

Action Exmo

Échanges de connaissance structurée médiatisés par ordinateur

Rhône-Alpes

THÈME 3A



*R*apport
*d'**A*ctivité

2000

Table des matières

| | | |
|-----------|--|-----------|
| 1 | Composition de l'équipe | 2 |
| 2 | Présentation et objectifs généraux | 2 |
| 3 | Fondements scientifiques | 3 |
| 3.1 | Sémantique de la representation de connaissance | 3 |
| 3.2 | Transformations et propriétés | 4 |
| 3.3 | Interprétation des flux d'information | 4 |
| 4 | Domaines d'applications | 5 |
| 4.1 | Panorama | 5 |
| 4.2 | Ingénierie des systèmes de transformations | 5 |
| 4.3 | Technologies pour un « web sémantique » | 6 |
| 5 | Logiciels | 7 |
| 5.1 | Panorama | 7 |
| 5.2 | DLML: description logic markup language | 7 |
| 6 | Résultats nouveaux | 7 |
| 6.1 | Textes et représentations formelles | 7 |
| 6.2 | XML et représentation de connaissance | 8 |
| 6.3 | Étude de formalismes de flux d'information | 9 |
| 6.4 | Intelligibilité des représentations | 9 |
| 7 | Contrats industriels (nationaux, européens et internationaux) | 9 |
| 7.1 | Génie | 9 |
| 7.2 | Ardeco | 10 |
| 8 | Actions régionales, nationales et internationales | 10 |
| 8.1 | Actions nationales | 10 |
| 8.1.1 | Action de recherche coopérative ESCRIRE | 10 |
| 9 | Diffusion de résultats | 11 |
| 9.1 | Animation de la communauté scientifique | 11 |
| 9.2 | Enseignement | 11 |
| 9.3 | Participation à des colloques, séminaires, invitations | 11 |
| 10 | Bibliographie | 12 |

1 Composition de l'équipe

Responsable scientifique

Jérôme Euzenat [CR Inria]

Assistante de projet

Françoise de Coninck

Chercheurs doctorants

Olivier Brunet [allocataire Ministère]

Pierre-Antoine Champin [CIFRE (Dassault-Systèmes) en co-direction avec Alain Mille (LISI)]

Raphaël Troncy [CIFRE (Institut National de l'Audiovisuel) depuis le 1er décembre 2000]

2 Présentation et objectifs généraux

Mots clés : représentation des connaissances, sémantique des représentations, transformations, ontologies, treillis, préservation de propriétés, web sémantique, représentation du contenu, sémiologie, TROEPS, WWW, DLML, XML, XSLT.

EXMO étudie l'échange de connaissance structurée et formalisée. La connaissance est représentée dans des langages formellement définis. Ils peuvent aller de XML – métalangage structuré mais sans sémantique – aux langages de représentation de connaissance – structurés, sémantiquement définis mais spécialisés.

Le but de l'action EXMO est le développement d'outils théoriques et logiciels pour aider à l'organisation, la manipulation, la composition et la présentation d'éléments de connaissance structurés lors de la communication entre humains. Dans le processus de communication, l'ordinateur peut introduire une plus-value à son rôle de médium et de mémoire en accomplissant des tâches comme le formatage, le filtrage, la catégorisation, le test de consistance ou la généralisation.

Assurer l'adéquation et l'intelligibilité de la connaissance pour les interlocuteurs nécessite le développement d'une compréhension abstraite des représentations et des transformations qui leur sont appliquées. Les travaux de l'action EXMO sont focalisés sur deux aspects. L'aspect transformation rend compte des modifications de la connaissance pendant la communication alors que l'aspect communication concerne la préservation de l'intelligibilité de la connaissance transformée.

L'aspect transformation a pour but d'élaborer une « théorie générale des transformations » fondée sur les propriétés satisfaites par les transformations plutôt que sur les transformations elles-mêmes. Parmi ces propriétés on trouve la préservation du contenu ou de la structure, la traçabilité des sources ou, au contraire, la confidentialité.

La partie communication cherche à contribuer à l'intelligibilité de la connaissance échangée entre les interlocuteurs. L'un de nos premiers travaux est la description de la sémantique des types de documents XML à l'instar de leur syntaxe. Ainsi, la forme des documents peut être altérée tout en préservant tout ou partie de leur interprétation.

Les applications anticipées sont l'ingénierie des systèmes de transformation et le « web sémantique » (serveurs de connaissance, représentation du contenu, transformation de représentations).

3 Fondements scientifiques

3.1 Sémantique de la représentation de connaissance

Depuis plusieurs années, la sémantique des langages de représentation de connaissance (logiques de descriptions, graphes conceptuels et représentations de connaissance par objets) a été étudiée ^[Neb90]. Elle est définie en général à l'aide de la théorie des modèles.

On définit inductivement sur la structure du langage une fonction d'interprétation (I) vers un ensemble nommé domaine d'interprétation (D). Cette fonction va refléter la construction du sens d'une expression en fonction de ses composants. Les assertions de ce langage vont être satisfaites par une interprétation si elles remplissent une condition (en général être interprétées dans un sous-ensemble particulier du domaine). Dans ce cadre un modèle est une interprétation dans laquelle tous les axiomes sont satisfaits. Une expression (δ) est alors une conséquence d'un ensemble d'axiomes (Δ) si elle est satisfaite par tous leurs modèles (ce qui est noté $\Delta \models \delta$).

Pour un ordinateur, le but consiste à déterminer si une expression particulière (par exemple une requête) est conséquence des axiomes (par exemple, la base de connaissances considérée). Pour cela on développe des systèmes exécutables, appelés démonstrateurs, qui peuvent être fondés sur l'interprétation de règles d'inférence ou sur des programmes plus classiques. Ils permettent de déduire des théorèmes (ce qui est noté $\Delta \vdash \delta$). Ces démonstrateurs vérifient la propriété de correction s'ils ne répondent positivement qu'en cas de conséquence et la propriété de complétude s'ils répondent positivement pour toutes les conséquences. Cependant, suivant le langage et sa sémantique, la décidabilité – c'est-à-dire la possibilité de créer de tels démonstrateurs – n'est pas garantie. Même dans le cas de langages décidables, leur complexité algorithmique peut être prohibitive.

Pour cela on est amené à réaliser un compromis entre l'expressivité d'un langage et la complexité ou la complétude des démonstrateurs associés. Ce choix s'est traduit par la définition de langages à l'expressivité bornée comme les graphes conceptuels ou les représentations par objets ou de familles de langages modulaires dont les démonstrateurs peuvent être étendus, comme les logiques de descriptions.

L'action EXMO s'appuie d'abord sur des langages dont la sémantique est ainsi définie afin d'établir les propriétés des manipulations informatiques appliquées aux représentations.

[Neb90] B. NEBEL, *Reasoning and revision in hybrid representation systems, Lecture Notes in Artificial Intelligence 422*, Springer Verlag, Berlin (DE), 1990.

3.2 Transformations et propriétés

Les traitements appliqués aux représentations sont qualifiés de transformations. On se concentrera sur des transformations obtenues par assemblage de transformations élémentaires dont on ne se préoccupe pas de la description. On connaît par contre ses entrées et sorties ainsi que des propriétés liant les entrées aux sorties.

On définit un système de transformation par l'ensemble de ses transformations élémentaires et l'ensemble de ses assembleurs de transformations. Un flot de transformation est un ensemble d'instances de transformations élémentaires assemblées et reliées par des canaux. Un flot de transformations est lui-même une transformation.

Plus concrètement nos travaux concernent des transformations syntaxiques de documents XML (« eXtensible Markup Language ») encodant des langages de représentation de connaissance. Nous cherchons à exploiter le langage de transformations XSLT (« XML Stylesheet Language Transformations »^[Cla99]) recommandé par le W3C, pour lequel nous avons proposé un langage d'assemblage de transformations.

Le but d'EXMO est d'étudier les propriétés des transformations et leur combinaison. Une propriété est simplement un prédicat booléen portant sur une transformation (par exemple, « préserver l'information » est un tel prédicat – il est vrai ou faux d'une transformation – et se vérifie s'il existe un procédé algorithmique permettant, pour toute représentation Δ , de retrouver Δ à partir de son transformé $\tau(\Delta)$). Les questions intéressantes sont alors celles de la décidabilité d'une propriété pour un système de transformations particulier et celui de vérifier, pour un flot de transformations donné qu'il vérifie bien une propriété.

Nous identifions deux types de propriétés : des propriétés de préservation et des propriétés à motifs qui consistent à tester si la présence d'un motif dans la représentation de départ entraîne la présence de ce motif dans la représentation d'arrivée.

Les propriétés de préservation nous intéressent plus particulièrement. Elles peuvent permettre de garantir la préservation d'une relation d'ordre entre les représentations d'entrée (Δ) et de sortie ($\tau(\Delta)$) comme son anti-préservation. Pour cela on peut identifier des propriétés :

- syntaxiques : comme la structure ou l'ordre des éléments;
- sémantiques : comme la préservation des conséquences ($\tau(\Delta) \Rightarrow \Delta$, c'est-à-dire $\forall \delta, \Delta \models \delta \Rightarrow \tau(\Delta) \models \tau(\delta)$);
- sémiologiques : comme la préservation des interprétations (soit Σ les règles d'interprétation des expressions et \models^i l'interprétation de l'individu i , $\forall \delta, \forall i, j, \Delta, \Sigma \models^i \delta \Leftrightarrow \tau(\Delta), \tau(\Sigma) \models^j \tau(\delta)$).

3.3 Interprétation des flux d'information

L'analyse des flux d'information est centrée autour d'une formalisation à base de structures mathématiques appelées *treillis* et qui permettent d'avoir une notion d'imprécision dans la description de l'information. Les treillis sont des ensembles partiellement ordonnés, et cet ordre est utilisé pour préciser si un état (un élément du treillis) contient plus d'information

[Cla99] « XSL Transformations (XSLT) », *rapport de recherche*, 1999.

qu'un autre (dans ce cas, il est plus grand pour l'ordre). On peut ainsi définir l'absence totale d'information (l'élément minimal de l'ordre), et un élément plus petit qu'un autre est une approximation de celui-ci (car il contient moins d'informations, mais ne le contredit pas).

Ainsi, considérant deux systèmes échangeant de l'information, on peut écrire de façon rigoureuse qu'un système connaît précisément telle information d'un autre système, ou qu'il ne connaît cette information que partiellement, ou encore pas du tout. En se basant sur cette approche, on étudie la façon dont se composent et s'assemblent les transformations. Cela se traduit pour les propriétés sur ces systèmes par d'autres transformations que l'on dérive des précédentes.

Cette étude se base sur la comparaison de ce formalisme avec d'autres travaux existant concernant la représentation de connaissance et les échanges d'information. Les principaux travaux considérés sont la théorie des flux d'information (*Information Flow Theory* [BS98]) et l'analyse de concepts [GW99].

4 Domaines d'applications

4.1 Panorama

EXMO applique son savoir-faire à plusieurs domaines d'applications tels que la mémoire technique (voir §7.1) ou la réutilisation en conception (voir §7.2). Mais les efforts du projet sont maintenant dirigés vers deux types d'applications particuliers : l'ingénierie des systèmes de transformation (§4.2) et les technologies pour un « web sémantique » (§4.3)

4.2 Ingénierie des systèmes de transformations

Résumé : *La diffusion et la transformation de documents structurés induisent une ingénierie des systèmes de transformations qu'il faudra pourvoir en logiciels et en méthodes formelles qu'EXMO cherche à construire.*

De plus en plus, l'informatisation et la mise en réseau des organisations les conduisent à échanger de l'information sous forme électronique. Le commerce électronique engendre un échange permanent de tels documents.

Comme les informations diffusées ne sont pas destinées ni adaptées à l'ensemble des membres d'une organisation, il est nécessaire de transformer leur structure et leur contenu. De manière similaire, les sites du Web sont de plus en plus souvent engendrés à partir de bases de données ou de fonds premiers et les documents du commerce électronique subissent de nombreuses transformations jusqu'à ce que le produit soit envoyé.

Les besoins d'interopérabilité ont conduit à la définition du langage de définition de documents structurés XML. Des langages, tels que XSLT ou Omnimark, permettent d'implémenter des transformations autonomes, mais il semble inévitable que dans l'avenir, il faudra gérer des systèmes de transformation complexes (voir §3.2).

[BS98] J. BARWISE, J. SELIGMAN, *Information Flow, the Logic of Distributed Systems*, Cambridge university press, Cambridge (UK), 1998.

[GW99] B. GANTER, R. WILLE, *Formal Concept Analysis*, Springer Verlag, Berlin (DE), 1999.

L'ingénierie des systèmes de transformations nécessitera des outils, des méthodologies et des méthodes formelles. EXMO est concerné par les outils et les méthodes formelles et cherche à associer les deux en proposant des solutions pour la création de véritables ateliers de conception de systèmes de transformation (voir §5.1).

4.3 Technologies pour un « web sémantique »

Résumé : *Dans un univers où les documents du web seront annotés formellement, il sera nécessaire d'importer et de manipuler ces annotations en accord avec leur sémantique et en accord avec leurs utilisations. Nous travaillons aux technologies permettant cela.*

Un intranet (ou, plus généralement, l'utilisation des technologies d'Internet) est l'opportunité pour les entreprises d'accéder et de partager la connaissance bien souvent difficilement accessible sous forme documentaire. Les documents numériques et numérisés peuvent être rendus accessibles de manière standard et transparente auprès de tous les utilisateurs concernés. L'ambition, à terme, est de réaliser de véritables serveurs de connaissances permettant la recherche et la manipulation des ressources de l'entreprise.

Cependant, les limites de cette approche apparaissent rapidement : l'organisation des sites se révèle une tâche coûteuse et la recherche en texte intégral peu efficace. La recherche et l'interrogation d'un site en s'appuyant sur le contenu des documents sont une nécessité et les formalismes de représentation de connaissance sont de bons candidats pour représenter ce contenu. La représentation du contenu permettra de le manipuler pour faire de la recherche par analogie, par spécialisation, par similitude, etc.

La communication de représentations formalisées est une partie de la vision d'un « web sémantique ». Deux aspects de nos travaux sont concernés.

Tout d'abord, *ESCRIRE* (voir §8.1) pour la représentation formelle de contenu qui est le cœur du web sémantique et pour lequel nous recherchons en permanence des solutions en liaison avec des utilisateurs.

Ensuite, l'ensemble des travaux d'EXMO ont généralement pour but de concourir à une meilleure appréhension des contenus. Cela concerne bien entendu les travaux sur l'intelligibilité des connaissances communiquées mais aussi les travaux sur la transformation de représentations formelles. En effet, les travaux sur le web sémantique se basent actuellement sur la notion d'ontologie (que l'on peut rapidement décrire comme un schéma conceptuel de base de connaissance). Même s'il existe un jour un langage commun de représentation, il faudra importer les ontologies décrites dans d'autres langages d'une façon qui respecte la sémantique des langages manipulés. Apporter des solutions à ce problème fait partie des ambitions d'EXMO. Nous l'avons abordé au travers d'expérimentations sur les logiques de descriptions et de langages de description de syllogismes (voir §5.2), du langage de description de sémantique et de contact pris avec d'autres groupes de recherche.

5 Logiciels

5.1 Panorama

L'équipe d'EXMO développait le logiciel TROEPS au sein du projet SHERPA (Année 1999, projet SHERPA, section logiciels, module troeps). Si TROEPS est toujours disponible il n'est plus l'objet de nouveaux développements.

Les travaux de l'action EXMO sont susceptibles de donner lieu au développement de logiciels. Nous avons conçu un cadre de développement de transformations fondé sur XSLT qui devrait permettre d'expérimenter nos travaux sur les transformations en permettant d'appréhender les systèmes de transformations à manipuler et les propriétés qui leur sont associées. Nous recherchons activement des partenaires pour réaliser une implémentation viable de ce système.

Enfin, dans le cadre d'expérimentation nous avons été amenés à développer des couches logicielles relativement légères. C'est le cas de DLML (voir §5.2) décrit ci-dessous.

5.2 DLML: description logic markup language

Participant : Jérôme Euzenat [Correspondant].

DLML n'est pas un langage mais plutôt un système modulaire de description de types de documents (DTD) permettant de décrire la syntaxe et la sémantique de nombreuses logiques de descriptions (voir §3.1). Une importante motivation de DLML est de pouvoir incorporer des représentations en langage formel (ici les logiques de descriptions) dans des documents XML en particulier pour l'action Ecrire (voir §8.1).

DLML tire parti de la conception modulaire des logiques de descriptions pour définir les logiques par assemblage de constructeurs élémentaires. Le système contient la spécification de la syntaxe (DTD) et de la sémantique (DSD) de plus de 40 opérateurs élémentaires. Ainsi, la définition d'une logique est donnée par la spécification, en XML, des opérateurs qu'elle autorise. À partir de cette déclaration, DLML est capable d'assembler les spécifications élémentaires pour engendrer les DTD et DSD de la logique elle-même. Plus de 25 logiques répertoriées sont décrites sur le site.

Le langage DLML est associé à un ensemble de transformations XSLT permettant de transformer des descriptions DLML d'une logique vers une autre (soit en normalisant, soit en appauvrissant l'information), d'un langage (par exemple, de syllogismes) vers une logique de descriptions, d'afficher une description à l'aide de LaTeX ou d'importer des descriptions d'un autre système (FaCT de l'université de Manchester).

Le système peut être utilisé depuis <http://co4.inrialpes.fr/xml/dlml/>.

6 Résultats nouveaux

6.1 Textes et représentations formelles

Participants : Farid Cerbah [Dassault aviation], Jérôme Euzenat [Correspondant].

Une partie de nos travaux concerne la liaison entre les textes et une représentation forma-

lisée de leur contenu. Dans ce cadre, nous travaillons sur trois fronts :

- L'intégration au sein de documents (hyper)textuels de représentations formelles qui est détaillée au §6.2.
- L'annotation manuelle de corpus de textes par une représentation formelle de sorte de retrouver les documents contenant certains éléments de connaissance (voir §8.1).
- La construction automatique de représentations du contenu à partir des textes et le maintien de liens entre les deux représentations qui sont présentés ici.

Dans le cadre de l'action de développement GÉNIE (voir §7.1). Nous avons développé une chaîne d'acquisition de taxonomies de classes à partir des informations tirées des textes. Cette chaîne est fondée sur le logiciel d'acquisition de terminologies XTERM (développé par Dassault aviation) et le logiciel de représentation de connaissance TROEPS (développé par l'ancien projet SHERPA, (Année1999, projetSHERPA, sectionlogiciels, moduletroeps). XTERM est capable d'extraire une terminologie (ensemble hiérarchisé de termes) à partir de textes à l'aide de techniques linguistiques. Cette terminologie peut être manipulée par un utilisateur afin de sélectionner les termes du domaine et de choisir les termes les plus représentatifs. Nous avons ajouté la possibilité d'engendrer directement des classes de représentations de connaissance (en TROEPS) à partir des éléments sélectionnés dans la terminologie.

Dans certains contextes (génération de squelettes d'application à partir de spécifications, formalisation d'un corpus préexistant en gestion de connaissance), un tel dispositif permet une traçabilité immédiate des éléments formalisés vers leurs sources textuelles [9, 7, 8].

En effet, la terminologie reste accessible à partir des classes engendrées de telle sorte qu'il est possible de naviguer de manière raisonnée des classes vers les termes et ultimement vers les classes. Il est aussi possible de naviguer dans le sens inverse comme il reste possible de naviguer entre chacun des outils.

6.2 XML et représentation de connaissance

Participant : Jérôme Euzenat [Correspondant].

La question de l'encodage de représentations de connaissance dans le format de documents XML se pose parce qu'il existe plusieurs façons de le réaliser.

Nous avons d'abord étudié les différentes manières d'encoder les représentations de connaissance par objets [12]. Cette étude a été appliquée au langage TROEPS et utilisée dans le cadre de l'action de développement GÉNIE. Cette première expérience a permis de développer rapidement l'encodage d'un ensemble de langages, les logiques de descriptions, dans le langage DLML (voir §5.2). Elle a par ailleurs été le support au développement des langages pivots adoptés dans l'action ESCRIRE (voir §8.1).

DLML a permis d'expérimenter la transformation de langages formalisés à l'aide du langage XSLT. Mais pour pouvoir s'assurer du bien fondé de ces transformations, il faut pouvoir le faire en fonction de la sémantique du langage utilisé. Pour cela, nous avons conçu la notion de DSD (« Description de Sémantique de Documents ») qui permet d'exprimer la sémantique des formalismes en théorie des modèles. La construction d'une logique en DLML construit simultanément sa DTD et sa DSD.

6.3 Étude de formalismes de flux d'information

Participant : Olivier Brunet [Correspondant].

Nous avons étudié différents formalismes de représentation de l'information et de ces transformations (ou flux), en vue d'édifier des méthodes de preuves concernant certaines propriétés de ces transformations. Nous avons montré la possibilité de traduction entre les classifications définies dans la théorie des flux d'information et le formalisme sous forme de treillis. De plus, nous avons prouvé que ce dernier formalisme est plus expressif que le premier [6, 5].

6.4 Intelligibilité des représentations

Participant : Jérôme Euzenat [Correspondant].

Les transformations d'un langage dans un autre, même si elles préservent la sémantique, ne préservent pas forcément l'intelligibilité pour un interlocuteur humain. En effet, il est possible que les règles utilisées par les individus pour interpréter les représentations diffèrent et cela même lorsque les représentations sont purement formelles. Nous avons exhibé les problèmes posés par cela lors de la communication et la transformation de représentations formelles (hiérarchies de classes, feuilles d'évaluation d'articles. . .) [10]. Nous avons proposé un schéma de prise en compte de ces problèmes lors de la communication de connaissances formalisées en exprimant les règles d'interprétations jusque là laissées implicites de façon à préserver ces interprétations lors des transformations (resituant ainsi le problème sémiotique dans le cadre des transformations et propriétés, voir §3.2) [11].

7 Contrats industriels (nationaux, européens et internationaux)

7.1 Génie

Participants : Jérôme Euzenat, Farid Cerbah.

Le projet EXMO, à la suite du projet SHERPA, a participé à la phase 2 de l'action de développement GÉNIE (Génération Nouvelle d'Ingénierie), sur le thème 3 « Mémoire technique ». Le travail concernant EXMO est l'intégration de connaissances modélisées et de connaissances textuelles pour la traçabilité. La liaison entre TROEPS et XTERM, logiciel d'extraction de terminologie à partir de textes développé par Dassault Aviation, a permis d'explorer l'intérêt de l'utilisation de connaissances terminologiques dans la traçabilité des objets vers les sources textuelles qui les décrivent. Ainsi, à partir des documents, il est possible de construire une terminologie puis d'alimenter une base TROEPS par des classes correspondant aux termes (voir §6.1). Les classes sont communiquées à TROEPS à l'aide de leurs descriptions XML (voir §6.2). Ce faisant, les éléments TROEPS sont liés aux termes qui ont permis de les définir et permettent de remonter aux documents de spécification. Enfin, il est aussi possible d'indexer dynamiquement une base TROEPS préexistante à l'aide d'une terminologie (assurant ainsi une « traçabilité a posteriori »).

7.2 Ardeco

Participants : Jérôme Euzenat [Correspondant], Pierre-Antoine Champin.

EXMO participe au programme PROSPER du CNRS dans le cadre du projet ARDECO «Extraction, modélisation et gestion des connaissances réutilisables de conception» qui rassemble le projet EIFFEL à Rocquencourt, le laboratoire LACO (Poitiers) et le LISI (Lyon), ainsi que la société Dassault Systèmes. Les objectifs du projet sont d'expérimenter des modèles et des mécanismes susceptibles de participer à la réutilisation des connaissances en conception : objets, points de vue multiples, modèles de tâches, protocoles de construction coopérative, etc.

Pour cela, les mécanismes de réutilisation en conception mécanique sont étudiés parallèlement sous l'aspect informatique (Dassault, LISI, EXMO) et sous l'aspect psycho-ergonomique (EIFFEL, LACO). Les deux points de vue seront ensuite articulés de manière à formuler les objets informatiques (modèles, plans, contraintes...) et circonstances méthodologiques (phases de conception, objets et opérateurs manipulés...) pour lesquels une aide peut être produite [4]. Les activités du concepteur ont été analysées dans un modèle de tâches et d'essais permettant de leur appliquer un raisonnement à partir de cas.

8 Actions régionales, nationales et internationales

8.1 Actions nationales

8.1.1 Action de recherche coopérative ESCRIRE

Participants : Jérôme Euzenat [correspondant], Raphaël Troncy.

Il existe différents formalismes de représentation de connaissance et nul ne connaît exactement leurs qualités respectives. Le but de l'ARC ESCRIRE consiste donc à comparer trois types de représentations de connaissance (graphes conceptuels, représentations de connaissance par objets et logiques de descriptions) du point de vue de la représentation du contenu de documents et de sa manipulation. Pour cela, l'action s'appuie sur les compétences dans chacune des représentations des projets ACACIA, EXMO et ORPAILLEUR respectivement. L'objectif de l'action consiste à comparer les apports de chacun des types de représentation pour la représentation du contenu dans les serveurs de connaissances.

La mise à l'épreuve de ces différents formalismes pour le traitement d'un jeu de documents a nécessité une réflexion méthodologique sur le passage des textes à leur représentation formelle (de façon suffisamment indépendante des formalismes employés) en lien avec le type d'accès que l'on veut avoir sur ces documents [13].

On a donc défini en XML un ensemble de langages pivots (c'est-à-dire indépendants des représentations de connaissance utilisées) pour décrire des ontologies et des documents, des requêtes et des réponses à ces requêtes. On a par ailleurs développé une interface permettant de poser graphiquement la même requête à plusieurs systèmes et de collecter les réponses à ces requêtes.

Un ensemble de requêtes défini de manière coordonnée sera évalué dans chacun des contextes. À l'issue de ce travail, les différents formalismes seront comparés entre eux (mais aussi à la

recherche en texte intégral) selon un protocole prédéfini. Celui-ci devra apprécier des critères tant qualitatifs (expressivité des requêtes, accessibilité/lisibilité des informations, etc.) que quantitatifs (temps de réponse à une requête, taux de précision/rappel des réponses, etc.). Cette évaluation proposera une grille d'analyse des avantages et inconvénients d'un langage de représentation formel vis-à-vis de la recherche d'informations sur le Web.

On trouvera plus d'informations sur *ESCRIRE* à <http://www.inrialpes.fr/exmo/cooperation/escrire/>.

9 Diffusion de résultats

9.1 Animation de la communauté scientifique

- Jérôme Euzenat est membre élu du bureau de l'Association Française d'Intelligence Artificielle (AFIA).
- Comité de rédaction de la revue «L'objet» (Jérôme Euzenat).
- Comité de programme des éditions 2000 des colloques «Practical applications of knowledge management», «rencontres des jeunes chercheurs en intelligence artificielle», du séminaire «Knowledge management and organizational memory» de l'ECAI 2000 et du séminaire «Temporal and spatial granularity» de AAAI 2000 (Jérôme Euzenat).

9.2 Enseignement

- Coordination du profil «Modèles et systèmes de connaissance», du DEA «Informatique, systèmes et communication» en 2000 (Jérôme Euzenat).
- «Sémantique des représentations de connaissance», DEA d'informatique, communication et système (ISC), université Joseph Fourier et INPG, Grenoble, 18h, Jérôme Euzenat.

9.3 Participation à des colloques, séminaires, invitations

- Participation à la table-ronde «Objets et points de vue» du colloque LMO 2001 (Mont-Saint Hilaire, CA, janvier 2000) et exposé sur «Enabling technologies for semantic interoperability» (Jérôme Euzenat).
- Participation au séminaire «Semantics for the Web» (Dagstuhl, DE, 19-24 mars 2000) où Jérôme Euzenat a coordonné le groupe de travail sur «Semantic interoperability and top-level ontologies».
- Participation au séminaire «Semantic web technologies» (Luxembourg, LU, 22-23 novembre 2000) et exposé sur «Enabling technologies for semantic interoperability» (Jérôme Euzenat).

10 Bibliographie

Ouvrages et articles de référence de l'équipe

- [1] R. DUCOURNAU, J. EUZENAT, G. MASINI, A. NAPOLI (éditeurs), *Langages et modèles à objets : état et perspectives de la recherche, collection "Didactique", 19*, INRIA, INRIA Rocquencourt (FR), 1998, <http://co4.inrialpes.fr/lmobook/>.
- [2] J. EUZENAT, *Représentations de connaissance: de l'approximation à la confrontation*, Habilitation à diriger des recherches, Université Joseph Fourier, Grenoble (FR), janvier 1999.

Articles et chapitres de livre

- [3] A. NAPOLI, J. EUZENAT, R. DUCOURNAU, «Les représentations de connaissances par objets», *Techniques et science informatique 19*, 1-3, 2000, p. 387–394.

Communications à des congrès, colloques, etc.

- [4] P. BOUGÉ, D. DENEUX, C. LERCH, J. EUZENAT, J. BARTHÈS, M. TOLLENAERE, «Localisation des connaissances dans les systèmes de production: approches multiples pour différents types de connaissance», in : *journées Gestion de connaissances, coopération, méthodologie de recherches interdisciplinaires, Toulouse (FR)*, p. 31–50, 2000, <http://www.univ-valenciennes.fr/PROSPER/serveur/doc/toulouse/Prosper-loc-formatted.zip>.
- [5] O. BRUNET, «Classifications et treillis», in : *Rencontres nationales de jeunes chercheurs en intelligence artificielle, Lyon (FR)*, p. 29–38, 2000, <http://www710.univ-lyon1.fr/~rjcia/papiers/Brunet.pdf>.
- [6] O. BRUNET, «Lattice approach to classifications», in : *European summer school logic, language and information, Birmingham (UK)*, p. 34–44, 2000.
- [7] F. CERBAH, J. EUZENAT, «Integrating textual knowledge and formal knowledge for improving traceability», in : *ECAI 2000 Workshop Knowledge Management and Organizational Memory, Berlin (DE)*, p. 10–16, 2000.
- [8] F. CERBAH, J. EUZENAT, «Integrating textual knowledge and forml knowledge for improving traceability», in : *international conference on knowledge engineering and knowledge management, Juan-les-Pins (FR)*, R. Dieng, O. Corby (éditeurs), 1937, p. 296–303, 2000.
- [9] F. CERBAH, J. EUZENAT, «Terminology extraction techniques for improving traceability from formal models to textual requirements», in : *international conference on applications of natural language to information systems, Versailles (FR)*, 2000.
- [10] J. EUZENAT, «Problèmes d'intelligibilité et solutions autour de XML», in : *séminaire Valorisation des données, Labège (FR)*, P. Kopp (éditeur), Toulouse (FR), 2000, <ftp://ftp.inrialpes.fr/pub/exmo/publications/euzenat2000c.ps.gz>.
- [11] J. EUZENAT, «Towards formal knowledge intelligibility at the semiotic level», in : *ECAI 2000 Workshop Applied Semiotics: Control Problems, Berlin (DE)*, p. 59–61, 2000, http://www.iitp.ru/asc2000/ps/12_EUZEN.PS.

-
- [12] J. EUZENAT, « XML est-il le langage de représentation de connaissance de l'an 2000? », *in : journées langages et modèles à objets, Mont Saint-Hilaire (CA)*, p. 59–74, 2000, <ftp://ftp.inrialpes.fr/pub/exmo/publications/euzenat2000a.ps.gz>.

Divers

- [13] R. TRONCY, *Intégration texte-représentation formelle pour la gestion de documents XML*, Dea d'informatique, Université Joseph Fourier-INPG, Grenoble (FR), 2000, <ftp://ftp.inrialpes.fr/pub/exmo/reports/dea-troncy.ps.gz>.