs techniques existent pour le recueil de données, héritées du travail effectué en acquisition des connaissances. Ces techniques alimentent la mémoire et le processus complet de modélisation incluant la construction de l'ontologie et de la description de la situation organisationnelle actuelle. Nous avons essentiellement utilisé trois de ces techniques : les entretiens semi-structurés, l'observation et l'analyse de documents.

Pour ce travail sur les entretiens, nous nous sommes inspirés de (Sebillotte, 1991) et (La France, 1992). Un entretien peut être extrêmement structuré (interrogation) ou complètement libre (expression spontanée) mais le type le plus commun est l'*entretien semi-structuré* qui se trouve n'importe où sur le continuum entre ces deux extrêmes. Nos entretiens semi-structurés se divisent ainsi : (a) Première partie non structurée : une discussion lancée par des questions très ouvertes et générales concernant habituellement les tâches et le rôle des personnes interviewées. L'interviewé parle et raconte son histoire jusqu'à ce qu'un long silence s'installe. Si nécessaire cette expression spontanée est encouragée par de courtes questions. (b) Retour en arrière et clarification : une fois arrivé à un silence terminal, une partie plus structurée peut commencer ce qui nécessite d'avoir pris des notes pendant la première partie et d'avoir déjà établi une certaine représentation personnelle afin de pouvoir identifier les points à clarifier ou détailler. C'est également dans cette partie que viennent les questions plus spécifiques préparées avant l'entretien et motivées par des informations recherchées. (c) Auto synthèse : une bonne pratique consiste à demander aux personnes interviewées de synthétiser, récapituler et analyser ce qu'elles ont dit pendant l'entretien et de les faire conclure. La généralisation et les regroupements qu'elles opèrent dans cette dernière partie sont particulièrement intéressants pour structurer l'ontologie.

Nous avons effectué un entretien avec une nouvelle employée, jeune secrétaire, de la société ATOS partenaire industriel du projet. L'entretien enregistré, retranscrit et exploité par notre collègue Alain Giboin¹ a, par exemple, souligné que la perception qu'une personne a de son rôle dans l'entreprise peut être différente de la définition officielle donnée par sa fiche de poste (importance du profil personnel). Nous avons aussi pu voir des exemples de définitions de tâches liées aux documents (ex : émission d'un bon de commande), ainsi que l'importance donnée au réseau de connaissances dans l'activité de tous les jours (utilité du modèle organisationnel).

Une autre technique de recueil que nous avons essayée est l'*observation*. L'observation peut s'intéresser aux personnes, la façon dont elles travaillent (on peut leur demander ou non de commenter celle-ci), sur une tâche réelle ou sur un scénario simulé, en temps réel, enregistré ou basé sur des traces de l'activité. Elle peut également se concentrer sur des indicateurs choisis (documents manipulés, organisation du bureau, réseau de connaissances...). Selon l'acteur et le scénario, la situation intéressante peut être très différente (un nouveau venu recherchant une

¹ Docteur en psychologie et chargé de recherche du projet ACACIA

information, un mentor expliquant, une personne de la veille technologique commentant un nouveau brevet...). Nous avons observé le lieu de travail de la jeune secrétaire par laquelle passe un grand nombre de documents internes. En s'appropriant son espace de travail, elle a réorganisé tous les documents et introduit certaines classifications. A titre d'exemple, elle a l'habitude d'utiliser deux critères pour étiqueter son 'bac de documents en cours': le type des documents (par exemple : formulaires de congés) et la prochaine tâche à faire sur ces documents (par exemple: à faire signer par Mr. X). Les 'post-it' révèlent différents types et utilisations d'annotations sur des documents (destinataire, date limite, mode d'utilisation d'un formulaire, tâches ou autres documents liés au document). Enfin nous avons également noté qu'il existe des documents (dans notre cas, il s'agissait simplement d'une carte de visite de la compagnie avec le numéro de téléphone et le numéro de fax) qu'elle ne souhaite pas trier ou ranger à part avec les autres (annuaire téléphonique, carnet d'adresses) car elle veut pouvoir y accéder instantanément, au premier coup d'œil, en cas de besoin. On peut en déduire qu'un système devrait dans ce cas pouvoir en offrir l'accès "au premier clic de souris". Ces observations aident à comprendre quels types d'annotations peuvent exister et révèlent tout un vocabulaire lié à l'utilisation et l'organisation de la mémoire documentaire ; elle sont cependant très consommatrices en temps ce qui nous empêche de répéter l'expérience pour faire une véritable généralisation.

La dernière technique de recueil que nous avons utilisée est l'*analyse de documents* collectés. Le rassemblement de documents typiques est une observation essentielle dans notre cas, étant donné que le projet CoMMA se concentre sur la recherche d'informations dans une mémoire documentaire d'entreprise. L'analyse de documents suppose une collecte systématique des documents pertinents pour les scénarios considérés. Cela inclut les documents textuels (lexiques, terminologies, rapports, etc.) qui peuvent faire l'objet d'une analyse de corpus par des outils de traitement de la langue naturelle, mais également les documents graphiques (formulaires, organigrammes...) exigeant une étude manuelle. La collecte des formulaires vides et remplis permet de voir leur utilisation, et leur suivi permet de découvrir leur cheminement dans l'organisation, et donc les flux de documents. Un exemple pour CoMMA est "la fiche d'itinéraire du nouvel employé" d'un de nos partenaires industriels : elle décrit ce qu'il faut faire, où aller, qui contacter, comment contacter et l'ordre dans lequel ces tâches doivent être effectuées par un nouvel employé. Ce document donne une idée des personnes impliquées dans le scénario d'intégration d'un nouvel employé, il indique l'existence d'un processus d'intégration dans l'entreprise et précise le vocabulaire requis pour décrire les différentes situations ou documents impliqués.

3.3 Réutilisation d'ontologies et autres sources d'expertise

Conception d'une ontologie pour un SMA de mémoire d'entreprise

Le recueil de données inclut également la réutilisation, la fusion et l'intégration d'*ontologies existantes*. Nous avons étudié les ontologies suivantes :

- Ontologie d'Entreprise (Uschold *et al.*, 1998)
- Ontologie de TOVE (TOVE, 2000)
- Ontologie Supérieure de Cyc® (Cyc, 2000)
- Ontologie PME (Kassel *et al.*, 2000)
- Ontologie CGKAT (Martin, 1996) & WebKB (Martin & Eklund, 2000)

La réutilisation d'ontologies est à la fois séduisante (elle devrait permettre d'économiser du temps et des efforts et elle favorise la normalisation) et difficile (les engagements et les conceptualisations doivent être réajustés entre l'ontologie réutilisée et l'ontologie désirée). Ces ontologies n'ont pas été importées directement ni traduites automatiquement. La meilleure façon de les réutiliser était d'analyser leur version informelle comme un document recueilli. Les différences entre les objectifs et les contextes de modélisation et d'utilisation de ces ontologies et ceux de notre ontologie, ainsi que les divergences dans la façon d'envisager une ontologie, nous ont amenés à penser qu'aucune importation automatique n'était possible et que la supervision humaine était obligatoire. Cependant les outils de traitement de la langue naturelle tels que ceux étudiés par la communauté TIA (Aussenac-Gilles *et al.*, 2000) pourraient aider l'analyse et l'examen de la version en langage naturel, et des traducteurs entre les différents formalismes permettraient de récupérer plus rapidement les pans identifiés comme intéressants et directement réutilisables.

Nous avons enrichi ces différentes contributions en considérant d'*autres sources informelles* fournissant une expertise dont nous ne disposions pas et qui n'était explicitée dans aucun des recueils précédents. Certaines sources très générales nous ont aidés à structurer les parties supérieures de certaines branches de notre ontologie. Par exemple des idées issues du livre "Using Language" d'Herbert H. Clark ont été utilisées pour structurer la branche des documents concernant les systèmes de représentation et pour injecter des réflexions issues de l'expertise sémiotique telles que la différenciation des systèmes de représentation iconique, indicielle et symbolique. D'autres sources très spécifiques nous ont permis de gagner du temps sur l'énumération de certaines feuilles de l'arbre taxinomique. Par exemple la norme MIME était une excellente source pour la description des formats électroniques et a permis d'énumérer les feuilles descendantes de la notion 'format de fichier de données'. Le Dublin Core a permis d'amorcer l'énumération des propriétés utilisées pour l'annotation d'un document dans sa globalité.

L'utilisation systématique des dictionnaires ou des lexiques disponibles est une bonne pratique. En particulier, les *méta-dictionnaires* se sont avérés extrêmement utiles. Ils permettent un accès simultané à tous les dictionnaires qu'ils répertorient qui contiennent la définition demandée et permettent ainsi de facilement comparer les différentes définitions et d'identifier ou de construire celle qui correspond à l'intension voulue. Nous avons intensivement utilisé le méta-dictionnaire en ligne OneLook (2000) qui a permis d'établir les premières expressions des intensions en anglais, le langage pivot du projet Européen.

Toutes les sources réutilisées ont été élaguées avant d'être intégrées dans l'ontologie pour supprimer les concepts que nous avons jugés inutiles. Les scénarios sont très utiles pour un tel élagage car ils capturent la portée de notre travail et une vision partagée par les différents membres du consortium. Ils peuvent être employés pour décider si une primitive est utile ou non en fournissant des exemples de situations où l'utilité de la primitive est manifeste. Une primitive qui n'est jamais utilisée est inutile, par exemple la relation de 'propriété' présente dans l'Ontologie d'Entreprise de AIAI n'a pas été réutilisée dans notre ontologie car elle n'était jamais exploitée dans nos scénarios.

3.4 Phase terminologique

Chaque terme dénotant un concept jugé utile pour le déroulement d'un scénario est candidat à l'analyse terminologique. Il entraîne avec lui tous les termes dénotant des notions proches et ses synonymes connus. Ainsi si nous considérons les termes *document*, *rapport* et *compte rendu d'analyse de tendance technologique* impliqués dans les scénarios de veille technologique, leurs candidatures sont très liées. Le point de départ peut être le terme *document* du fait d'une étude de la partie haute d'ontologies existantes, ou le terme *rapport* apparu lors de l'entretien avec la jeune assistante ou enfin le terme *compte rendu d'analyse de tendance technologique* à la suite de l'observation d'une instance particulière de ce document (ex: "compte-rendu d'analyse de tendance technologique intitulé 'Capitalisation des expériences du WAP et passage à la génération UMTS' "). Les termes candidats sont rassemblés dans un jeu de tableaux, passant ainsi à une structure de recueil semi-informelle. ACACIA a proposé des définitions en langage naturel pour chacun de ces termes candidats. Cette première version a été présentée aux membres du consortium et des extensions de bas niveau (termes et définitions) sont faites par les partenaires industriels, par exemple:

"Réfèrent de Domaine : Observateur responsable de l'expertise d'un domaine et qui est chargé de gérer un groupe d'observateurs contributeurs pour ce domaine."

L'intérêt d'avoir d'un côté la contribution d'un ingénieur de la connaissance pour la méthodologie et les concepts de haut niveau, et de l'autre celle des industriels pour les concepts spécifiques est flagrant.

L'étude terminologique est au cœur de l'ingénierie ontologique et le principal travail sur les termes identifiés est la production des définitions consensuelles qui expriment l'intension de chaque concept. Il existe trois cas de figure :

- Un terme correspondant à une et une seule notion : le terme non ambigu devient le libellé de la notion.
- Plusieurs termes correspondant à une notion : les termes sont des synonymes, on garde la liste des synonymes et le terme le plus communément utilisé devient le libellé de la notion.
- Un terme correspondant à plusieurs notions : il y a ambiguïté, le terme est noté comme étant ambigu. Les différentes expressions d'intension sont

définies et des termes non ambigus (souvent des termes composés) sont choisis comme libellés.

Etiqueter des concepts avec des termes est à la fois commode et dangereux. C'est une source importante de "rechute d'ambiguïté" où nous retombons rapidement et souvent inconsciemment dans l'ambiguïté en utilisant les libellés selon la définition que nous leurs associons naturellement, et non pas selon la définition qui leur a été réellement associée au cours de l'engagement sémantique. Comme le remarquent Uschold et Gruninger (1996) une bonne pratique pour des vérifications ponctuelles durant le travail de modélisation est de remplacer les libellés par des identificateurs sans signification. Cependant, utiliser des identificateurs sans signification pour représenter les intensions c'est rendre l'ontologie illisible pour les spécialistes du domaine comme pour le concepteur, pour lesquels la langue naturelle reste le premier moyen d'accès aux concepts et aux connaissances. Si du point de vue du système, les concepts tirent leur signification de l'expression formelle de leur intension et de leur position dans l'ontologie, il n'en est pas de même du point de vue l'utilisateur pour qui le libellé linguistique est la représentation privilégiée.

La représentation explicite du niveau terminologique et l'existence de fonctionnalités de gestion de ce niveau dans les outils assistant l'ontologiste sont donc réellement nécessaires. Tout au long de l'évolution de l'ontologie, nous essayons de capturer et de maintenir le niveau terminologique et le niveau intensionnel. La communauté du traitement automatique de la langue naturelle pourrait proposer beaucoup d'améliorations sur ces derniers points, en particulier en s'intéressant aux interfaces exploitant les relations entre les termes et les intensions pour désambiguer l'échange avec l'utilisateur.

3.5 Structuration de l'ontologie

Les concepts obtenus sont organisés en taxonomie. Les principes derrière cette structuration remontent à Aristote qui définissait une espèce en donnant son genre (genos) et son différentia (diaphora). Le genre est un genre plus général sous lequel l'espèce considérée se place. Le différentia est ce qui caractérise l'espèce dans ce genre. Ainsi nous avons commencé à regrouper des concepts premièrement de façon intuitive, puis en raffinant itérativement la structure suivant les principes aristotéliens étendus donnés par Bachimont (2000). Ces principes poussent à éliminer le multi-héritage au profit de la multi-instanciation. Ceci nous pose un problème avec l'introduction de 'rôles', dans l'ontologie qui ont tendance, à cause de RDFS, à rendre le multi-héritage nécessaire. Une solution serait d'explicitement des points de vue dans la taxonomie (Rivière, 1999) et de limiter l'application de ces principes étendus à l'intérieur d'un point de vue. Une approche est proposée dans (Kassel *et al.*, 2000) introduisant la notion d'axes sémantiques pour regrouper sous un genre les différentes dimensions ou natures de critères suivant lesquelles le différentia s'énonce. Les principes étendus de Bachimont peuvent alors s'appliquer pour des concepts issus des mêmes axes sémantiques. Il nous faudrait étudier comment une telle approche pourrait s'implanter en RDF. De même, le travail de Guarino et de Welty (Guarino, 1992 ; Guarino & Welty, 2000) contribue à poser les bases fondamentales de l'ingénierie

ontologique fournissant un cadre théorique, des définitions et des contraintes basées sur ces dernières permettant de vérifier la taxonomie : un ontologiste se fondant sur ces définitions peut ainsi contrôler certains aspects de validité de ses liens de subsomption. Le seul problème est que, pour autant que nous sachions, aucun outil n'est disponible pour aider un ontologiste à effectuer ce travail facilement et indépendamment d'un langage de formalisation ; cela peut devenir un travail titanesque d'appliquer ces théories à de grandes ontologies. Ces travaux semblent plus adaptés à la validation de la partie haute de l'ontologie qui, par extension, assure une cohérence minimale dans le reste de l'ontologie.

La façon de concevoir une ontologie est encore sujette à beaucoup de discussions dans la communauté. Notre compréhension des différentes contributions faites jusqu'à maintenant est qu'il y a une tendance à distinguer trois options de construction :

- Approche ascendante : L'ontologie est construite par généralisation en partant des concepts des basses couches taxinomiques. Cette approche encourage la création d'ontologies spécifiques et adaptées.
- Approche descendante : L'ontologie est construite par spécialisation en partant de concepts des hautes couches taxinomiques. Cette approche encourage la réutilisation d'ontologies.
- Approche centrifuge : La priorité est donnée à l'identification de concepts centraux à l'application que l'on va ensuite généraliser et spécialiser pour compléter l'ontologie. Cette approche encourage l'émergence de domaines thématiques dans l'ontologie et favorise la modularité.

Notre première idée était qu'il serait intéressant d'essayer une approche mixte (ascendante et descendante) alliant les atouts d'être spécifique et permettant de réutiliser d'autres ontologies. Après expérience nous ne sommes pas convaincus qu'il existe réellement une approche purement ascendante, descendante ou centrifuge. Pour nous, il s'agit de trois perspectives complémentaires d'une méthodologie complète. Lors de la structuration taxinomique, la spécialisation d'un concept générique, ou la généralisation d'un concept spécifique sont présentes de façon concourante à chaque niveau de profondeur (en bas, au centre ou en haut) et à différents niveaux de granularité (au niveau des concepts, au niveau de groupes de concepts). La nature holistique de la connaissance implique la nature holistique des ontologies, et la nature holistique des ontologies mène à la nature holistique des méthodologies pour les établir. Nous ne prétendons pas que pour une application donnée, la construction ne va pas se faire principalement selon une perspective (par exemple des ontologies de substances chimiques font une utilisation intensive de l'approche ascendante). Cependant nous ne pensons pas qu'il faille opposer les différentes approches, mais plutôt considérer qu'elles correspondent à trois perspectives combinées en ingénierie ontologique : quand un ontologiste conçoit une ontologie il doit effectuer les tâches définies dans ces trois perspectives en parallèle.

Pour CoMMA, nous avons effectivement effectué certaines tâches en parallèle dans différentes perspectives.

- Nous avons utilisé l'approche descendante quand nous avons étudié des ontologies de haut niveau et les couches hautes d'ontologies existantes pour structurer notre couche supérieure et réutiliser des distinctions déjà existantes, par exemple la distinction chose / entité / situation.
- Nous avons utilisé l'approche centrifuge quand nous avons étudié différentes branches, domaines, et micro-théories d'ontologies existantes ainsi que les thèmes principaux identifiés pendant le recueil pour identifier les domaines principaux dont nous avons besoin et regrouper les concepts.
- Nous avons utilisé l'approche ascendante quand nous avons exploité les rapports de l'analyse par scénarios et les traces du recueil pour identifier les concepts spécifiques et les regrouper par généralisation.

Les différents candidats dans chaque perspective (concept de haut niveau, concept central, concept spécifique) sont à l'origine de sous-taxonomies partielles. L'objectif est alors de faire se rejoindre ces différentes sous-taxonomies partielles ; chaque approche influence les autres, un événement dans une perspective déclenche des vérifications et des tâches dans les autres.

3.6 Du semi-informel au semi-formel

A partir de la terminologie informelle nous avons commencé à séparer les attributs et les relations des concepts et nous avons obtenu trois tableaux (figure 4b). Ces tableaux ont évolué depuis une simple représentation semi-informelle (tableaux terminologiques terme/notion) vers une représentation semi-formelle (relations taxinomiques, signatures des relations) dont la structure finale était:

- le libellé des concepts, relations ou attributs.
- les concepts liés par les relations ou le concept et le type élémentaire liés par l'attribut.
- le concept central le plus proche du concept considéré (personne, document...) ou les champs thématiques que la relation considérée relie.
- les liens d'héritage du concept ou de la relation.
- les termes synonymes du libellé.
- une expression de l'intension en langage naturel.
- les propriétés de la relation : symétrique, anti-symétrique, transitive, réflexive.
- la source de recueil ayant déclenché l'ajout.

Cette dernière colonne introduit le principe de traçabilité des concepts, des relations ou des attributs et elle est intéressante en particulier pour abstraire une méthodologie du travail effectué. Elle permet de savoir quel type de contribution a influencé une partie donnée de l'ontologie et d'évaluer l'efficacité et l'impact de la réutilisation ou d'une source de recueil particulière. Cependant cela est loin d'être suffisant et la logique de conception de l'ontologie devrait être capturée parce qu'elle explique ce qui a motivé sa forme actuelle et ceci peut aider la compréhension, l'adaptation et même l'adhésion à l'ontologie.

De cette approche résulte une première ontologie contenant 346 concepts organisés approximativement en trois couches :

- Une partie supérieure très générale qui ressemble plus ou moins à la partie supérieure d'autres ontologies.
- Une couche centrale qui tend à se diviser en deux zones distinctes : (1) une liée au domaine de la mémoire d'entreprise en général (documents, organisation, personnes...) (2) l'autre consacrée aux sujets du domaine d'application (par exemple pour les télécommunications: technologies sans fil, technologies réseau...).
- Une couche d'extension qui tend à être spécifique aux scénarios et à l'entreprise, avec des concepts complexes propres à l'organisation (rapport d'analyse de tendances, fiche d'itinéraire du nouvel employé...).

Nous étudions actuellement la possibilité de développer une nouvelle branche concernant le vocabulaire nécessaire aux agents pour se décrire, décrire les services, le système, etc. ainsi que la possibilité d'isoler l'ontologie du domaine voire de la formaliser différemment.

3.7 Formalisation de l'ontologie

Le travail de formalisation ne consiste pas à substituer une version formelle à une version informelle. Il s'agit d'augmenter une version informelle avec la correspondance formelle, des aspects sémantiques intéressants et pertinents de l'ontologie informelle afin d'obtenir une ontologie documentée (description en langage naturel éventuellement augmentée par des capacités de navigation introduites par la description formelle) et opérationnelle (description formelle des aspects sémantiques appropriés nécessaires aux opérations du système envisagé). L'ontologiste stoppe ce processus d'augmentation dès qu'il a atteint le niveau formel nécessaire et suffisant pour son système.

La version informelle de l'ontologie n'est donc pas une trace intermédiaire qui disparaît après la formalisation. L'ontologie formelle doit inclure les définitions en langue naturelle, les commentaires, les remarques, qui seront exploitées par les personnes essayant de s'appropriier l'ontologie. "Les ontologies doivent être intelligibles aux ordinateurs comme aux humains " (Mizoguchi & Ikeda, 1997). Ceci joue un rôle important dans la réutilisabilité des ontologies, le reengineering et le reverse-engineering. Les tableaux décrits précédemment ont évolué depuis une représentation semi-informelle vers une représentation semi-formelle par structuration itérative jusqu'à ce que la taxonomie soit suffisamment explicite pour être traduite en RDF (figure 4c). Un jeu de script utilisant les libellés non-ambigus nous a permis alors de basculer de la représentation semi-formelle des tableaux vers une représentation formelle en RDF incluant tous les aspects de la représentation semi-informelle mais reposant sur la structure forte et explicite donnée par les relations taxinomiques. Ce point de basculement ne change rien à l'ontologie en elle-même, il change uniquement sa représentation interne.

La figure 3 montre comment le langage RDF est utilisé pour implanter les différents niveaux introduits précédemment :

Conception d'une ontologie pour un SMA de mémoire d'entreprise

- le niveau terminologique où s'organisent les termes recueillis. Les relations entre le niveau intensionnel et le niveau terminologique dénotent les libellés possibles pour chaque intension (propriété `rdfs:label`) Une intension ayant des liens avec différents termes (ex. dans la figure 3 : I₄) est caractéristique de la synonymie de ces termes. Un terme ayant des liens avec différentes intensions (ex. dans la figure 3 : T₃) est caractéristique de l'ambiguïté de ce terme.
- le niveau intensionnel où s'inscrit la structure intensionnelle formalisant l'ontologie. Les relations entre le niveau intensionnel et le niveau extensionnel représentent l'instanciation de chaque classe de concept. Les faisceaux de liens relient une intension à son extension (ex. dans la figure 3 : I₈ et E₈).
- le niveau extensionnel où s'organise la mémoire factuelle (les annotations des documents, la description de la situation organisationnelle et les profils utilisateurs). Une extension reliées à plusieurs intensions (ex. dans la figure 3 : I₆ et I₇) est caractéristique de la multi-instanciation

Grâce à la technologie XML nous pouvons régénérer en permanence toutes les vues informelles ou semi-formelles précédentes en passant par des feuilles de style XLST (figure 4d):

- une feuille de style recrée le tableau terminologique initial, présentant une sorte de lexique de la mémoire.
- une feuille de style présente la table des concepts à partir du schéma RDFS et une seconde présente celle des relations et des attributs.
- un jeu de feuilles permet de rechercher des termes et de naviguer au niveau terminologique ou au niveau intensionnel en proposant en permanence de passer de l'un à l'autre. On peut ainsi naviguer dans la taxonomie des intensions de concepts ou de relations, représentées par les ensembles de synonymes. On peut consulter les signatures des relations portant sur un concept et demander à voir l'extension d'une intension. Un échantillon d'extensions joue le rôle d'exemplification automatique et permet d'augmenter la compréhension du concept.
- une dernière feuille propose une nouvelle vue présentant un arbre indenté des intensions. Leur définition associée apparaît dans une fenêtre popup suivant la souris.

L'aspect terminologique permet une meilleure navigation ainsi qu'un accès multilingue à la même structure des intensions et la même mémoire des extensions. Les feuilles de styles sont aussi utilisées pour documenter le résultat des requêtes en utilisant les définitions et les termes associés aux intensions. L'affichage des définitions dans de petites fenêtres popup est une première tentative pour désambigüer par anticipation la navigation ou les requêtes. Avant même que l'utilisateur ne clique, le système affiche sa définition en langage naturel du concept sur lequel se trouve la souris; cette fenêtre popup invite l'utilisateur à comparer sa définition personnelle avec la définition employée par le système et donc à éviter un malentendu.

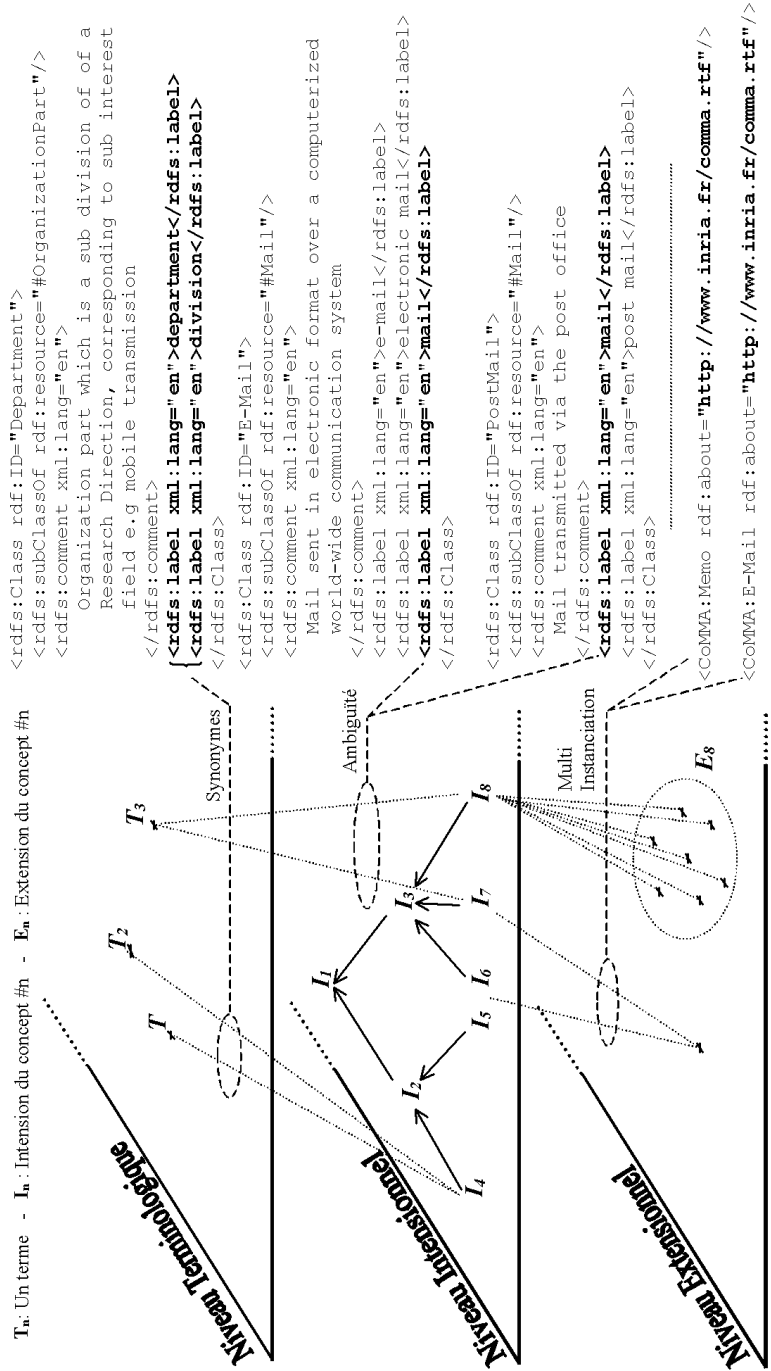


FIG. 3 - Niveaux terminologique, intensionnel et extensionnel dans la représentation d'une ontologie

(a) Scénarios et Recueil

Scenarios Observations Entretiens Documents

(b) Du semi-informel au semi-formel

Relation	Domain	Range	View	Super Relation	Other Terms	Natural Language Definition	Sy	Tr	Re	Pr	
Manage	Organizational Entity;	Organizational Entity;	Organization	Relation		Relation denoting that an Organizational Entity (Domain) is in charge/control of another			Tr		EO
Attribute	Domain	Range Type	View	Super Relation	Other Terms	Natural Language Definition	Pr				
Designation	Thing;	literal (string)	*			Identifying word or words by which a thing is called and classified or distinguished from others	Us				
Family Name	Person;	literal (string)	Person;	Designation;	Last Name;	The name used to identify the	Us				
Class	View	Super class	Other Terms	Natural Language Definition			Pr				
Thing	Top-Level;					Whatever exists animate, inanimate or abstraction.	Us				
Event	Top-Level; Event;	Thing;				Thing taking place, happening, occurring; usually recognized as important, significant or unusual	Us				
Gathering	Event;	Event;				Event corresponding to the social act of a group of Persons assembling in one place	Us				

(c) RDF

```

<rdfs:Class rdf:ID="Dictionary">
  <rdfs:subClassOf
    rdf:resource="#ReferenceDocument"/>
  <rdfs:comment xml:lang="en">Reference
    Document in which words are listed
    alphabetically and their
    meanings</rdfs:comment>
  
```

(d) Navigation et Utilisation

designatation : designatation.

Possible Matches :

The search was done in the terminology (📖).
Click on the 🔍 to get to see the corresponding concept.
Click on the 🔗 to get to see the corresponding relation.

Number of matches for the term "person": 5

- 🔍 has for personal interest. has for personal interest.
- 🔍 PDA / P.D.A. P.D.A., personal digital assistant.
- 🔍 person. person.
- 🔍 personal homepage. personal homepage.
- 🔍 salesperson. salesperson.

🔍 Gathering formally arranged for a particular purpose, usually in a dedicated room and/or around a table

- 🔍 social reunion
- 🔍 party event
- 🔍 SocialGathering
- 🔍 EntertainmentEvent

FIG. 4 - Etapes de conception de l'ontologie

3.8 Discussion

L'ontologie actuelle repose sur une taxinomie utilisant le multi-héritage et contenant 346 concepts, 51 relations et d'une profondeur maximale de 12 subsomptions. En ce qui concerne l'équilibre utilisabilité/réutilisabilité de l'ontologie, la partie haute très abstraite et la première partie de la couche centrale décrivant les concepts assez communs d'une mémoire organisationnelle (ex: personne, employé, document, rapport, groupe département, etc.) semblent réutilisables dans d'autres scénarios. La deuxième partie de la couche centrale est liée au domaine d'application (dans notre cas, les télécom), elle ne serait réutilisable que pour des scénarios dans un même domaine d'application. La dernière couche, qui étend les deux parties précédentes avec des concepts très spécifiques ne semble en aucun cas réutilisable dès que l'organisation ou le domaine d'application change.

Les partenaires industriels confrontés à l'ontologie actuelle trouvent sa complexité bien trop grande, ce qui la rend complètement hermétique. En particulier le niveau supérieur de l'ontologie introduit des distinctions philosophiques tout à fait intéressantes d'un point de vue de la modélisation (le haut de l'ontologie fournit des briques saines permettant de démarrer la modélisation et d'assurer certaines vérifications de la cohérence) mais très hermétiques et souvent inutiles pour l'utilisateur typique du système. En outre, plus on monte dans la taxonomie, plus le consensus est difficile. Deux collègues de travail seront plus facilement d'accord sur la modélisation de concepts qu'ils manipulent et échangent tous les jours (ex: une news est un type de document présentant de nouvelles informations) que sur le haut de l'ontologie qui demande un grand niveau de généralisation et d'abstraction et porte sur des concepts que nous ne sommes pas habitués à discuter et manipuler tous les jours (ex: les choses se divisent entre les entités et les situations, les entités étant des choses capables de jouer un rôle dans une situation). Le haut niveau relève souvent d'une culture voire de croyances personnelles et ses concepts sont utiles pour des manipulations internes au système (généralisation de requêtes, structuration des couches hautes) mais pas pour l'interaction directe avec l'utilisateur. D'un point de vue ergonomique, les interfaces utilisant l'ontologie doivent donc proposer des représentations filtrées et pédagogiques si l'on souhaite que l'utilisateur adopte le système. Un utilisateur qui se perd dans les méandres de l'ontologie ou qui se noie dans des considérations philosophiques à chaque fois qu'il doit formuler une requête ou écrire une annotation, ne restera pas longtemps un utilisateur. Nous envisageons d'utiliser le profil utilisateur pour noter les points d'entrée préférés dans l'ontologie, ce qui revient à écrêter l'ontologie et choisir des concepts centraux (ex : personne, document, organisation) à partir desquels l'utilisateur peut commencer sa navigation. L'aspect apprentissage symbolique cherche aussi à repérer les concepts les plus souvent utilisés afin d'améliorer la présentation de résultats.

Enfin, avec cette première version de l'ontologie, se posent maintenant les problèmes de son cycle de vie et de la maintenance, non seulement de l'ontologie mais aussi de ce qui a été construit au-dessus (annotations, modèles, inférences...). Il nous faut encore étudier comment gérer ce cycle de vie global.

4 CONCLUSION

Dans cet article, nous avons présenté une vision de la mémoire d'entreprise ainsi que l'approche évaluée dans le projet CoMMA basée sur les systèmes multi-agents, l'apprentissage symbolique et le Web sémantique. Nous avons montré que l'ontologie est une clef de voûte d'un tel système d'information multi-agents et nous avons détaillé chaque étape de sa conception. Nous avons notamment expliqué notre démarche de recueil et l'utilité de la définition de scénarios d'utilisation. Nous avons ensuite décrit comment d'autres ontologies et d'autres sources d'expertise ont influencé notre travail en soulignant qu'une réutilisation immédiate n'était pas possible. Nous avons détaillé le travail sur la terminologie en insistant sur la nécessité de préserver cet aspect dans l'ontologie. Enfin nous avons exposé comment le développement et la formalisation de l'ontologie se sont déroulés par structuration itérative et comment RDF et les feuilles de styles nous permettent de représenter et de naviguer dans l'ontologie au travers des niveaux terminologique, intensionnel et extensionnel. Tout au long de notre travail nous avons fait appel à de multiples outils informatiques ou conceptuels ; il ne fait aucun doute que l'ingénierie d'ontologie demande une convergence de toutes ces contributions vers un environnement intégré permettant de suivre tout le cycle de vie de l'ontologie.

5 REMERCIEMENTS

Nous remercions le consortium CoMMA pour les discussions fructueuses que nous avons eues et la Commission Européenne qui subventionne ce projet.

REFERENCES

- AUSSENAC-GILLES N., BIEBOW B, SZULMAN S (2000). Revisiting Ontology Design : a Method Based on Corpus Analysis, In Proc. EKAW'2000, Juan-les-Pins, p172-188.
- BACHIMONT B. (2000) Engagement sémantique et engagement ontologique: conception et réalisation d'ontologies en ingénierie des connaissances. In Jean Charlet, Manuel Zacklad, Gilles Kassel, Didier Bourigault, Ingénierie des connaissances Evolutions récentes et nouveaux défis. Eyrolles, ISBN 2-212-09110-9
- BERNERS-LEE T., HENDLER J., LASSILA O. (2001), The Semantic Web, In Scientific American, May 2001, p35-43
- CAROLL J. M. (1997) Scenario-Based Design, In M. Helander, T.K. Landauer, P. Prabhu Handbook of Human-Computer Interaction. 2nd edition, Ch. 17, Elsevier Science B.V.
- CoMMA (2000) Corporate Memory Management through Agents, In Proc. E-Work & E-Business, Madrid, p383-406, Octobre 2000
- CYC (2000) Ontology <http://www.cyc.com/cyc-2-1/cover.html> Cycorp Inc. Cyc®
- DIENG R., CORBY O., GIBOIN A., GOLEBIOWSKA J., MATTA N., RIBIERE M. (2000) Méthodes et outils pour la gestion des connaissances, Dunod
- FERNANDEZ M., GOMEZ-PEREZ A., JURISTO. N. (1997) METHONTOLOGY: From Ontological Arts Towards Ontological Engineering. In Proc. AAAI97 Spring Symposium Series on Ontological Engineering, p33-40, Stanford, USA

- GOMEZ-PEREZ A., FERNANDEZ M., DE VICENTE A., (1996) Towards a Method to Conceptualize Domain Ontologies, In Proc. Workshop on Ontological Engineering ECAI'96. p41-51
- GUARINO N, WELTY C, (2000) Towards a methodology for ontology-based model engineering. In Proc. ECOOP-2000 Workshop on Model Engineering. Cannes
- GUARINO N., GIARETTA P. (1995) Ontologies and Knowledge Bases: Towards a Terminological Clarification. In Mars, Towards very large knowledge bases, IOS Press.
- GUARINO N. (1992) Concepts, Attributes, and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases. In Data and Knowledge Engineering 8: p249-261
- KASSEL G., ABEL M.-H., BARRY C., BOULITREAU P., IRASTORZA C., PERPETTE S., (2000) Construction et exploitation d'une ontologie pour la gestion des connaissances d'une équipe de recherche. In Proc. IC'2000, p251-260, Toulouse
- LA FRANCE M. (1992) Questioning Knowledge Acquisition, In Lauer, ThoSMA W, Peacock Eileen, Graesser, Arthur C, Questions And Information Systems. Editions: L. Erlbaum associates, ISBN: 0-8058-11018-8.
- MARTIN P. (1996) Exploitation de Graphes Conceptuels et de documents structurés et Hypertextes pour l'acquisition de connaissances et la recherche d'informations. Thèse en informatique univ. Nice Sophia-Antipolis
- MARTIN P., EKLUND P., (2000) Knowledge Retrieval and the Word Wide Web. IEEE Intelligent Systems special issue Knowledge Management and the Internet, p18-25
- MIZOGUCHI RIICHIRO, IKEDA MITSURU, (1997) Towards Ontology Engineering, In Proc. of Joint 1997 Pacific Asian Conference on Expert systems / Singapore International Conference on Intelligent Systems, p259-266
- ONELOOK (2000) Online meta-dictionary available at <http://www.onelook.com/>
- RIBIERE M., (1999) Représentation et gestion de multiples points de vue dans le formalisme des graphes conceptuels. Thèse en informatique univ. Nice Sophia-Antipolis
- SEBILLOTTE S. (1991) Décrire des tâches selon les objectifs des opérateurs de l'Interview à la Formalisation. In Rapport de recherche de l'INRIA n°125
- STEELS L. (1993) Corporate Knowledge Management. In Barthès J. P., Proc. of ISMICK'93, p. 9-30, Compiègne
- TOVE (2000) <http://www.eil.utoronto.ca/tove/ontoTOC.html> Enterprise Integration Laboratory (EIL), Department of Industrial Engineering University of Toronto
- USCHOLD M., KING M., MORALEE S., ZORGIOS Y., (1998) The Enterprise Ontology. In Uschold & Tate The knowledge engineering review special issue on putting ontologies to use Vol. 13. Also available as AIAI-TR-195
- USCHOLD M. & GRUNINGER M. (1996) Ontologies: Principles, methods and applications. Knowledge Engineering Review, (11)2 p93-136,. Also available as AIAI-TR-191
- WOOLDRIDGE M., JENNINGS N.R., KINNY D. (1999) A Methodology for Agent-Oriented Analysis and Design. In Proc of Autonomous Agents '99 Seattle, ACM 1-58113-066-x/99/05