POLYPUBLIE
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

Titre: Title:	Production ascendante d'ontologies légères sur le web sémantique : une application au référencement de sections de cours
Auteur: Author:	Yan Bodain
Date:	2007
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Bodain, Y. (2007). Production ascendante d'ontologies légères sur le web sémantique : une application au référencement de sections de cours [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/7998/

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/7998/
Directeurs de recherche: Advisors:	Jean-Marc Robert
Programme: Program:	Unspecified

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal

https://publications.polymtl.ca

UNIVERSITÉ DE MONTRÉAL

PRODUCTION ASCENDANTE D'ONTOLOGIES LÉGÈRES SUR LE WEB SÉMANTIQUE : UNE APPLICATION AU RÉFÉRENCEMENT DE SECTIONS DE COURS

YAN BODAIN
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIAE DOCTOR (PH.D.)
(GÉNIE ÉLECTRIQUE)
DÉCEMBRE 2007

© Yan Bodain, 2007.

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-37122-0
Our file Notre référence
ISBN: 978-0-494-37122-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

CONCEPTION ET VALIDATION D'UNE MÉTHODOLOGIE POUR LA PRODUCTION ASCENDANTE D'ONTOLOGIES LÉGÈRES SUR LE WEB SÉMANTIQUE : UNE APPLICATION AU RÉFÉRENCEMENT DE SECTIONS DE COURS

présentée par: BODAIN Yan

en vue de l'obtention du diplôme de: Philosophiae Doctor

a été dûment acceptée par le jury d'examen constitué de:

M. DESMARAIS Michel, Ph.D., président

M. ROBERT Jean-Marc, Doctorat, membre et directeur de recherche

M. GAGNON Michel, Ph.D., membre

M. PAQUETTE Gilbert, Ph.D., membre externe

Lorsque j'étais jeune, mon père me disait souvent que j'avais le défaut de tout savoir. Aujourd'hui je suis, moi aussi, devenu trop vieux pour tout savoir.

Je sais toutefois que je serais à tout jamais reconnaissant envers mes parents pour les sacrifices consentis pour assurer mon éducation.

REMERCIEMENTS

Mes remerciements vont d'abord à ceux qui ont accepté de participer à ce jury. Je remercie Monsieur Michel Desmarais qui me fait l'honneur de présider mon jury de thèse. Je remercie Messieurs Michel Gagnon et Gilbert Paquette du temps consacré à la lecture de cette thèse. J'ai sollicité leur concours en connaissant la qualité scientifique de leurs travaux et les éclairages multiples qu'ils étaient susceptibles de donner sur mes propres recherches.

Je témoigne de ma reconnaissance à mon directeur de thèse, Monsieur Jean-Marc Robert pour la confiance et la liberté qu'il m'a accordées. Je lui suis notamment reconnaissant pour la correction des quelques 200 pages que contient cette thèse.

Je tiens spécialement à remercier les gens qui ont accepté de participer à l'expérimentation du prototype logiciel et qui ont fourni de précieux commentaires sur la manière d'améliorer celui-ci.

Je remercie tous ceux sans qui cette thèse ne serait pas ce qu'elle est, aussi bien par les discussions que j'ai eu la chance d'avoir avec eux, que leurs suggestions ou contributions. Je pense ici en particulier à Katia Costa, Walter Cybis et Gracia Gingras.

Je tiens également à faire part de ma reconnaissance à Suzanne Guindon et Diane Bernier, pour leur gentillesse et leur professionnalisme. Sans leur aide précieuse, je n'aurais certainement pas pu faire face aussi facilement aux difficultés administratives rencontrées durant la réalisation de cette thèse.

Je suis finalement reconnaissant envers les organismes qui ont assuré le soutien financier de cette thèse, notamment la compagnie Solutions Orca inc. et la fondation McConnell (par le biais d'une bourse de recherche accordée au projet du "Sac d'école électronique" financé par l'Université de Montréal, l'École Polytechnique de Montréal et HEC Montréal).

RÉSUMÉ

Plusieurs initiatives de métadonnées ont été proposées pour améliorer le repérage des contenus sur le Web. Malgré leur niveau de sophistication élevé, aucune des initiatives de métadonnées ne réussit à s'adapter complètement au caractère atomique des objets d'apprentissage. La nature complexe des objets d'apprentissage demande, en effet, l'utilisation de structures de métadonnées flexibles pouvant être définies sur mesure pour répondre au contexte particulier de chaque objet. La définition d'une structure de métadonnées commune exige toutefois un effort de coordination important entre différents acteurs pour définir, maintenir et mettre à jour les éléments de descriptions. Le consensus nécessaire entre ces intervenants rend lui-même encore plus difficile la personnalisation des éléments de descriptions.

Il est toutefois possible de réaliser une description fine des objets d'apprentissage en insérant directement des annotations à l'intérieur des contenus Web. Une annotation est une note, une explication, ou tout autre type de remarque externe pouvant être attachée à un document sans toutefois être nécessairement insérée dans ce document.

Il est aussi possible de préciser la sémantique d'une annotation en utilisant des descriptions RDF (*Resource Description Framework*). Le RDF est une recommandation du W3C pour la description des ressources Web. Il s'agit d'un modèle de données pour la description de ressources qui peut donc être considéré, à ce titre, comme un modèle de métadonnées (ou méta-métadonnée). Les expressions RDF se présentent comme des triplets composés d'un sujet, d'un prédicat et d'un objet de relation. Les éléments de triplets RDF servent à indiquer qu'une ressource possède une propriété et une valeur donnée.

Une expression RDF peut faire référence à des ontologies pour préciser le sens d'une ressource Web. Une ontologie définit d'une manière formelle les connaissances communes d'un domaine particulier partagées entre différents utilisateurs. Les ontologies jouent ainsi le rôle d'une langue universelle, une sorte d'*interlingua*, qui

permet à des gens ou des applications d'échanger des informations sur une base commune. Ces informations concernent aussi bien les concepts que les rapports qui existent entre les différents éléments de connaissance d'un domaine

La conception d'ontologies reste une opération complexe qui demande un travail de réflexion important. Des ontologies réalisées de manière isolée par des individus différents peuvent ainsi donner naissance à des descriptions très différentes d'un même domaine. Une solution pour réduire l'hétérogénéité structurale et sémantique des ontologies consiste à mettre en place des équipes de travail qui réalisent ensemble la sélection et la définition des éléments d'une ontologie commune. Le recours à ces équipes spécialisées implique toutefois les mêmes inconvénients que ceux rencontrés dans la construction des descriptions proposées par les grandes initiatives de métadonnées.

Nous pensons qu'il est toutefois possible de réaliser des ontologies consensuelles sans nécessairement impliquer une équipe de conception spécialisée. Sur la base de notre propre expérience, nous avons constaté que lorsqu'un concepteur de cours récupère un contenu déjà annoté, celui-ci se montre aussi généralement intéressé à conserver la valeur des annotations récupérées. Ces mêmes annotations sont aussi souvent réutilisées de nouveau pour réaliser des descriptions supplémentaires. Nous pensons donc qu'il est ainsi possible de favoriser la construction d'ontologies en permettant simplement à des concepteurs de cours d'échanger librement des contenus annotés entre eux tout en permettant à chacun de rajouter/retrancher les descriptions sémantiques rattachées aux annotations récupérées. Nous croyons que les éléments d'ontologies récupérés par chacun seront ainsi systématiquement réutilisés pour favoriser la construction d'ontologies de plus en plus importantes.

Nous faisons ainsi l'hypothèse que les emprunts d'annotations réalisés successivement par différents concepteurs de cours se traduisent toujours par un bilan positif entre les ajouts et les retraits réalisés par chacun d'eux, réalisant ainsi un effet de levier positif sur

la production globale des annotations. Autrement dit, nous croyons que le nombre de descriptions augmente au fur et à mesure de l'implication d'un nouvel intervenant dans une chaîne de partage.

Pour vérifier cette hypothèse, nous avons réalisé une expérience avec huit sujets pour vérifier le taux de réutilisation des annotations et le taux de réutilisation des classes d'ontologie associées à ces mêmes annotations au travers des échanges successifs de contenus entre concepteurs de cours.

Nous avons construit un prototype logiciel qui permet la construction et l'échange d'annotations RDF associées à des ontologies OWL (*Web Ontology Language*). Les huit sujets avaient pour consigne d'échanger entre eux des contenus de cours et de modifier, si nécessaire, les annotations déjà réalisées par les autres. Les actions des différents utilisateurs étaient enregistrées par le logiciel. En étudiant le fichier de journalisation généré par le logiciel, nous avons démontré que le taux de réutilisation des annotations est de 88% alors que celui des classes d'ontologies échangées atteint 99%. Nous avons ainsi découvert l'existence d'un effet de levier important dans la conception de contenus annotés qui facilitera certainement la mise en place définitive du Web sémantique.

Les gains de cette découverte sont nombreux : notamment de ne plus être dépendant d'équipes spécialisées pour la production d'ontologies consensuelles, de réduire substantiellement la nécessité d'avoir à recourir à des techniques complexes d'alignement d'ontologies et de favoriser la capture des connaissances directement au niveau des concepteurs de contenu.

ABSTRACT

Several metadata initiatives have been proposed to improve the retrieval of content on the Web. Despite their high level of sophistication, metadata initiatives fail to adapt their description to the atomic nature of learning objects. The complex nature of learning objects requires the use of flexible metadata structures that can be tailored to meet the specific context of each object. The definition of a common metadata requires coordination between different actors to define, maintain and update the elements of descriptions. The necessary consensus between these stakeholders makes it even more difficult to customize the elements of descriptions.

However, it is possible to achieve a fine description of the learning objects by adding annotations directly within the web content. An annotation is a note, an explanation, or any other type of external remark that can be attached to a document without necessarily being included in this document.

It is also possible to specify the semantics of an annotation by using RDF descriptions (*Resource Description Framework*). The RDF is a W3C recommendation for describing web resources. It is a data model for the description of resource that can be considered as such, as a metadata model (or meta-metadata). The RDF descriptions are expressed as triples with a subject, a predicate and an object relationship. Elements of RDF triples are used to indicate that a resource has a property and a given value.

A RDF expression can refer to one or many ontologies to clarify the meaning of a web resource. An ontology defines a formal common knowledge in a particular field shared among different users. Ontologies are acting as a universal language, a kind of *interlingua*, which allows people or applications to share information on a common basis. This information relates to both the concepts and the relation between the different elements of knowledge in a domain.

The design of ontologies is a complex operation. Ontologies that are done in isolation by different individuals can give rise to very different descriptions of the same area. One way to reduce the structural and semantic heterogeneity of ontologies is to set up specialized teams that make the selection and definition of the elements of a common ontology. The use of these teams, however, involves the same disadvantages as those encountered in the construction of descriptions offered by the major metadata initiatives.

We believe that it is possible to achieve a consensus in ontologies without necessarily implying a specialized team. Based on our own experience, we have found that when a course designer retrieves content that is already annotated, he is also generally interested in retaining the value of annotations that were recovered. These annotations are often reused again to make additional descriptions. We therefore believe that it is possible to support the construction of ontologies by allowing course designers to freely exchange contents already annotated and to modify them if necessary. We believe that the ontology elements will be systematically reused to promote the construction of increasingly more important ontologies.

We make the hypothesis that the annotations being successively exchanged by different course designers will always be translate into a positive difference between the additions and the withdrawals made by each of them, thereby making a positive leverage effect on the overall production of annotations. In other words, we believe that the number of descriptions increases as the involvement of new users in a cycle of sharing gets more important.

To test this hypothesis, we conducted an experiment with eight subjects to test the rate of reuse of annotations and the rate of reuse in ontology classes associated with these annotations through successive content exchanges between designers.

We built a software prototype that allows the construction and exchange of RDF annotations associated with OWL ontologies (*Web Ontology Language*). The eight subjects were instructed to exchange course content between them and, if necessary, to

modify the annotations that have been achieved by others. The actions of individual users were recorded by the software. By studying the log file generated by the software, we have shown that the rate of reuse of annotations is 88% while that of ontology classes that had been exchanged reached 99%. We discovered the existence of an important leverage effect in the design of annotated contents which will certainly facilitate the definitive establishment of the Semantic Web.

Gains of this discovery are numerous: in particular no longer be dependent on specialized teams for the production of consensual ontologies, reduce substantially the need to resort to complex alignment techniques for ontologies, and encourage the capture of knowledge directly at the level of content developers.

TABLE DES MATIÈRES

DÉDICACE ..... **IV**
REMERCIEMENTS ..... **V**
RÉSUMÉ ..... **VI**
ABSTRACT ..... **IX**
TABLE DES MATIÈRES ..... **XII**
LISTE DES TABLEAUX ..... **XIV**
LISTE DES FIGURES ..... **XV**
LISTE DES ANNEXES ..... **XVII**
CHAPITRE 1 : INTRODUCTION **1**
CHAPITRE 2 : LES INITIATIVES DE MÉTADONNÉES **7**
2.1 INITIATIVES DE COLLABORATION 8
2.2 MÉTADONNÉES PÉDAGOGIQUES 10
2.3 OBJETS D'APPRENTISSAGE 14
2.4 PROBLÈME DE LA PERSONNALISATION DES STRUCTURES DE MÉTADONNÉES 16
CHAPITRE 3 : LES ANNOTATIONS **18**
3.1 ANNOTATIONS ÉLECTRONIQUES 21
3.1.1 Altered Vista 26
3.1.2 Annotea 26
3.1.3 Edutella 27
3.1.4 SMORE 28
3.1.5 Melita 29
3.1.6 KIM 30
3.2 CATÉGORISATION 31
CHAPITRE 4 : LES ONTOLOGIES **34**
4.1 CONTEXTE D'UTILISATION 37
4.2 LOGIQUE DE DESCRIPTION 38
4.3 STRUCTURE D'ONTOLOGIE 40
4.4 NIVEAUX DE DESCRIPTION 43
4.5 PROCESSUS DE DÉVELOPPEMENT 44
4.5.1 Méthodologie formelle de développement 46
4.5.2 Acteurs 50
4.5.3 Alignement des connaissances 52
CHAPITRE 5 : LE WEB SÉMANTIQUE **66**
5.1 RESOURCE DESCRIPTION FRAMEWORK (RDF) 67
5.2 SCHÉMA 68
5.3 OWL 71
5.4 ARCHITECTURE 72

CHAPITRE 6 : PROBLÈME DE DESCRIPTION DES CONTENUS 75

6.1 STRUCTURE DE DESCRIPTION DES MÉTADONNÉES 75
6.2 ANNOTATIONS MANUELLES 76
6.3 HÉTÉROGÉNÉITÉ DES DESCRIPTIONS RDF 78
6.4 ALIGNEMENT DES ONTOLOGIES 79
6.5 ACTEURS 82
6.6 QUESTIONS DE RECHERCHE 83

CHAPITRE 7 : MÉTHODOLOGIE 86

7.1 PROTOTYPE LOGICIEL 86
7.1.1 Annotation robuste 87
7.1.2 Requis 89
7.1.3 Démarche de conception 91
7.1.4 Preuve de concept 91
7.1.5 Éditeur d'ontologie 93
7.1.6 Générateur de pages Web 95
7.1.7 Architecture 97
7.2 EXPÉRIMENTATION 101
7.2.1 Sujets 101
7.2.2 Tâche expérimentale 102
7.2.3 Matériel 103
7.2.4 Plan de l'expérience 104
7.2.5 Procédure 107
7.2.6 Consignes 108
7.2.7 Questionnaire 109

CHAPITRE 8 : ANALYSE DES RÉSULTATS 110

8.1 DURÉE 110
8.2 RÉUTILISATION DES ANNOTATIONS 111
8.3 ONTOLOGIES PRODUITES 114
8.4 RÉUTILISATION DES ONTOLOGIES 122
8.5 ANNOGRAMME 127
8.6 FACILITÉ D'UTILISATION 132
8.7 COMPORTEMENT ET ATTITUDE 134
8.8 DISCUSSION 136
8.8.1 Qualité des ontologies produites 140

CHAPITRE 9 : CONCLUSION 143

9.1 CONTRIBUTION À L'AVANCEMENT DES CONNAISSANCES 144
9.2 SUITE DES RECHERCHES 146

BIBLIOGRAPHIE 148

ANNEXES 163

LISTE DES TABLEAUX

Tableau 2.1 : Initiatives de normalisation des métadonnées 13
Tableau 3.1: Caractéristiques des principaux outils d'annotation. 22
Tableau 3.2: Classement des outils d'annotation. 33
Tableau 4.1 : Catégories d'erreurs et de glissements de sens. 55
Tableau 7.1: Caractéristiques physiques des différents sujets. 102
Tableau 7.2: Description technique de la machine de table servant aux tests. 104
Tableau 8.1: Durée des séances. 110
Tableau 8.2: Données brutes sur l'utilisation des annotations. 111
Tableau 8.3: Taux de réutilisation des annotations. 113
Tableau 8.4: Répartition des classes pour chaque ontologie créée. 123
Tableau 8.5: Données brutes sur l'utilisation des classes d'ontologie. 123
Tableau 8.6: Analyse du taux de réutilisation des classes d'ontologie entre sujet. 124
Tableau 8.7: Distribution des réponses au questionnaire. 133

LISTE DES FIGURES

Figure 3.1 : Exemple de texte annoté. 18
Figure 3.2 : Récupération de pages annotées avec Annotea. 27
Figure 3.3 : Interface du logiciel d'annotation SMORE. 29
Figure 3.4 : Suggestions d'annotation proposées par Melita. 30
Figure 3.5 : Annotation automatique d'une page HTML avec KIM. 31
Figure 4.1: Ontologie comme langue universelle 37
Figure 4.2 : Comparaison de modèles. 40
Figure 4.3 : Modèle d'information basé sur trois ontologies distinctes 40
Figure 4.4 : Description des connaissances d'une entreprise 41
Figure 4.5: Découpage en quatre niveaux d'ontologies. 42
Figure 4.6 : Échelle de progression formelle des ontologies 44
Figure 4.7: Méta processus de connaissance 47
Figure 4.8: Processus de connaissance. 48
Figure 4.9: Deux processus pour réaliser la gestion des connaissances 49
Figure 4.10 : Acteurs impliqués dans le développement d'un système commun 51
Figure 4.11: Acteurs impliqués dans le développement d'un système de connaissance .52
Figure 4.12 : Taxonomie des problèmes associés au pairage des ontologies. 54
Figure 4.13 : Différences sémantiques et structurales entre deux ontologies. 56
Figure 4.14 : Processus de pairage 57
Figure 4.15 : Taxonomie des méthodes de pairage de schéma. 58
Figure 4.16 : Classification des méthodes de pairage de schéma 60
Figure 4.17 : Ancrage sur une ontologie de référence 61
Figure 4.18 : Validation structurale d'ontologies. 62
Figure 4.19 : Alignement d'ontologies avec COMA++. 63
Figure 4.20 : Alignement d'ontologies avec PROMPT. 64
Figure 5.2 : Schéma RDF partiel de l'élément *TITLE* 69
Figure 5.4 : Contenu partiel du schéma RDFS. 70
Figure 5.6 : Architecture du Web sémantique. 72

Figure 5.7 : Calendrier de lancement technologique du Web sémantique 73
Figure 6.2 : Différences entre ontologies selon leurs concepts généraux 81
Figure 7.1: Manipulations réalisées sur les annotations. 88
Figure 7.2 : Vue globale de l'interface de KATIA. 93
Figure 7.3: Menu contextuel pour l'édition de classe 94
Figure 7.9 : Exemple d'une page de cours à annoter. 103
Figure 7.10 : Répartition des contenus 105
Figure 8.1: Réutilisation des annotations. 112
Figure 8.2 : Ontologie réalisée sur papier par le premier sujet. 114
Figure 8.3 : Illustration de l'ontologie produite par le sujet 1 et le sujet 8. 115
Figure 8.4: Dispositif.owl 117
Figure 8.5: Clavier.owl 118
Figure 8.6: Stylet_lumineux.owl 119
Figure 8.7: Souris.owl 120
Figure 8.8: Microphone.owl 121
Figure 8.9: Réutilisation des classes d'ontologies. 125
Figure 8.10: Répartition des différentes actions des sujets. 125
Figure 8.12 : La spirale de connaissance du modèle SECI 138
Figure 9.1 : Pyramide méthodologique. 143

LISTE DES ANNEXES

ANNEXE A : CERTIFICAT DU COMITÉ D'ÉTHIQUE..... 163
ANNEXE B : FORMULAIRE DE CONSENTEMENT DES SUJETS..... 164
ANNEXE C : QUESTIONNAIRE..... 166
ANNEXE D : CONTENU DE COURS..... 169
ANNEXE E : ONTOLOGIES..... 174
ANNEXE F : ANNOGRAMME..... 184

CHAPITRE 1 : INTRODUCTION

L'une des grandes difficultés des concepteurs de cours est de repérer les contenus déjà disponibles sur le Web. Il existe des repositoires qui permettent l'entreposage des contenus de cours mais ces systèmes imposent des structures d'indexation particulièrement rigides qui s'adaptent mal aux besoins des utilisateurs Web. La recherche à l'intérieur de ces repositoires donne généralement des résultats décevants. Les repositoires de contenus n'ont ainsi jamais connu une grande popularité alors qu'un bon nombre d'utilisateurs se tournent plutôt vers le Web pour réaliser leurs recherches.

La recherche de contenu de cours sur le Web ne donne pourtant pas nécessairement de meilleurs résultats que la recherche de contenu dans les repositoire de cours. Les moteurs de recherche sur le Web s'appuient sur des méthodes d'analyse syntaxiques et structurales. Un bon exemple de méthode syntaxique consiste à utiliser des mots clés à l'intérieur d'un index ou d'un champ d'information pour effectuer des recherches. Il s'agit actuellement de la forme de recherche la plus couramment utilisée sur le Web alors que celle-ci ne donne généralement que des résultats imprécis sous la forme de centaines, voire de milliers, de réponses n'ayant que très peu de ressemblance avec le sujet cherché.

En raison du caractère non structuré et hétérogène des ressources Web, chaque site Web utilise une terminologie différente pour décrire des choses similaires. L'insertion de métadonnées à l'intérieur des pages Web n'améliore généralement que très peu la recherche d'information car ces mots clés ne possèdent souvent aucune signification contextuelle particulière. L'utilisation de mots clés à partir d'une liste de termes contrôlés n'est pas pour autant plus avantageuse puisque cette méthode demande un effort important d'indexation et ne garantit par pour autant l'expression fidèle des différents contextes d'utilisation des contenus Web.

Le rapprochement des contenus Web peut être réalisé au moyen d'une correspondance sémantique. Il est en effet possible d'ajouter directement une sémantique aux pages Web au moyen d'un cadre de description de ressource appelé RDF (*Resource Description Framework*). Le RDF a été initialement proposé par le consortium Web (W3C) pour modéliser des métadonnées. Il est aujourd'hui utilisé comme un système général pour modéliser l'information sur le Web.

Le modèle de métadonnées du RDF s'appuie sur la description des ressources au moyen d'assertions prenant la forme d'expressions sujet-attribut-objet, appelées des triplets dans la terminologie RDF. Le sujet dénote une ressource, alors que l'attribut dénote des traits ou des aspects de cette ressource et exprime un rapport entre le sujet et l'objet. Par exemple, une façon d'exprimer qu'un « document a pour auteur Yan » en RDF consiste à établir un triplet sous cette forme : un sujet dénotant l'adresse URI du document en question, un attribut « author », et un objet dénotant «Yan». Les triplets RDF permettent ainsi la description sémantique d'une ressource Web sans toutefois nécessairement recourir à un modèle ontologique.

Une « ontologie » est une formulation d'un schéma conceptuel complet et rigoureux d'un domaine. Les ontologies se présentent généralement comme une structure hiérarchique de données contenant tous les éléments nécessaires à la définition d'un domaine. Les concepts de base d'une ontologie sont : les classes, les propriétés de classes et les relations entre classes. Les ontologies peuvent être exprimées en RDF au moyen d'un langage appelé OWL (*Web Ontology Language*).

Pour résumer, le RDF peut directement être utilisé pour décrire les ressources Web alors que OWL peut être utilisé pour exprimer des concepts ontologiques reliés aux descriptions RDF. L'utilisation de RDF et de OWL reste néanmoins problématique parce qu'il n'existe actuellement aucune adoption répandue de ces normes pour les créateurs de pages et les outils de gestion de site Web. Ces normes doivent être employées avant même que les agents logiciels destinés à exploiter ces descriptions sémantiques puissent

être conçus. De plus, rien ne prouve que les contenus ainsi décrits pourront être facilement classés, catalogués, ou extraits pour supporter la mise en place d'un « Web sémantique ».

Le concept du Web sémantique marque une avancée importante en terme de précision et de rapidité d'utilisation des contenus Web. Le Web sémantique fournit un cadre d'interopérabilité commun qui permet à des données d'être partagées et réutilisées à travers différentes applications, entreprises et domaines tout en traversant les frontières érigées entre différentes communautés d'utilisateurs. Ce projet d'un Web sémantique universel est actuellement supporté par le W3C, le DARPA (*Defense Advanced Research Projects Agency*) et un nombre important de chercheurs et d'associés industriels.

Le fonctionnement actuel des engins de recherche repose sur des méthodes syntaxiques et structurales. Le défi actuel est de favoriser l'inclusion de descriptions sémantiques à l'intérieur des pages Web pour mieux décrire les contenus tout en répondant aux exigences des propositions actuelles du Web sémantique. Ces descriptions sémantiques devraient idéalement relier les contenus Web à des ontologies pour faciliter le repérage et la compréhension des concepts utilisées. Il existe une multitude d'ontologies déjà disponibles sur le Web et il devrait théoriquement être facile d'utiliser ces ontologies pour améliorer la description des contenus Web. Plusieurs ontologies différentes peuvent toutefois être utilisées pour décrire une même entité. Il est donc important de développer des moyens de faciliter l'alignement des concepts équivalents issus de différentes ontologies lorsque ceux-ci sont utilisés pour décrire la même chose. Il existe donc actuellement un triple problème dans la conception et l'utilisation de descriptions pour le Web sémantique. D'abord, nous ignorons encore comment réaliser l'intégration des descriptions sémantiques à l'intérieur de la tâche générale de conception de contenu Web. Ensuite, nous ignorons comment nous assurer de l'alignement des ontologies utilisées pour décrire des concepts similaires. Finalement, nous ignorons encore

comment appuyer le développement d'ontologies consensuelles dans un contexte décentralisé.

La thèse que nous présentons ici propose une méthodologie originale pour réaliser la production de contenu pour le Web sémantique en favorisant l'intégration de la tâche de description sémantique à même la tâche de conception de contenu, et en favorisant l'alignement des concepts utilisés pour la description de ces contenus tout en s'assurant que ces descriptions sémantiques pourront évoluer de manière parallèle sans jamais compromettre l'alignement général des ontologies utilisées.

La contribution certainement la plus importante de cette thèse reste toutefois la présentation d'une méthodologie nouvelle pour la conception d'ontologie qui offre l'avantage de ne pas exiger la mise en place d'équipes spécialisées pour la production d'ontologies consensuelles, de réduire substantiellement la nécessité d'avoir à recourir à des techniques complexes d'alignement d'ontologies et de favoriser la capture des connaissances directement au niveau des concepteurs de contenu.

Cette thèse est construite de la manière suivante :

- Le chapitre 2 présente l'état de l'art sur l'utilisation des métadonnées pour la description de contenus de cours par le biais d'une revue des principales initiatives de métadonnées reliées au domaine de l'éducation. Une définition formelle des objets d'apprentissage y est aussi proposée.
- Le chapitre 3 présente l'état de l'art sur la conception d'annotation en présentant les différents systèmes actuellement disponibles pour réaliser l'annotation de document sur le Web. Un classement formel des différents systèmes d'annotation y est présenté.
- Le chapitre 4 présente l'état de l'art sur la conception d'ontologie. Une définition formelle du concept d'ontologie est présentée. Les principes de conception d'ontologies sont résumés brièvement ainsi qu'une liste des moyens mis en oeuvre pour réaliser l'alignement des ontologies. Il est aussi question de la

composition des équipes de conception nécessaires à la mise en place d'ontologies consensuelles.

- Le chapitre 5 fait l'état de l'art des technologies du Web sémantique en présentant le principe de fonctionnement des descriptions RDF, des schémas RDFS ainsi que celui des ontologies DAML+OIL et OWL. L'architecture générale du Web sémantique est aussi présentée de manière graphique.
- Le chapitre 6 présente les nombreux problèmes associés à la description des contenus sur le Web. Ces problèmes concernent les difficultés d'utilisation des métadonnées, les difficultés de production des annotations manuelles, les problèmes causés par l'hétérogénéité des descriptions RDF, la difficulté de réaliser l'alignement des ontologies et la nécessité de simplifier le rôle des différents acteurs impliqués dans la création des ontologies consensuelles. Finalement une question de recherche est proposée pour voir s'il existe un moyen d'atténuer ces différents problèmes en donnant directement la possibilité aux concepteurs de contenus de récupérer des contenus annotés et d'utiliser ceux-ci pour réaliser la diffusion et la production indirecte d'ontologies sur le Web sémantique.
- Le chapitre 7 présente la méthodologie utilisée pour résoudre la question de recherche. Cette méthodologie consiste en la création d'un prototype logiciel et à son utilisation avec huit sujets pour évaluer le taux de réutilisation de contenus de cours annotés.
- Le chapitre 8 présente les résultats d'expérience qui sont à la fois issus de l'observation directe des ontologies crées par les différents sujets et à la fois issus des traces d'activités enregistrées par le logiciel. L'analyse de ces résultats permet ainsi de confirmer l'hypothèse émise précédemment en précisant le taux exact d'annotations et de classes d'ontologies récupérées entre les différents sujets.

- Le chapitre 9 présente finalement la conclusion en résumant la contribution de cette thèse à l'avancement des connaissances.

CHAPITRE 2 : LES INITIATIVES DE MÉTADONNÉES

La structure du Web est particulièrement bien adaptée à la navigation entre les éléments mais cette structure pose néanmoins des problèmes particuliers pour le repérage des contenus. Le Web présente une structure réticulaire où chaque nœud (document) est relié aux autres par le biais d'hyperliens. Cette disposition en graphe favorise une navigation aisée entre les éléments mais cette disposition rend aussi, en même temps, plus difficile le repérage et l'indexation des ressources sur le Web.

Les engins de recherche réussissent actuellement à donner une image organisée du Web bien que leur logique d'indexation repose principalement sur un traitement lexical de l'information. Cette façon de procéder ne protège en aucune manière les engins de recherche des problèmes associés à la complexité des contenus Web, à l'augmentation constante du nombre de documents et à la volatilité relative des informations repérées. Pour [Lawrence, 1999] et [Risvik, 2002], la progression constante du Web entraînera inévitablement une diminution de la couverture des engins de recherche:

- Les documents d'origine personnelle augmentent considérablement alors que le pourcentage de contenu approuvé par une forme de politique éditoriale diminue de manière proportionnelle. Il devient ainsi de plus en plus difficile d'évaluer la qualité et la véracité des informations trouvées.
- L'utilisation de formats variés rend les tâches de mises à jour de plus en plus complexes. Cette complexité a, à son tour, pour effet de diminuer la fréquence de rafraîchissement des contenus et de diminuer, encore plus, la valeur des informations trouvées.
- L'insertion de métadonnées (balise HTML <meta>) à l'intérieur des pages Web ne permet généralement pas d'améliorer le classement de l'information. Ces balises sont, la plupart du temps, utilisées d'une manière peu rigoureuse et certaines données sont parfois même volontairement faussées pour influencer l'ordre de classement des engins de recherche.

- La logique de fonctionnement des engins de recherche n'est pas toujours transparente pour les utilisateurs. La plupart des engins de recherche utilisent des algorithmes qui augmentent artificiellement la visibilité de certaines pages Web selon des paramètres propriétaires; il devient ainsi de plus en plus difficile pour les utilisateurs d'évaluer la qualité des recherches effectuées et d'anticiper les critères de sélection qui fonctionneront le mieux avec ces engins de recherche.

Le Web facilite les recherches des étudiants mais, d'une manière tout à fait paradoxale, n'améliore aucunement le repérage des contenus de cours dans un contexte de réutilisation entre professeurs. Pour un professeur, le repérage des contenus doit souvent correspondre à des objectifs pédagogiques, des contraintes de temps ou d'équipement qui ne sont pas nécessairement importantes pour un étudiant qui parcourt librement le Web. Les différents formats utilisés pour la production de documents numériques (Word, Powerpoint, PDF, ...) facilitent encore moins la recherche de ces contenus. Plus grave encore, les concepteurs de cours doivent généralement valider eux-mêmes la qualité et la véracité des éléments trouvés en réalisant souvent de nouvelles recherches complémentaires.

2.1 Initiatives de collaboration

Plusieurs initiatives ont été proposées pour améliorer le repérage et la récupération des contenus de cours. Ces initiatives prennent la forme d'un système de description commun qui facilite la comparaison entre les éléments. Il existe plusieurs initiatives de descriptions communes :

- Aux États-Unis, une division du *National Science Foundation* fournit un appui financier aux institutions qui valorisent un rapprochement entre collèges et universités (*Educational System Reform*[1]). Un portail appelé SMET[2] (*Science, Math, Engineering and Technology*) permet déjà d'y déposer des ressources

[1] http://www.ehr.nsf.gov/esr/
[2] http://www.smete.org/

académiques pour les rendre disponibles à autrui. Certaines universités américaines possèdent aussi leur propre portail de cours; le MIT a notamment rendu publique l'ensemble de ses contenus de cours (*MIT OpenCourseWare*[3]) au moyen d'une architecture informatique dédiée à cette usage (*DSpace*[4], *Open Knowledge Initiative*[5]).

- En Europe, la fondation ARIADNE[6] encourage le développement d'entrepôts communs pour favoriser la réutilisation des ressources pédagogiques au sein de l'union européenne. Le projet ARIADNE compte ainsi déjà plus de 31 partenaires différents.
- Au Canada, CANARIE soutient le projet CanCore (*Canadian Core Learning Object Metadata Application Profile*[7]) qui fait la promotion d'une norme d'échange commune entre les différentes institutions d'éducation du pays. CANARIE finance notamment l'université d'Alberta pour le développement d'un repositoire de cours appelé CAREO[8] (*Campus Alberta Repository of Educational Objects*).
- Il existe aussi plusieurs autres projets tels que EdnA, HEAL, iLumina, Learn-Alberta, Lydia, MERLOT, etc.

La majorité de ces initiatives ont en commun l'utilisation de métadonnées pour décrire les contenus de cours. Cette description porte sur des informations aussi diverses que la nature des contenus, leur usage distinctif ou les informations particulières sur leur mode d'accès. Contrairement aux balises <meta> insérées à l'intérieur des pages Web, les métadonnées utilisées dans ces projets présentent une structure spécialement adaptée au domaine de l'éducation.

---

[3] http://ocw.mit.edu/
[4] http://dspace.org/
[5] http://web.mit.edu/oki/
[6] http://www.ariadne-eu.org/
[7] http://www.cancore.org/
[8] http://careo.netera.ca/

2.2 Métadonnées pédagogiques

Le terme "métadonnée" signifie littéralement "données sur les données" et représente une donnée qui fournit de l'information sur la nature d'une autre donnée. Dans un contexte pédagogique, les métadonnées sont utilisées pour décrire la nature et le contexte d'usage des objets d'apprentissage. Cette description porte aussi bien sur l'objet même de la représentation que sur les propriétés et les éléments d'interaction qui leurs sont associés [Merril, 1998].

Les métadonnées peuvent être séparées des ressources qu'elles décrivent par une certaine distance physique. Par exemple, le résumé d'un livre peut se retrouver directement sur la couverture de l'ouvrage mais il peut aussi se retrouver sur une fiche bibliographique totalement isolée de celui-ci. Suivant cette même analogie, différents types de métadonnées peuvent être utilisés conjointement pour décrire un même contenu et une distance physique plus ou moins grande peut exister entre les métadonnées et l'objet qu'elles décrivent.

Les métadonnées peuvent ainsi prendre différentes formes. Pour [Kashyap, 1996] et [Shah, 1998], les métadonnées se divisent en deux catégories distinctes :

1) **métadonnées indépendantes du contenu**: qui ne possèdent aucune dépendance au contenu représenté (ex., adresse, date de création);
2) **métadonnées dépendantes du contenu**: qui capturent l'information du contenu représenté. Cette catégorie se subdivise elle-même en deux :
    a) **métadonnées liées au contenu même**: qui dépendent directement du contenu présent (ex., mot clé apparaissant directement dans un document);
    b) **métadonnées liées à la description du contenu**: qui utilisent les éléments de descriptions associés à l'objet représenté et qui ne font pas directement partie du contenu représenté (ex., annotation textuelle

associée à une image). Ce type de métadonnées se subdivise à son tour en deux sous-catégories :

i. **métadonnées indépendantes du domaine**: qui capturent l'information présente de manière indépendante au domaine d'information représenté (ex., la définition DTD d'un document XML);

ii. **métadonnées dépendantes du domaine**: qui capturent l'information présente d'une manière spécifique au domaine d'information. Le choix du vocabulaire devient alors une considération importante et le vocabulaire doit être choisi en suivant les règles particulières du domaine (ex., définition d'une ontologie qui utilise des termes et une structure particulières à un domaine de connaissances).

Le niveau de dépendance plus ou moins grand des métadonnées envers leur contenu peut rendre plus difficiles les opérations de maintenance et de mise à jour des métadonnées. À titre d'exemple, le choix d'un vocabulaire spécialisé demande la mise en place d'un consensus entre les différents intervenants pour définir précisément la valeur et l'usage de chaque élément. La sélection d'un vocabulaire commun est, en plus, une tâche complexe qui exige des efforts de collaboration de longue durée entre différents types de spécialistes pour assurer à la fois la définition, la mise à jour et la révision des métadonnées:

> *«Effective research progress on metadata will need to involve intense collaboration between metadata specialists and communities trying to solve functional problems, such as rights management, resource discovery, archiving and preservation, or the organization and management of data specific to various disciplines.*

> *The definition and maintenance of metadata standards over time is a complex social process requiring negotiation, consensus-building, and iteration. Learning to manage such processes effectively and to coordinate the ever-growing activities of many disparate communities of interest is clearly a long-term research undertaking involving complex economic, technical, and social questions».* [EU-NSF Working Group on Metadata, 1999]

La mise en place d'une structure de métadonnées à l'intérieur d'une organisation demande des ressources et des efforts continus. La promotion d'une structure de métadonnées à l'intérieur d'une industrie demande toutefois des efforts certainement encore plus importants.

Le Tableau 2.1 présente les principales initiatives de métadonnées reliées au domaine de l'éducation. Ces initiatives de métadonnées se distinguent à la fois par leur lieu d'origine (Amérique vs. Europe) et leur domaine d'application (scolaire, militaire, aviation).

Les projets ARIADNE (initiative universitaire européenne), IMS (initiative de compagnies américaines) et ADL (initiative militaire américaine) ont été instaurés de manière indépendante par des groupes d'intérêt différents. Le *Learning Object Metadata* [LOM IEEE, 2002] se distingue toutefois des précédents en agissant comme élément fédérateur pour reprendre les points essentiels des initiatives IMS, ARIADNE, SCORM (ADL) et AICC (certaines parties applicables) et favoriser la convergence de ces éléments de description.

Tableau 2.1 : Initiatives de normalisation des métadonnées pour le domaine de l'apprentissage.

1996	**ARIADNE** http://www.ariadne-eu.org (éducationnel européen)	*Alliance of Remote Instructional Authoring and Distribution Networks for Europe* ARIADNE est un consortium européen qui développe des outils et des méthodologies nécessaires pour produire, contrôler et réutiliser des éléments pédagogiques sur ordinateur et programmes de formation en ligne. Les normes ARIADNE et ARIADNE II présentent à la fois un modèle de vocabulaire et une architecture pour l'entreposage des données.
1997	**IMS** http://www.imsproject.org (éducationnel international)	*Instructional Management Systems* Ce consortium développe et favorise des caractéristiques ouvertes pour le développement de la formation en ligne. Plusieurs compagnies importantes font partie de ce consortium dont notamment Apple, Cisco, IBM, Microsoft et Sun.
	ADL http://www.adlnet.org (éducationnel militaire)	*Advanced Distributed Learning* a été établi par le département de la défense américaine pour développer une stratégie globale sur l'utilisation des technologies d'apprentissage au sein des forces armées et pour favoriser une coopération accrue entre les gouvernements, les industries et les maisons d'éducation. La norme proposée est actuellement connue sous le vocable de SCORM (*Sharable Content Object Reference Model*).
1998	**AICC** http://www.aicc.org (éducationnel international)	*Aviation Industry CBT (Computer Based Training) Committee.* Cette organisation élabore des directives pour l'industrie de l'aviation pour favoriser le développement, la livraison et l'évaluation des technologies de formation en ligne.
1999	**DCMI** http://dublincore.org/ (éducationnel international)	*Dublin Core Metadata Initiative* a originellement pris naissance en 1994 mais un groupe spécialisé a toutefois été formé en 1999 pour étudier spécifiquement l'utilisation des métadonnées dans le domaine de l'apprentissage. Un ensemble réduit de métadonnées a ainsi été retenu pour favoriser la découverte de ressources électroniques (*Metadata for Electronic Resources*).
	LTSC http://ltsc.ieee.org (éducationnel international)	IEEE Learning Technology Standards Committee Placé sous l'égide du *IEEE Computer Society Standards*, le LTSC développe et accrédite des guides, des normes techniques et des recommandations pratiques pour les technologies d'apprentissage. La désignation officielle du standard préconisé est *IEEE 1484.12.1-2002 Learning Object Metadata* (LOM).

ARIADNE comprend 40 éléments répartis en sept catégories distinctes. IMS comprend 90 éléments répartis en neuf catégories distinctes alors que LOM comprend à lui seul 101 éléments répartis en huit catégories distinctes. Ces 101 éléments touchent aussi bien à la description des objets eux-mêmes que leur cycle de vie, leur contexte technique, les droits d'auteurs et les relations aux autres objets. LOM permet ainsi l'utilisation d'annotations et d'éléments de classification générale pour relier ces objets à différents systèmes de référencement.

Malgré son niveau de sophistication élevé, LOM n'est pas réellement efficace pour la description d'objets de taille réduite. LOM a été conçu pour la description d'objets de la taille d'un cours ou une leçon entière. Les objets de la taille d'une page ou d'un paragraphe unitaire exigent un effort de descriptions beaucoup trop important pour rendre LOM attrayant aux yeux des utilisateurs. De plus, LOM se montre particulièrement mal adapté à la description d'objets d'apprentissage composés dynamiquement à partir d'éléments unitaires de tailles et de formats variés.

2.3 Objets d'apprentissage

Il n'existe pas de définition rigoureuse d'un « objet d'apprentissage » et certains auteurs considèrent que ce terme décrit simplement un élément technologique utilisé dans un contexte d'apprentissage. Un objet d'apprentissage peut ainsi prendre la forme d'un cours, d'un module logiciel, d'un site Web ou d'une institution entière [Linecar, 1994] [Siviter, 1999] [IEEE LOM, 2002] :

> *"A Learning Object is defined here as any entity, digital or non-digital, which can be used, re-used or referenced during technology-supported learning.*
>
> *[...] Examples of Learning Objects include multimedia content, instructional content, instructional software and software tools, referenced during technology supported learning. In a wider sense, learning objects could even include learning objectives, persons, organizations, or events."* [IEEE LOM, 2002]

Selon cette définition, les objets d'apprentissage sont des éléments de taille plus ou moins importante. Certains auteurs considèrent toutefois que les objets d'apprentissage se distinguent par leur taille spécifiquement réduite et qui les rend particulièrement propices à une réutilisation subséquente [Wiley, 2000]. Dans cette optique, les objets d'apprentissage s'apparentent plus volontiers à une leçon ou une page d'information unitaire. La taille réduite de ces objets leur permet ainsi d'être réutilisés pour composer des objets de plus grandes dimensions, un peu à la manière d'un atome composé à partir de particules élémentaires. À titre d'exemple, une leçon peut être considérée comme un objet d'apprentissage de grande taille composé d'objets internes plus petits tels qu'une discussion, un exemple ou une démonstration. Ces objets peuvent, à leur tour, être décomposés en éléments plus fins pour finalement isoler des objets de la taille d'un paragraphe, d'un mot ou d'une image.

Certains auteurs considèrent que les objets d'apprentissage se distinguent à la fois par leur taille réduite et leur capacité particulière de pouvoir être transférés par Internet. Les objets d'apprentissage s'inscrivent ainsi dans un contexte d'échange où la taille devient un enjeu important pour favoriser l'échange et la réutilisation de leurs composantes :

> *"Learning objects are elements of a new type of computer-based instruction grounded in the object-oriented paradigm of computer science. Object-orientation highly values the creation of components (called "objects") that can be reused [Dahl & Nygaard, 1966] in multiple contexts. This is the fundamental idea behind learning objects: instructional designers can build small (relative to the size of an entire course) instructional components that can be reused a number of times in different learning contexts.*
> *Additionally, learning objects are generally understood to be digital entities deliverable over the Internet, meaning that any number of people can access and use them simultaneously (as opposed to traditional instructional media, such as an overhead or video tape, which can only exist in one place at a time). Moreover, those who incorporate learning objects can collaborate on and benefit immediately*

from new versions. These are significant differences between learning objects and other instructional media that have existed previously." [Wiley, 2000, p.3]

D'une manière globale, il est possible de définir les objets d'apprentissage comme suit: un objet d'apprentissage est un élément structuré pouvant lui-même être composé d'éléments de format varié (texte, image, éléments audio/vidéo ou tout autre format binaire) et qui possède la capacité de pouvoir être délivré par des moyens numériques. Ces objets d'apprentissage peuvent être regroupés ensemble pour former une entité plus importante ou être insérés les uns dans les autres pour former une suite variée d'éléments de tailles et de dimensions différentes.

Cette définition sommaire met déjà en lumière les problèmes associés à la description des objets d'apprentissage et la définition des structures de métadonnées équivalentes.

2.4 Problème de la personnalisation des structures de métadonnées

Le repérage des objets d'apprentissage sur Internet est une opération difficile à réaliser et plusieurs initiatives de métadonnées ont été proposées pour normaliser la description de ces objets. L'utilisation de métadonnée reste une tâche difficile à réaliser [Friesen, 2004]. Cette tâche exige des qualifications techniques et des connaissances qui sont généralement hors de portée des principaux utilisateurs. Cette tâche exige aussi un investissement en temps et en argent important.

Malgré leur niveau de sophistication élevé, aucune des initiatives de métadonnées ne réussit à s'adapter complètement au caractère atomique des objets d'apprentissage. La nature complexe des objets d'apprentissage demande, en effet, l'utilisation de structures de métadonnées flexibles pouvant être définies sur mesure pour répondre au contexte particulier de chaque objet. La définition d'une structure de métadonnées commune exige toutefois un effort de coordination important entre différents acteurs pour définir, maintenir et mettre à jour les éléments de descriptions. Le consensus nécessaire entre ces

intervenants rend lui-même encore plus difficile la personnalisation des éléments de descriptions.

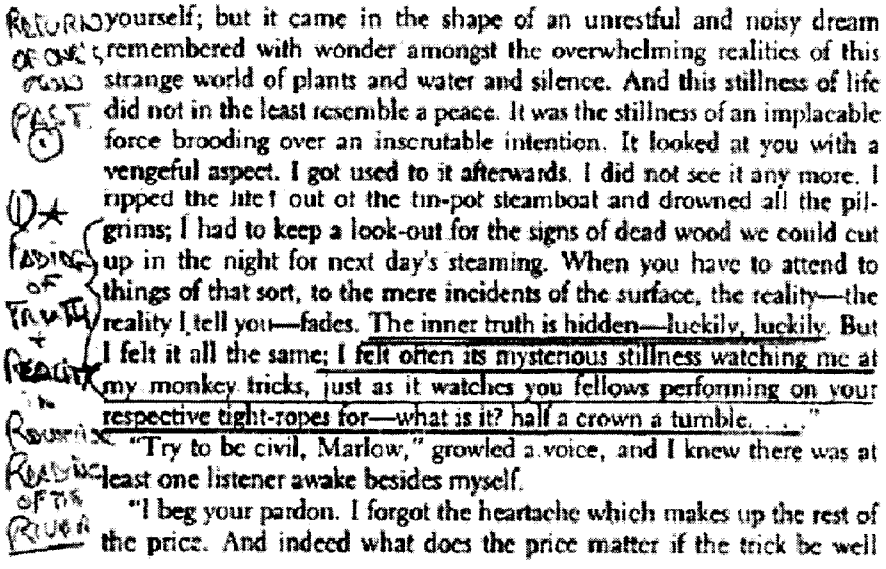
Les grandes initiatives de métadonnées (Tableau 2.1) offrent un certain niveau de flexibilité en permettant l'ajout d'éléments supplémentaires autour d'un noyau fixe d'éléments communs. L'ajout de ces éléments supplémentaires ne demande généralement pas l'imposition d'un consensus commun entre les intervenants et peut le plus souvent être réalisé de manière individuelle. Cette solution présente l'avantage de permettre l'ajout d'éléments supplémentaires à l'intérieur d'une structure de descriptions mais présente toutefois l'inconvénient de ne pas réellement permettre la modification intrinsèque de la structure de métadonnées.

La personnalisation des structures de métadonnées est un problème complexe qui préoccupe plusieurs chercheurs à travers le monde. Une solution intéressante pour améliorer la description des contenus Web consiste à assigner ces descriptions non pas à une page entière mais à un élément spécifique de la page. L'assignation des descriptions à un élément de page est réalisée grâce à des annotations.

CHAPITRE 3 : LES ANNOTATIONS

Une annotation est une note, une explication, ou tout autre type de remarque externe pouvant être attachée à un document sans toutefois être nécessairement insérée dans ce document. Une annotation peut aussi bien être réalisée en format papier qu'en format électronique.

Les annotations sont produites par des auteurs pour aider leurs lecteurs à comprendre des informations, indiquer l'existence de données complémentaires ou tout simplement identifier des ressources particulières [Marshall, 1997]. Les lecteurs produisent, quant à eux, des annotations pour supporter leur mémoire, localiser un problème, ou enregistrer les traces de leurs activités (Figure 3-1).



yourself; but it came in the shape of an unrestful and noisy dream remembered with wonder amongst the overwhelming realities of this strange world of plants and water and silence. And this stillness of life did not in the least resemble a peace. It was the stillness of an implacable force brooding over an inscrutable intention. It looked at you with a vengeful aspect. I got used to it afterwards. I did not see it any more. I ripped the life† out of the tin-pot steamboat and drowned all the pilgrims; I had to keep a look-out for the signs of dead wood we could cut up in the night for next day's steaming. When you have to attend to things of that sort, to the mere incidents of the surface, the reality—the reality I tell you—fades. The inner truth is hidden—luckily, luckily. But I felt it all the same; I felt often its mysterious stillness watching me at my monkey tricks, just as it watches you fellows performing on your respective tight-ropes for—what is it? half a crown a tumble. . . ."

"Try to be civil, Marlow," growled a voice, and I knew there was at least one listener awake besides myself.

"I beg your pardon. I forgot the heartache which makes up the rest of the price. And indeed what does the price matter if the trick be well

Figure 3-1 : Exemple de texte annoté [auteur inconnu].

En se basant sur ses propres observations, [Mille, 2005] établit la taxonomie de tâches suivante:

1. Restructurer :
   1.1. Donner un titre;
   1.2. Hiérarchiser;
   1.3. Synthétiser;
   1.4. Reformuler;
2. Ajouter une remarque personnelle :
   2.1. Critiquer :
      2.1.1. Positivement;
      2.1.2. Négativement;
   2.2. Exprimer une idée connexe;
   2.3. Développer :
      2.3.1. Compléter;
      2.3.2. Ajouter un exemple;
      2.3.3. Résoudre un problème;
      2.3.4. Expliquer textuellement/graphiquement;
   2.4. Faire référence à un autre document;
3. Catégoriser :
   3.1. Objectivement :
      3.1.1. Par type prédéfini;
      3.1.2. Par similarité de forme;
   3.2. Subjectivement :
      3.2.1. Par valeur d'importance (du passage, pas de l'annotation elle-même);
      3.2.2. Par similarité de sens;
4. Créer une relation entre deux passages;
5. Planifier une action :
   5.1. Approfondir;
   5.2. Réviser :
      5.2.1. Supprimer un passage;
      5.2.2. Insérer un passage;
      5.2.3. Reformuler un passage;
      5.2.4. Déplacer un passage;
6. Soutenir l'attention.

Ces actions sont réalisées de manières différentes selon le contexte particulier dans lequel elles prennent place. [Wolfe, 2001] constate ainsi que le contexte de lecture influe directement sur le type et la quantité des annotations produites. La forme des annotations serait ainsi modulée selon le contexte de communauté (pré-requis), le contexte de tâche (but de la lecture), et le contexte matériel (lieux, moyens).

Le contexte de communauté se distingue par des comportements d'annotation parfois similaires et parfois totalement différents entre différents groupes d'utilisateurs. Il existe ainsi des patrons individuels de lecture et d'annotation qui feront en sorte qu'un même utilisateur adoptera toujours un même format d'annotation tout en continuant à faire évoluer de manière personnelle sa propre manière de sélectionner et de réaliser des annotations [Marshall, 1999].

Le contexte de tâche influence aussi le comportement d'annotation en favorisant des similitudes entre les annotations réalisées par différents utilisateurs. Par exemple, en examinant les annotations papiers réalisées par des étudiants, [Marshall, 1998] constate que « les étudiants ont tendance à mettre en valeur les mêmes idées ».

Le contexte de tâche agit directement sur la manière de produire des annotations en favorisant des annotations de facture graphique ou de format textuel pour soutenir les buts subséquents de la lecture.

Le contexte matériel modifie aussi la manière de réaliser des annotations en modifiant la tâche d'annotation en fonction de la disponibilité ou non d'outils particuliers (ex: surligneur, crayon, règles, etc.) [Marshall, 2000].

3.1 Annotations électroniques

Les annotations de format papier sont réalisées sur un medium fixe qui change peu dans le temps. Ces annotations peuvent être permanentes ou mobiles (signet, plis de page, note autocollante, insertion papier, etc.) et utiliser un codage personnel ou explicite pour les autres lecteurs tout en étant placées directement dans le texte ou à côté de celui-ci [Marshall, 1997].

Les annotations de format électronique reprennent, pour leur part, les mêmes attributs que les annotations de format papier mais en rajoutant cependant la possibilité d'être facilement utilisables par un logiciel et de pouvoir supporter une mise en relation automatique avec d'autres données de format numérique.

Le navigateur Mosaic 1.2 (1993) a été le premier logiciel à permettre la création d'annotations électroniques sur le Web. Plusieurs autres logiciels ont par la suite été proposés pour répondre aux besoins particuliers des utilisateurs: AktiveDoc [Lanfranchi, 2005], Altered Vista [Recker, 2001], Annotea [Kahan, 2001], Annozilla [Obendorf, 2003] , Edutella [Nejdl, 2001], IBM annotation tool [Hori, 2003], KIM [Popov, 2003], Magpie [Dzbor, 2003], Mangrove [McDowell, 2003], Melita [Ciravegna, 2002], MnM [Vargas-Vera, 2002], OLR3 [Brase, 2003], OntoGloss [Mostowfi, 2005], OntoMat [Handschuh, 2002], PDFTab [Eriksson, 2007], Semantic Markup Tool [Kettler, 2005], SMORE [Kalyanpur, 2003], etc.

Le Tableau 3.1 présente les principaux d'outils d'annotation disponibles. Ce tableau a été colligé à partir des travaux initiaux réalisés par [Prié, 2004] et [Uren, 2006] ainsi que de nos propres recherches sur le sujet.

Tableau 3.1: Caractéristiques des principaux outils d'annotation actuellement disponibles.

Outil	Format standard	Interface graphique	Métadonnées ou ontologie	Format de document	Évolution de document	Entreposage des annotations
AeroSWARM (AeroDAML)	DAML+OIL, OWL	Service Web	Ontologies locales	HTML		Incorporé aux pages Web
AktivDoc	HTML, RDF	Application		HTML		Repositoire RDF
Amaya	RDF(S,) XLink, XPointer	Navigateur	Schéma RDF	HTML, XHTML, XML	XPointer	Local, serveur d'annotation
Annotea	RDF, XPointer	Service Web	Schéma RDF	HTML, HTML, XML, SVG	XPointer	Serveur HTTP dédié
Armadillo	RDF(S)	Service Web		HTML		Repositoire RDF
COHSE annotator	DAML+OIL	Module externe pour Mozilla & IE	Serveur d'ontologies	HTML (via DOM)	XPointer	Serveur d'annotation, fichier local
FOAF project (RDFWeb)	RDF	Service Web	Schéma RDF, Propriété DAML	HTML		Serveur
h-TechSight	DAML + OIL, RDF	Portail KM	Éditeur d'ontologie, métriques dynamiques	HTML		Incorporé aux pages Web, serveur
KIM	RDF(S), OWL	Navigateur	KIMO, ontologies de haut niveau PROTON	HTML		RDF(S)
KnowItAll	HTML					
Lixto	(wrapper)					
Magpie	HTML, OCML	Navigateur		HTML		

Outil	Format standard	Interface graphique	Métadonnées ou ontologie	Format de document	Évolution de document	Entreposage des annotations
Mangrove	RDF	Application		HTML, email		Repositoire RDF (avec Jena)
Melita	RDF(S), DAML + OIL	Navigateur IE	Local, ontologies éditables	HTML, texte	Expression régulière	
MnM	RDF(S), DAML+OIL, OCML	Navigateur	Serveur d'ontologies	HTML, texte	Annotations sauvegardées	Incorporé aux pages Web
M-OntoMat-Annotizer	XML, RDF(S), DOLCE	Client Java		MPEG-7		Serveur d'annotation
OntoAnnotate	RDF(S)	Application Java		HTML		
OntoMat-Annotizer	RDF(S), DAML + OIL, OWL, SQL, XPointer, F-Logic	Client Java	Serveur d'inférence OntoBroker	HTML	XPointer, pattern matching	Serveur d'annotation, incorporé aux pages Web ou encore fichier séparé
OnTeA	Texte		Comparaison des termes avec les scores de notoriété de Google			Repositoire RDF (Sesame)
Open Ontology Forge	RDF(S), XML, Xlink, XPointer, Dublin Core	Navigateur Web	Local, ontologies éditables	HTML, texte, images (SVG)	XPointer	Fichier RDF ou XML local
PANKOW	HTML	CREAM	Comparaison des termes avec les scores de notoriété de Google			

Outil	Format standard	Interface graphique	Métadonnées ou ontologie	Format de document	Évolution de document	Entreposage des annotations
Parmenides	XML		Regroupement pour suggérer des additions			
PDFTab	Texte, PDF, OWL	Module externe de Protégé	Ontologies de haut niveau.	PDF		Annotations insérées à l'intérieur des fichiers PDF
PhotoStuff	RDF, OWL	Client Java, service Web	Ontologies locales	HTML, SVG		Local ou serveur
Rainbow Project	RDF(S), DAML+OIL	Navigateur Web	Ontologies de haut niveau	HTML		Repositoire RDF (Sesame)
SemanticWord	DAML+OIL	Microsoft Word		Word	Étiquettes attachées à des régions	
SemTag	RDF(S)		Liste de données, ontologies	HTML		Serveur HTTP dédié, PICS
SHOE Knowledge annotator	SHOE	Client Java	Serveur d'ontologies	HTML		Incorporé aux pages Web
SmartWeb	RDF, RDF(S), OWL					Repositoire RDF
SMORE	RDF(S), (Scraper)	Éditeur, navigateur Web	Serveur d'ontologies, édition d'ontologies	HTML, texte, email, images		
Thresher	RDF	Navigateur Web	Personnalisation d'ontologies	HTML		

Outil	Format standard	Interface graphique	Métadonnées ou ontologie	Format de document	Évolution de document	Entreposage des annotations
Trellis	HTML, XML, RDF, DAML+OIL	Navigateur Web	Dublin Core	HTML, texte		
Vannotea	XML, RDF, OWL	Application		MPEG-2, JPEG2000, Direct3D		Serveur d'annotation
WebKB	RDF	Service Web	Ontologies construites à partir de WordNet	HTML		Serveur
WiCKOffice	Microsoft Smart Documents	Environnement Microsoft Office (fournit un support pour l'édition de formulaire)		Microsoft Office		Repositoire d'annotation (avec triplets)

Les logiciels *Altered Vista, Annotea, Edutella, SMORE, Melita* et *KIM* représentent les outils d'annotation les mieux connus. Ces outils font chacun partie d'une catégorie différente de système d'annotation.

3.1.1 Altered Vista

Le logiciel *Altered Vista* [Recker, 2001] permet de recueillir les recommandations personnelles de différents utilisateurs en fonction de catégories précises de description. Les informations saisies peuvent aussi bien prendre la forme de texte libre que d'appréciations qualitatives annotées au moyen d'une échelle de Likert graduée de "faible" à "excellent". Chaque catégorie de descriptions regroupe ainsi plusieurs appréciations différentes que les utilisateurs peuvent eux-mêmes sélectionner en fonction d'un contexte donné.

3.1.2 Annotea

Le logiciel *Annotea* a été proposé par le W3C pour supporter l'insertion directe d'annotations à l'intérieur de pages Web. Ce logiciel est directement utilisé sur un poste client (navigateur) alors que les annotations finales sont déposées sur un serveur public de manière à les rendre facilement accessibles à autrui.

Annotea permet aussi bien de rédiger des annotations que de réviser des annotations réalisées par d'autres [Kahan, 2001]. Ce logiciel n'est toutefois pas, en soi, un système réservé à la description des contenus mais peut néanmoins être utilisé à cette fin. Il est ainsi possible d'utiliser *Annotea* pour décrire des contenus Web et rajouter ces descriptions à une liste de descriptions déjà produites par d'autres [Figure 3-2].

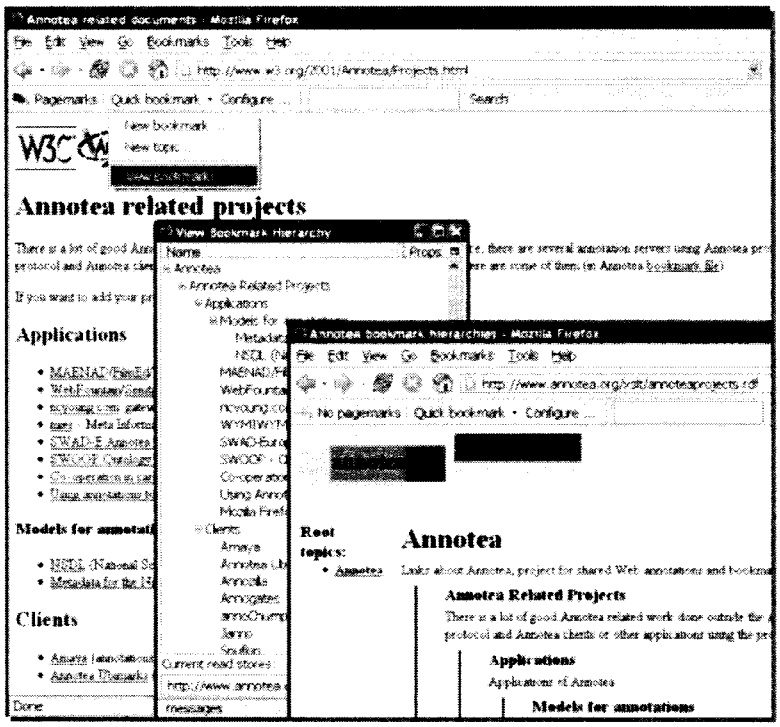


Figure 3-2 : Récupération de pages annotées avec Annotea.

Annotea ne supporte pas l'utilisation de structure complexe de métadonnées et il serait difficile de le considérer comme un véritable système de métadonnées en soi. D'autres développements réalisés à partir *Annotea* tentent toutefois de combler cette lacune en améliorant la gestion des structures de métadonnées qui font actuellement défaut au logiciel.

3.1.3 Edutella

Le logiciel *Edutella* reprend l'idée des annotations personnelles en utilisant toutefois une architecture poste à poste (*P2P*) qui favorise le partage des descriptions entre différents utilisateurs [Nejdl, 2001].

Un peu à la manière d'*Annotea*, le logiciel *Edutella* permet à chaque utilisateur d'ajouter ses propres observations sur une page Web. *Edutella* utilise toutefois une structure de métadonnées particulière qui oblige les utilisateurs à respecter un classement logique.

Les informations échangées entre utilisateurs ne concernent que la description des ressources et la récupération de ces descriptions s'effectue par le biais d'une URI précisant leur lieu d'origine. Cette particularité permet notamment d'éviter les problèmes associés à la gestion des droits d'auteur et aux droits d'accès aux ressources en dissociant les contenus Web des descriptions correspondantes.

Edutella est un logiciel spécifique au domaine de l'éducation qui reprend fidèlement le modèle de descriptions des grandes initiatives de métadonnées. Les éléments de descriptions sont ainsi organisés en couches concentriques en utilisant *Dublin Core* au centre et *IMS/IEEE-LOM* en périphérie. La couche de description externe reprend les paquets de contenus proposés par IMS et une certaine partie du modèle de descriptions proposé par *SCORM*. D'autres catégories supplémentaires de métadonnées peuvent aussi être rajoutées aux descriptions périphériques pour compléter la description des contenus.

3.1.4 SMORE

Le logiciel *SMORE* (*Semantic Markup, Ontology and RDF Editor*) est un environnement intégré pour la création de pages Web, des courriers électroniques, et d'autres types de fichier textuel [Kalyanpur, 2003]. L'éditeur graphique de *SMORE* permet d'associer des éléments textuels ou graphiques à des ontologies (Figure 3-3). Ces ontologies peuvent être directement récupérées à partir du Web. Ce logiciel possède aussi la capacité de reconnaître certaines balises HTML et de récupérer automatiquement les ontologies correspondantes pour faciliter l'annotation des documents.

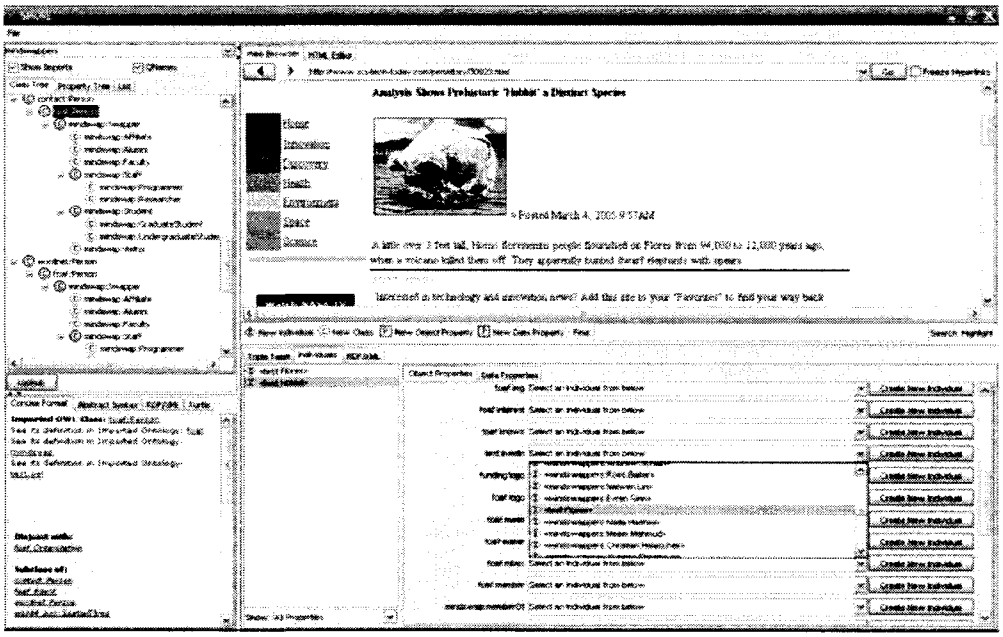


Figure 3-3 : Interface du logiciel d'annotation SMORE.

3.1.5 Melita

Melita est une application qui permet l'annotation directe de documents HTML au moyen d'une interface graphique [Ciravegna, 2002]. Cette application possède aussi toutefois la capacité d'extraire elle-même des annotations et de les suggérer à l'utilisateur afin de valider ses propres stratégies d'extraction (Figure 3-4). Les annotations approuvées par l'utilisateur servent ainsi à consolider le comportement de *Melita* qui finit par apprendre à mieux extraire les contenus proposés en fonction d'ontologies données. *Melita* se définit ainsi comme un système d'annotation semi-automatique.

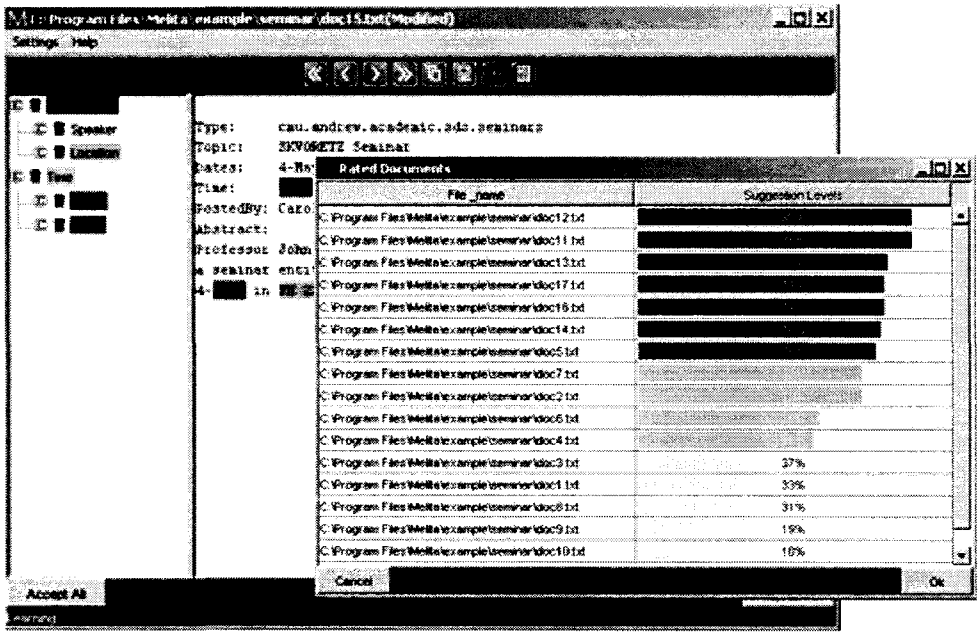


Figure 3-4 : Suggestions d'annotation proposées par Melita.

3.1.6 KIM

KIM est un module externe qui s'installe à l'intérieur d'un navigateur pour réaliser l'annotation de pages Web [Popov, 2003]. Contrairement aux outils cités précédemment, *KIM* effectue une annotation automatique en extrayant lui-même les termes importants et en les associant par lui-même à une ontologie de haut niveau (Figure 3-5). *KIM* se définit ainsi comme un système d'annotation entièrement automatique.

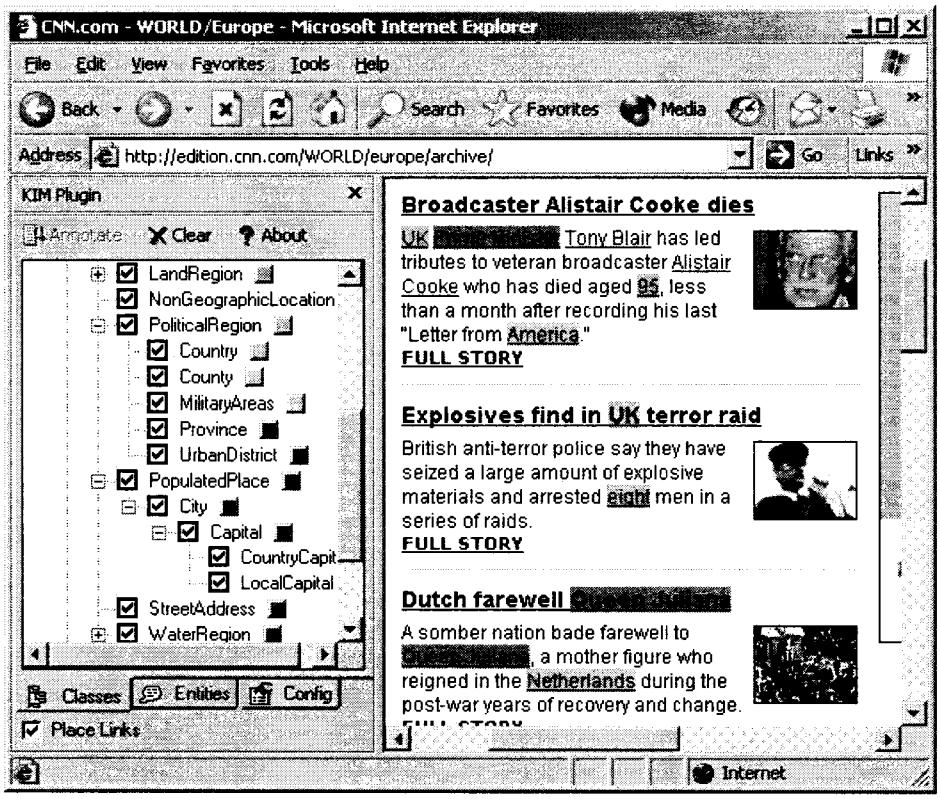


Figure 3-5 : Annotation automatique d'une page HTML avec KIM.

3.2 Catégorisation

Les logiciels *Annotea, Edutella, SMORE, Melita* et *KIM* représentent cinq manières différentes de réaliser des annotations. Ces logiciels se distinguent chacun des autres au moyen de deux critères distincts: 1) par le caractère plus ou moins structuré des annotations et 2) par le caractère plus ou moins automatique des annotations produites.

Le caractère plus ou moins structuré des annotations se définit par l'utilisation de texte libre par l'utilisateur ou, encore, par l'imposition de métadonnées ou d'ontologies particulières. Le caractère plus ou moins automatique des annotations produites se manifeste, quant à lui, soit par la nécessité d'une intervention directe de l'utilisateur (annotation manuelle), soit par la suggestion de valeurs par défaut réalisée par le logiciel (annotation semi-automatique) ou soit par la production autonome d'annotations

(annotation automatique). Chacune des catégories précédentes possèdent évidemment ses propres avantages et inconvénients.

Les annotations réalisées de manière automatique possèdent l'avantage de pouvoir être produites plus rapidement et à plus faible coût que les annotations réalisées manuellement. En contrepartie, les annotations manuelles possèdent l'avantage de représenter plus fidèlement le modèle cognitif de l'utilisateur bien qu'elles aient aussi le désavantage d'exiger un effort humain plus important. Les annotations semi-automatiques forment une catégorie intermédiaire qui tend à mitiger les impacts des deux autres catégories.

Aucune des catégories différentes ne représente en soi une panacée au problème de conception des annotations et chacune de ces catégories trouve son utilité propre. Les annotations manuelles, semi-manuelles et automatiques représentent donc tout simplement des moyens différents d'obtenir des annotations de qualité et de formes variées.

Il est possible d'ajouter une autre dimension aux catégories précédentes en précisant la capacité d'une annotation à être traitée automatiquement par une machine. [Azouaou, 2004] propose ainsi d'établir une distinction entre une annotation « sémantique » (qui supporte le sens d'une description) et une annotation « non-sémantique » (qui est le fruit d'un commentaire ou une observation de l'utilisateur). Une annotation « sémantique » peut elle-même être décomposée en une annotation « cognitive » (qui supporte le modèle de l'utilisateur) ou une annotation « computationnelle » (qui supporte un traitement logique réalisé par la machine). Le Tableau 3.2 illustre le classement des outils d'annotation selon leur capacité à traités plus ou moins automatiquement et la valeur plus ou moins sémantique de leur contenu.

Tableau 3.2: Classement des outils d'annotation [Azouaou, 2004].

Semantics	Author / Adressee	Manual	Semi-automatic	Automatic
Non semantic annotation	Cognitive and non computational annotation	Imarkup, Acrobat, Web-Notes, CoNote, WebAnn, Epost		Google's ToolBar
	Non cognitive and computational annotation	Manual index in libraries	MyAlbum Annotate	Google search engine
	Cognitive and computational annotation	Knowledge Pump, Xlibris		Cached Google Links
Semantic annotation	Cognitive and non computational annotation	Annotea + Amaya, Yawas, ThirdVoice Mark-Up		
	Non cognitive and computational annotation	Edutella, OntOmat, SHOE, HTML-A, WebKB, Karina		AeroDAML
	Cognitive and computational annotation	Mangrove, SMORE	MnM,Melita, Teknowledge, IMAT	KIM, MnM, Magpie, COHSE

Les logiciels qui permettent la production d'annotations électroniques se divisent en systèmes manuels, semi-automatiques et automatiques. Aucun de ces systèmes n'est réellement meilleur que les autres. Il est néanmoins possible de réaliser une subdivision à l'intérieur de ces trois grandes catégories pour reconnaître les systèmes qui produisent des annotations sémantiques par rapport aux systèmes qui produisent des annotations dites « non-sémantiques ».

Les annotations sémantiques cognitives et computationnelles (représentées par la derrière ligne du tableau) se distinguent ainsi des autres catégories par leur capacité à supporter le sens des contenus annotés tout en permettant à la fois l'interprétation de ces contenus par un humain et une machine. L'interprétation des annotations faisant partie de cette catégorie est réalisée par le biais d'ontologies qui précisent le sens de ces annotations. Le chapitre qui suit présente les principes de construction de ces ontologies.

CHAPITRE 4 : LES ONTOLOGIES

Le terme ontologie réfère à un champ d'étude de la philosophie qui porte sur le sujet de l'existence. Ce terme a été repris en ingénierie des connaissances selon différents vocables:

1. Un langage de description précisant la construction d'un vocabulaire propre à un sujet:

   « An ontology defines the basic terms and relations comprising the vocabulary of a topic area, as well as the rules for combining terms and relations to define extensions to the vocabulary. » [Neches, 1991]

2. Une représentation des connaissances exprimée par un langage de description de haut niveau :

   « An ontology is an explicit specification of a conceptualization. [...] When the knowledge of a domain is represented in a declarative formalism, the set of objects that can be represented is called the universe of discourse. This set of objects, and the describable relationships among them, are reflected in the representational vocabulary with which a knowledge-based program represents knowledge. [...] Formally, an ontology is the statement of a logical theory. » [Gruber, 1993]

3. Une représentation de la connaissance possible d'un agent :

   « An (AI-) ontology is a theory of what entities can exist in the mind of a knowledgeable agent. » [Wielinga, 1993]

4. Une taxonomie de concepts permettant l'interprétation sémantique des connaissances liées à une tâche ou un domaine :

« An ontology for a body of knowledge concerning a particular task or domain describes a taxonomy of concepts for that task or domain that define the semantic interpretation of the knowledge. » [Albert, 1993]

5. Une conceptualisation consensuelle et partagée des connaissances :

   « Ontologies are agreements about *shared* conceptualizations. Shared conceptualizations include conceptual frameworks for modeling domain knowledge; content-specific protocols for communication among inter-operating agents; and agreements about the representation of particular domain theories. In the knowledge sharing context, ontologies are specified in the form of definitions of representational vocabulary. A very simple case would be a type hierarchy, specifying classes and their subsumption relationships. Relational database schemata also serve as ontologies by specifying the relations that can exist in some shared database and the integrity constraints that must hold for them. » [Gruber, 1994]

6. Une description explicite d'une conceptualisation :

   « An ontology is an explicit, partial account of a conceptualization. » [Guarino, 1995]

7. Une spécification explicite d'une conceptualisation de haut niveau :

   « An ontology is an explicit, partial specification of a conceptualization that is expressible as a meta-level viewpoint on a set of possible domain theories for the purpose of modular design, redesign and reuse of knowledge-intensive system components. » [Schreiber, 1995]

8. Une spécification d'une conceptualisation :

   « An ontology is a explicit knowledge level specification of a conceptualization, [...] which may be affected by the particular domain and task it is intended for. » [van Heijst, 1996]

9. Une spécification explicite permettant la description des contenus d'une base de connaissances :

   « An ontology provides the means for describing explicitly the conceptualization behind the knowledge represented in a knowledge base. » [Bernaras, 1996]

10. Une théorie logique inhérente au modèle d'un langage logique :

    « An ontology is a logical theory that constrains the indended models of a logical language. » [Guarino, 1997]

11. Une hiérarchie structurée de termes permettant la construction d'une base de connaissances propre à un domaine:

    « An ontology is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base. » [Swartout, 1997]

12. Une spécification formelle d'une conceptualisation partagée :

    « An ontology is a formal specification of a shared conceptualization. » [Borst, 1997]

13. Une théorie logique détaillant le sens d'un vocabulaire formel :

    « An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualisation of the world. » [Guarino, 1998]

14. Une spécification formelle et explicite d'une conceptualisation partagée:

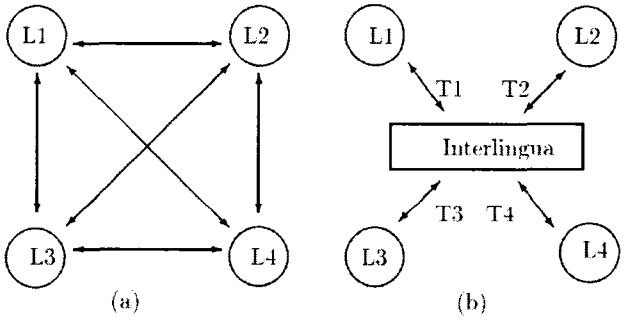
    « An ontology is a formal, explicit specification of a shared conceptualization. » [Studer, 1998]

Avec cette dernière définition, Studer reprend les termes mêmes de Gruber en précisant toutefois le concept de partage (aussi initialement proposé par Gruber). Selon cette dernière définition, une ontologie définit d'une manière formelle les connaissances communes d'un domaine particulier partagées entre différents utilisateurs.

Dans la définition précédente, le terme « connaissance » joue un rôle très particulier et ne devrait pas être confondu avec le terme « données » ou « information ». Une donnée est un signal non interprétée. Une information est une donnée associée à une signification particulière. Dans le contexte des ontologies, la connaissance est ainsi considérée comme une « information sur l'information » [Schreiber, 2000].

4.1 Contexte d'utilisation

Les ontologies jouent le rôle d'une langue universelle, une sorte d'*interlingua*, qui permet à des gens ou des applications d'échanger des informations sur une base commune. Ces informations concernent aussi bien les concepts que les rapports qui existent entre les différents éléments de connaissance d'un domaine (Figure 4-1).



To translate from language L_i to L_j and vice versa, a translator is required between L_i and the inter-lingua and another between the inter-lingua and L_j. Thus, given n languages, only $O(n)$ translators are required, not $O(n^2)$.

Figure 4-1: Ontologie comme langue universelle [Uschold, 1996].

Cette *interlingua* joue un rôle fédérateur dans le partage de connaissances entre différents intervenants. La position centrale des ontologies permet de:

- partager une compréhension commune entre des personnes ou des agents logiciel;
- rendre explicites les éléments de connaissance d'un domaine;
- favoriser la réutilisation des connaissances d'un domaine;
- établir une séparation entre les connaissances d'un domaine et les connaissances opérationnelles;
- favoriser l'analyse des connaissances d'un domaine.

La capacité des ontologies à jouer un rôle fédérateur dans l'expression des connaissances d'un domaine est due à la forme particulière que prennent ces connaissances. Les ontologies contiennent des définitions qui sont interprétables par des machines et qui portent tout autant sur les concepts de base d'un domaine que sur les relations qui s'établissent entre ces concepts [Noy, 2001]. Ces relations permettent à leur tour de réaliser des inférences logiques ou de réaliser des traitements symboliques sur les connaissances.

4.2 Logique de description

Au début des années 70, [Quillian, 1967] avait imaginé le concept des « réseaux sémantiques » qui se présente comme une série de concepts associés entre eux par le biais de liens significatifs. Minsky reprit cette même idée en y ajoutant la notion d'élément prototype appelé « frames ». Les frames utilisent un modèle emprunté à cognition humaine pour représenter des éléments de connaissances:

> « Here is the essence of the theory: When one encounters a new situation (or makes a substantial change in one's view of the present problem) one selects from memory a structure called a Frame. This is a remembered framework to be adapted to fit reality by changing details as necessary. »

> « A frame is a data-structure for representing a stereotyped situation, like being in a certain kind of living room, or going to a child's birthday party. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed. »
>
> [Minsky, 1974, p.212]

Les idées de Quillian et de Minsky ont à leur tour favorisé la naissance d'un formalisme de description des connaissances appelé « logique de description » (*Description logics,* plus communément désignée par l'acronyme *DL*). Le terme « description » réfère ici à la description des concepts alors que le terme « logique » réfère à l'expression de relations logiques.

La logique de description est un modèle de classification de concepts et d'instanciation d'objets qui utilise la notion de rôle pour décrire des relations binaires entre des individus. Par exemple, en utilisant les concepts de négation (¬), d'intersection (∏) et d'existence (∃), [Baader, 2003] affirme que :

Femme ≡ Personne ∏ Feminin
(une femme est une personne de type féminin)

Homme ≡ Personne ∏ ¬ Femme
(un homme est une personne qui n'est pas une femme)

Mère ≡ Femme ∏ ∃ possèdeEnfant.Personne
(une mère est une femme qui a une personne comme enfant)

Père ≡ Homme ∏ ∃ possèdeEnfant.Personne
(un père est un homme qui a une personne comme enfant)

La logique de description permet la construction d'ontologies en utilisant des déclarations de classes organisées selon une hiérarchie « is-a » déjà connue des langages orientés objets. Une ontologie se présente ainsi comme une description formelle et explicite des concepts d'un domaine (« classes » ou « instances »), des propriétés de

concept décrits par des caractéristiques et attributs (« attributs », « rôles » ou « propriétés ») et des restrictions sur les attributs (« facettes » ou « restrictions de rôles ») d'un domaine particulier.

La Figure 4-2 illustre une comparaison entre la logique de description, le modèle des frames, la programmation orientée objet et les descriptions RDF (que nous traiterons plus en détail au chapitre suivant) proposée par [Lassila, 2001] :

OOP Systems	Frame Systems	Description Logics	RDF
instance	frame, instance, individual	instance, individual	resource
attribute, instance variable	slot	role, attribute	property
value	filler	filler	property value
class, type	frame, schema	class, concept	class

Figure 4-2 : Comparaison de modèles [Lassila, 2001].

4.3 Structure d'ontologie

Dans un contexte d'entreprise, [Abecker, 1998] suggère que tous les éléments d'information et de connaissance soient décrits en utilisant un certain nombre d'attributs faisant partie d'un méta modèle d'information. Ce méta modèle décrit les contenus d'information, le domaine ainsi que le contexte d'application. Les concepts nécessaires à la description de connaissance peuvent ainsi être regroupés selon trois ontologies distinctes (Figure 4-3):

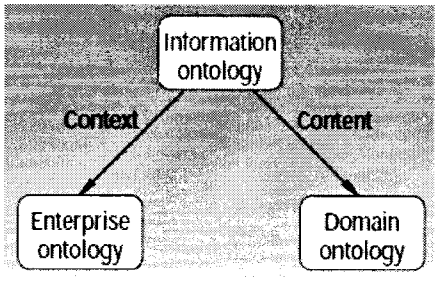


Figure 4-3 : Modèle d'information basé sur trois ontologies distinctes [Abecker, 1998].

- **Ontologie d'information :** fournit le vocabulaire nécessaire au méta modèle d'information pour décrire les différents type émetteurs d'informations avec leur structure, leur mode d'accès ainsi que leur format respectif. Il pourrait aussi bien

s'agir de connaissances descriptives sur certains produits que de connaissances prescriptives concernant la manière de réaliser certaines choses.

- **Ontologie du domaine :** est utilisée pour la description des contenus.
- **Ontologie d'entreprise :** décrit le contexte de création et le contexte d'utilisation prévu des éléments de connaissance.

La Figure 4-4 illustre l'exemple des éléments d'information reliés à une entreprise (compagnie, département, employés, etc.), aux informations véhiculées (document, livre, titre, etc.) et aux éléments de domaine de l'entreprise (contenu, phrase, mot clé, etc.).

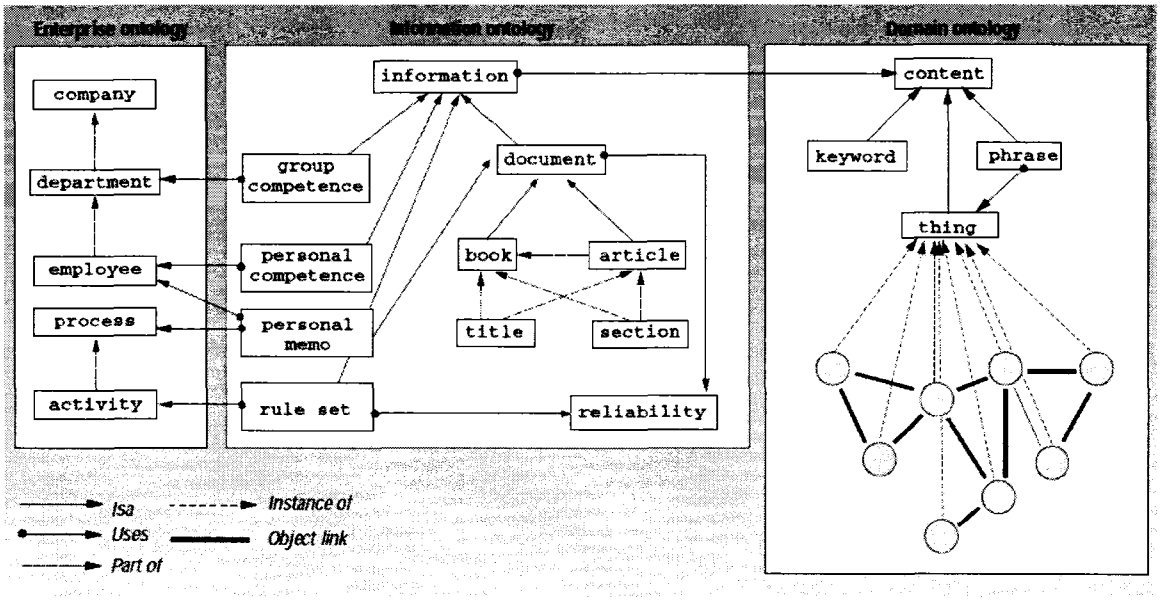


Figure 4-4 : Description des connaissances d'une entreprise [Abecker, 1998].

Cette façon de détailler le domaine d'une entreprise par l'utilisation de trois ontologies distinctes soutient l'idée qu'il est nécessaire d'établir une distinction entre les concepts généraux et les éléments pratiques d'un domaine.

Dans un contexte de description plus général, [Guarino, 1998] propose de réaliser un découpage à quatre niveaux pour mieux circonscrire les éléments de connaissance d'un domaine (Figure 4-5):

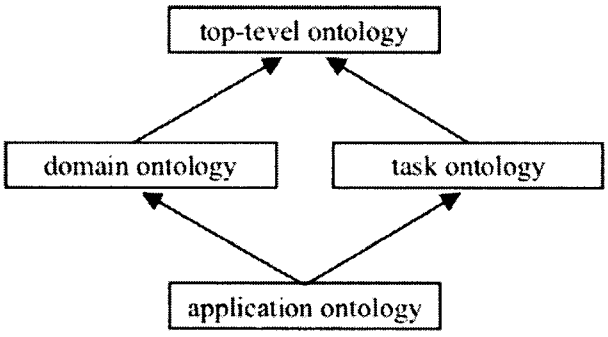


Figure 4-5: Découpage en quatre niveaux d'ontologies [Guarino, 1998].

Ce découpage se présente de la manière suivante :

- **Ontologie de haut niveau**: décrit les concepts généraux indépendants d'un problème ou d'un domaine particulier (comme l'espace, le temps, les événements). Ce type d'ontologie possède l'avantage de pouvoir être partagé par une grande communauté d'utilisateurs.
- **Ontologie du domaine**: décrit le vocabulaire lié à un domaine générique en spécialisant les concepts déjà utilisés dans l'ontologie de haut niveau.
- **Ontologie de tâche**: décrit le vocabulaire lié à une tâche ou une activité générique en spécialisant aussi les concepts déjà utilisés dans l'ontologie de haut niveau.
- **Ontologie d'application**: décrit les concepts correspondant aux rôles joués par des acteurs du domaine dans un contexte d'activité précis.

Ce découpage en quatre niveaux d'ontologies permet d'isoler les éléments de connaissance propres à un domaine tout en assurant d'une certaine interopérabilité avec les ontologies déjà existantes par le bais d'une référence à une ontologie commune de haut niveau.

Ce découpage offre ainsi l'avantage de pouvoir faire varier le niveau de détail des différentes ontologies selon le domaine à couvrir.

4.4 Niveaux de description

Les ontologies doivent leur existence à la présence de certains éléments obligatoires [Lassila, 2001] :

Présence obligatoire :

1. Vocabulaire contrôlé (extensible) fini;
2. Interprétation non ambiguë des relations de classes et de termes;
3. Rapports hiérarchiques stricts de sous-classes entre les classes.

Typiques mais non obligatoires:

4. Spécifications de propriétés au niveau des classes;
5. Inclusion individuelle à l'intérieur des ontologies;
6. Spécifications de restriction de valeur au niveau des classes.

Souhaitables mais non obligatoires ni même typiques:

7. Spécifications de classes disjointes, de rapports inverses, de rapports partie-entier;
8. Spécifications des rapports logiques arbitraires entre les termes.

Les ontologies ne possèdent pas toutes le même niveau de description. Il est ainsi possible d'établir une distinction entre les ontologies « légères » qui se présentent sous la forme d'une hiérarchie d'abstraction et les ontologie « lourdes » qui utilisent une structure de description logique [Studer, 1998]. Il est même possible de réaliser une classification encore plus fine des ontologies selon leur niveau de complexité respectif (taxonomie de concepts plus ou moins approfondie, axiomes formels, notion de disjointure, etc.).

La Figure 4-6 illustre la classification proposée par [Lassila, 2001]. Cette classification se présente sous la forme d'une ligne droite qui illustre une échelle de progression entre différents systèmes de description qui débutent par une simple liste de termes et se termine avec les ontologies lourdes qui exploitent des contraintes logiques.

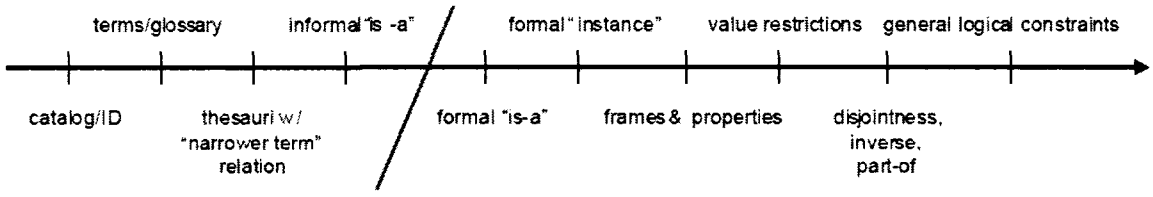


Figure 4-6 : Échelle de progression formelle des ontologies [Lassila, 2001].

Entre ces deux extrêmes, il est possible de reconnaître une panoplie de systèmes de description différents. Une barre oblique placée sur l'échelle de progression établit toutefois une distinction entre les systèmes de classement pour marquer le début des ontologies légères. Les ontologies se distinguent ainsi des glossaires, des thésaurus et des systèmes de classement hiérarchique informel par un formalisme de description qui précise de manière non ambiguë la position et les relations qu'entretiennent les classes entre elles. Une ontologie est une spécification des éléments de connaissance d'un domaine exprimée dans un langage formel.

Les lecteurs attentifs remarqueront que la Figure 4-6 ne fait aucunement mention à l'existence d'une quelconque « base de connaissances ». L'absence de ce concept est reliée au fait que l'ensemble des instances individuelles des classes d'une ontologie forme en soi une « base de connaissances ». Ainsi, à toute fin pratique, la distinction entre une ontologie et une base de connaissance reste ténue [Noy, 2001].

4.5 Processus de développement

[Gruninger, 1995] propose de débuter le processus de construction d'une ontologie avec l'établissement d'une liste de questions auxquelles l'ontologie terminée devrait pouvoir répondre. Cette liste, aussi appelée « questions de compétence », est utilisée pour

évaluer si la granularité et le niveau de détail des éléments de l'ontologie produite permettront de couvrir de manière suffisante un domaine en fonction du contexte d'application choisi.

Par exemple, dans le contexte d'une ontologie destinée à supporter la construction d'un cours Web en IHO (Interface Humain Ordinateur), ces questions pourraient prendre la forme suivante :

- Quels critères devrait-on considérer dans le choix d'une technique d'analyse ergonomique pour mesurer l'efficience d'une tâche?
- Est-ce qu'une analyse hiérarchique de tâche (AHT) permet de mesurer le temps réel d'exécution d'une tâche?
- Existe-t-il une méthode pour mesurer le taux d'erreur dans l'accomplissement d'une tâche?

Les questions de compétence facilitent assurément le démarrage d'une ontologie mais ne permettent pas en soi de valider le niveau d'achèvement d'une ontologie. La forme finale que devrait avoir une ontologie est une question à laquelle il est difficile de répondre parce qu'il n'existe pas qu'une seule manière de modéliser un domaine de connaissances. Il existera toujours des alternatives viables aux ontologies produites.

[Noy, 2001] soutient que la forme de l'ontologie dépend généralement de l'application que l'on désire mettre en place et des extensions que l'on anticipe utiliser. Le développement d'une ontologie serait donc un processus nécessairement itératif qui alterne entre la conception et l'utilisation de l'ontologie. Les concepts utilisés par l'ontologie devraient ainsi directement se rapprocher des objets (physiques ou logiques) et des relations du domaine d'intérêt. Ces éléments seront très certainement des noms (objets) ou des verbes (relations) faisant partie de phrases qui décrivent le domaine d'intérêt que l'on désire représenter.

Toujours selon [Noy, 2001], une ontologie devrait être modelée en fonction du (des):

- Domaine à couvrir;
- Contexte d'application;
- Contexte d'utilisation;
- Acteurs impliqués dans le maintien et l'utilisation de l'ontologie.

Indépendamment des conseils prodigués par les auteurs précédents, il existe des méthodologies formelles qui facilitent le développement des ontologies.

4.5.1 Méthodologie formelle de développement

De manière pratique, le développement d'une ontologie implique:

1. La définition des classes de l'ontologie;
2. L'arrangement des classes en une hiérarchie taxinomique (sous-classe, superclasse);
3. La définition des attributs et la description de leurs valeurs permises;
4. L'ajout des valeurs reliées aux attributs d'instance.

Il existe toutefois un certain nombre d'approches différentes pour développer une hiérarchie de classes [Uschold, 1996] :

- Approche descendante qui débute par la définition des concepts les plus généraux du domaine;
- Approche ascendante qui débute par la définition des classes les plus spécifiques;
- Approche hydride qui débute par la définition des concepts les plus saillants suivi par une généralisation et une spécialisation correspondante.

Ces différentes approches de développement peuvent être matérialisées au travers d'une multitude de méthodologies formelles différentes [Schreiber, 2000] [Fernández-López ,2002]. Il est toutefois possible de résumer la plupart de ces méthodologies de développement en utilisant l'approche de [Staab, 2001] qui consiste à employer simultanément deux processus distincts pour réaliser le développement d'une ontologie.

Ces deux processus sont respectivement :

- Un méta processus de connaissance : qui s'adresse aux aspects liés à la présentation et au maintien d'une nouvelle solution de gestion de connaissances à l'intérieur d'une organisation;
- Un processus de la connaissance : qui s'adresse à la manipulation d'un système de gestion de connaissance déjà implanté.

Le « méta processus de connaissance » débute par une étude de faisabilité suivie d'une phase de démarrage (Figure 4-7). Les éléments de connaissances sont, par la suite, assemblés les uns aux autres suivant des étapes de raffinement et d'évaluation répétées en boucle. L'ontologie résultante est finalement utilisée sur le terrain avant de connaître une nouvelle phase d'évolution.

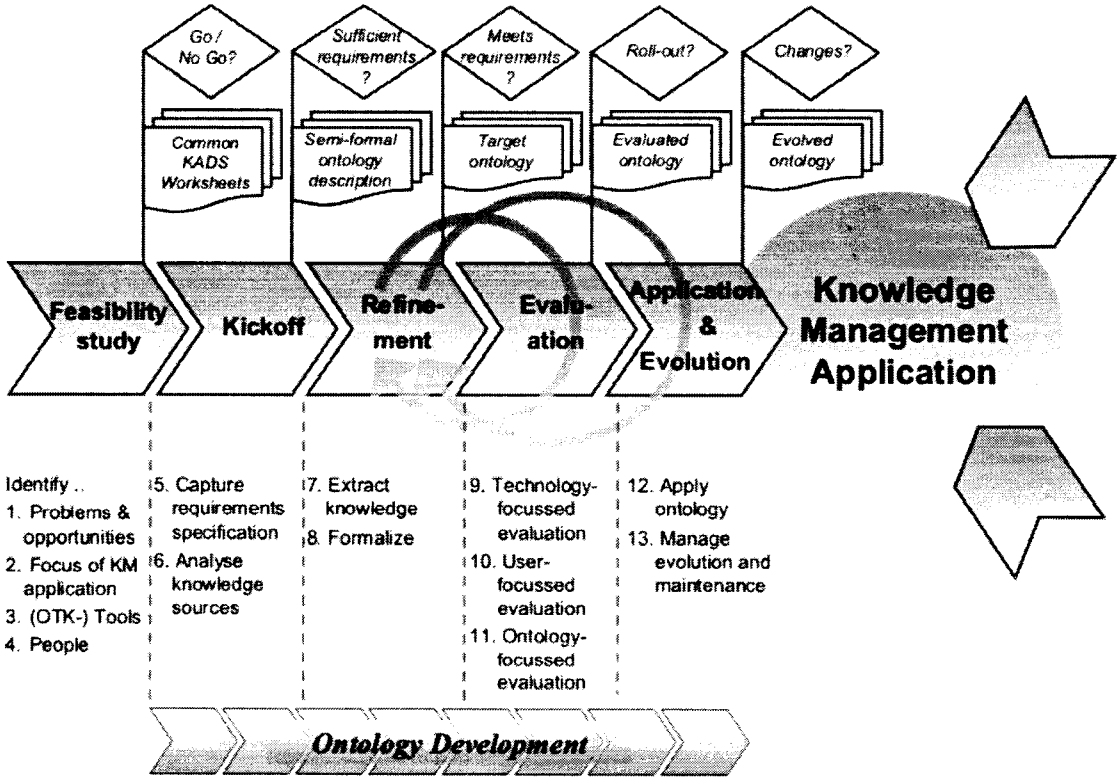


Figure 4-7: Méta processus de connaissance [Staab, 2001].

Les éléments de la Figure 4-7 se lisent comme suit :

- Le tracé principal indique des étapes (phases) qui guident la construction d'une application de gestion de connaissances. Ces phases sont : « Étude de faisabilité», « Démarrage », « Amélioration », « Évaluation » et « Application et évolution ».
- Les sous-étapes les plus importantes sont esquissées au-dessous de chaque phase.
- Les drapeaux au-dessus d'une phase indiquent les principaux résultats de l'étape.
- Le losange situé au-dessus des drapeaux représente les décisions principales qui doivent être prises afin d'accéder à la phase suivante.
- Les phases « application et évolution - amélioration - évaluation - application et évolution » et « évaluation - amélioration - l'évaluation » pourraient s'exécuter selon des cycles itératifs.

Le second processus, appelé « Processus de connaissance », débute lorsque qu'une application de gestion de connaissance commence enfin à être implantée dans une organisation.

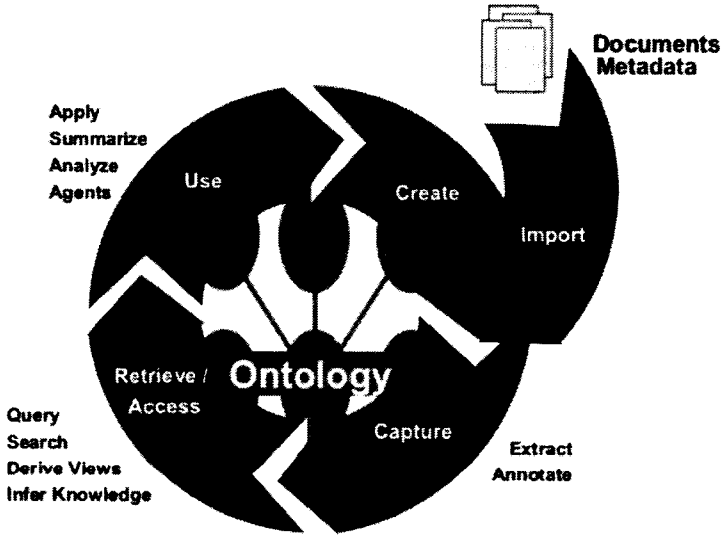


Figure 4-8: Processus de connaissance [Staab, 2001].

Ce processus gravite autour de quatre étapes répétées en boucle:

- **Création de connaissances** et/ou importation de documents et de métadonnées;
- **Capture** des éléments de connaissances afin d'élucider leur importance ou interdépendance;
- **Récupération et accès** aux connaissances pour satisfaire les demandes courantes des utilisateurs;
- **Utilisation** des éléments de connaissances et traitement supplémentaire pour les appliquer à un contexte d'usage particulier.

Le « méta processus de connaissance » et le « processus de la connaissance » devraient être combinés ensemble tout en étant toujours utilisés de manière séparée afin de maintenir l'intégrité respective de chaque processus. La Figure 4-9 illustre ainsi la position orthogonale que ces deux processus devraient maintenir l'un envers l'autre.

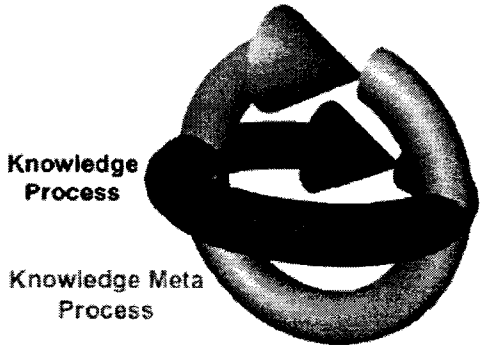


Figure 4-9: Deux processus orthogonaux pour réaliser la gestion des connaissances [Staab, 2001].

Ces deux processus s'utilisent de pair pour réaliser le développement et le maintien des éléments de connaissance. Ces processus sont toutefois réalisés par des acteurs différents qui assurent la définition, la conception et la validation des ontologies correspondantes.

4.5.2 Acteurs

Le nombre et le type d'acteurs impliqués dans le cycle de vie des systèmes de connaissances peuvent varier selon l'importance d'un projet particulier. Un système de connaissance peut toutefois être réalisé à l'intérieur d'une organisation avec la mise en place d'une équipe restreinte formée de gestionnaires, d'experts de contenu et de spécialistes.

[Garwood , 2005] reconnaît quatre rôles clés :

- **Développeur de schéma** : responsable de développer un modèle de données répondant aux exigences d'une large communauté d'utilisateurs.
- **Ingénieur de la connaissance** : responsable de la conception et du maintien des ontologies correspondant au schéma.
- **Fournisseur de données** : responsable de la génération des jeux de données qui correspond au schéma et qui utilise les ontologies.
- **Consommateur de données** : responsable de faire un usage efficace et systématique des éléments de connaissance produits.

Chacun de ces acteurs interagit avec les autres selon une logique particulière (Figure 4-10). Les ingénieurs de la connaissance utilisent le vocabulaire fourni par des utilisateurs. Ces éléments de connaissances utilisent des schémas réalisés par un comité de standardisation. Ce comité exploite lui-même les informations fournies par des fournisseurs de données pour développer et améliorer les schémas de description. Tous ces acteurs travaillent ainsi en boucle pour alimenter le travail des autres.

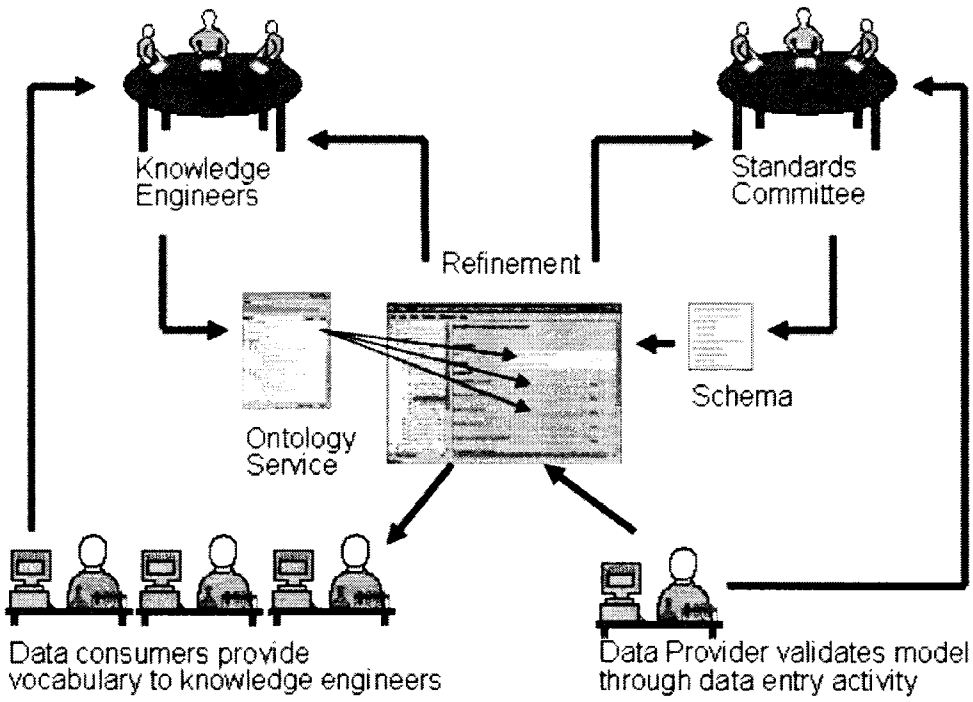


Figure 4-10 : Acteurs impliqués dans le développement d'un système commun [Garwood, 2005]

La Figure 4-10 illustre de manière globale le rôle joué par les différents acteurs sans nécessairement tenir compte du nombre réel des intervenants impliqués dans de tels projets. De manière générale, la conception des systèmes de connaissances est une entreprise complexe qui demande l'emploi d'un ou plusieurs gestionnaires de projets. L'équipe de conception est ainsi formée de cinq ou six acteurs qui travaillent de manière conjointe à la réalisation du système de connaissances (Figure 4-11).

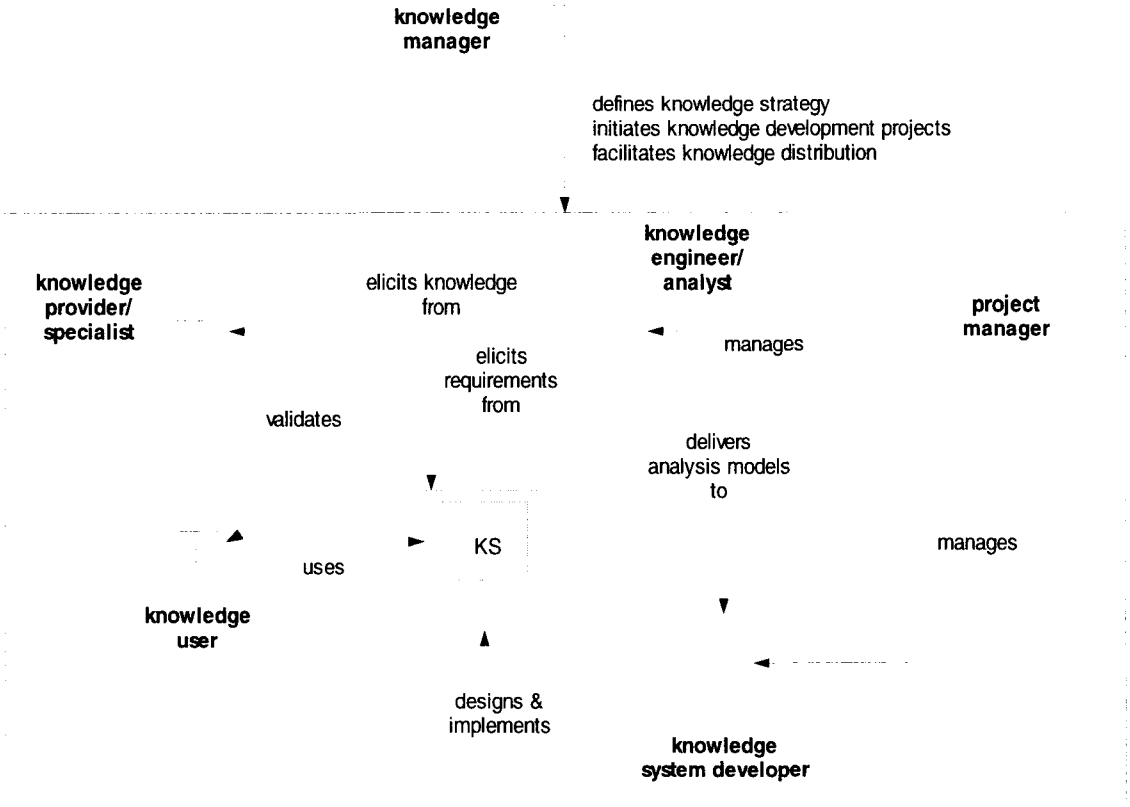


Figure 4-11: Acteurs impliqués dans le développement d'un système de connaissance [Schreiber, 2000].

L'ingénieur des connaissances joue un rôle central dans le développement des ontologies. L'une des tâches les plus difficiles que celui-ci doit pouvoir réaliser concerne la mise à jour et la réutilisation des connaissances. Cette mise à jour implique généralement la fusion et l'alignement de concepts entre différentes ontologies.

4.5.3 Alignement des connaissances

La maintenance des ontologies demande l'utilisation d'une série d'opérations différentes. [Klein, 2001] décrit ces opérations comme suit:

- **Aligner**: « bring two or more ontologies into mutual agreement, making them consistent and coherent ».

- **Combiner**: « using two or more different ontologies for a task in which their mutual relation is relevant »;
- **Faire correspondre**: « relating similar (according to some metric) concepts or relations from different sources to each other by an equivalence relation. A mapping results in a virtual integration ».
- **Fusionner**: « creating a new ontology from two or more existing ontologies with overlapping parts, which can be either virtual or physical »;
- **Gérer les versions**: « a method to keep the relation between newly created ontologies, the existing ones, and the data that conforms to them consistent ».
- **Traduire**: « changing the representation formalism of an ontology while preserving the semantics ».
- **Transformer**: « changing the semantics of an ontology slightly to make it suitable for purposes other than the original one ».

La maintenance des ontologies peut nécessiter une réutilisation d'ontologies déjà existantes, l'expansion de nouvelles sections d'ontologies ou encore la fusion et l'alignement de concepts avec d'autres ontologies existantes. L'alignement d'ontologies consiste à établir des liens de correspondance entre différentes ontologies de manière à faciliter la mise en relation de leurs concepts. La fusion d'ontologies consiste, quant à elle, à combiner ensemble plusieurs ontologies dans le but d'en produire une seule.

L'alignement et la fusion d'ontologie sont des opérations souvent difficiles à réaliser parce que les données manipulées peuvent elle-même comporter des erreurs plus ou moins importantes [Appriou, 2001]. Ces problèmes peuvent prendre la forme d'ambiguïtés, d'incertitudes, d'imprécisions, d'imperfections, de flous et de contradictions.

De plus, ces problèmes se présentent rarement de manière isolée de sorte qu'il n'est pas rare de devoir faire face à plus d'un problème en même temps. Par exemple, les

problèmes d'ambiguïté peuvent être causés par un groupe d'expressions qui traitent des mêmes sujets alors que d'autres expressions sont en même temps en contradiction avec d'autres éléments. Il devient alors difficile de savoir si le problème provient de la contradiction elle-même ou si c'est l'ambiguïté qui est à l'origine de cette contradiction. Il est parfois même impossible de régler ces problèmes sans détruire des éléments au passage.

La Figure 4-12 illustre la taxonomie de problèmes associés à la maintenance des ontologies proposée par [Klein, 2001].

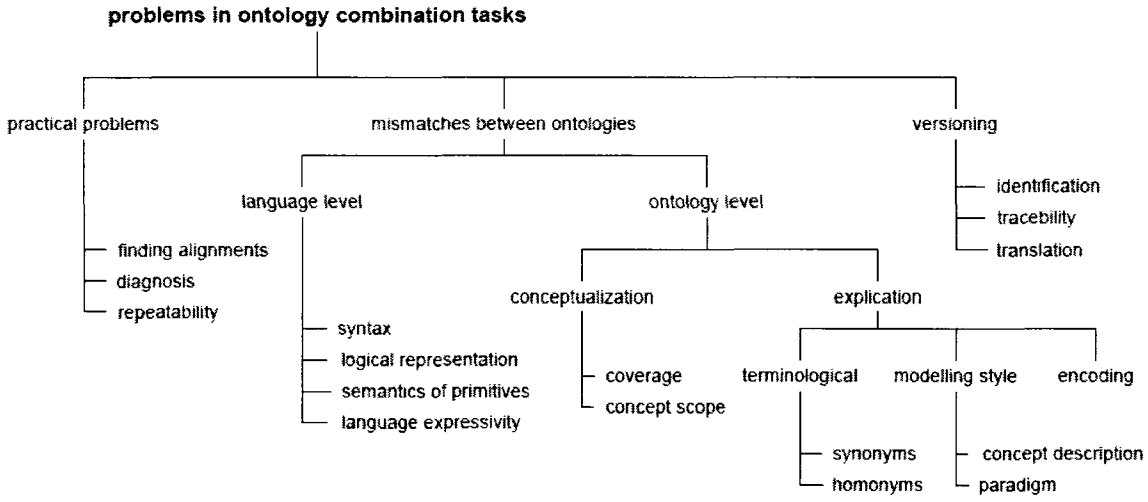


Figure 4-12 : Taxonomie des problèmes associés au pairage des ontologies [Klein, 2001].

Différents problèmes peuvent se présenter à l'ingénieur des connaissances dont, notamment, celui de repérer les éléments à aligner entre les différentes ontologies. Les ambiguïtés qui existent entre les différentes ontologies peuvent aussi bien se situer au niveau du langage utilisé pour la description des ontologies qu'au niveau même de la structure de description de ces ontologies. Le Tableau 4.1 énumère ainsi les différentes catégories d'erreurs qu'il est possible de rencontrer.

Tableau 4.1 : Catégories d'erreurs et de glissements de sens.

Antonyme :	décrit une relation de sens contraire (par exemple: grand est l'antonyme de petit). Son antonyme est synonyme.
Holonyme :	décrit une relation de tout à partie (par exemple: le corps est un holonyme de bras, de même que maison est un holonyme de toit). Son antonyme est méronyme.
Homographe :	décrit une relation de similitude dans l'orthographe. Son hyperonyme est homonyme.
Homonyme :	décrit une relation de similitude formelle entre des mots de sens différent (par exemple: le canon d'une armée et le canon de la messe sont homonymes). Ses hyponymes sont homographe est homophone.
Homophone :	décrit une relation de similitude phonétique entre des mots de sens différent (par exemple: veau et vos sont deux homophones). Son hyperonyme est homonyme.
Hyperonyme :	décrit une classe d'une instance spécifique. X est l'hyperonyme de Y si Y est un (une sorte de) X (par exemple: animal est l'hyperonyme de chien). Son antonyme est hyponyme.
Hyponyme :	X est l'hyponyme de Y si Y est un (une sorte de) X (par exemple "chien" est l'hyponyme de "animal"). Son antonyme est hyperonyme.
Méronyme :	décrit une relation de partie au tout (par exemple: le bras est un méronyme de corps, de même que toit est un méronyme de maison). Son antonyme est holonyme.
Paronyme :	décrit la ressemblance phonétique à un autre mot (par exemple: bailler/bâiller et chasse/châsse sont des paronymes).
Synonyme :	décrit une relation de même sens (par exemple: long, large, vaste, haut sont des synonymes de grand). Son antonyme est antonyme.

Par exemple, la Figure 4-13 illustre une différence sémantique et structurale occasionnée par une synonymie et une polysémie entre les concepts de deux ontologies différentes:

- **Synonymie**: avec l'utilisation de différentes terminologies (les termes « booking » et « reservation » sont synonymes l'un de l'autre).
- **Polysémie** : avec des significations différentes assignées au même mot (le terme « airline » est à la fois utilisé pour décrire une réservation de compagnie aérienne et une réservation de vol).

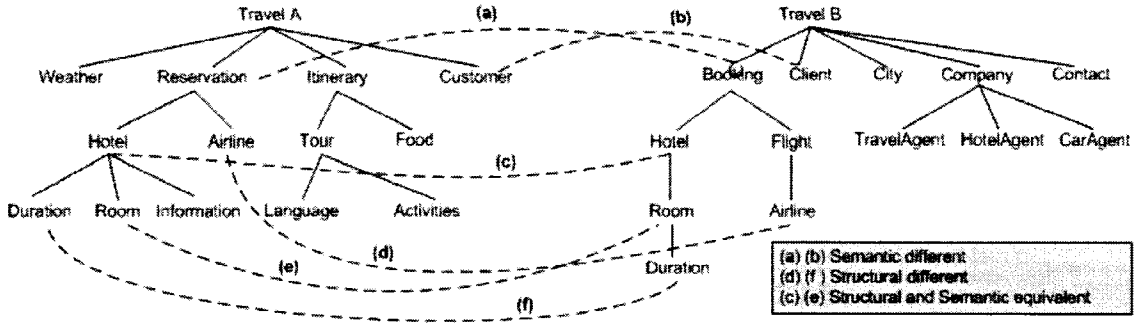


Figure 4-13 : Différences sémantiques et structurales entre deux ontologies [Kiu, 2006].

L'utilisation de taxonomies différentes occasionne des hétérogénéités structurales entre les ontologies [Euzenat, 2004] [Noy, 2000] [Stumme, 2001]. L'utilisation de schémas de description différents aggrave ce problème en rendant encore plus difficile la comparaison entre les ontologies. Il est toutefois possible de réaliser un pairage des schémas pour faciliter l'alignement des ontologies.

Un pairage de schémas consiste à trouver des correspondances sémantiques entre différents éléments de schémas différents. Le processus de pairage peut être modélisé de la manière suivante (Figure 4-14) :

- L'opération de pairage détermine l'alignement (A') pour une paire de d'ontologies (o et o').
- Il existe quelques autres paramètres qui peuvent prolonger la définition du processus de pairage, à savoir :
    1. L'utilisation d'un alignement d'entrée (A) qui doit être accompli par le processus;
    2. Les paramètres de pairage, p (par exemple, poids, seuils);
    3. Les ressources externes utilisées par le processus de pairage, r (par exemple, thesaurus).

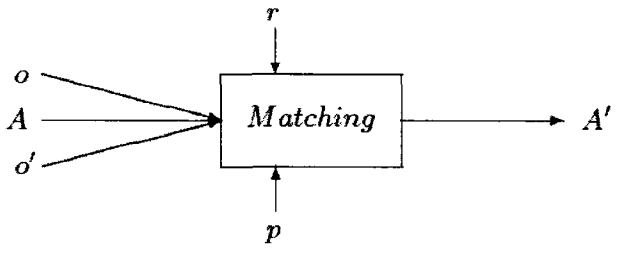


Figure 4-14 : Processus de pairage [Shvaiko, 2005].

La Figure 4-15 présente une taxonomie des méthodes de pairage de schéma [Rahm, 2001]. Il existe ainsi une distinction entre les méthodes basées sur les schémas, sur les instances, les éléments, la structure, le langage et les contraintes.

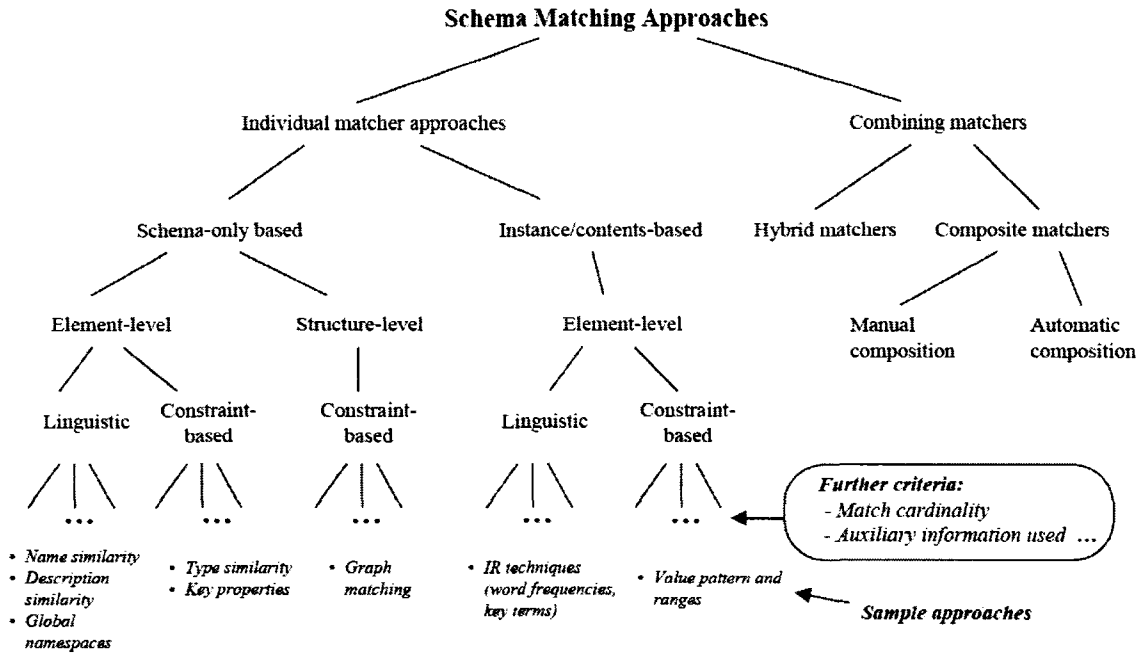


Figure 4-15 : Taxonomie des méthodes de pairage de schéma [Rahm, 2001].

Ces critères de différenciation de cette taxonomie sont basés sur une classification orthogonale entre :

- **Instance** vs. **schéma** : les approches de pairage peuvent considérer les données d'instance (c.-à-d., contenu de données) ou seulement les informations au niveau du schéma.
- **Élément** vs. **structure** : le pairage peut être exécuté pour différents éléments de schéma, tels que des attributs, ou pour des combinaisons d'éléments, tels que des structures complexes de schéma.
- **Langue** vs. **contrainte** : le pairage peut exploiter une approche linguistique (par exemple, basé sur des noms et des descriptions textuelles des éléments de schémas) ou une approche basée sur l'utilisation de contraintes (par exemple, basé sur des clefs et des relations).

- **Cardinalité** : le résultat global de pairage peut relier un ou plusieurs éléments d'un schéma à un ou plusieurs éléments d'un autre selon l'une ces quatre dispositions : 1-1, 1-n, n-1 et n-m. De plus, chaque élément de pairage peut relier un ou plusieurs éléments des deux schémas. En outre, il peut aussi exister différentes cardinalités de pairage au niveau des instances.
- **Information auxiliaire** : la plupart des pairages se fondent non seulement sur les schémas d'entrée mais aussi sur l'information auxiliaire comme les dictionnaires, les schémas globaux, les décisions précédentes de pairage et les entrées réalisées par les utilisateurs.

Cette même taxonomie de méthodes de pairage de schémas a été reprise par [Shvaiko, 2005] pour introduire deux nouveaux éléments (Figure 4-16) :

- **Granularité / interprétation d'entrée**: cette classification est basée sur 1) la granularité du pairage (c.-à-d., soit au niveau de l'élément lui-même ou soit au niveau de la structure), et 2) sur la manière dont ces techniques de pairage interprètent généralement l'information d'entrée ;
- **Type d'entrée** : cette classification est basée sur le genre d'entrée qui est employée par les différentes techniques de pairage.

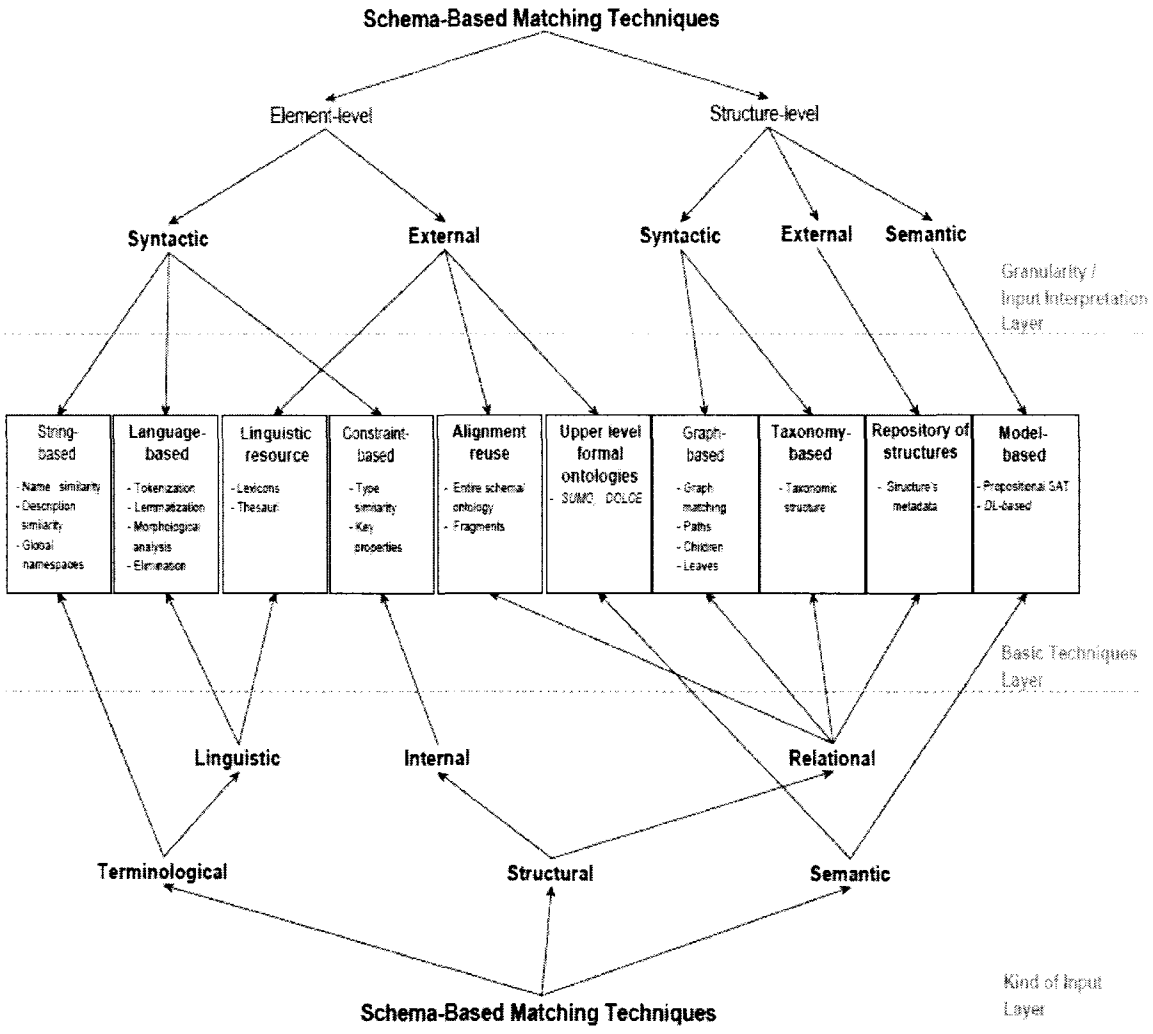


Figure 4-16 : Classification des méthodes de pairage de schéma [Shvaiko, 2005].

Concrètement, l'alignement des ontologies peut être réalisé en utilisant l'une des quatre approches suivantes:

1. **Ontologie partagée**: où une seule ontologie est utilisée. Les inconvénients à utiliser une seule ontologie partagée sont exactement les mêmes que l'utilisation d'une norme commune obligatoire [Visser, 1998].

2. **Ontologie intermédiaire** : où une ontologie externe sert de référence pour réaliser un pairage sémantique entre des ontologies qui sont syntaxiquement différentes (Figure 4-17). Malheureusement de telles ontologies de référence

n'existent pas toujours pour décrire tous les domaines et même lorsqu'elles existent, celles-ci sont souvent peu susceptibles de couvrir l'ensemble des pairages prévus des entres ontologies d'entrée [Sabou, 2006].

3. **Approche linéaire**: où un ensemble de règles est fourni pour chaque ontologie de manière à permettre une communication avec les autres ontologies sans être obligé de recourir à une ontologie intermédiaire. Par exemple, l'alignement des ontologies peut être réalisé en utilisant les définitions de WordNet comme élément de référence ou, encore, en utilisant le score de popularité de Google pour pondérer l'alignement approximatif des concepts entre différentes ontologies [Gligorov, 2007]. L'alignement peut être réalisé au moyen de règles conditionnelles [Chang, 1998], de fonctions [Chawathe, 1994], d'opérations logiques [Guha, 1991], de procédures [Weinstein, 1998] ou tout autre système procédural. Cette approche possède toutefois le désavantage d'exiger une grande complexité de calcul.

4. **Regroupement d'ontologies**: où les ressources sont regroupées ensemble sur la base de leurs similitudes. Des grappes d'ontologies peuvent ainsi être organisées selon une hiérarchie particulière [Visser, 1999].

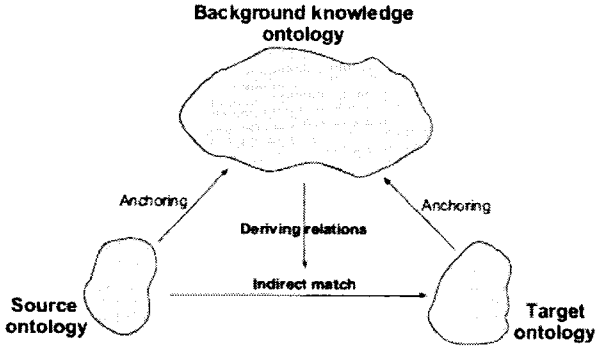


Figure 4-17 : Ancrage sur une ontologie de référence [Aleksovski, 2006].

Indépendamment de la méthode choisie, l'alignement des ontologies devrait permettre à l'ingénieur des connaissances de réaliser une validation structurale des ontologies d'entrée en comparant la position respective de leurs concepts (Figure 4-18).

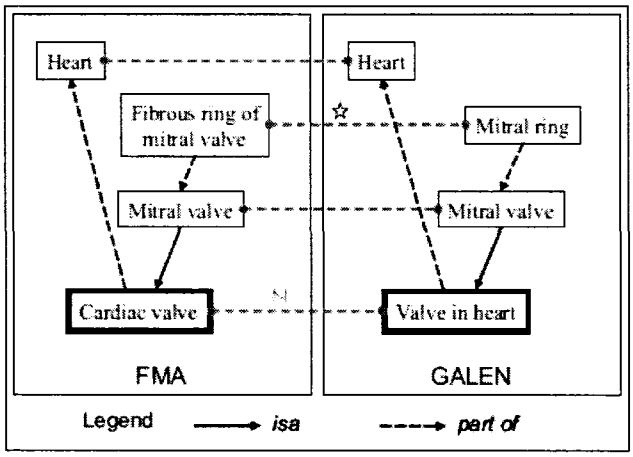


Figure 4-18 : Validation structurale d'ontologies [Zhang, 2006].

Cette validation structurale s'effectue en 1) identifiant les ancres lexicales (c.-à-d., les concepts partagés), 2) en établissant les relations sémantiques implicites et explicites entre les éléments, et 3) en identifiant les ancres structurelles. Cette opération peut ainsi mettre en évidence des différences structurales ou lexicales qui demandent à être éclaircies.

Les opérations de gestion, de fusion, d'alignement et de validation structurale peuvent être réalisées avec différents logiciels [Gómez-Pérez, 2002] [Gargantilla, 2004][Youn, 2006]. L'alignement et la validation structurale des ontologies peuvent notamment être réalisés à l'aide des logiciels Agreement Maker [Cruz, 2007], Autoplex [Berlin, 2001], Automatch [Berlin, 2002], Clio [Miller, 2001], COMA [Do, 2002], Cupid [Madhavan, 2001], Delta [Clifton, 1996], DIKE [Palopoli, 2000], EJX [Li, 1994], FCA-Merge [Stumme, 2001], GLUE [Doan, 2003], HCONE-Merge [Kotis, 2006], LSD [Doan, 2001], MOMIS [Castano, 1999], PROMPT [Noy, 2000], SemInt [Li, 2000], SKAT [Mitra, 1999], Similarity Flooding [Melnik, 2002] et TranScm [Milo, 1998].

La Figure 4-19 et la Figure 4-20 illustrent deux de ces logiciels. La Figure 4-19 présente COMA++, un outil générique qui permet de réaliser le pairage entre des schémas ou des ontologies exprimés dans les langages SQL, XML ou OWL.

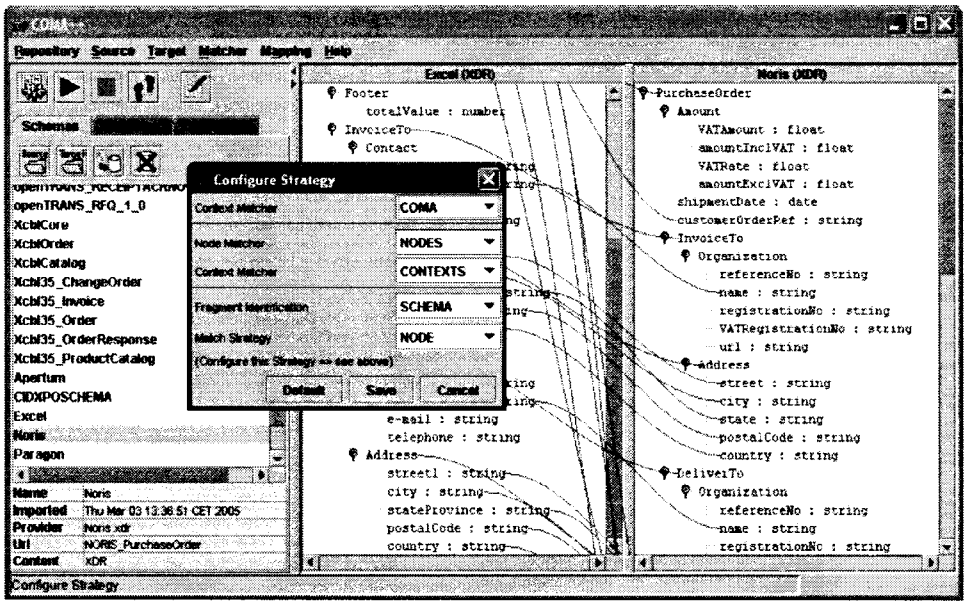


Figure 4-19 : Alignement d'ontologies avec COMA++ [Aumueller, 2005].

Le processus automatique de pairage de COMA++ se déroule en plusieurs étapes. Les schémas ou les ontologies doivent tout d'abord être importés pour être transformés dans une représentation graphique générique. Un arbre est construit en affichant, à la position de chaque branche, le nom des classes et des propriétés selon leurs différents niveaux d'agrégation ou de spécialisation respectif. Le pairage est effectué automatiquement par le logiciel en exploitant différentes stratégies commandées par l'utilisateur. Les similitudes trouvées par le logiciel sont ainsi représentées par une ligne rouge qui relie directement chaque paire d'éléments entre elles. Le résultat du pairage peut être par la suite sauvegardé dans un fichier externe.

La Figure 4-20 illustre le module externe Prompt utilisé à l'intérieur de la plateforme de développement d'ontologies Protégé.

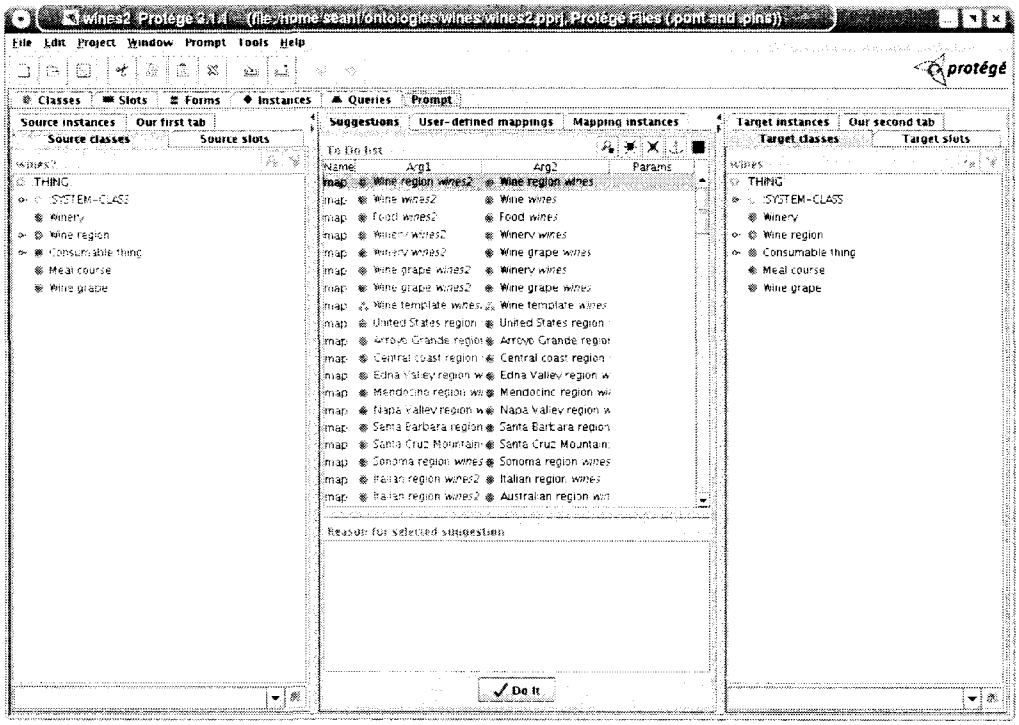


Figure 4-20 : Alignement d'ontologies avec PROMPT [Noy, 2000].

PROMPT utilise une approche semi-automatique pour réaliser la fusion et l'alignement d'ontologies. Ce logiciel prend deux ontologies en entrée et guide l'utilisateur dans la création d'une ontologie fusionnée en sortie. PROMPT détermine les contradictions possibles entre les ontologies et établit automatiquement des suggestions sur la manière de remédier à ces contradictions.

Prompt utilise des mesures de similitude linguistique pour réaliser le pairage des ontologies. Le logiciel arrive ainsi à identifier les conflits de nom, les références brisées, les redondances à l'intérieur de la hiérarchie de classe, ainsi que les restrictions d'attribut qui violent l'héritage de classe.

Dans ce chapitre, nous avons présenté une définition formelle du concept d'ontologie. Nous avons aussi résumé les grands principes de conception d'ontologies ainsi qu'une liste des moyens mise en oeuvre pour réaliser l'alignement de ces ontologies. Nous avons présenté la composition des équipes de conception nécessaires à la mise en place

d'ontologies consensuelles. Finalement nous avons résumé la liste des outils capables de réaliser le pairage entre des schémas ou des ontologies exprimés dans différents langages.

Ces différents outils sont directement utilisés pour la construction d'ontologies sur le Web sémantique. Le chapitre suivant présente le concept du Web sémantique en expliquant son principe de fonctionnement.

CHAPITRE 5 : LE WEB SÉMANTIQUE

Le Web sémantique est un réseau parallèle au Web actuel qui utilise un système de description de ressource, appelé RDF, pour représenter des relations entre ressources. Le RDF peut aussi bien être exprimé en XML qu'en tout autre format (Turtle, Notation 3, etc.).

Le Web sémantique aura, dans quelques années, un impact déterminant sur la manière de concevoir et d'utiliser des contenus d'information.

> *«The semantic web will facilitate the development of automated methods for helping users to understand the content produced by those in other scientific disciplines. On the semantic web, one will be able to produce machine-readable content that will provide, say, automated translation between the output of a scientific device and the input of a datamining package used in some other discipline, or a self-evolving translator that allows one group of scientists to directly interact with the technical data produced by another.*
>
> *These new products will allow users to create relationships that allow communication when the commonality of concept has not (yet) led to a commonality of terms. The semantic web will provide unifying underlying technologies to allow these concepts to be progressively linked into a universal web of knowledge, and will therefore help to break down the walls erected by lack of communication, and allow researchers to find and understand products from other scientific disciplines. The very notion of a journal of medicine separate from a journal of bioinformatics, separate from the writings of physicists, chemists, psychologists and even kindergarten teachers, will someday become as out of date as the print journal is becoming to our graduate students.»*
>
> [Berners-Lee, 2001]

5.1 Resource Description Framework (RDF)

Le RDF est une recommandation du W3C pour la description des ressources Web. Il s'agit d'un modèle de données pour la description de ressources qui peut donc être considéré, à ce titre, comme un modèle de métadonnées (ou méta-métadonnée).

Les expressions RDF se présentent comme des *triplets* composés d'un sujet, prédicat et objet de relation. Lorsqu'ils sont représentés graphiquement, ces éléments prennent la forme conventionnelle d'un cercle pour indiquer une ressource, d'un carré pour indiquer un texte et d'un arc pour en indiquer le sens de la relation.

Les éléments qui composent les triplets RDF servent à indiquer qu'une ressource possède une propriété et une valeur donnée. Ces éléments peuvent utiliser n'importe quelle ressource identifiable par le biais d'une URI[9]. Un même élément peut aussi être repris par un autre triplet pour former une chaîne de relations plus complexes. L'enchevêtrement de ces triplets forme ainsi une « toile de relations sémantiques » désignée sous l'appellation anglaise "*semantic web*" [Berners-Lee, 1998].

Le RDF est généralement exprimé sous la forme d'un document XML bien qu'il puisse aussi être représenté sous forme graphique. L'exemple suivant (Figure 5.1) illustre une expression RDF qui indique qu'une certaine ressource possède comme titre «Conception et validation d'une méthodologie...». Une référence à l'initiative de métadonnées *Dublin Core* est ici représentée par le biais d'un nom qualifié *dc:title* pour indiquer le sens de la relation.

[9] Une URI (*Uniform Resource Identifier*) est une chaîne de caractères qui identifie de manière non ambiguë une ressource Web. Les URI peuvent prendre la forme d'une URL ou d'une URN. Les URL (*Uniform Resource Locator*) identifient les ressources en précisant le protocole de communication et le nom des ressources. Un URN (*Uniform Resource Name*) utilise plutôt un nom logique qui ne précise pas directement la position des ressources sur Internet.

```
<?xml version="1.0"  encoding="ISO-8859-1"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  xmlns:dc="http://dublincore.org/2002/08/13/dces#" >
    <rdf:Description  rdf:about="http://www.polymtl.ca/sujet_theseYan.html">
          <dc:title>Conception et validation d'une méthodologie...</dc:title>
    </rdf:Description>
</rdf:RDF>
```

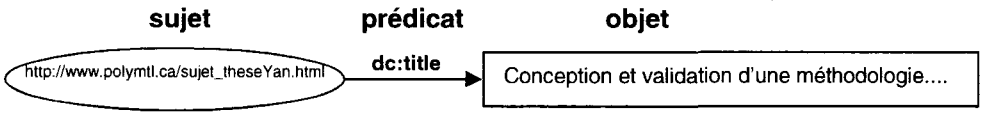


Figure 5.1: Expressions RDF et représentation graphique correspondante.

Les balises RDF peuvent être associées à des pages Web pour améliorer le repérage des contenus. Par exemple, dans un contexte académique, l'utilisation de balises RDF peut améliorer la personnalisation et favoriser une traduction automatique des contenus Web [Berners-Lee, 2001].

5.2 Schéma

Un *schéma* est un formalisme de descriptions qui sert à décrire l'ensemble des usages possibles d'un concept ainsi que la manière de l'utiliser. Les schémas peuvent directement être exprimés en RDF. Ils sont généralement regroupés dans un fichier isolé de manière à favoriser une plus grande réutilisation.

La Figure 5-2 présente le schéma défini par le *Dublin Core Metadata Initiative* pour la définition de la propriété *title* (utilisée précédemment dans l'exemple de la Figure 5.1).

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xmlns:eor="http://dublincore.org/2000/03/13/eor#"  >
    <rdf:Property  rdf:about="http://dublincore.org/2002/08/13/dces/title">
         <rdfs:label>Title</rdfs:label>
         <rdfs:comment>A name given to the resource.</rdfs:comment>
         <eor:comment>Typically, a Title will be a name by which the resource is formally known.</eor:comment>
         <rdfs:isDefinedBy rdf:resource = "http://dublincore.org/2002/08/13/dces/" />
    </rdf:Property>
</rdf:RDF>
```

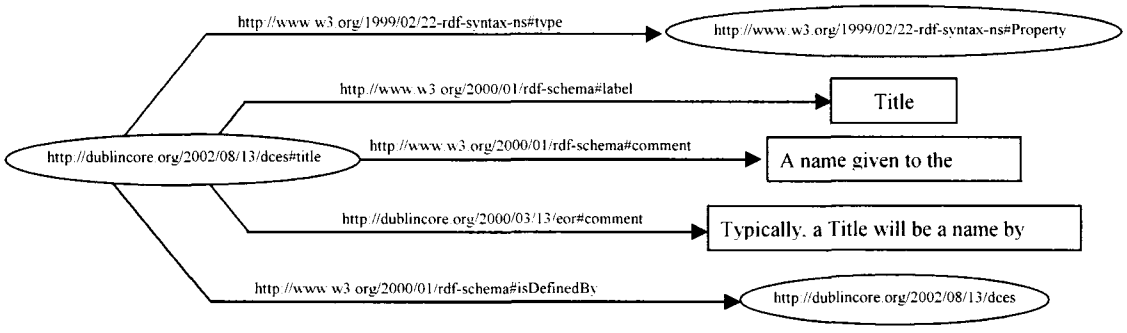


Figure 5-2 : Schéma RDF partiel de l'élément *TITLE* du *Dublin Core Metadata Initiative.*

Ce schéma indique que la propriété *title* possède pour nom *Title* et que cette propriété est attribuée à une ressource pour en décrire le titre. La propriété *title* est elle-même définie (*isDefinedBy*) par une autre ressource localisée à *http://dublincore.org/2002/08/13/dces/.*

Les schémas RDF peuvent ainsi faire référence à d'autres schémas RDF. Le concept d'héritage permet à une classe mère d'étendre ses caractéristiques à des classes dérivées. La Figure 5-3 reprend la ressource «http://www.polymtl.ca/these.html» pour indiquer que cette ressource est aussi une instance de la classe « DoctoralThesis », qui est elle-même une spécialisation de la classe « Publication ». Autrement dit, la ressource «http://www.polymtl.ca/these.html» est décrite de manière non ambiguë comme une forme de publication d'un type particulier appartenant à la classe des thèses de doctorat.

```
<?xml version="1.0"  encoding="ISO-8859-1"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xmlns:dc="http://dublincore.org/2002/08/13/dces#"  >
  <rdf:Description  rdf:about="http://www.polymtl.ca/these.html">
         <dc:title>Conception et validation d'une méthodologie...</dc:title>
         <rdf:type rdf:resource="#DoctoralThesis"/>
   </rdf:Description>
  <rdfs:Class rdf:ID="DoctoralThesis">
         <rdfs:label>Thèse de doctorat</rdfs:label>
         <rdfs:subClassOf  rdf:resource="#Publication"/>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Publication">
         <rdfs:label>Publication scientifique</rdfs:label>
         <rdfs:subClassOf rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
   </rdfs:Class>
</rdf:RDF>
```

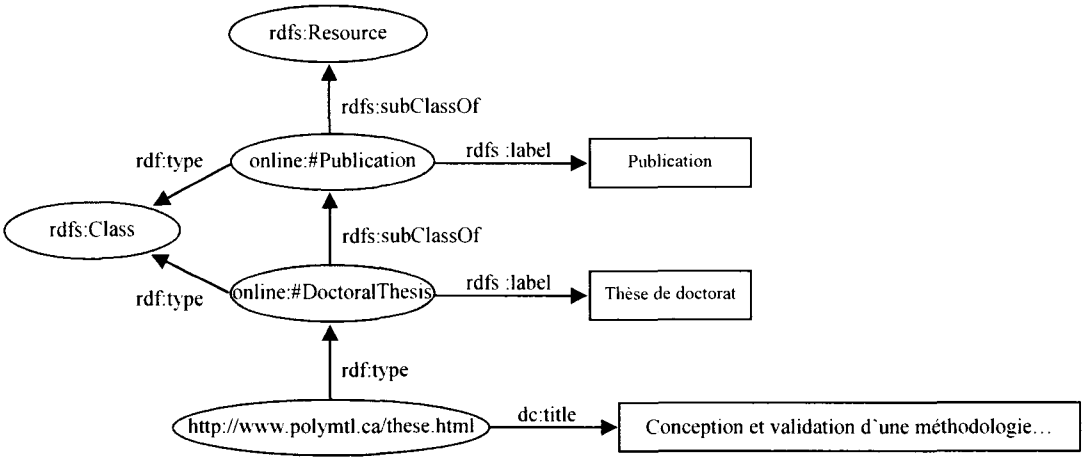


Figure 5-3 : Description RDF exploitant un schéma local.

L'exemple précédent utilise la propriété *rdfs:subClassOf* identifiée par l'espace de nom *rdfs* réservé au schéma de base (*RDF Schema*). Ce schéma de base contient l'ensemble des structures de descriptions nécessaires à la définition courante des éléments. L'exemple suivant présente une vue partielle du schéma RDFS:

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >
    <rdfs:Class rdf:ID="Resource">
                <rdfs:label xml:lang="en">Resource</rdfs:label>
                <rdfs:label xml:lang="fr">Ressource</rdfs:label>
                <rdfs:comment>The most general class</rdfs:comment>
    </rdfs:Class>
    <rdf:Property ID="subClassOf">
                <rdfs:label xml:lang="en">subClassOf</rdfs:label>
                <rdfs:label xml:lang="fr">sousClasseDe</rdfs:label>
                <rdfs:comment>Indicates membership of a class</rdfs:comment>
                <rdfs:range rdf:resource="#Class"/>
                <rdfs:domain rdf:resource="#Class"/>
     </rdf:Property>
</rdf:RDF>
```

Figure 5-4 : Contenu partiel du schéma RDFS.

Le schéma de base RDFS permet l'expression de relations de classe et d'héritage. Le niveau de détail contenu dans ce schéma ne permet toutefois pas l'expression de relations complexes telles que l'inclusion, la jointure ou les contraintes de classe. Des travaux supplémentaires ont donc été entrepris par le DARPA[10] et le W3C pour élargir le

[10] Le DARPA est historiquement responsable des premiers développements du réseau Internet en 1968 (*Defense Advanced Research Projects Agency*, http://www.darpa.mil).

schéma de base RDFS pour permettre l'expression de relations complexes entre les éléments.

5.3 OWL

Le W3C a proposé, en 2002, un nouveau langage d'ontologie spécifiquement adapté au Web. Le *Web Ontology Language* (*OWL*) offre les améliorations suivantes :

- relations taxonomiques entre classes;
- instances de classes et instances de propriété;
- types de propriétés et descriptions d'attributs pour les éléments de classes;
- propriétés d'objet et descriptions de relation entre éléments de classes.

L'exemple suivant (Figure 30) illustre l'utilisation d'une référence à une ontologie OWL indiquant que la ressource «http://www.bodain.ca/these.html» est dérivée de la classe «http://www.cs.umd.edu/...#DoctoralThesis ».

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xmlns:dc=" http://dublincore.org/2002/08/13/dces#"
         xmlns:owl="http://www.ontology.org/2001/03/owl#" >
  <rdf:Description rdf:about="http://www.bodain.ca/these.html">
        <dc:title>Conception et validation d'une méthodologie...</dc:title>
        <rdfs:subClassOf rdf:resource="http://www.cs.umd.edu/owl/cs1.1.owl#DoctoralThesis" />
        <dc:creator rdf:parseType="Resource">
            <rdf:value>Yan Bodain</rdf:value>
            <owl:equivalentTo rdf:resource="http://www.polymtl.ca/personel.rdf#Bodain" />
        </dc:creator>
        <dc:relation>
            <dc:isPartOf>
                <rdf:value rdf:resource="http://www.polymtl.ca/these.html"/>
            </dc:isPartOf>
        </dc:relation>
  </rdf:Description>
</rdf:RDF>
```

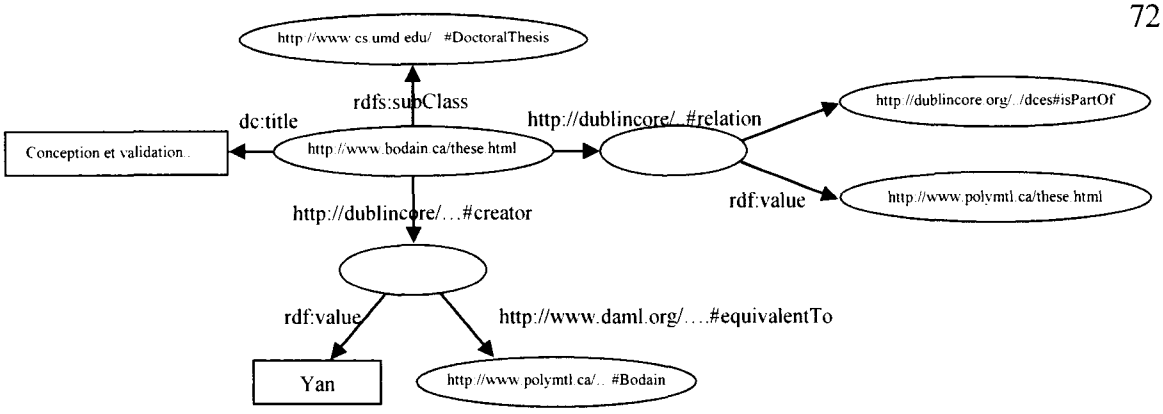


Figure 30 : Description RDF utilisant une référence à une ontologie OWL externe.

5.4 Architecture

Le RDF, RDFS et OWL forment ainsi la base du Web sémantique où chaque technologie joue un rôle particulier dans la description des contenus Web: le RDF permet la définition des ressources, le RDFS réalise la définition des schémas et OWL supporte des définitions sémantiques dans un format d'ontologie adapté au Web (Figure 5-6).

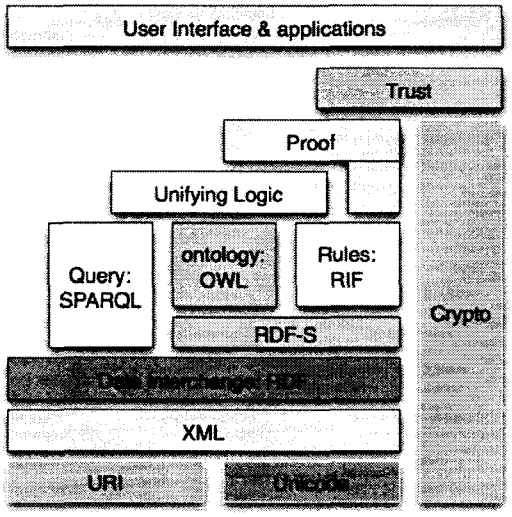


Figure 5-6 : Architecture du Web sémantique (http://www.w3.org/DesignIssues/).

La Figure 5-7 présente le calendrier de lancement technologique du Web sémantique proposé par le W3C. Le Web sémantique y est décrit comme un réseau parallèle au Web

actuel qui utilise un format de données XML pour réaliser la description de ressources Web.

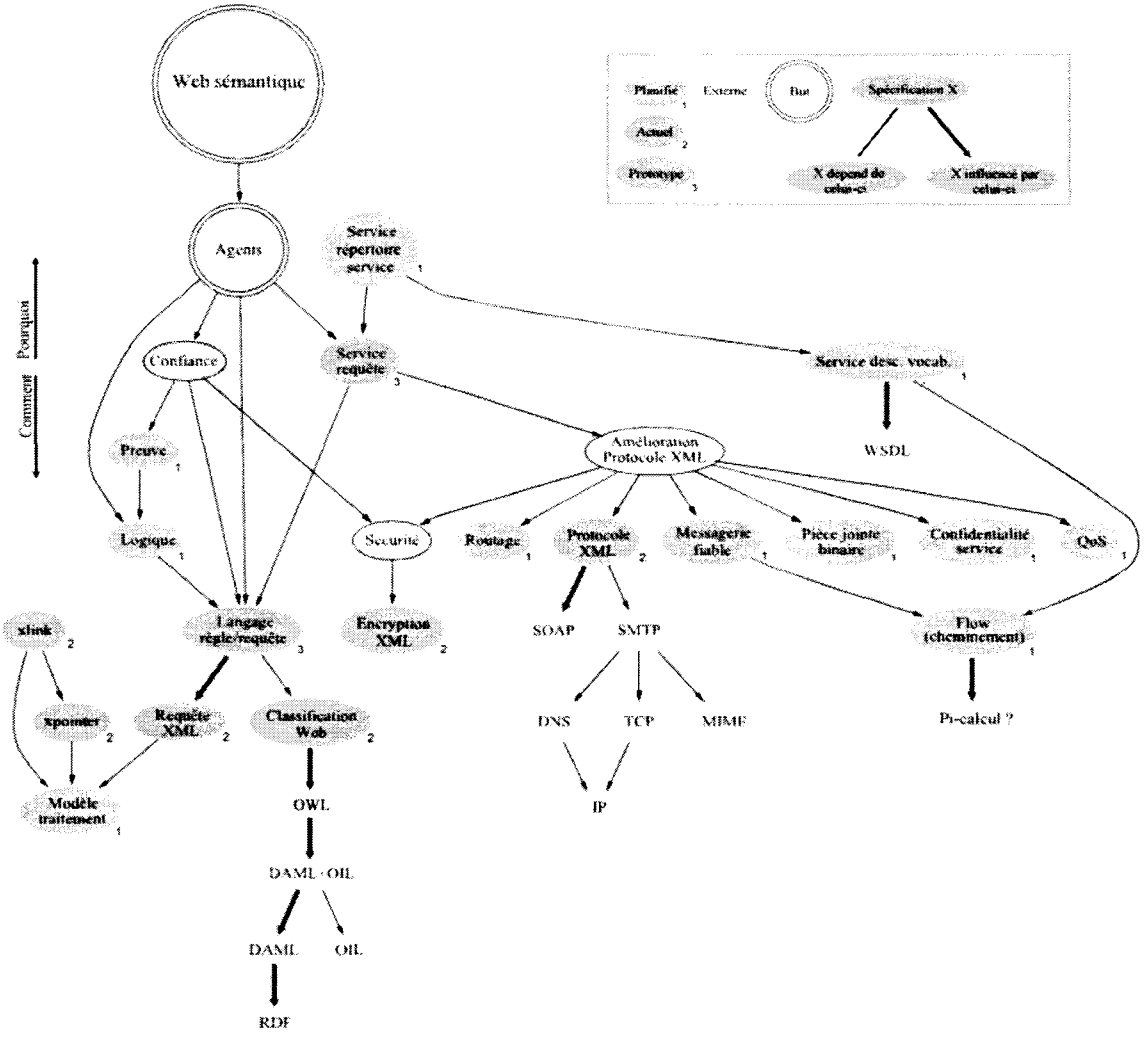


Figure 5-7 : Calendrier de lancement technologique du Web sémantique adapté à partir du calendrier officiel (http://www.w3.org/2001/sw).

Ces descriptions sont réalisées en RDF. Le RDF est lui-même utilisé pour supporter la mise en place du DAML. Le DAML et OIL forment ensemble un langage de description d'ontologie appelé DAML+OIL. Les éléments de ce langage sont utilisés pour supporter le langage d'ontologie OWL (qui fait actuellement figure de norme commune pour la description d'ontologie sur le Web).

Le langage d'ontologie OWL permet, à son tour, la mise en place d'un système de classification Web basé sur l'utilisation de concepts ontologiques. Ce système est exploité par un langage de requête nommé SPARQL (*Simple Protocol and RDF Query Language*) qui est encore actuellement en cours de développement. Ce langage de requête permet à son tour la mise en place d'un mécanisme de preuve et de logique. Le mécanisme de logique exploite les inférences permises par les ontologies pour valider les descriptions trouvées. Le mécanisme de preuve permet, quant à lui, de tracer ou d'expliquer les étapes utilisées pour réaliser ce raisonnement logique. Finalement ces mécanismes de logique et de preuve sont utilisés pour supporter un mécanisme de confiance qui permet à des agents informatiques de discriminer les ressources trouvées en fonction d'information de référence (qui seraient éventuellement sécurisées au moyen d'un algorithme d'encryption et d'authentification particulier). Le Web sémantiques se présente ainsi comme un système automatique où les agents utiliseraient à la fois les informations du Web sémantique et des services Web pour supporter les actions des utilisateurs.

Dans ce chapitre, nous avons fait l'état de l'art des technologies du Web sémantique en présentant le principe de fonctionnement des descriptions RDF, des schémas RDFS ainsi que celui des ontologies DAML+OIL et OWL. Dans le chapitre suivant, nous présentons les problèmes reliés à la conception et la description des contenus sur le Web sémantique.

CHAPITRE 6 : PROBLÈME DE DESCRIPTION DES CONTENUS

Les objets d'apprentissage numériques se distinguent par leur format particulier qui facilite leur réutilisation sur Internet. Un peu à la manière d'une molécule formée d'atomes, les objets d'apprentissage offrent la possibilité de pouvoir être assemblés/désassemblés pour former des objets de structure plus ou moins complexe. La dimension variable des objets d'apprentissage pose ainsi le problème particulier de leur description en exigeant notamment la prise en compte des différentes relations qui interviennent lorsque ces objets d'apprentissage sont décomposés et recomposés de nouveau. Contrairement à un livre papier qui forme un tout unitaire, les objets d'apprentissage demande une description fine qui tient compte des différents niveaux de granularité des éléments qui les composent.

6.1 Structure de description des métadonnées

Les repositoires de cours qui réalisent le classement de contenu en suivant le modèle des grandes initiatives de métadonnées n'ont jamais connu de grand succès auprès des utilisateurs. Les raisons sont nombreuses : perte de droits d'auteur et de propriété intellectuelle, facteurs financiers, absence de principe de réciprocité dans l'échange de contenus [Campbell, 2001], difficulté de la tâche, absence de reconnaissance [Koppi, 2003], absence d'intérêt, résistance au changement [McNaught, 2003], difficulté d'utilisation, couverture inégale des sujets, interface trop complexe, faible qualité des métadonnées, absence de valeur ajoutée [Heery, 2005] ou encore barrière culturelle ou organisationnelle [Harris, 2006].

Il est vrai que les utilisateurs n'apprécient guère d'être obligés de réaliser l'indexation de leur contenus. Les repositoires de cours n'améliorent pas réellement le repérage des contenus recherchés par les humains et n'ont ainsi jamais réussi à déclasser les engins de recherche pour réaliser la recherche de contenus sur le Web.

Dans une étude récente, [Godby, 2004] soulève le problème récurrent d'une utilisation superficielle des structures de métadonnées par les utilisateurs avec une sémantique qui reste beaucoup trop générale pour bien décrire les contenus. Ce problème pourrait théoriquement être imputable à une résistance des utilisateurs à se plier à l'obligation de réaliser la description des contenus. Ce problème pourrait aussi toutefois être imputable à l'incapacité même des structures de métadonnées de répondre aux besoins intrinsèques de la description des objets d'apprentissage. Les grandes initiatives de métadonnées utilisent, en effet, un niveau de description global qui tient rarement compte des caractéristiques particulières de chaque objet d'apprentissage. Cette description reprend, en quelque sorte, le modèle des fiches bibliographiques qui décrivent toujours le contenu des livres de manière homogène en spécifiant le titre, l'auteur et autres informations d'intérêt général. Un livre publié en format papier n'est toutefois aucunement comparable à un objet d'apprentissage élaboré de manière numérique. Les grandes initiatives de métadonnées proposent ainsi des systèmes de description résolument statiques alors que la nature même des objets d'apprentissage permet une agrégation dynamique de leur contenu.

6.2 Annotations manuelles

Il est possible de réaliser une description plus fine des objets d'apprentissage en utilisant des annotations insérées directement à l'intérieur des contenus. Il existe différentes sortes d'annotations, mais les annotations « sémantiques cognitives et computationnelles » (illustrées à la derrière ligne du Tableau 3.2 de la page 33) offrent des avantages particuliers en supportant le sens des contenus annotés tout en permettant à la fois l'interprétation de ces contenus par un humain et une machine.

La génération plus ou moins automatique de ce type d'annotation implique des enjeux et des stratégies de mise en œuvre différents :

- Les annotations réalisées automatiquement impliquent l'utilisation d'algorithmes plus ou moins sophistiqués tels que, par exemple, le recours aux indices d'indexation de WordNet ou de Google pour prioriser les termes annotés (par exemple : OnTeA, PANKOW).
- Les annotations réalisées manuellement impliquent l'effort soutenu d'un humain pour sélectionner les termes et réaliser leur classement à l'intérieur d'une ontologie ou d'un système de référence particulier.

La majorité des recherches actuelles se concentrent principalement sur l'amélioration des systèmes d'annotation automatiques avec l'utilisation d'algorithmes sophistiqués ou encore par l'utilisation de stratégies particulières qui exploitent l'intelligence collective présente sur le Web (par le biais d'informations communes accessibles sur les blogues, wiki, ou moteurs de recherche similaires à Google).

La tendance générale à considérer les annotations manuelles comme une voie de recherche non viable se fonde en bonne partie sur l'idée que la tâche d'annotation est difficile à réaliser pour un humain et que les humains seront, de toute façon, toujours plus lents que les machines. Il est vrai que les systèmes d'annotation automatiques sont capables de produire des centaines, voire des milliers d'annotations à la seconde mais la qualité et la valeur sémantique de ces annotations sont difficilement comparables aux annotations réalisées par un humain.

Les machines sont encore incapables de rivaliser de compréhension avec les humains en ce qui a attrait aux concepts de la vie courante. Les systèmes d'annotations semi-automatiques, qui allient le travail précis d'indexation des humains à celui de la production rapide des machines, semblent donc être une voie de recherche beaucoup plus prometteuse à court terme que les systèmes d'annotation entièrement automatiques. Il se pose tout de même la question de savoir comment supporter l'humain dans la

réalisation d'annotations « sémantiques cognitives et computationnelles » alors que les structures de description des initiatives de métadonnées lui font encore défaut. La réponse réside probablement dans l'utilisation nouvelle du Web sémantique.

6.3 Hétérogénéité des descriptions RDF

Le Web sémantique est le résultat d'un effort de collaboration mené conjointement par le W3C et un nombre important de chercheurs et d'associés industriels (MIT, Stanford University, Hewlet Packard, Nokia, ...). Malgré l'importance des ressources impliquées, il n'existe encore aujourd'hui aucun outil informatique capable de supporter pleinement la conception et l'indexation de contenus de cours destinés au Web sémantique.

Le RDF permet une définition arbitraire des ressources Web. Chaque élément d'une expression RDF peut ainsi être modelé sur mesure pour répondre au besoin ponctuel d'une description. À titre d'exemple, la Figure 6.1 illustre la description RDF d'un objet d'apprentissage prenant la forme d'un paragraphe (« *Une analyse de tâche...* ») nommé « *AHT* » et faisant partie d'une section de cours appelée « *Analyse_ergonomique* ». Cette section de cours est elle-même associée au domaine générique de « *Human Factor* ». Les différentes technologies du Web sémantique (RDF, RDFS, OWL) peuvent ainsi être utilisées de manière personnalisée pour réaliser la construction sur mesure de relations, de schémas ou d'ontologies nécessaires à la description des ressources Web.

Cette flexibilité de description rend l'utilisation du RDF problématique pour la description des objets d'apprentissage car cette pratique s'oppose au modèle des initiatives de métadonnées (Tableau 2.1, page 13) qui privilégient l'utilisation d'une structure de description commune entre partenaires. Le RDF favorise une description arbitraire des ressources Web qui contredit ainsi le modèle centralisé des initiatives de métadonnées.

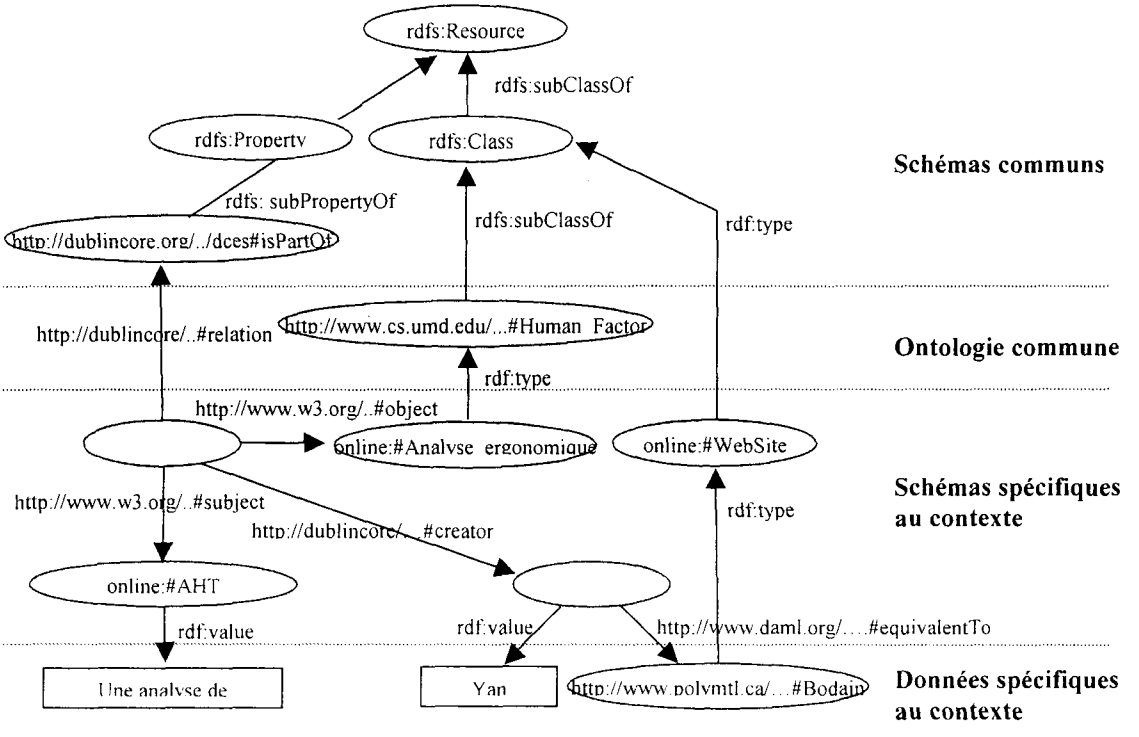


Figure 6.0: Description RDF d'un paragraphe de cours.

Paradoxalement, la flexibilité de description du RDF favorise une plus grande interopérabilité que les modèles de description des initiatives de métadonnées. La présence d'ontologies communes entre les différentes expressions RDF permet, en effet, de réaliser un pont entre les différentes descriptions de manière à assurer une correspondance sémantique entre les éléments (même lorsque la syntaxe utilisée diffère entre chacune des expressions). L'utilisation d'ontologies différentes pour décrire des concepts similaires posent toutefois le problème d'alignement sémantique entre les ontologies utilisées.

6.4 Alignement des ontologies

La conception d'une ontologie est une opération complexe qui demande un travail de réflexion important. Des ontologies réalisées de manière isolée par des individus différents peuvent ainsi donner naissance à des descriptions très différentes d'un même domaine. Ces différences pourront principalement être attribuables à des différences

dans la généralisation ou la spécialisation des classes, à des niveaux de généralisation différents et à des différences de nomenclature.

Il existe toutefois des précautions permettant de limiter ces problèmes [Noy, 2001] :

1- Généralisation/spécialisation des classes :

- Une sous-classe représente le concept de « est une sorte de» concept représenté par une super-classe. On dit qu'il existe une relation de spécialisation lorsqu'une classe B est une sous-classe de A si chaque instance de B est également une instance de A.
- Au moment de réaliser une spécialisation, il est important de s'assurer que les sous-classes possèdent entre elles une relation transitive. Il existe une relation de transitivité lorsque B est une sous-classe de A et C est une sous-classes de B, alors C est une sous-classe de A (par exemple : « analyse ergonomique » comme super-classe de « test d'utilisabilité » elle-même le parent de « analyse heuristique »).
- Il est aussi important d'éviter les cycles (boucles) dans une hiérarchie de classe. On dit qu'il existe un cycle quand une classe A possède une sous-classe B et qu'en même temps B est une super-classe de A. Créer un cycle dans une hiérarchie équivaut à déclarer que les classes A et B sont équivalentes.

2- Niveau de généralisation :

- Les classes filles d'une même mère devraient être à un même niveau de généralité. Les concepts de sens contraire (antonymes) forment des classes à part entière qu'il faut savoir placer au bon niveau de généralisation.
- Lorsqu'une classe mère ne possède qu'une seule sous-classe directe, il est alors possible de supposer l'existence d'un problème de modélisation.
- Inversement, s'il existe plus d'une douzaine de sous-classes pour une classe donnée, il serait alors nécessaire de s'interroger sur la pertinence d'ajouter une catégorie intermédiaire.

3- Nomenclatures :

- Les classes représentent des concepts d'un domaine et non pas des mots désignant ces concepts. Le nom des classes pourrait ainsi changer en utilisant une terminologie différente (par exemple : « charge mentale de travail » pourrait être aisément remplacée par « charge cognitive »).
- Les synonymes d'un même concept ne représentent pas des classes différentes (« interface humain-ordinateur » n'est pas une classe différente de « interface humain-machine » ou «human-computer interaction »). Une liste de synonymes devrait donc toujours être associée aux noms des classes le plus communes pour limiter ce problème.

Malgré toutes ces précautions, il arrive souvent que des ontologies présentent des différences structurales importantes entre elles. Par exemple, la Figure 6-2 illustre les concepts de haut niveau utilisés pour la représentation de l'univers selon quatre modèles différents (CYC, Wordnet, GUM, Sowa).

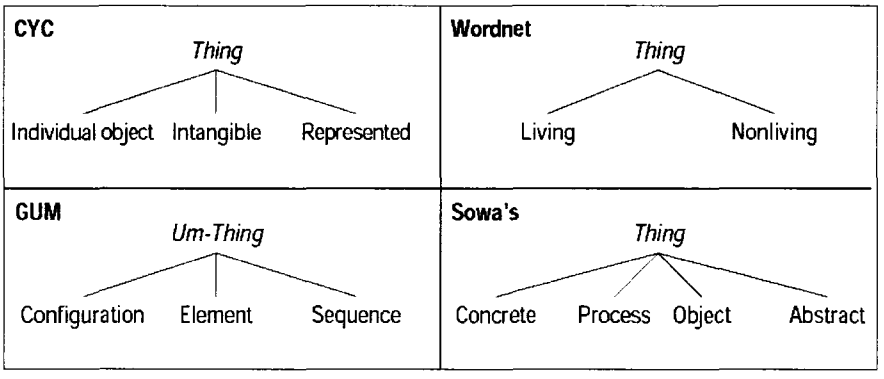


Figure 6-2 : Illustration des différences entre ontologies selon leurs concepts les plus généraux [Chandrasekaran, 1999].

La différence entre ces quatre modèles n'est pas symptomatique en soi mais correspond plutôt à des contextes d'usage différents.

> *« Some of these differences arise because not all of these ontologies are intended to be general-purpose tools, or even explicitly to be ontologies. Another reason for the differences is that, in principle, there are many different taxonomies. »*
>
> [Chandrasekaran, 1999]

Malgré les différences structurales entre les ontologies, il est généralement possible d'établir des ponts entre les concepts en utilisant différentes techniques de pairage (Figure 4-16, page 60). Ces techniques ne réussissent néanmoins pas toujours car différents problèmes peuvent toujours venir s'opposer à la mise en commun des concepts entre différentes ontologies (Figure 4-12, page 54).

Une solution pour réduire l'hétérogénéité structurale et sémantique des ontologies consiste à mettre en place des équipes de travail qui réalisent ensemble la sélection et la définition des éléments d'une ontologie commune.

6.5 Acteurs

Les ontologies sont généralement construites par des équipes spécialisées qui travaillent à établir un consensus dans la définition des termes. Les acteurs impliqués dans la conception d'ontologie sont parfois nombreux (Figure 4-11, page 52).

Le nombre plus ou moins important d'acteurs dans le processus de développement est souvent considéré comme une condition nécessaire à la mise en place d'un consensus commun au sein d'une communauté. Le recours à des équipes de travail entraîne toutefois des problèmes similaires à ceux du développement des initiatives de métadonnées : plus le nombre d'intervenants augmente, moins rapide sera la sélection et la définition des termes. Les descriptions produites seront ainsi toujours en retard sur les besoins actuels des différents utilisateurs.

Cette situation n'est pas en soi problématique pour la production d'ontologie de haut niveau puisque, dans ce type d'ontologie, les concepts de base restent toujours les mêmes. Par contre, cette situation est problématique pour les ontologies de bas niveau qui doivent répondre aux besoins particuliers de certaines applications (Figure 4-5, page 42).

Se pose ainsi la question : comment favoriser la production d'ontologie de bas niveau sans impliquer un nombre important d'acteurs différents mais tout en s'assurant, néanmoins, de la mise en place d'un consensus dans la définition même des termes?

6.6 Questions de recherche

La conception d'annotations « sémantiques cognitives et computationnelles » est une opération fastidieuse à réaliser pour un humain. Il est toutefois probablement possible d'alléger la tâche de conception de ces annotations en favorisant un partage et une réutilisation accrus de contenus annotés entre concepteurs de cours.

Sur la base de notre propre expérience, nous avons constaté que lorsqu'un concepteur de cours récupère un contenu déjà annoté, celui-ci se montre aussi généralement intéressé à conserver la valeur des annotations récupérées. Ces mêmes annotations sont aussi souvent réutilisées pour réaliser des descriptions supplémentaires. Nous pensons donc qu'il est ainsi possible de favoriser la construction d'ontologies en permettant simplement à des concepteurs de cours d'échanger librement des contenus annotés entre eux tout en permettant à chacun de rajouter/retrancher les descriptions sémantiques rattachées aux annotations récupérées. Nous croyons que les éléments d'ontologies récupérés par chacun seront ainsi systématiquement réutilisés pour favoriser la construction d'ontologies de plus en plus importantes. Les ontologies créées à travers de ces emprunts seront, de ce fait, des ontologies essentiellement consensuelles. En d'autres termes, nous croyons qu'il est possible de mettre en place une méthodologie de production ascendante et décentralisée d'ontologies sur le Web sémantique en favorisant directement des échanges de contenus annotés entre concepteur de cours. Les ontologies

produites de cette manière auront ainsi l'avantage de ne pas nécessiter le recours à des équipes spécialisées pour réaliser la mise en place d'un consensus à travers différentes communautés d'utilisateurs.

Cette idée de permettre à chacun de modifier le contenu des autres n'est pas toutefois absente de tout problème. Dans une situation où chaque concepteur de cours fait partie d'une boucle d'échange qui lui permet de récupérer le matériel des autres pour le modifier à son tour, il est facile de comprendre que les modifications réalisées par les uns affecteront la valeur des annotations récupérées par les autres. Ainsi, par exemple, si chacun retire une annotation, il s'en suit une réduction rapide des ontologies récupérées à l'aide de ces annotations par les autres. De même, si chacun retire une définition de l'ontologie récupérée, il s'en suit une diminution tout aussi rapide des ontologies produites. Inversement, si chacun ajoute de nouvelles annotations, il s'en suit un effet de levier dans la diffusion des ontologies récupérées à l'aide de ces annotations par les autres. Incidemment, si chacun ajoute aussi une nouvelle définition à l'ontologie récupérée, il s'en suit un effet de levier encore plus important dans la production d'ontologies communes

Nous faisons l'hypothèse que les emprunts d'annotations réalisés successivement par différents concepteurs de cours se traduisent toujours par un bilan positif entre les ajouts et les retraits réalisés par chacun d'eux, réalisant ainsi un effet de levier positif sur la production globale des annotations. Autrement dit, nous croyons que le nombre de descriptions augmente au fur et à mesure de l'implication d'un nouvel intervenant dans une chaîne de partage.

Pour vérifier cette hypothèse, nous proposons une expérience sur le terrain pour vérifier 1) le taux de réutilisation des annotations et 2) le taux de réutilisation des classes d'ontologie associées à ces mêmes annotations au travers des échanges successifs de contenus entre concepteurs de cours.

Des mesures seront prises pour comptabiliser le nombre d'annotations et d'ontologies créées pour chacun des sujets participant à l'expérimentation. Un fichier de journalisation (« log ») sera aussi utilisé pour quantifier le type et le nombre d'opérations réalisées sur les ontologies par les différents sujets. Les ontologies seront aussi analysées pour évaluer le niveau de découpage des classes entre les différents fichiers utilisées.

Nous espérons ainsi pouvoir démontrer l'accroissement systématique du nombre d'annotations utilisées entre chaque cycle d'emprunt de contenus de cours. Nous pensons aussi prouver que le nombre d'ontologies (rattachées à ces mêmes annotations) ne cessera d'augmenter au fur et à mesure des emprunts de contenus entre les différents sujets.

CHAPITRE 7 : MÉTHODOLOGIE

La méthodologie utilisée pour valider cette thèse consiste en la réalisation d'un prototype logiciel qui permet l'échange de contenus annotés et l'utilisation de ce logiciel avec huit sujets différents pour mesurer l'ampleur des emprunts réalisés entre chacun d'eux :

1. **Conception et développement d'un prototype logiciel.** Élaboration d'un modèle RDF nécessaire pour supporter la description des contenus. Conception d'un logiciel. Conception d'une interface graphique. Test sommaire avec des utilisateurs. Développement d'un système de journalisation (« log ») pour enregistrer les opérations réalisées par les utilisateurs.
2. **Expérimentation.** Récupération de contenus de cours traitant de l'ergonomie des interactions humain-ordinateur (IHO). Sélection de 8 sujets pour réaliser l'annotation de ces contenus. Leur tâche consiste à ajouter des annotations aux contenus de cours à l'aide du logiciel pour leur assigner différentes classes d'ontologies. Les sujets peuvent aussi récupérer des contenus déjà annotés par leurs collègues pour les modifier si nécessaire. Ils peuvent aussi construire de toute pièce de nouvelles ontologies ou récupérer des ontologies déjà existantes sur le Web.

7.1 Prototype logiciel

Les contenus destinés au Web sémantique se distinguent des contenus Web conventionnels par l'utilisation d'annotations servant à relier ces contenus à des ontologies particulières. Sur le Web sémantique, les concepteurs de cours produisent des pages HTML à l'aide de logiciels d'édition conventionnels (Frontpage, WebExpert, HotMetal, etc.) alors que les annotations sont réalisées à l'aide de logiciels spécialisés (AktiveDoc, Annotea, Annozilla, etc.). Au moment de réaliser la conception et la mise à jour des cours, certains contenus doivent être ajoutés/retranchés aux pages HTML tout

en déplaçant ces contenus entre différentes sections du site Web. Ces adaptations demandent généralement des corrections équivalentes aux annotations et aux ontologies associées aux contenus déplacés. Il devient ainsi nécessaire d'utiliser simultanément un logiciel d'édition HTML, un logiciel d'annotation et un éditeur d'ontologie. Les opérations de copier/coller deviennent ainsi difficiles à réaliser.

7.1.1 Annotation robuste

Les annotations qui résistent à l'édition de texte tout en gardant intacte leur description peuvent être définies comme des « annotations robustes ». Les annotations robustes sont vitales au développement du Web sémantique afin de permettre aux auteurs de pouvoir apporter des changements à leurs contenus sans être obligés de se soucier de la perte possible d'annotations au moment de déplacer ces contenus entre différentes sections de texte ou différentes pages de leur site Web.

Différentes stratégies ont été utilisées jusqu'ici pour rendre les annotations robustes [Bernheim, 2002]:

1. Calculer une signature à partir des contenus;
2. Autoriser seulement les annotations à des positions prédéfinies;
3. Utiliser une combinaison de texte annoté et de contenus avoisinants pour reconstruire l'ancrage des annotations.

Ces stratégies donnent de bons résultats pour les tâches d'édition simples mais échouent lorsqu'il s'agit de réaliser des modifications importantes. La Figure 7-1 illustre la complexité du problème.

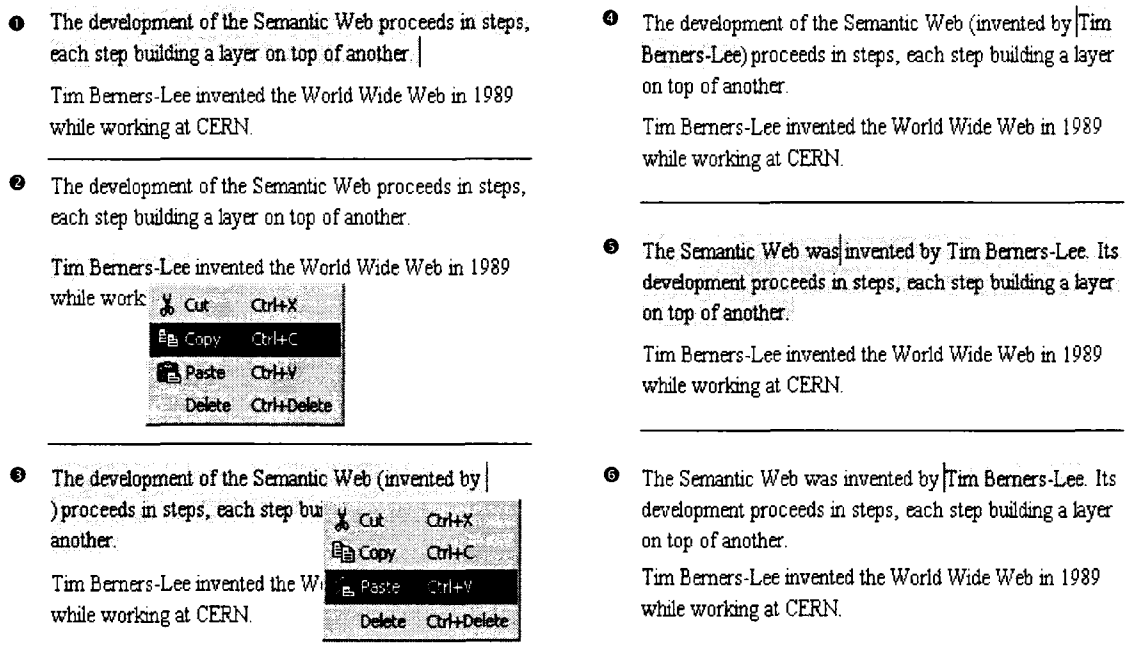


Figure 7-1: Manipulations réalisées sur les annotations.

Dans cet exemple, l'emplacement d'une annotation est révélée par le texte grisé qui apparaît et disparaît lorsque le curseur est déplacé à l'intérieur du document (❶). Dans la section ❷, l'utilisateur sélectionne une partie du texte déjà annotée pour la coller dans une section liée à une autre annotation (❸). Dans cette situation, les deux annotations devraient coexister ensemble de manière à ce que, lorsque l'utilisateur déplace le curseur, l'annotation interne puisse être toujours visible (❹) sans pour autant réaliser d'interférence avec l'annotation qui l'englobe. Chacune des annotations devrait aussi permettre des modifications subséquentes réalisées au texte qui les supporte (❺) sans jamais toutefois compromettre les annotions avoisinantes (❻). Pour rendre ce problème encore plus complexe, l'ensemble de ces manipulations devraient pouvoir être réalisées sur différents documents à l'intérieur d'une ou de plusieurs pages Web localisées sur des sites parfois différents.

Les systèmes d'annotation actuels tels que Annotea, Mangrove, Melita ou OntoMat, (Tableau 3.1, page 22) permettent la production de descriptions RDF, mais ne parviennent pas à produire des annotations robustes résistantes aux modifications

apportées par les auteurs. Avec ces systèmes, les opérations aussi simples que "copier et coller" pourraient altérer ou endommager irrémédiablement le contenu des annotations. De plus, les systèmes tels que Mangrove, MnM et QBLS utilisent des documents XML qui limitent la granularité des annotations à la disposition même des noeuds XML, rendant ainsi la tâche d'annotation encore plus restrictive.

L'utilisation d'annotations robustes peut parfois poser des problèmes. Des risques de glissement de sens sont possibles lorsque les contenus sont modifiés alors que la valeur même des annotations reste inchangée. Les modifications successives de texte à l'intérieur d'un document pourraient ainsi éventuellement modifier la sémantique d'une annotation alors que le sens même du texte est modifié. Il n'existe pas de réponse facile à ce problème et la meilleure solution consiste probablement à donner les moyens aux auteurs de pouvoir toujours vérifier l'alignement constant entre le texte et l'annotation correspondante durant tout le processus d'édition.

L'un des plus grands défis dans la production de contenu pour le Web sémantique consiste donc à donner aux auteurs les moyens de mettre à jour leurs documents avec une flexibilité (lexicale et syntaxique) sans jamais affecter la sémantique des annotations.

7.1.2 Requis

Les annotations robustes ne peuvent pas être directement obtenues par l'application d'un algorithme logiciel. Une approche globale doit être privilégiée pour permettre aux auteurs de bien comprendre l'évolution des descriptions lorsqu'ils réalisent l'édition de contenus annotés. Nous soutenons que cinq conditions sont requises pour réduire les risques d'un éventuel désalignement entre les contenus et les annotations correspondantes :

1. Ancres robustes: les annotations ne devraient jamais être détruites lorsque le contenu est déplacé ou lorsque des insertions sont réalisées à l'intérieur des pages Web. Les

annotions devraient ainsi toujours rester liées à leur contenu cible même lorsque ces contenus sont coupés / copiés / déplacés vers d'autres parties du document ou d'autres sites Web.

2. Transparence: au moment de réaliser ces modifications, les auteurs devraient pouvoir rester conscients de l'impact de leur action sur le modèle de description des données. Cela ne signifie pas pour autant que les descriptions RDF devraient être constamment visibles aux auteurs. Bien au contraire, il serait plutôt préférable de rendre ces descriptions moins visibles de manière à réduire la complexité de la tâche et de rendre l'interface plus facile à utiliser.
3. Ontologies multiples: le Web fonctionne en « permettant à tout le monde de pouvoir (techniquement) dire n'importe quoi sur n'importe quoi » [Berners-Lee, 1998-2]. Conséquemment, les auteurs devraient avoir la liberté d'utiliser des ontologies, ou des parties d'ontologies, sans avoir à souscrire à l'une d'entre elles en particulier. Les auteurs devraient ainsi être en mesure d'utiliser des ontologies et de pouvoir adapter leurs contenus pour répondre à leurs besoins immédiats.
4. Granularité: les annotations ne devraient pas être restreintes à une taille particulière et devraient aussi bien pouvoir être appliquées à de petites unités de texte que des grandes unités de documentation. Par exemple, dans le cadre d'un cours universitaire, les annotations devraient pouvoir être présentes au niveau du cours entier, d'un document particulier, ainsi qu'au niveau d'un paragraphe ou d'un mot isolé.
5. Mise à jour dynamique: chaque nouvelle modification apportée aux contenus par les auteurs devraient être automatiquement reportée aux éléments connexes afin de préserver intact l'alignement général des éléments Le lien de dépendance entre les ressources n'est pas toujours limité à la frontière d'une page unique et certaines modifications pourraient également devoir être appliquées à d'autres documents aussi. Pour cette raison, les systèmes d'annotation devraient toujours réaliser une supervision de toutes les modifications apportées par les auteurs et de corriger, si nécessaire, l'alignement des éléments en fonction des modifications apportées.

7.1.3 Démarche de conception

Notre objectif est de concevoir, de développer et de tester un système d'annotation robuste qui peut facilement être utilisé par toute personne en répondant aux cinq conditions précédentes. Pour atteindre cet objectif, nous avons suivi une démarche de conception qui se développe en cinq étapes distinctes :

1. Générer un identifiant unique pour chaque nouveau paragraphe inséré dans le document et utiliser un algorithme pour étendre ou réduire la portée de cet identifiant en fonction des modifications (couper / copier / coller / insérer) apportées au texte. Un nouvel identifiant est également créé pour chaque nouveau texte (paragraphe, mot ou une lettre) annotée par l'utilisateur.
2. Mettre au point une interface qui présente une rétroaction immédiate sur les changements apportés par les utilisateurs en mettant en évidence les limites des blocs annotés et en montrant directement les éléments qui s'y rattachent.
3. Développer un éditeur d'ontologies sous la forme d'un arbre hyperbolique qui permet aux utilisateurs d'ajouter/modifier/supprimer directement des éléments à l'aide de leur souris.
4. Mettre en place un arbre de navigation donnant accès à l'ensemble des dossiers liés à un site Web. Cet arbre permet ainsi aux utilisateurs de modifier le contenu des pages ainsi que la structure entière du site Web.
5. Implanter un générateur automatique de fichiers pour régénérer les pages HTML, mettre à jour les liens entre les différentes sections et relier les différents contenus HTML à un fichier RDF externe destiné à décrire leur contenu sur le Web sémantique.

7.1.4 Preuve de concept

La Figure 7-2 présente KATIA (*Knowledge Annotation Through Inbuilt Anchors*), un système d'annotation qui prend la forme d'une application de traitement de texte conventionnelle. Katia permet la conception et l'échange d'annotations entre différents

documents. Ce logiciel permet aussi de récupérer les ontologies associées à l'annotation récupérée en même temps que celle-ci.

La partie gauche de l'écran présente un arbre de navigation qui permet la création et la manipulation de fichiers à l'intérieur d'un site Web local. La partie centrale de l'application permet la création/édition de contenus HTML. Les annotations sont réalisées en sélectionnant directement une région de texte (ou image) et en choisissant l'élément d'ontologie correspondant dans l'arbre hyperbolique de droite.

KATIA permet d'annoter des documents selon différents niveaux de granularité, en partant de la structure entière d'un site Web jusqu'au niveau fin d'une portion de texte. Les annotations peuvent ainsi être directement attachées à un document entier, un paragraphe, une phrase, un mot ou une lettre isolée. Une annotation est réalisée en sélectionnant directement un élément de texte (ou image) et en choisissant une classe correspondante dans l'éditeur d'ontologie. Les positions de départ et d'arrivée de l'élément sélectionné sont utilisées pour insérer un bloc d'identification unique dans le document.

Un algorithme modifie par la suite la portée de ce bloc en réponse aux actions posées par l'utilisateur. Lorsque des sections de textes sont copiées ou déplacées, l'algorithme modifie alors automatiquement la frontière de ces blocs pour préserver intact l'alignement des annotations aux contenus. Ce mécanisme automatique libère ainsi les utilisateurs des contraintes généralement associées aux éditeurs d'annotation conventionnels en leur permettant de réaliser des manipulations directes (couper/copier/coller) entre différentes pages, voire même différents sites Web, sans prendre de précautions particulières.

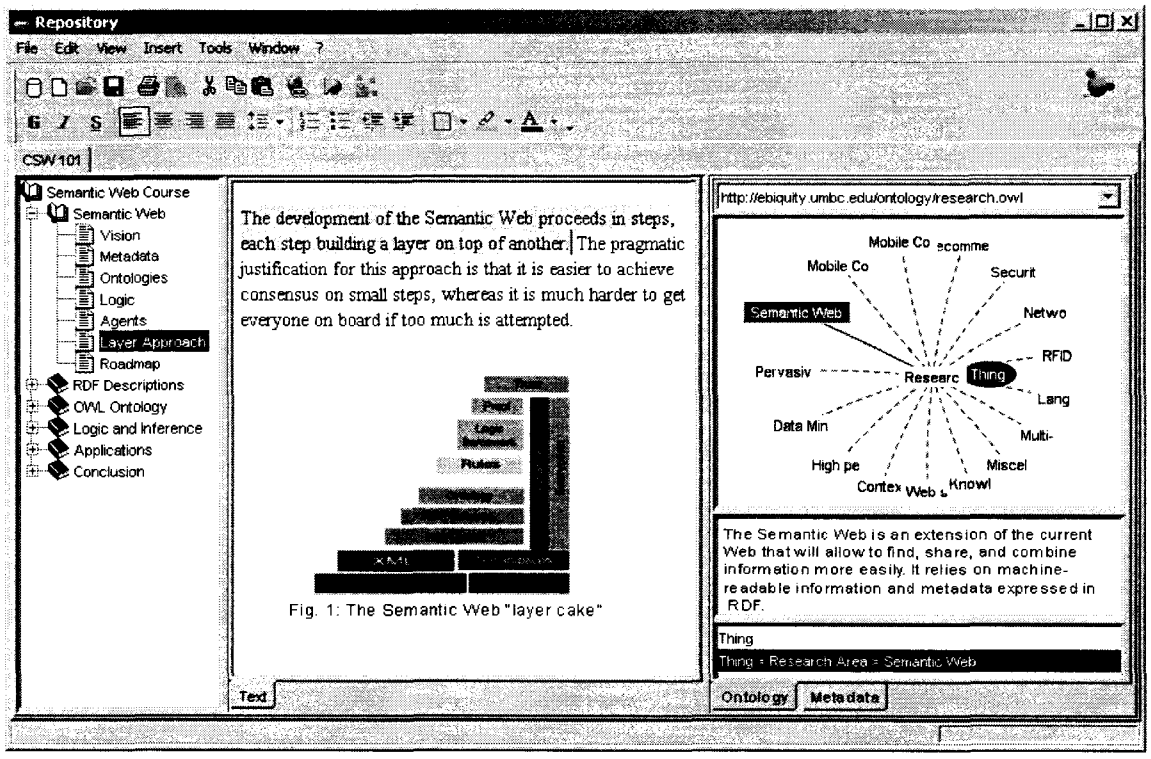


Figure 7-2 : Vue globale de l'interface de KATIA avec son système de navigation de fichiers (gauche), son panneau d'édition (centre) et son éditeur d'ontologie (droite).

Pour chaque nouvelle insertion réalisée dans document, KATIA produit une description RDF qui décrit la date de création ainsi que le nom et les coordonnées de l'auteur correspondant. Une description RDF est aussi automatiquement produite lorsqu'un dossier est inséré ou déplacé à l'intérieur du site Web, de manière à préciser la position courante de cette ressource à l'intérieur de la structure globale du site Web.

7.1.5 Éditeur d'ontologie

La partie droite de KATIA (Figure 7-2) présente un panneau d'ontologies. Ce panneau est composé d'une liste éditable (*comboBox*) qui permet la sélection directe d'ontologies OWL sur le Web. L'ontologie sélectionnée est représentée sous la forme d'un arbre hyperbolique.

Le choix d'utiliser ici un arbre hyperbolique s'appuie sur une étude [Pirolli, 2003] qui démontre les avantages indéniables de ce type de représentation pour les tâches de recherche et de récupération de contenu: un arbre hyperbolique permet, en effet, à un utilisateur d'examiner un plus grand volume de nœuds et ce, à une plus grande vitesse, que les arbres conventionnels.

L'ontologie peut être directement manipulée par l'utilisateur à l'aide d'une souris. Un menu contextuel (Figure 7-3) permet notamment de réaliser des opérations élémentaires (couper, copier, coller, insérer, supprimer, renommer) sur les éléments d'ontologie. Les classes sélectionnées sont représentées par un rectangle bleu qui encadre leur nom ainsi que par une branche continue qui relie les classes sélectionnées à la racine de départ de l'arbre hyperbolique.

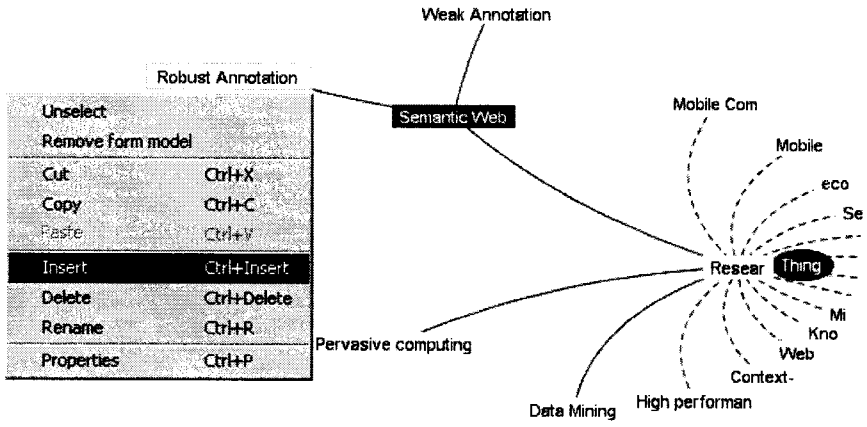


Figure 7-3: Menu contextuel pour l'édition de classe

KATIA permet à un utilisateur d'étendre les classes de base d'une ontologie en ajoutant directement des noeuds aux différentes branches de l'arbre hyperbolique. Les classes sélectionnées ou rajoutées par l'utilisateur sont facilement reconnaissables au dessin de leur ligne pleine. Une liste située directement sous l'arbre hyperbolique (Figure 7-2) permet de retracer les nœuds sélectionnés lorsque le comportement de rotation de l'arbre hyperbolique rend temporairement invisible ces noeuds à l'écran.

L'ontologie modifiée peut, par la suite, être sauvegardée dans un fichier local, ou dans une base de données SQL, avant de servir à son tour à la construction d'annotations RDF. Cette ontologie peut aussi être récupérée et modifiée à l'aide d'un logiciel d'annotation (tel que Protégé) avant d'être utilisée de nouveau par KATIA.

7.1.6 Générateur de pages Web

Une option du logiciel permet la construction et le transfert automatique de pages HTML sur un serveur public (Figure 7.4). Les pages sont directement construites en HTML de manière à permettre leur transfert sur tout type de serveur Web, indépendamment de la présence ou non de scripts dynamiques (ASP, JSP, PHP, etc.).

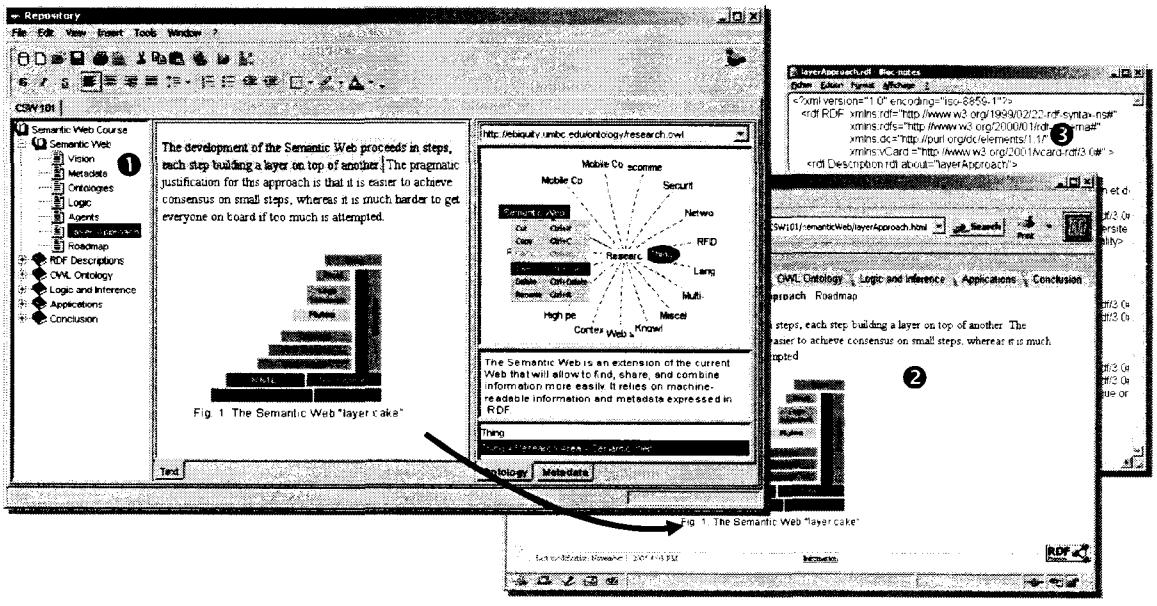


Figure 7.4: Génération de pages Web.

Les pages construites (❷) reprennent la disposition des contenus présentés dans la liste arborescente de KATIA (❶ et aussi la Figure 7-6) alors que les annotations sont enregistrées dans un fichier RDF externe (❸ et aussi la Figure 7-7).

La correspondance entre les pages HTML et les fichiers RDF est réalisée par le bais d'une information insérée directement dans l'en-tête des pages HTML :

```
<head>
<link rel="meta" type="application/rdf+xml" href="layerApproach.rdf" />
</head>
```

Cette information permet aux agents informatiques, présents sur le Web sémantique, de retracer les descriptions Dublin Core (sujet, description, date, format, langue, créateur, droits) et les ontologies correspondantes aux différents contenus de chaque page Web.

Le logiciel permet la création d'un index (Figure 7-8) qui résume la totalité des annotations disponibles sur le site. Cet index prend la forme d'une hiérarchie de concepts en énumérant la position de chaque concept à l'intérieur du site Web. L'index est construit automatiquement par le logiciel en reprenant les classes d'ontologie reliées aux annotations et en énumérant la position des pages Web où chacune de ces annotations a lieu. Les classes d'ontologie sont représentées en ordre, en débutant par des concepts généraux et en se terminant par des concepts particuliers sous la forme d'une liste hiérarchique. L'extrémité inférieure de chaque liste arborescente contient les mots associés à chaque annotation ainsi qu'un lien permettant d'accéder la page correspondante.

7.1.7 Architecture

Ce logiciel est développé en Java et exploite la librairie Jena 2 [Jena, 2003] pour réaliser la lecture/écriture de triplets RDF à l'intérieur d'une base de données SQL (Figure 7.5). L'utilisation d'une base de données centrale permet ainsi la mise en place d'un repositoire commun qui favorise le partage de contenus entre différentes communautés d'utilisateurs.

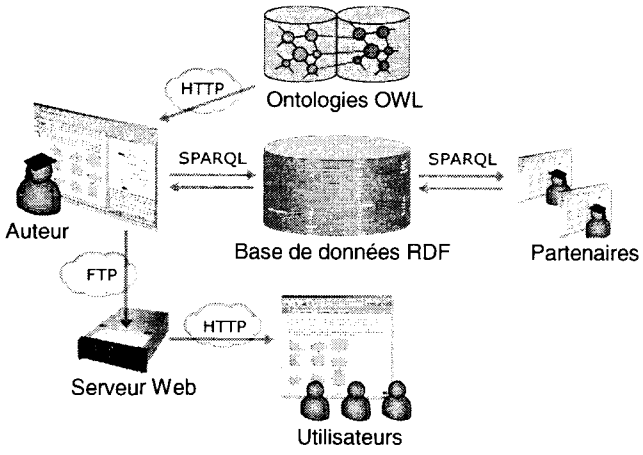


Figure 7.5: Architecture logicielle de KATIA.

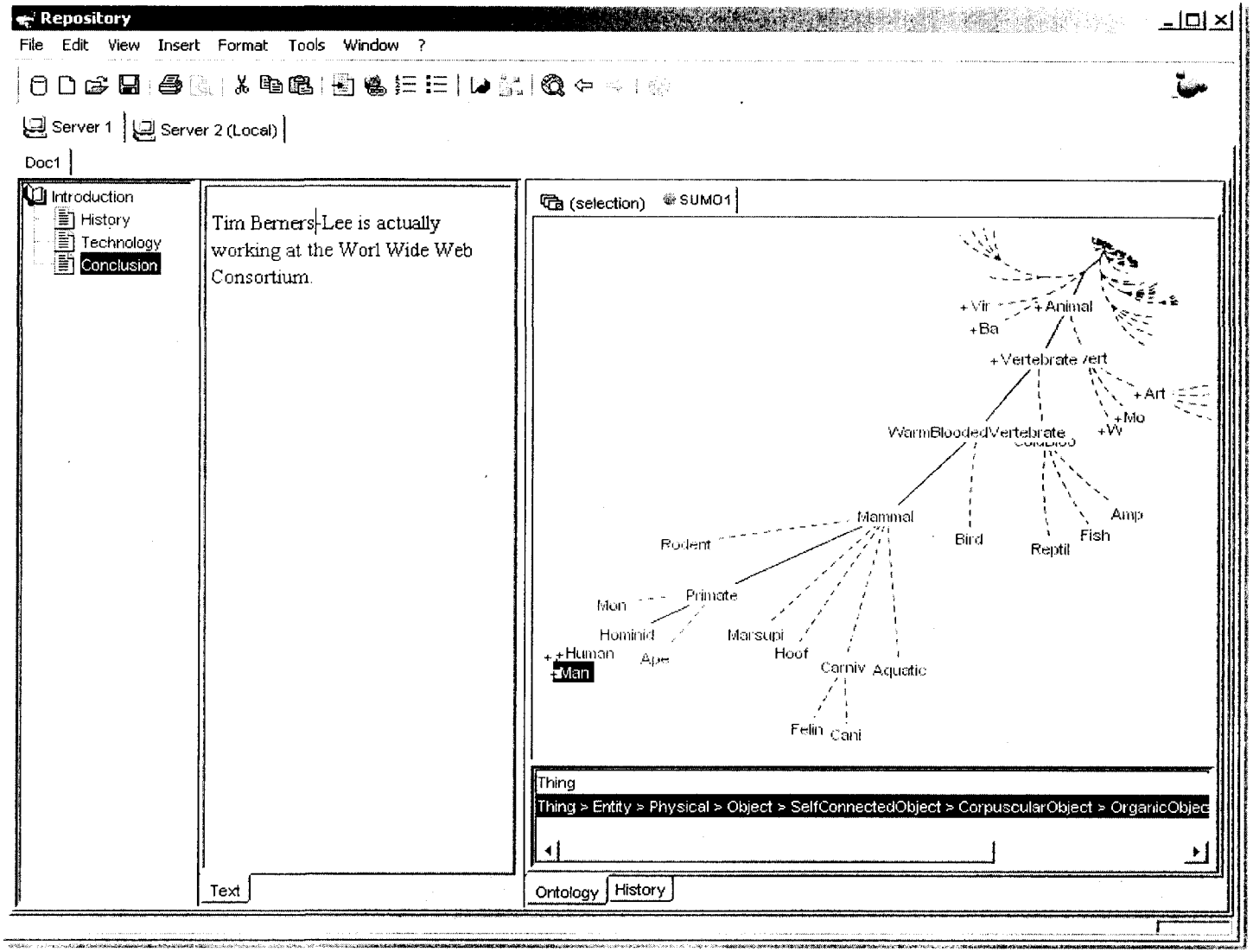


Figure 7-4: Interface du prototype logiciel.

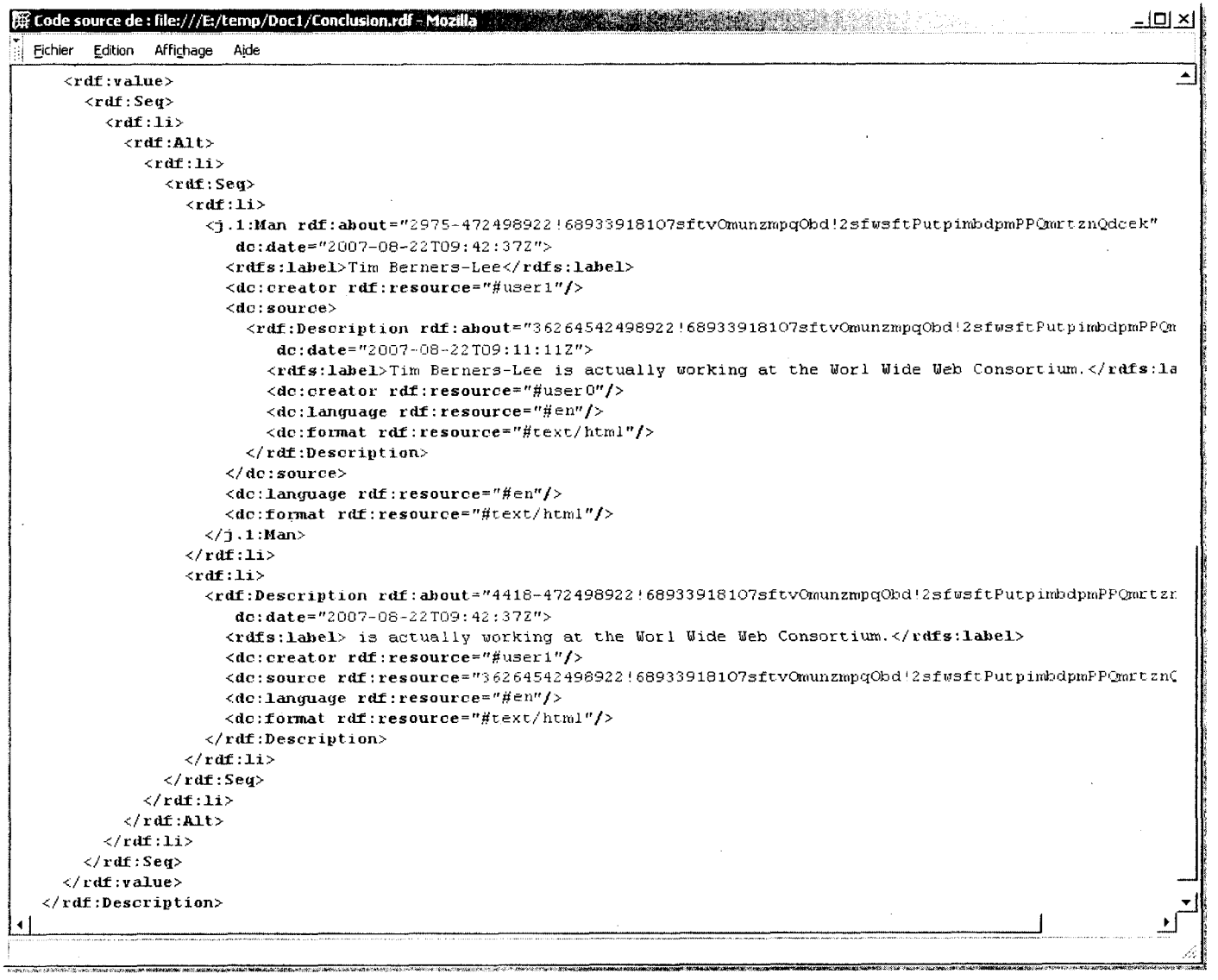


Figure 7-5: Descriptions RDF générées par le logiciel.

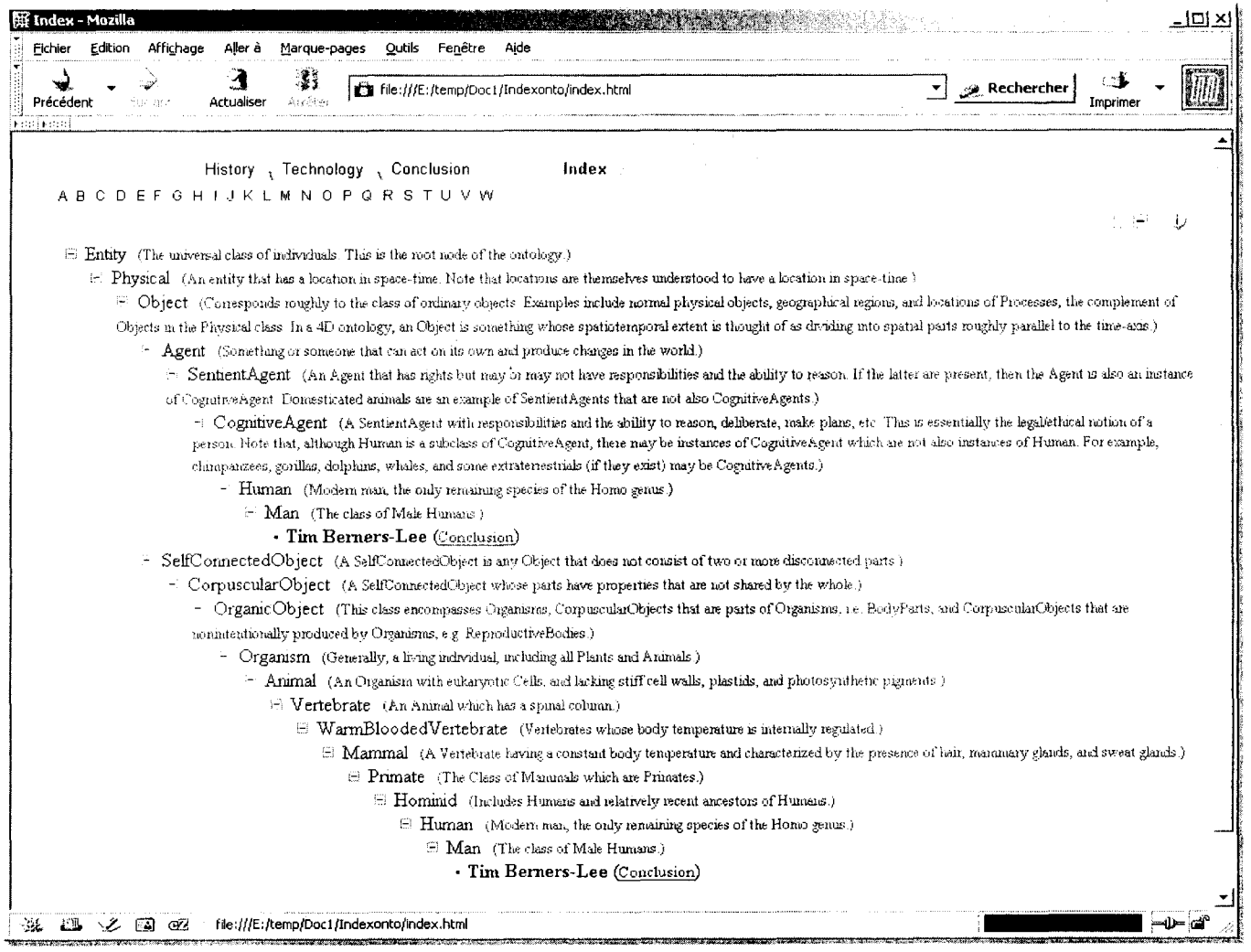


Figure 7-6: Page d'index créée à partir des annotations.

7.2 Expérimentation

L'expérimentation consiste à demander à des sujets de récupérer des contenus de cours déjà existants à l'aide du logiciel et d'y apporter des annotations pour produire un index représentatif du cours. Une partie du matériel proposé possède déjà des annotations alors que d'autres parties en sont totalement absentes. L'objectif est de mesurer le taux d'annotations réutilisées entre chaque sujet en quantifiant les emprunts par le biais d'un fichier de journalisation.

7.2.1 Sujets

Le recrutement des sujets était basé sur les conditions suivantes :

- être un étudiant aux cycles supérieurs en science ou en ingénierie (ou posséder un diplôme dans ces domaines);
- connaître les notions de base en IHO;
- être familier avec l'environnement graphique Windows.

Les sujets ont été recrutés par le biais d'un professeur enseignant dans le domaine des IHO et par des contacts personnels. Huit sujets ont acceptés de participer à l'expérience. Ce nombre était suffisant pour couvrir les 53 pages de cours à annoter. Les sujets étaient âgés entre 25 et 55 ans et un seul sujet était une femme (Tableau 7.1). Ces huit sujets présentaient tous un bagage académique et professionnel légèrement différent les uns des autres (consultant en utilisabilité, étudiant en informatique, chargé de cours, etc.). Tous étaient familiers avec l'environnement Windows.

Tableau 7.1: Caractéristiques physiques des différents sujets.

Sujet	Age	Sexe	Occupation	Scolarité	Familiarité à Windows
1	20-29	F	Étudiante	Bac. (Génie informatique)	Excellente
2	30-39	H	Chargé de cours	Bac. (Administration)	Excellente
3	50-59	H	Étudiant	M. Sc. A. (Génie logiciel)	Excellente
4	20-29	H	Étudiant	Bac. (Génie informatique)	Excellente
5	20-29	H	Étudiant	Bac. (Génie télécommunication)	Excellente
6	20-29	H	Consultant	Bac. (Communication)	Excellente
7	30-39	H	Associé de recherche	Bac. (Informatique)	Excellente
8	20-29	H	Consultant	Bac. (Génie informatique)	Excellente

7.2.2 Tâche expérimentale

Les sujets étaient invités à lire le texte d'un cours en ligne. Chaque sujet avait ensuite pour consigne d'annoter les contenus de texte de manière à ce que l'index produit par le logiciel puisse permettre à d'autres utilisateurs de repérer les contenus importants du cours. Le contenu à annoter était composé de deux sections différentes : une première section était composée de matériel de cours déjà annoté par d'autres sujets alors que la seconde section était exempte de toute annotation. Chaque sujet était ainsi libre de rajouter, retrancher ou modifier les annotations déjà présentes.

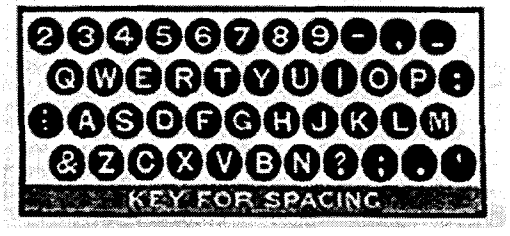
Le processus d'annotation consistait à sélectionner une région de texte et à l'assigner à une classe d'ontologie. Les sujets étaient libres de télécharger de nouvelles ontologies directement à partir du Web, d'en créer d'autres de toute pièce ou encore de reprendre les ontologies déjà utilisées par leurs prédécesseurs. Les classes figurant à l'intérieur de ces ontologies pouvaient aussi être effacées, renommées, ajoutées ou déplacées à l'intérieur de l'ontologie ou entre différentes ontologies choisies par le sujet.

Les expériences se sont déroulées sur une période totale de trois à six heures en une ou deux séances distinctes selon le choix et la disponibilité de chaque sujet.

7.2.3 Matériel

Le matériel de cours a été fourni par un professeur universitaire possédant un doctorat dans le domaine des IHO. Ce matériel consistait en 53 pages Web exposant les différences existant entre les dispositifs spécialisés d'entrée-sortie pour les interfaces utilisateurs (Figure 7-9). Le texte de chaque page était accompagné d'images que les sujets n'avaient toutefois pas à annoter. Un échantillon de ce matériel de cours a été mis en annexe (page 169).

Le premier clavier commercial a été inventé par Christopher Latham Sholes en 1873, qui présentait déjà la disposition relative de touches QWERTY. Celle-la n'a pas été choisie en fonction de la commodité de l'utilisateur face à sa tâche, mais plutôt en raison du fonctionnement du système de barres de frappe dans les machines à écrire mécaniques. Des anecdotes associent l'origine de cette configuration au besoin de changer la disposition des touches des prototypes de façon à ralentir la vitesse de frappes des opérateurs, ce qui bloquait les barres mécaniques. Il y a aussi une version des faits, par laquelle cette disposition visait à favoriser l'écriture du mot TYPEWRITER, ce qui était important lors des démonstrations de vente de machine à taper. Voir un peut plus de ces histoires sur le site Web MYTHS about QWERTY.



Le nom QWERTY désigne les lettres de la première rangée supérieure de touches alphabétiques, en partant de la gauche sur le clavier. Cette disposition est utilisé en Amérique du Nord et présente des déclinaisons régionales comme AZERTY (à droite), utilisée en France et QWERTZU, utilisée en Allemagne. Le site Web de WikiPedia présente d'autres dispositions de clavier utilisées dans le monde, comme le clavier arabe (à droite).

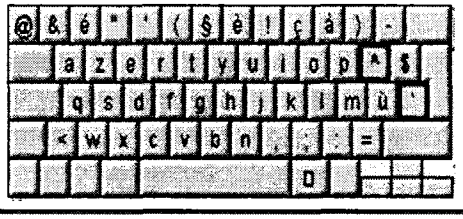


Figure 7-9 : Exemple d'une page de cours à annoter.

Le logiciel d'annotation était installé sur un ordinateur de table dont les caractéristiques techniques sont présentées au Tableau 7.2.

Tableau 7.2: Description technique de la machine de table servant aux tests.

Système d'exploitation :	Window XP
Processeur :	Intel Pentium 4
Vitesse :	2.66 GHz
Mémoire :	2 Go RAM
Résolution d'écran :	1280 par 1024 pixels
Taille d'écran :	21 pouces de diagonale

7.2.4 Plan de l'expérience

Le cours a été séparé en quatre sections distinctes (identifiés par A B C D sur la Figure 7-10) de manière à fournir un contenu nouveau à annoter pour chaque sujet de rang impair (Figure 7-10). Le sujet 1 a reçu un document de 10 pages à annoter (identifié par A). Une fois l'expérience terminée, le matériel annoté a été présenté au sujet 2 qui avait le choix de réutiliser le matériel déjà annoté par le sujet précédent ou de repartir à zéro avec un matériel encore non annoté.

Lorsque le sujet 2 a finalement terminé son travail d'annotation, son matériel annoté à été présenté à au sujet 3 qui avait le choix de réutiliser le matériel du sujet 1 ou le matériel du sujet 2. Le sujet 3 avait aussi pour tâche d'annoter un nouveau matériel (identifié par B) qui correspondait à 23 pages de texte. Ce plan d'expérience a été repris ainsi de suite pour englober les 8 sujets qui ont participés à l'expérience.

Le nombre de pages traités entre les différents sujets était inégal parce que nous avons tenu compte de la taille de chaque page pour équilibrer la charge de travail entre chaque sujet (Figure 7-10). Le nombre de pages traitées par chaque sujet variait ainsi largement (par exemple, 23 pages pour le contenu B et 10 pages pour le contenu D) mais la somme totale des mots traités par chacun d'eux était parfaitement similaire (1000 mots répartis sur environ 20 paragraphes différents).

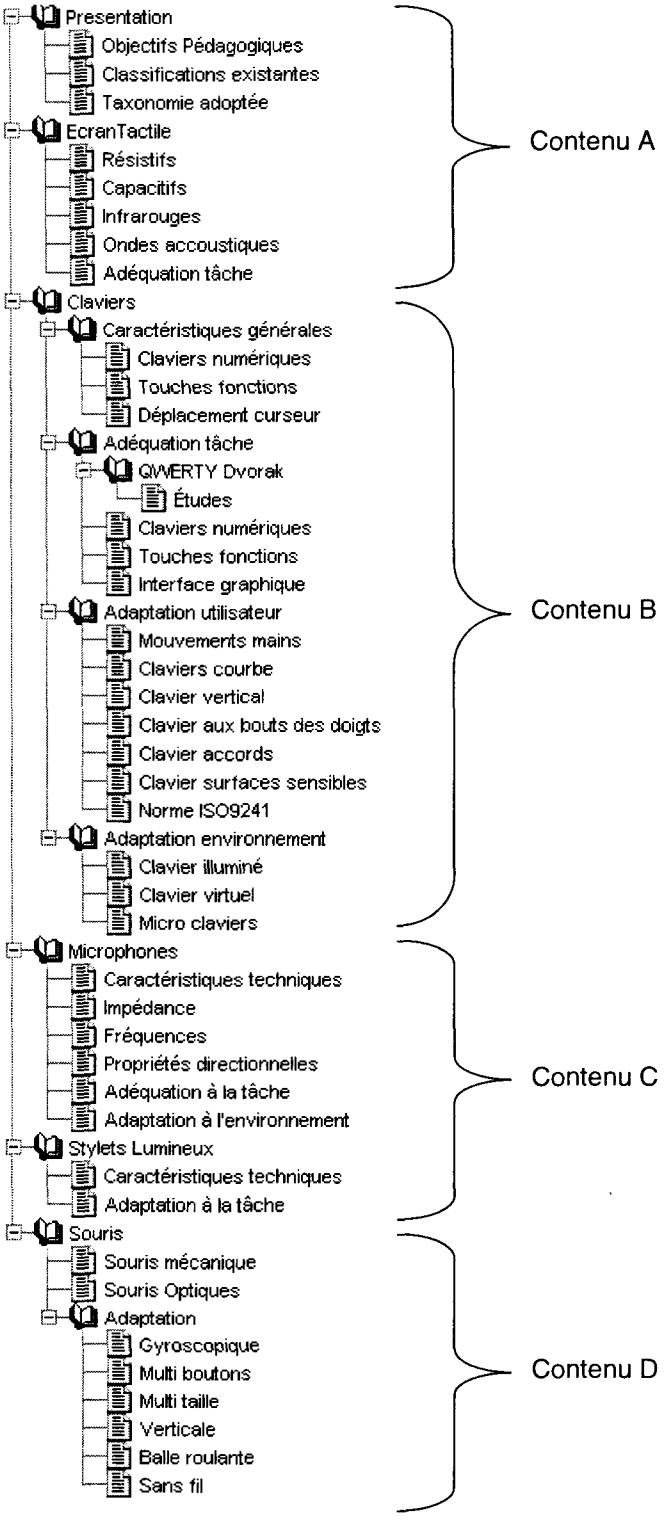


Figure 7-10 : Répartition des contenus.

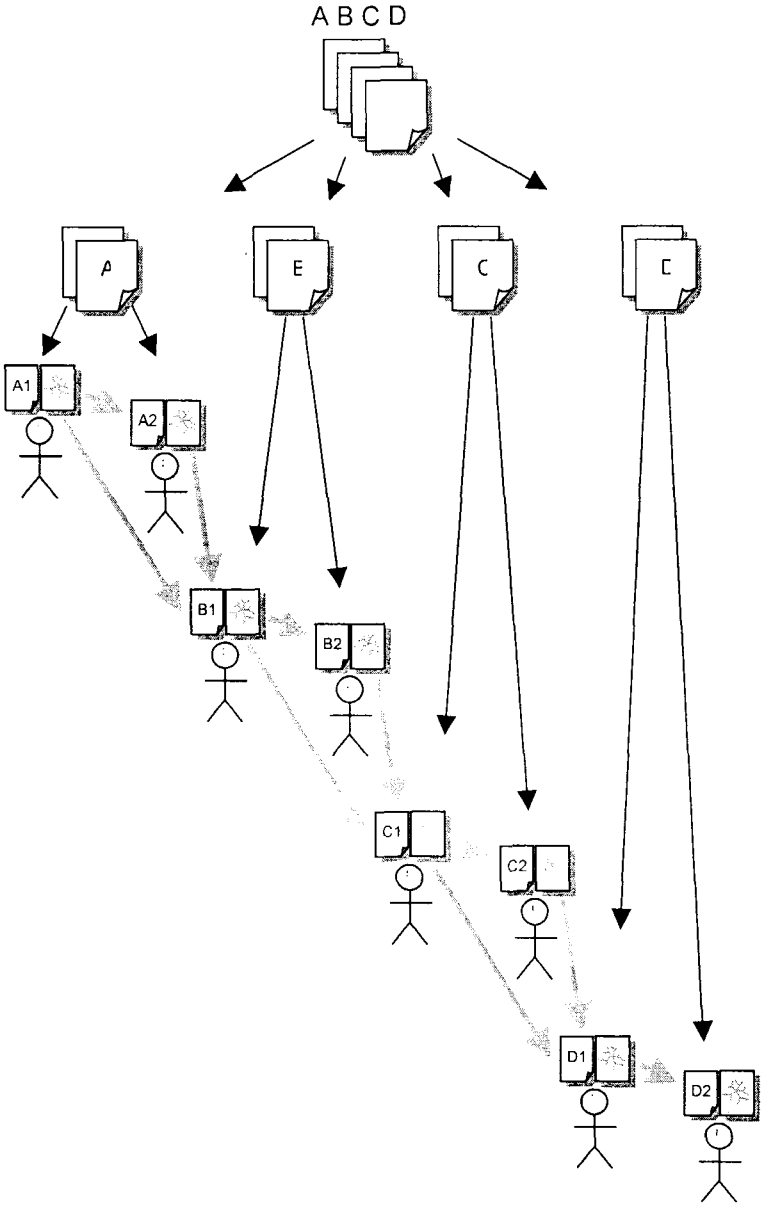


Figure 7.8: Plan d'expérimentation.

Ce cycle d'expérience a été reproduit plusieurs fois de manière à permettre à chaque sujet d'annoter le matériel de cours des sujets précédents et de créer de nouvelles annotations si nécessaire. La configuration de l'expérience a ainsi été choisie pour mettre en évidence les sources d'emprunts de chaque sujet et maximiser la diversité des sources de contenus annotés et non annotés.

Le logiciel laisse aux utilisateurs le soin de récupérer des annotations et de décider s'ils désirent conserver les classes récupérées avec celles-ci. Ils peuvent aisément copier / déplacer / fusionner des ontologies qu'ils jugent conflictuelles. Les contradictions entre les ontologies sont ainsi réglées par les utilisateurs eux-mêmes.

7.2.5 Procédure

Les sujets ont reçu chacun une courte formation (20 minutes) sur la manière d'utiliser le logiciel d'annotation. Cette formation a été donnée par l'auteur de cette thèse. Il leur a été ensuite été demandé de faire un exercice simple pour vérifier leur compréhension de l'interface. Chaque sujet devait ainsi démontrer qu'il savait construire une nouvelle page Web, récupérer du matériel de cours, réaliser des annotations, créer de toute pièce une nouvelle ontologie, récupérer des ontologies déjà existantes sur le Web (à l'aide du moteur de recherche Swoogle), modifier des ontologies et réaliser diverse opérations de maintenance. Les sujets pouvaient ensuite commencer l'expérience. Durant toute la durée de la séance d'expérience, il leur était toujours possible de poser des questions pour éclaircir certaines fonctionnalités du logiciel.

L'expérience a eu lieu dans un bureau isolé avec la présence d'un modérateur. Les sujets étaient encouragés à parler à voix haute pour permettre au modérateur de prendre des notes. Les premiers sujets ont été filmés sur vidéo mais cette procédure a rapidement été abandonnée parce que la qualité des enregistrements ne permettait pas de reproduire fidèlement les mouvements réalisés sur l'interface du logiciel. Ce problème technique était relié à la densité d'information présente à l'écran et à la difficulté de relire après coup ces informations sur un moniteur vidéo. Toutes les manipulations de texte ont été néanmoins enregistrées par le logiciel dans un fichier de journalisation (« log »). Chaque ligne du fichier de log contenait ainsi l'heure et le détail particulier des actions de l'utilisateur.

Les sujets pouvaient en tout temps demander une pause et il était toujours possible de reporter une partie de l'expérience au lendemain si la fatigue se faisait sentir. Cette

disposition a notamment été utilisée par les deniers sujets qui ont eu à réaliser l'annotation de l'ensemble des différentes pages du cours.

À la fin de l'expérience, les sujets ont été invités à remplir un questionnaire et à fournir sur papier des remarques ou des commentaires généraux sur le déroulement de l'expérience. Au moment de remettre leur questionnaire, chaque sujet a reçu une somme équivalente à 15$ de l'heure pour sa participation à l'expérience. Les modalités de l'expérience ont toutes été approuvées au préalable par le comité d'éthique de l'École Polytechnique de Montréal (le certificat de conformité est présenté en annexe à la page 163).

7.2.6 Consignes

Avant de débuter leur travail, les sujets ont chacun reçu une consigne verbale leur mentionnant l'objectif à atteindre :

> « Vous devrez concevoir une section de cours Web en récupérant du matériel déjà créé et en associant les éléments qui vous semblent importants à des classes d'ontologies dans l'arbre hyperbolique.
>
> Pour vous aider dans votre tâche, vous êtes libre de récupérer des ontologies déjà existantes sur le Web. Il vous est aussi possible de les créer de toute pièce à l'aide du prototype logiciel.
>
> Vous êtes libre d'ajouter, modifier ou supprimer les annotations et les ontologies qui accompagnent les contenus de cours. Votre seul objectif est de vous assurer que les annotations présentes dans le texte seront suffisantes pour permettre le repérage des contenus de cours les plus significatifs. Nous vous invitons à vérifier la qualité de ce repérage en générant automatiquement l'index du cours à l'aide du logiciel et en vous assurant que les concepts présents dans cet index respectent une stricte hiérarchie de classe. »

7.2.7 Questionnaire

À la fin de chaque expérience, un questionnaire était fourni aux sujets pour leur demander d'évaluer le niveau de difficulté éprouvée durant l'expérience. Le questionnaire contenait quatre questions d'ordre démographique (sexe, groupe d'âge, niveau de scolarité, familiarité à l'environnement Windows) et 20 questions concernant l'ergonomie du logiciel et le niveau de difficulté éprouvée durant l'expérience. Les questions ont été bâties selon le modèle du « *IBM Computer Usability Satisfaction Questionnaires* » [Lewis,1995]. Ce questionnaire est présenté en annexe (page 166).

L'objectif de ce questionnaire était d'évaluer l'effort des sujets et de signaler des défauts importants de l'interface qui auraient demandés une correction immédiate de l'application.

CHAPITRE 8 : ANALYSE DES RÉSULTATS

Le fichier de journalisation et les ontologies produits par les différents sujets ont été analysés pour mettre en évidence la durée de l'expérience, le taux de réutilisation des annotations, le nombre de classes échangées et le taux de réutilisations de ces classes.

8.1 Durée

L'expérience a eu lieu entre les mois d'avril et d'octobre 2007 selon la disponibilité des différents sujets. La durée de chaque séance variait entre deux et six heures. Le tableau suivant illustre le temps consacré à rechercher des ontologies avec le moteur de recherche Swoogle (http://swoogle.umbc.edu/) ainsi que la durée de chaque séance.

Tableau 8.1: Durée des séances.

	Utilisation de Swoogle	Durée de la séance
Sujet 8	0	4 heures
Sujet 7	0	4 heures
Sujet 6	0	6 heures
Sujet 5	2 minutes	5 heures
Sujet 4	3 minutes	2,5 heures
Sujet 3	3 minutes	2 heures
Sujet 2	5 minutes	2 heures
Sujet 1	5 minutes	2,5 heures

Les quatre premiers sujets ont travaillé pendant deux heures. La durée des séances a toutefois commencé à doubler à partir du cinquième sujet pour passer à une durée moyenne de cinq heures.

Cette augmentation de temps est probablement attribuable à la configuration même de l'expérience (Figure 7.11) qui imposait au sujet 5 de récupérer les contenus A et B avant d'annoter à son tour un nouveau matériel. Il est néanmoins curieux de remarquer que le sujet 7 affiche une réduction du temps par rapport à son prédécesseur malgré une configuration d'expérience semblable à celle du sujet 5. Ces écarts de temps pourraient être donc simplement attribuables à des différences particulières entre les sujets.

8.2 Réutilisation des annotations

Le fichier de journalisation généré par le logiciel d'expérimentation a permis de réaliser un décompte manuel des annotations produites. Le Tableau 8.2 présente le décompte des annotations en fonction des différents utilisateurs.

Tableau 8.2: Données brutes sur l'utilisation des annotations.

	Réutilisées	Retirées	Créées	(Total)
Sujet 8	702	51	227	878
Sujet 7	604	2	163	765
Sujet 6	535	24	112	623
Sujet 5	410	22	147	535
Sujet 4	293	10	145	428
Sujet 3	121	17	189	293
Sujet 2	48	10	100	138
Sujet 1			48	48

Ce tableau indique que le sujet 1 a créé 48 annotations à l'intérieur des différentes pages Web qui lui étaient attribuées. Ces 48 annotations ont par la suite toutes étés réutilisées par le sujet 2. Celui-ci a toutefois décidé d'en retirer 10. Il a néanmoins créé 100 nouvelles annotations qu'il a rajoutées au document. Le nombre d'annotations présentes dans les 10 pages Web du contenu A a ainsi atteint le chiffre total de 138 annotations.

Le sujet 3 a lui aussi récupéré les annotations du sujet 2 mais en exerçant toutefois une sélection plus particulière. Il n'a récupéré que 121 annotations sur un total possible de 138. Par la suite, il a décidé d'en retirer 17 autres supplémentaires. La configuration du test prévoyait aussi que le sujet 3 réalise l'annotation d'un nouveau matériel. Il a ainsi créé 189 annotations réparties à la fois dans son propre contenu et dans les contenus récupérés du sujet 2.

Le Tableau 8.2 indique aussi que le nombre total d'annotations créées suit une progression constante malgré la configuration particulière de l'expérience qui imposait seulement du nouveau matériel au sujet de rang impair (1, 3, 5, 7). Les sujets de rang pair auraient très bien pu se contenter de simplement réutiliser le matériel produit sans nécessairement ajouter d'annotations supplémentaires. De manière très surprenante, tous

les sujets ont travaillé de manière égale pour produire un nombre relativement important d'annotations (Figure 8-1).

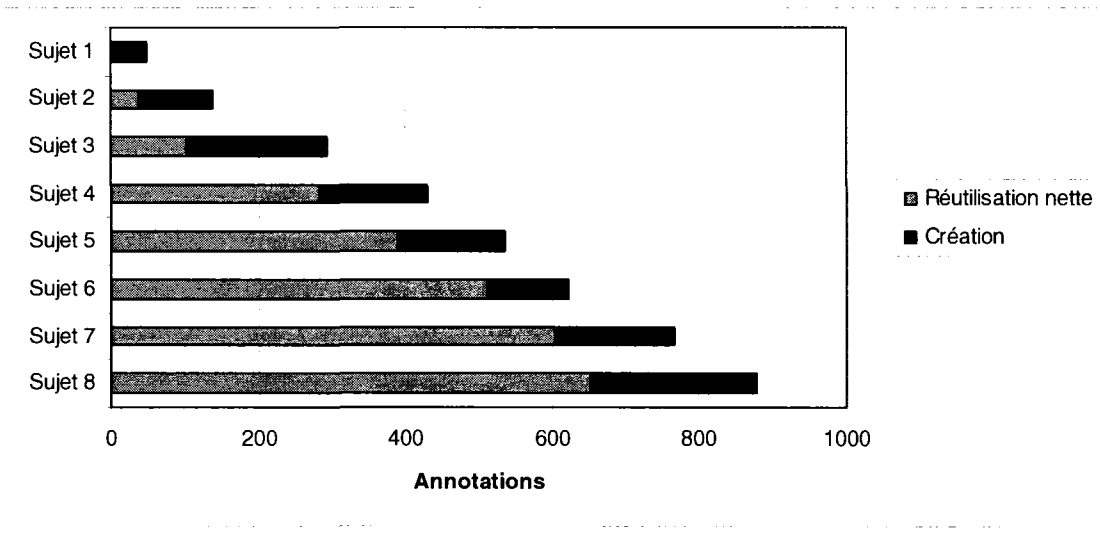


Figure 8-1: Réutilisation des annotations.

La Figure 8-1 montre que la progression des emprunts est constante. Cette progression est aussi bien due à la réutilisation des contenus des sujets précédents qu'au travail de chacun à produire de nouvelles annotations.

En utilisant les données du Tableau 8.2, il est possible de calculer le nombre de «réutilisations nettes » réalisées par chaque sujet en soustrayant les annotations retirées du nombre des annotations récupérées. Le Tableau 8.3 reprend ce calcul pour trouver le taux de réutilisation réel des annotations entre sujets.

Tableau 8.3: Taux de réutilisation des annotations.

	Réutilisation nette	Création	Total d'annotations	Réutilisation
Sujet 8	651	227	878	85 %
Sujet 7	602	163	765	97 %
Sujet 6	511	112	623	96 %
Sujet 5	388	147	535	91 %
Sujet 4	283	145	428	97 %
Sujet 3	104	189	293	75 %
Sujet 2	38	100	138	79 %
Sujet 1	0	48	48	
				88 %

Ce tableau indique que, par exemple, le sujet 8 n'a utilisé que 651 annotations sur un total de 765 annotations rendues disponibles par le sujet 7. Le taux de réutilisation est donc de (651/765) 85%. En faisant la moyenne de tous ces taux, nous obtenons une moyenne de réutilisation de 88%. Autrement dit, le taux de rejet des annotations entre chaque sujet ne dépasse pas 12%.

Ces chiffres montrent aussi qu'il existe un gain constant dans le nombre d'annotations réutilisées par chaque sujet. Ce taux de réutilisation de 88% permet ainsi la réalisation d'un effet de levier important entre les sujets en permettant à chacun de bénéficier d'un nombre de plus en plus importants de contenus annotés à chaque étape du cycle d'emprunt.

Il est intéressant de noter que le taux de réutilisation des annotations reste indépendant de la position de chaque utilisateur. Cette constatation s'explique directement par l'adoption d'un plan d'expérimentation similaire pour chaque groupe de sujets. Elle confirme aussi que ce phénomène n'est pas isolé à un utilisateur particulier et que le comportement de réutilisation est partagé par l'ensemble des sujets testés.

8.3 Ontologies produites

Le sujet 1 a commencé son travail en réalisant un plan papier de l'ontologie qu'il désirait produire pour décrire le texte de départ qui lui avait été fourni (Figure 8-2). Les autres sujets ont plutôt préféré travailler directement à l'ordinateur en profitant, le plus possible, des ontologies déjà créées par leurs prédécesseurs.

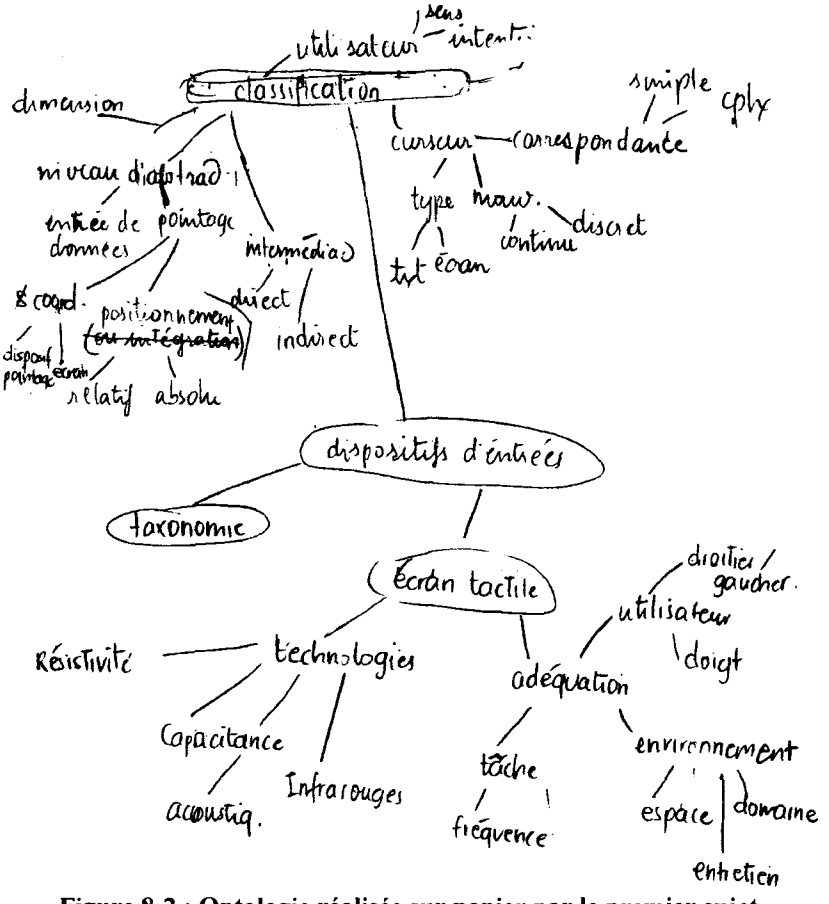


Figure 8-2 : Ontologie réalisée sur papier par le premier sujet.

L'ontologie du sujet 1 a été successivement réutilisée par les sujets 2 à 8 en apportant chacun de légères corrections à la structure et à la nomenclature des classes. La Figure 8-3 présente l'une à côté de l'autre l'ontologie réalisée par le sujet 1 et l'ontologie finale produite à la fin du cycle de réutilisation entre les huit sujets (une liste détaillée des classes produites est aussi disponible en annexe).

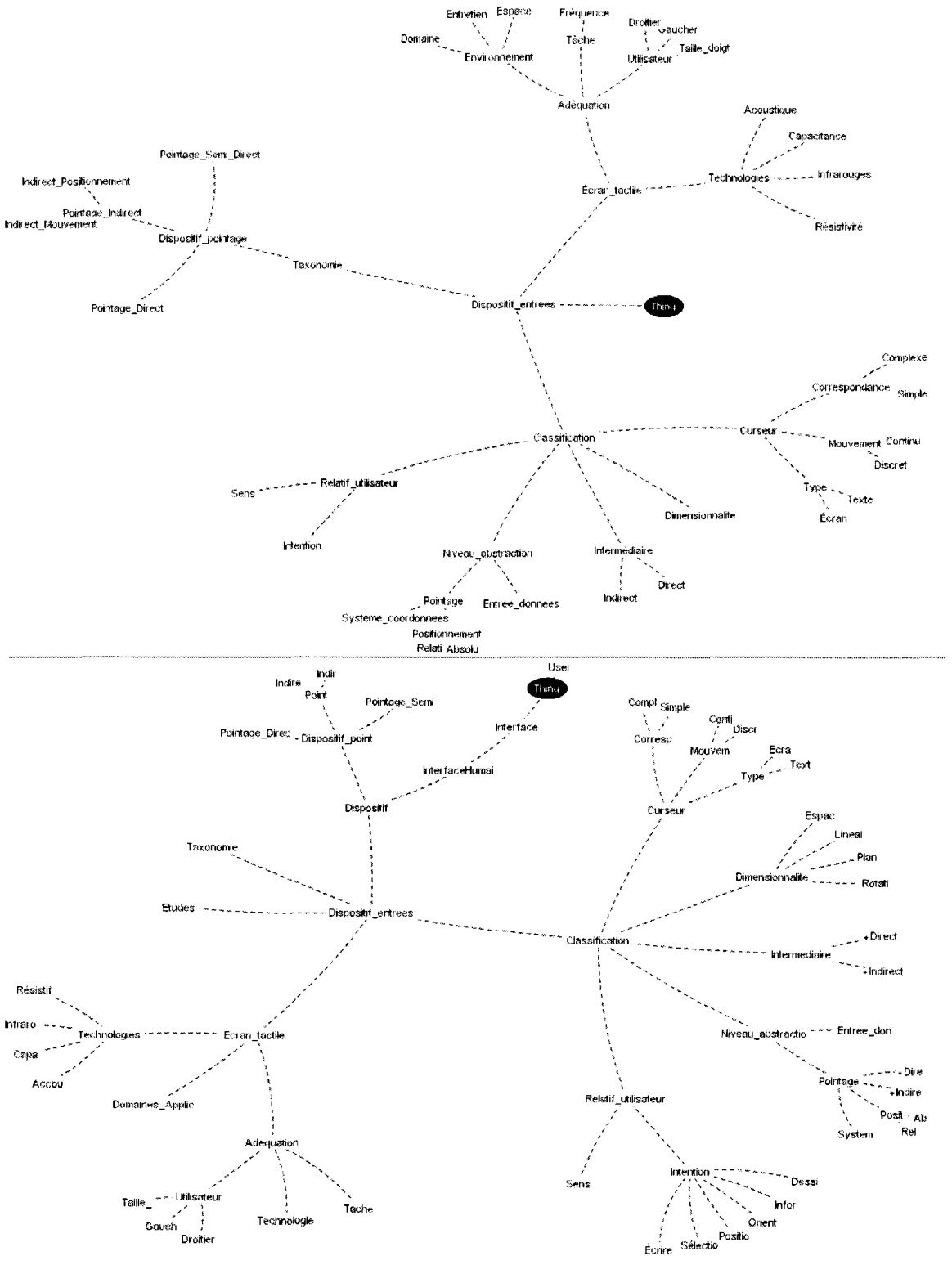


Figure 8-3 : Illustration de l'ontologie produite par le sujet 1 (en haut) et ontologie résultante après le passage du sujet 8 (en bas).

Le nombre total de classes entre ces deux ontologies est relativement similaire (52 classes pour la première et 64 pour la secondes). La nomenclature des classes n'a pas, non plus, beaucoup changée. La structure des ontologies présente toutefois des disparités relativement importantes. Par exemple, la classe «Pointage_Semi_Direct» qui figure dans la première ontologie est située au niveau de l'arborescence suivante :

Thing > Dispositif_entrees > Taxonomie > Dispositif_pointage > Poitage_Semi_Direct

Alors que dans la seconde ontologie, cette classe se situe au niveau de :

Thing > Interface > InterfaceHumain > Dispositif > Dispositif_pointage > Poitage_Semi_Direct

Il est ainsi intéressant de noter que ce changement de structure n'implique pas en soi la disparition des classes intermédiaires. Celles-ci ont été le plus souvent réutilisées ailleurs dans l'ontologie pour donner naissance à d'autres catégories à part entière. Par exemple, la classe « Taxonomie » qui figure au troisième niveau de la première ontologie se retrouve désormais au cinquième niveau de la seconde ontologie :

Thing > Dispositif_entrees > Taxonomie

Thing > Interface > InterfaceHumain > Dispositif > Dispositif_entrees > Taxonomie

Ces ontologies ont ainsi été modifiées par des utilisateurs qui ne se sont jamais rencontrés et qui n'ont jamais convenu d'une nomenclature commune. Il n'est donc pas surprenant de constater qu'il existe une certaine hétérogénéité dans le format des noms. Par exemple, les noms composés sont parfois construits en juxtaposant directement deux termes ensemble (« InterfaceHumain ») alors qu'ils sont parfois construits en utilisant un élément de séparation (« Poitage_Semi_Direct »). La case des lettres du second terme varie aussi en adoptant parfois une lettre minuscule (« Dispositif_pointage ») et parfois une lettre majuscule (« Domaine_Application »).

Les figures suivantes présentent les cinq ontologies créées par les différents sujets. Ces ontologies présentent toutes ces mêmes variations de nomenclature. La liste complète des classes utilisées par ces ontologies a été mise en annexe à la page 174.

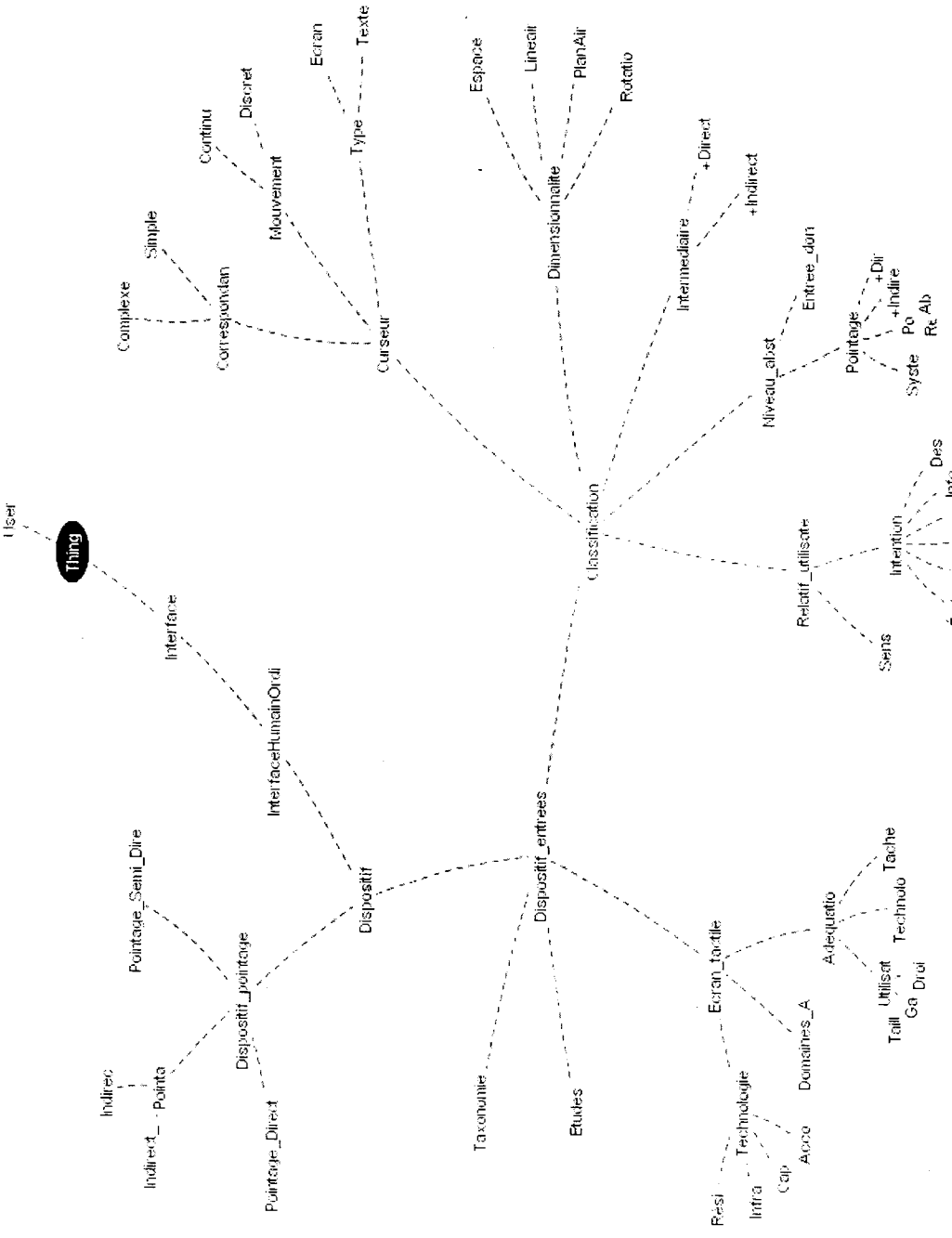


Figure 8-4: Dispositif.owl

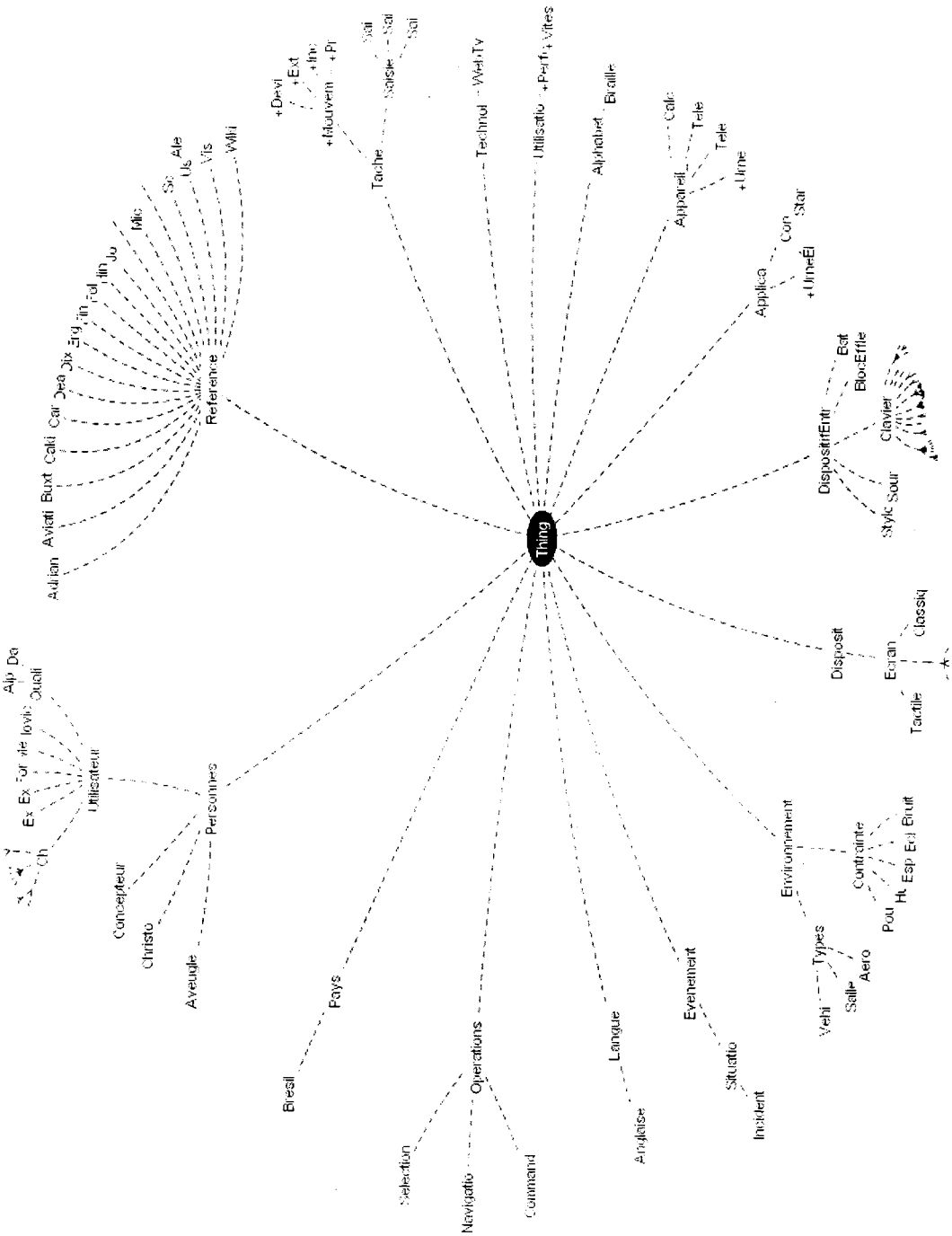


Figure 8-5: Clavier.owl

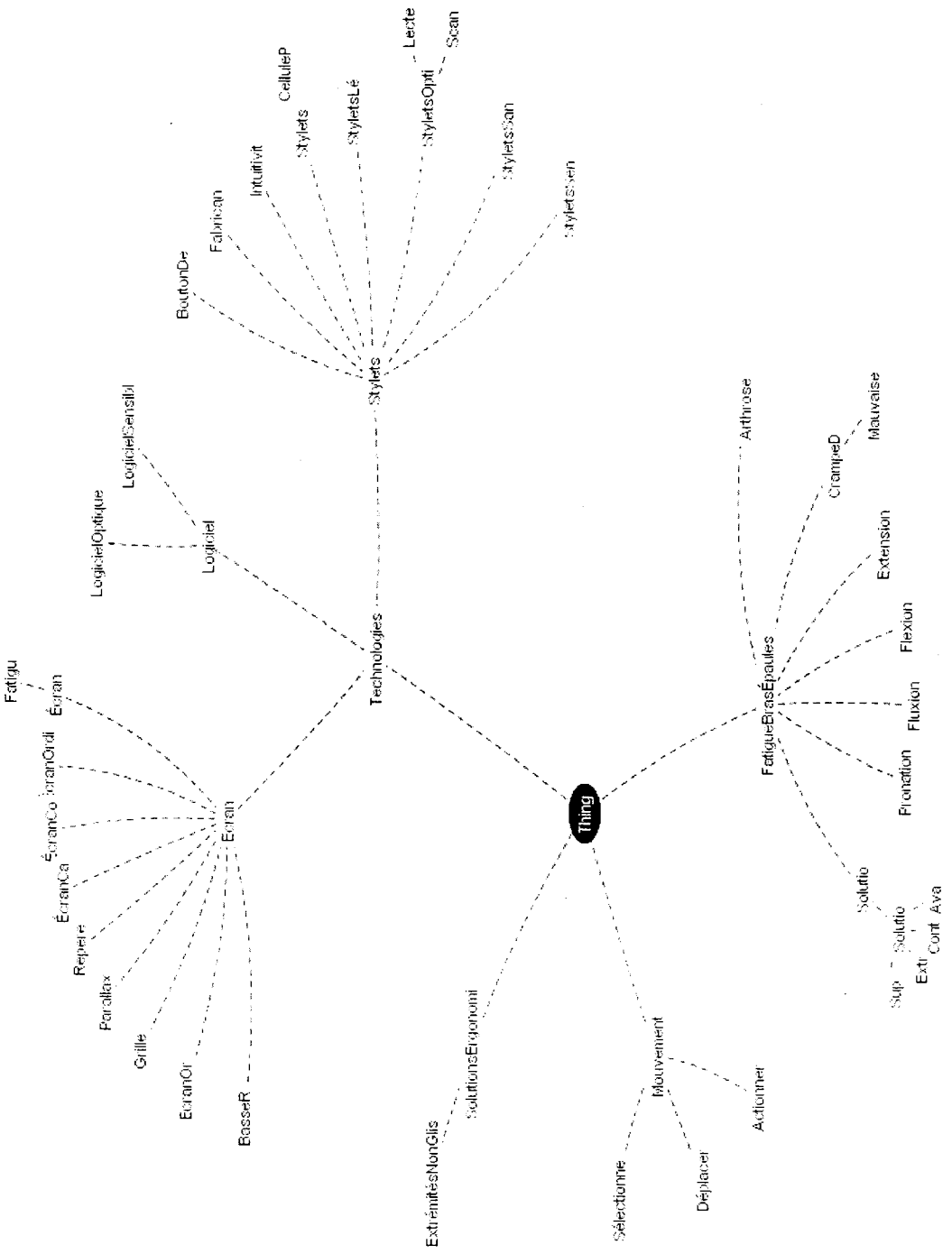


Figure 8-6: Stylet_lumineux.owl

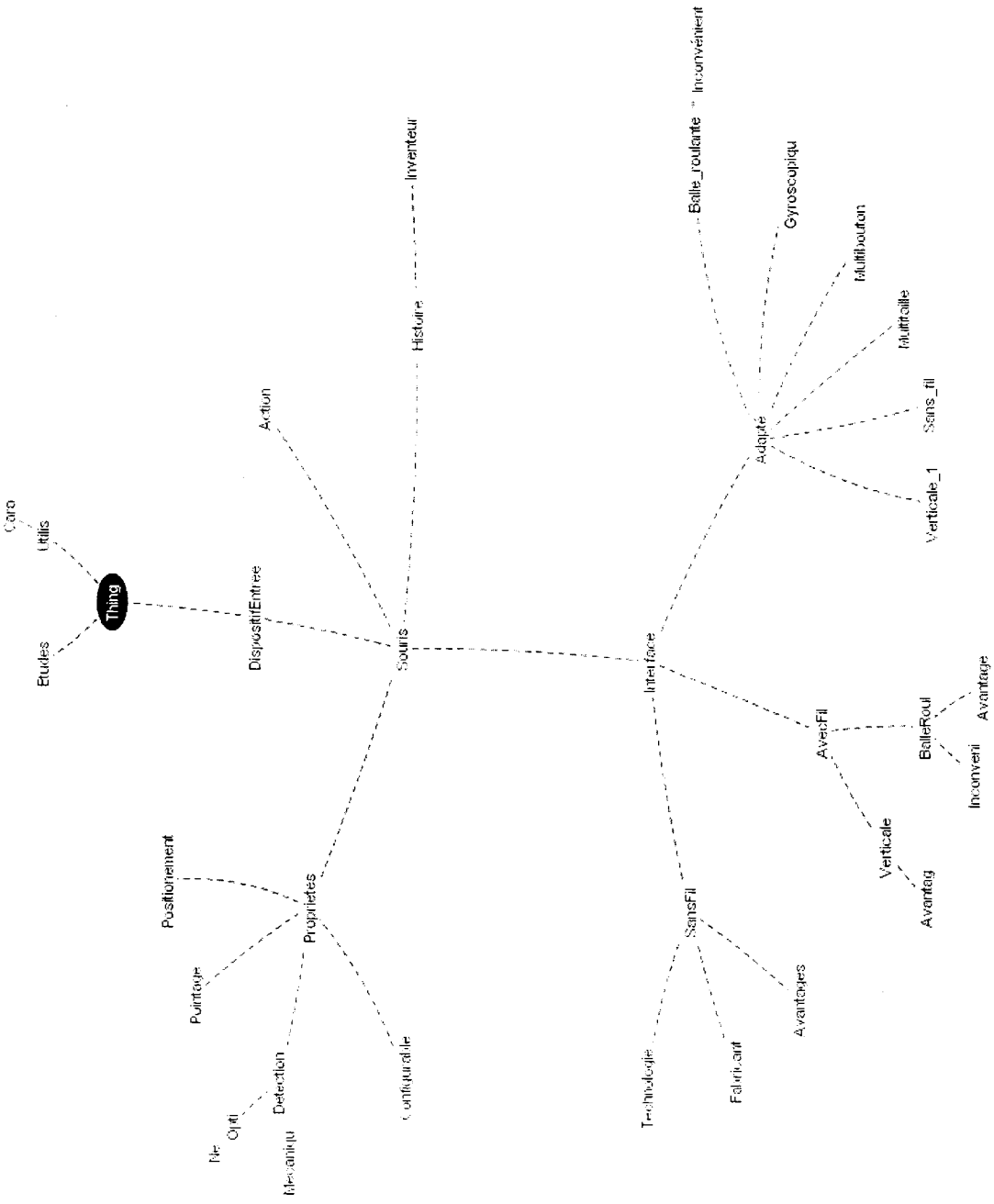


Figure 8-7: Souris.owl

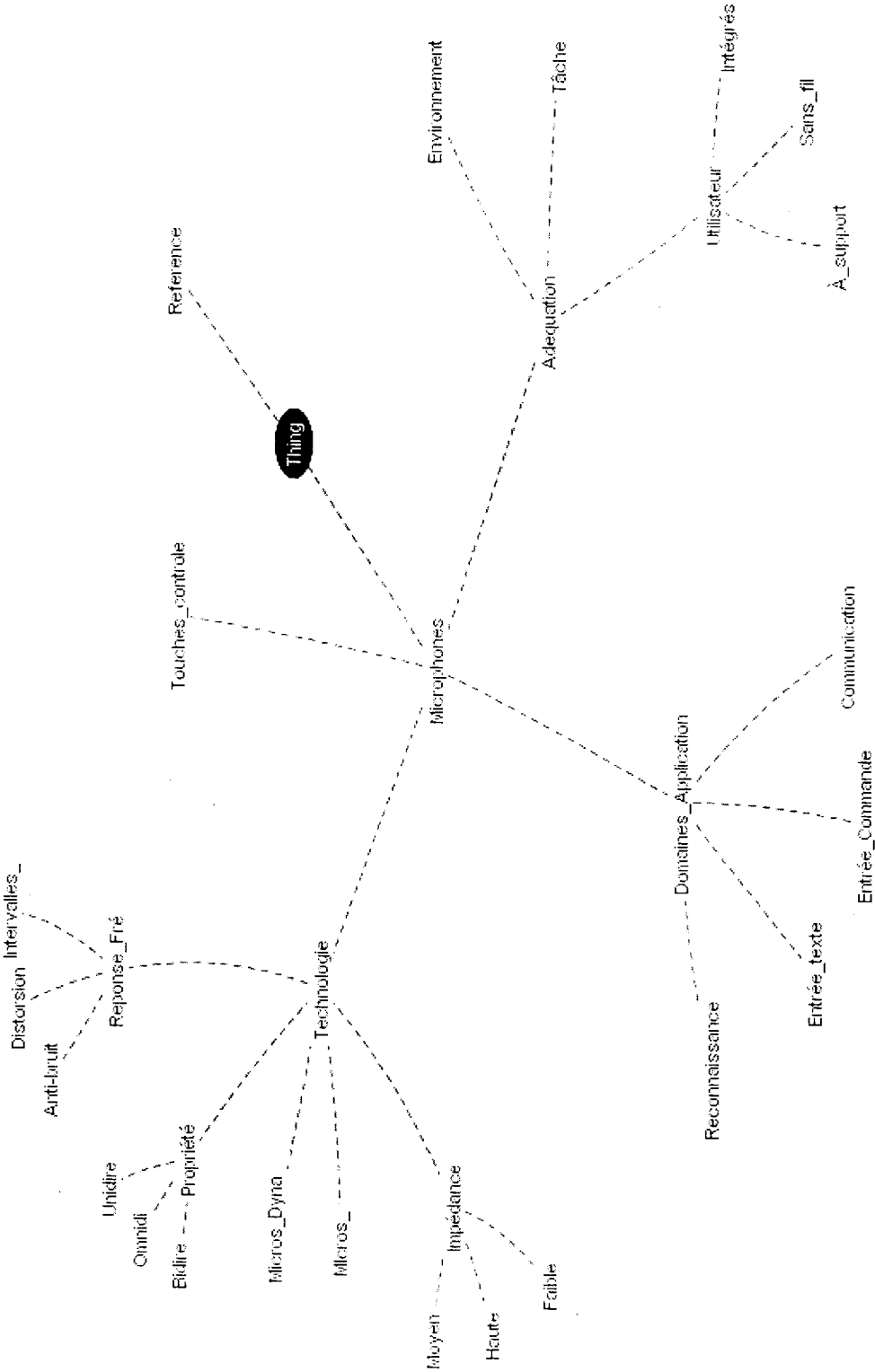


Figure 8-8: Microphone.owl

Les sujets ont eu recours à plusieurs ontologies différentes. Le nom de ces différentes ontologies (clavier, dispositif, microphone, souris, stylet lumineux) correspond ainsi directement au contexte de leur création. Lorsque les sujets ne réussissaient pas à trouver de classes intéressantes à l'intérieur des ontologies existantes, ceux-ci préféraient toujours démarrer à partir d'une nouvelle ontologie vide plutôt que d'être obligés de créer directement une nouvelle classe au niveau de la racine de l'une des ontologies déjà présentes. Ce comportement est aussi vrai pour les classes réalisées par les autres que les classes réalisées par soi-même. Par exemple, le contenu « A » a été assigné à l'ontologie « dispositif .owl ». Le contenu « B » a favorisé la création de l'ontologie « clavier.owl ». Le contenu « C » a donné naissance à l'ontologie « microphone.owl ». Toutefois, puisque l'auteur de cette ontologie ne trouvait pas de classes intéressantes à l'intérieur de celle-ci pour décrire un stylet lumineux, il a alors décidé de créer une nouvelle ontologie nommée « stylet_lumineux.owl » pour supporter ces nouvelles descriptions. De manière similiare, le contenu « D » a donné naissance à une ontologie distincte appelée « souris.owl ».

8.4 Réutilisation des ontologies

Chacune des cinq ontologies a été analysée pour identifier le nombre de classes qu'elles comportaient. L'analyse a été réalisée à l'aide de l'éditeur d'ontologies Protégé (version 4) produit par l'université Stanford. Ce logiciel permet notamment de générer des statistiques automatiques sur la composition des ontologies. Nos ontologies possédaient des classes temporaires (identifiées par la propriété « deprecated ») que Protégé ne pouvait pas distinguer des classes ordinaires. Ces classes dépréciées venaient fausser le décompte du logiciel. Nous avons donc utilisé les données premières fournies par Protégé pour ensuite soustraire à ce nombre les classes dépréciées identifiées à l'aide d'un décompte manuel. Le tableau suivant présente le résultat final de ce décompte.

Tableau 8.4: Répartition des classes pour chaque ontologie créée.

	Dispositif	Clavier	Microphone	Stylet	Souris	(Total)
Sujet 8	64	168	32	58	37	359
Sujet 7	57	164	32	58	29	340
Sujet 6	64	162	29	55		310
Sujet 5	65	133	29	55		282
Sujet 4	53	133				186
Sujet 3	51	121				172
Sujet 2	51					51
Sujet 1	52					52

La disposition en escalier des données correspond aux moments de créations différents des ontologies. Il est intéressant de noter que le nombre de classes ne varie relativement pas pour les sujets de rang pair qui avaient immédiatement récupéré les ontologies créées par leur prédécesseur. Par exemple, l'ontologie « Dispositif » contient 52 classes différentes au moment de sa création. Après avoir récupéré cette ontologie, le sujet 2 n'y a apporté qu'une seule modification (un retrait). Les autres sujets se sont aussi comportés de la même manière en apportant très peu de modifications aux ontologies déjà existantes. Ainsi, à la fin du cycle d'emprunt, les cinq ontologies comportaient un nombre total de 359 classes différentes.

Le tableau suivant résume les manipulations réalisées par les différents sujets. Il indique le décompte des classes qui ont été réutilisées, créées, retirées, déplacées, restructurées (qui indique le déplacement d'une classe à l'intérieur d'un autre arbre) et renommées. Ces données proviennent du fichier de log généré automatiquement par le logiciel.

Tableau 8.5: Données brutes sur l'utilisation des classes d'ontologie.

	Réutilisées	Créées	Retirées	Déplacées	Restructurées	Renommées	(Nb. d'évènement)
Sujet 8	340	19	0	0	58	34	451
Sujet 7	310	37	7	0	2	41	397
Sujet 6	282	29	1	0	18	58	388
Sujet 5	186	96	0	5	20	151	458
Sujet 4	172	14	0	1	0	21	208
Sujet 3	51	121	0	0	68	135	375
Sujet 2	52	0	1	0	48	7	108
Sujet 1		52					52

Comme nous l'avons fait précédemment, nous avons calculé la «réutilisation nette » des classes en soustrayant les classes retirées du nombre de classes réutilisées. Le Tableau 8.6 reprend ce calcul pour évaluer le taux de réutilisation des classes entre sujets.

Tableau 8.6: Analyse du taux de réutilisation des classes d'ontologie entre sujet.

	Réutilisation nette	Création	Total de classes	Réutilisation
Sujet 8	340	19	359	100 %
Sujet 7	303	37	340	98 %
Sujet 6	281	29	310	100 %
Sujet 5	186	96	282	100 %
Sujet 4	172	14	186	100 %
Sujet 3	51	121	172	100 %
Sujet 2	51	0	51	98 %
Sujet 1	0	52	52	
				99 %

Le taux de réutilisation des classes entre sujets est réalisé en prenant le nombre de réutilisations nettes et en le divisant par le total des classes disponibles au niveau précédent. Par exemple, le sujet 8 a utilisé 340 classes sur un total de 340 rendues disponibles par le sujet précédent. Le taux de réutilisation est donc de 100%. En faisant la moyenne de tous ces taux, nous obtenons une moyenne de réutilisation de 99%. Autrement dit, le taux de rejet des classes entre sujets est presque nul.

La Figure 8-9 illustre le niveau de réutilisation de classes pour chaque cycle d'emprunt entre sujets. Il apparaît clairement que les classes produites sont systématiquement reprises par les sujets suivants en y rajoutant constamment de nouvelles classes supplémentaires. Les sujets qui avaient été exposés à du nouveau matériel (sujet 1, 3, 5, 7) ont toutefois produits beaucoup plus de classes que les sujets exposés aux classes produites par leurs prédécesseurs (sujet 2, 4, 6, 8).

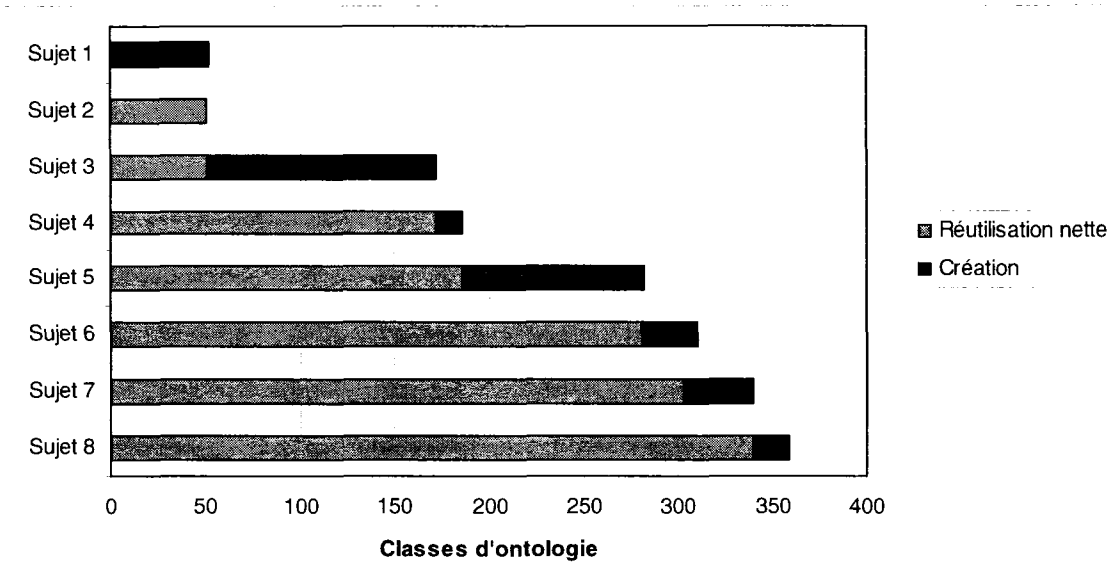


Figure 8-9: Réutilisation des classes d'ontologies.

Le travail des sujets de rang impair n'est pas pour autant négligeable. La Figure 8-10 illustre le détail des manipulations réalisées par chaque sujet. Cette figure a été réalisée à partir des données brutes du Tableau 8.5 sur l'utilisation des différentes classes d'ontologies.

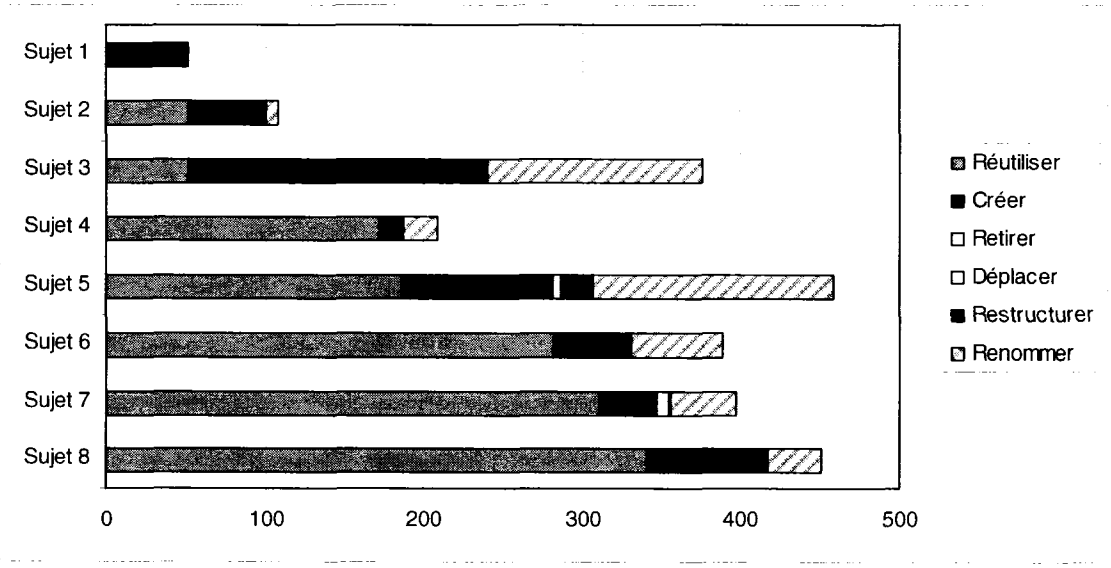


Figure 8-10: Répartition des différentes actions des sujets.

Le travail des sujets de rang pair se décrit comme suit :

- Le sujet 2 n'a pas créé de nouvelles classes. Il a toutefois réalisé une restructuration importante des classes récupérées de son prédécesseur.
- Le sujet 4 a créé quelques classes pour compléter celles déjà produites par le sujet 3. Il a aussi contribué au travail de ses prédécesseurs en renommant certaines classes déjà présentes.
- Le sujet 6 a créé de nouvelles classes pour compléter celles déjà produites par le sujet 5. Il a aussi contribué en restructurant certaines de ces classes. Sa contribution la plus grande a été de renommer un certain nombre de classes.
- Le sujet 8 a contribué à la forme finale des ontologies en restructurant un nombre important de classes. Il aussi contribué en renommant une bonne partie d'entre-elles.

L'implication de chaque sujet variait ainsi à la fois selon le type d'action entreprise et par le nombre de classes touchées par ces différentes actions. Il ne semble pas exister de patron (*pattern*) d'action particulier mais il est néanmoins possible de déceler deux motifs récurrents :

- L'opération qui consiste à renommer les classes est récurrente pour l'ensemble des sujets, indépendamment de la position de chacun d'eux dans la boucle d'échange.
- Le déplacement ou la restructuration de l'arborescence est une opération récurrente pour les sujets de rang pair qui étaient immédiatement exposés aux ontologies nouvellement produites par leur prédécesseur.

Il est tout de même important de souligner que, malgré l'ampleur des modifications apportées par chacun aux contenus de leurs pairs, les classes nouvellement créées par les sujets ont été maintenues à travers l'ensemble des échanges (avec un taux de réutilisation de 99%).

8.5 Annogramme

Nous avons construit un logiciel d'analyse pour nous permettre de mieux comprendre les cycles d'échange d'annotations entre utilisateurs en mettant en évidence le rôle de chaque sujet ainsi que les modifications apportées par chacun d'eux aux ontologies correspondantes.

Ce logiciel d'analyse a été réalisé en Java et fonctionne de manière autonome sans recourir à une connexion directe au logiciel d'annotation. Il utilise toutefois le fichier de journalisation en entrée pour produire, en sortie, un histogramme des emprunts (Figure 8-11).

Le fichier de journalisation présente les événements associés aux annotations dans un ordre chronologique. Le traitement du logiciel d'analyse consiste donc à isoler une annotation particulière et à repérer les lignes suivantes indiquant la présence d'un nouvel événement associé à cette même annotation. Pour chaque nouvel événement trouvé, le logiciel relève des informations sur la nature et l'auteur de l'évènement. Le logiciel reporte ensuite ces informations sur un diagramme pour illustrer la position de chaque événement en fonction de la position de chaque utilisateur dans le cycle d'emprunt des annotations. Puisque ce diagramme représente le cycle de vie des annotations, nous avons décidé de baptiser celui-ci « annogramme ».

Un annogramme illustre les transformations apportées à une annotation par les différents utilisateurs en précisant le moment de :

- Création: une annotation est créée en associant une section de texte à une classe d'ontologie particulière;
- Réutilisation : lorsqu'une annotation est reprise par un tiers;
- Retrait : lorsque la classe associée à l'annotation est volontairement retirée ou que l'annotation est elle-même détruite;

- Non réutilisation : lorsque la classe associée à l'annotation n'est plus reprise par personne.
- Changement de nom : de la classe associée à l'annotation;
- Restructuration : lorsque la classe associée à l'annotation est déplacée à l'intérieur de l'ontologie;
- Consolidation : lorsque la classe associée à l'annotation est déplacée vers une autre ontologie.

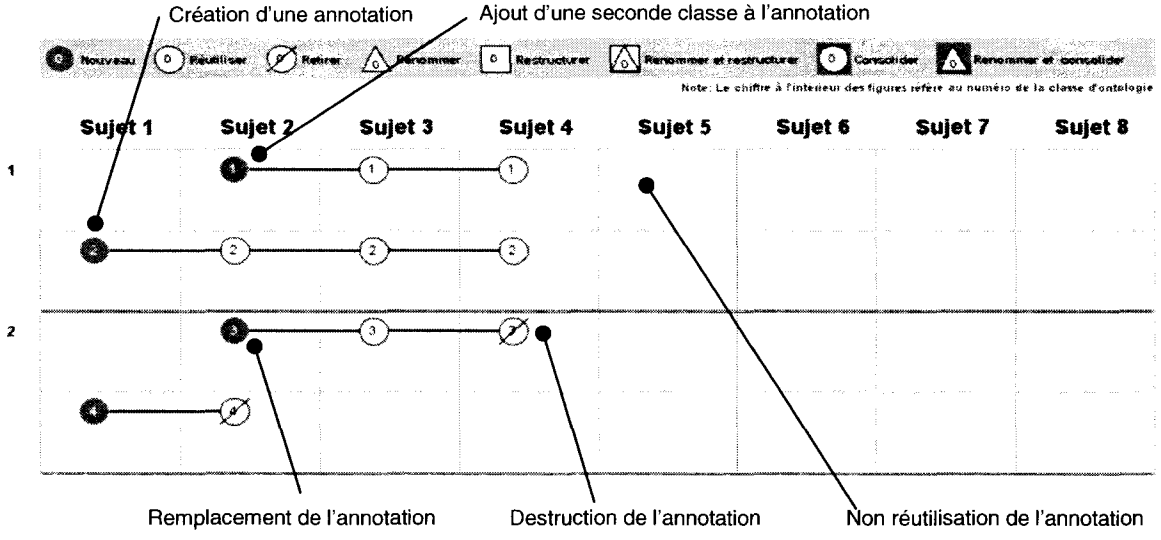


Figure 8-11: Exemple d'annogramme.

L'exemple de la Figure 8-11 indique que l'annotation 1 a été associée simultanément à deux classes d'ontologies différentes. Cette annotation a été crée par le sujet 1 en l'associant à la classe d'ontologie « dispositif.owl#Taille_doigt » (le nombre 2 qui figure à l'intérieur du cercle correspond à une classe dont l'adresse est donnée en annexe à la page 184). Le sujet 2 a réutilisé cette même annotation en y ajoutant toutefois la classe « dispositif.owl#Tache » (identifiée par le chiffre 1 à l'intérieur du cercle). Les sujets 3 et 4 ont aussi repris cette annotation sans toutefois y apporter de changement particulier. Cette annotation a toutefois été délaissée par le sujet 5 qui ne l'a tout simplement pas utilisée.

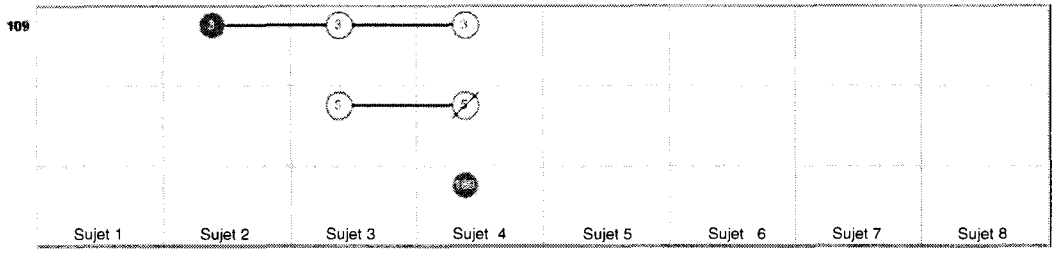
La deuxième ligne de la Figure 8-11 indique que l'annotation 2 a été associée successivement à deux classes d'ontologies différentes. L'annotation a d'abord été créée par le sujet 1 en l'associant à la classe « dispositif.owl#Domaine ». Le sujet 2 a repris cette même annotation mais en remplaçant cette classe avec une classe plus spécifique appelée « dispositif.owl#Environnement ». Le sujet 3 a repris cette annotation sans y apporter de changement supplémentaire. Le sujet 4 a repris, à son tour, cette même annotation mais pour finalement décider de la retirer.

Nous avons utilisé notre logiciel d'analyse pour générer un annogramme illustrant les différentes manipulations réalisées par les huit sujets qui ont participé à notre expérience. Cet annogramme a été mis en annexe à la page 184.

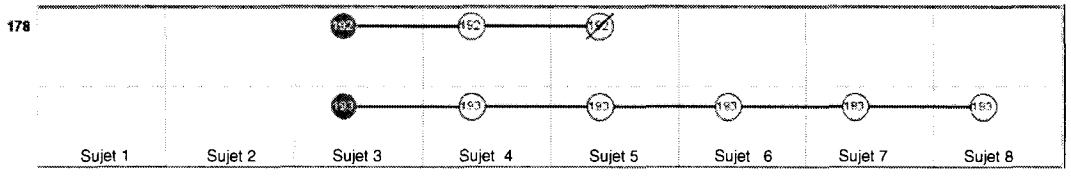
L'annogramme que nous avons généré devrait toutefois être utilisé avec précaution parce que certaines données sont manquantes. Nous avons, en effet, eu des problèmes techniques assez importants avec le prototype logiciel qui devait permettre aux sujets de réaliser l'annotation des documents (KATIA). Le prototype a notamment effectué un arrêt de traitement non prévu au moment même de réaliser la sauvegarde des données du sujet 4. Les identifiants qui devaient être associés aux annotations pour faciliter leur repérage subséquent ont ainsi été irrémédiablement perdus. Conséquemment, l'annogramme généré à partir du fichier de journalisation indique une interruption des emprunts à partir du quatrième sujet, ce qui ne correspond pas du tout à la réalité de l'expérience.

L'annogramme que nous avons généré à partir du fichier de journalisation nous a néanmoins permis d'identifier certains comportements d'utilisation. Nous avons ainsi remarqué que certaines annotations avaient été immédiatement acceptées par les sujets suivants alors que d'autres ont subi une multitude de changements successifs.

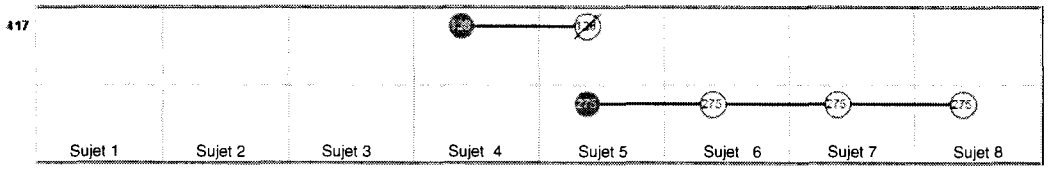
Voici quelques exemples choisis parmi l'ensemble des emprunts observés à l'intérieur de l'annogramme généré par notre logiciel d'analyse :



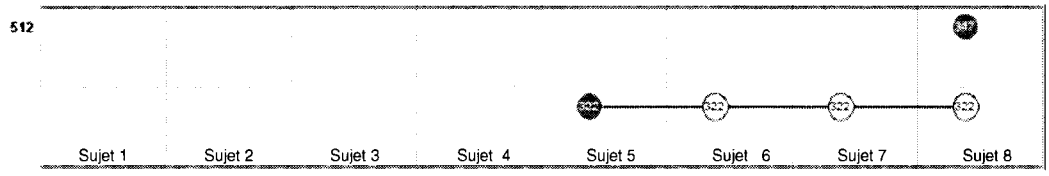
- L'annotation 109 a été créée par le sujet 2. Le sujet 3 a repris cette annotation pour y rajouter une nouvelle classe. Le sujet 4 a repris, à son tour, l'annotation modifiée par le sujet 3 mais en retirant toutefois la classe ajoutée par ce dernier pour la remplacer par une classe de son propre cru. Le sujet 4 a néanmoins décidé de conserver la classe d'origine du sujet 2. Le sujet 5 a toutefois décidé de ne pas reprendre cette annotation.



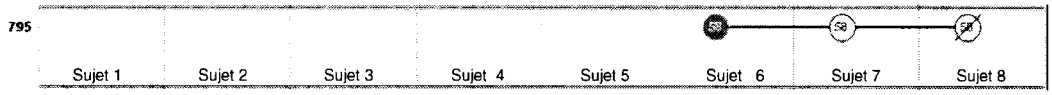
- L'annotation 178 a été créée par le sujet 3 en lui associant deux classes d'ontologie différentes. Le sujet 4 a réutilisé cette annotation sans y apporter de changement particulier. Le sujet 5 a repris cette même annotation en y retirant toutefois l'une des deux classes. Les sujets suivants ont par la suite repris cette même annotation sans y apporter aucun changement supplémentaire.



- L'annotation 417 a été créée par le sujet 4 en l'associant à une classe d'ontologie seulement. Le sujet 5 a repris cette annotation pour la remplacer par une nouvelle. Les sujets suivants ont ensuite réutilisé cette même annotation sans y apporter de changements nouveaux. En agissant de cette manière, les sujets 6, 7 et 8 ont ainsi, en quelques sorte, validé l'intervention réalisée par le sujet 5.



- L'annotation 512 a été créée par le sujet 5. Le sujet 6 et 7 n'ont pas trouvé bon de modifier cette annotation. Le dernier sujet a néanmoins trouvé important d'associer cette annotation à une deuxième classe d'ontologie pour mieux en préciser le sens.



- L'annotation 795 a été créée par le sujet 6. Elle a été directement reprise par le sujet suivant. Le dernier sujet a toutefois jugé que cette annotation était superflue et a décidé de la détruite.

Ces exemples illustrent bien le principe de filtrage successif qui a permis le retrait des annotations qui ne répondaient pas à l'ensemble du groupe pour ne laisser passer que les annotations correspondant au consensus commun. Nous pouvons donc ici parler d'un principe de « construction ascendante » du consensus.

Les sujets qui ont décidé de supprimer des annotations auraient très bien pu aussi décider de supprimer les classes d'ontologie associées à ces mêmes annotations. Paradoxalement, 99% des classes ont survécu aux changements apportés par les utilisateurs alors que 12% des annotations ont été carrément détruites. Nous pouvons

donc aussi ici parler d'un principe de « consensus ascendant » dans la construction des ontologies.

8.6 Facilité d'utilisation

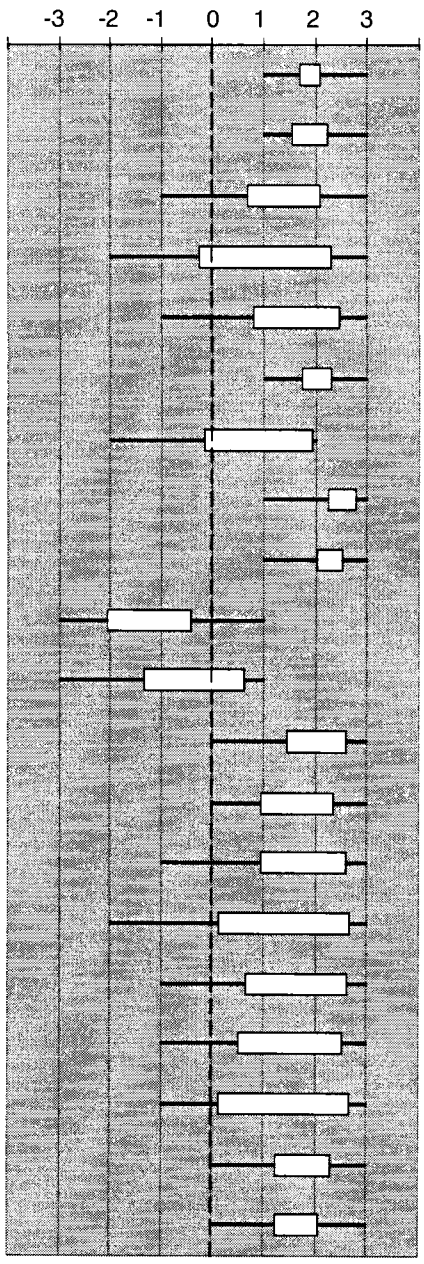
À la fin de l'expérimentation, chaque sujet a été soumis à un questionnaire pour tenter d'évaluer l'effort demandé et identifier les difficultés dans la tâche qui leur avait été imposée. Une copie de ce questionnaire a été mise en annexe à la page 166.

Les résultats des huit sujets ont été compilés avec Excel pour produire un diagramme en boîtes de dispersion (Tableau 8.7). Ce diagramme illustre les valeurs minimales et maximales pour chaque réponse ainsi que la valeur médiane des réponses. Le premier et le troisième quartile sont représentés par une boite fermée fixée aux échelles correspondantes.

Les résultats compilés indiquent clairement que tous les participants ont éprouvé de la difficulté à trouver et récupérer des ontologies sur le Web (questions 10-11). Certains sujets ont, en effet, passé 5 minutes à réaliser des recherches avec Swoogle (Tableau 8.1, page 110) pour essayer de repérer des ontologies adéquates à la description des contenus de cours en IHO. Aucun des sujets n'a toutefois réussi à trouver une ontologie utile avec Swoogle. Ces résultats ne sont pas en soi surprenants puisque [Wang, 2006] a déjà démontré que la plupart des ontologies présentement sur le Web sont généralement peu expressives et ne sont pas en soi suffisamment développées pour être vraiment utiles. D'autre part, il aurait été très surprenant que nos sujets trouvent une ontologie vraiment adaptée à leur besoin car les ontologies présentes sur le Web sont généralement construites comme des ontologies de haut niveau alors que nos sujets avaient ici besoin d'ontologies de bas niveau adaptées au contexte particulier de leur tâche. Finalement, cette situation illustre bien la nécessité de proposer des alternatives aux méthodes actuelles de développement d'ontologie pour favoriser le développement d'ontologie de bas niveau (tel que nous nous proposons de le faire ici).

Tableau 8.7: Distribution des réponses au questionnaire.

1. L'organisation de l'information est claire.
2. L'information disponible contribue à soutenir ma tâche.
3. Ce système est simple à utiliser.
4. L'interface de ce système est agréable.
5. Je suis satisfait de la facilité d'utilisation de ce système.
6. J'ai eu beaucoup de facilité à apprendre comment utiliser ce système.
7. L'utilisation de ce système demande peu d'effort.
8. J'ai été en mesure de produire facilement des pages Web.
9. J'ai été en mesure d'assigner facilement des annotations aux pages Web.
10. J'ai été en mesure de trouver facilement des ontologies sur le Web.
11. J'ai été en mesure de récupérer facilement des ontologies sur le Web.
12. J'ai été en mesure de modifier facilement la structure des ontologies.
13. J'ai été en mesure de construire rapidement ma propre ontologie.
14. Je me sens à l'aise avec ce système.
15. Ce système est agréable à utiliser.
16. J'ai été en mesure de compléter rapidement ma tâche avec ce système.
17. Je crois que ce système améliore la qualité de mon travail.
18. Je crois que ce système améliore ma productivité (rapport en quantité produite et coût mis en œuvre).
19. Je crois que ce système améliore mon efficacité (qualité de produire des résultats utiles).
20. Je crois que ce système simplifie mon travail.



Le questionnaire comportait deux questions similaires (question 4 et 15) placées vers le début et vers la fin du questionnaire pour essayer d'évaluer la constance des répondants. Les réponses données à chacune de ces questions sont similaires et il est donc possible d'en détruire que les participants ont réellement pris de temps d'évaluer les questions avant d'y répondre. Par contre, la réponse donnée à ces questions indique clairement que

certains participants n'ont pas trouvé ce système agréable à utiliser. Pourtant la plupart d'entre eux s'entendent pour affirmer que le système est simple utiliser (question 3). Nous pensons toutefois pouvoir expliquer cette apparente contradiction entre les réponses :

- Le logiciel d'annotation que nous avons utilisé était un prototype aux fonctions très limitées. De plus, ce prototype était miné par de nombreux défauts qui ont souvent demandé le redémarrage de l'application.
- La machine utilisée pour réaliser les tests était particulièrement mal adaptée aux exigences techniques du prototype logiciel (Tableau 7.2, page 104). La sauvegarde des quelques 120 000 triplets RDF générés par le logiciel a ainsi demandé plusieurs minutes de traitement, paralysant du même coup, l'interface utilisateur.

Certains sujets ont aussi souligné que l'utilisation du système demandait des efforts importants (question 7). Cette affirmation peut sans doute s'expliquer de la même manière que la précédente. Il faut néanmoins rajouter que la manipulation de plusieurs centaines d'annotations dans un délai de temps aussi court (2 à 6 heures) impliquait nécessairement un effort de concentration particulièrement important de la part des sujets. Ce travail aurait dû être exécuté sur des périodes de temps plus espacées de manière à justement limiter la nécessité d'un effort soutenu.

8.7 Comportement et attitude

Les sujets ont tous bien apprécié leur participation à l'expérience malgré l'importance de l'effort de concentration qui leur était demandé. Ils ont notamment apprécié le fait de pouvoir bénéficier du travail des autres pour accélérer leur propre travail.

- « Je suis chanceux parce qu'ils ont fait un gros travail. » Sujet 3

En observant l'ensemble les classes produites par les autres, un participant s'est même interrogé sur l'identité des auteurs précédents :

- « Je voudrais bien savoir qui a fait l'ontologie précédente.... pour savoir quel niveau de crédibilité lui donner. » Sujet 5

Cette question renvoie ainsi à la nécessité d'instaurer un mécanisme de confiance (*trust*) pour réaliser la validation des contenus trouvés. Et c'est exactement ce qui est déjà prévu dans l'architecture technologique du Web sémantique (Figure 5-7, page 73).

Un participant s'est aussi mis à douter de la qualité même des ontologies produites en soulignant que leur structure reprenait trop directement la structure même des paragraphes décrits :

- « Ce n'est pas une ontologie mais un résumé de cours hiérarchisé. » Sujet 4

Un autre sujet a aussi soulevé une même remarque en exprimant sa difficulté à visualiser la structure globale de l'ontologie.

- « Je n'arrive pas à classer les mots dans le sens que je voudrais lui donner... J'y vais page par page, mais j'ai l'impression qu'il me manque une vue globale. » Sujet 8

Le problème soulevé par le sujet 8 est particulièrement intéressant. Il souligne en effet que les ontologies ont été construites pour répondre à un besoin précis en partant d'un texte très particulier. Lorsqu'une ontologie est relativement petite, les défauts de classement ne sont généralement pas très apparents. Lorsqu'une ontologie grossit, ces défauts causent, par contre, des problèmes beaucoup plus importants. Le sujet 8 commence ainsi à vouloir « s'élever » dans le niveau d'abstraction des ontologies et désirerait plutôt manipuler des ontologies de niveau intermédiaire (par opposition à des ontologies de bas niveau). Ce problème avait d'ailleurs été soulevé très tôt par le sujet 1:

- « Je pense que je vais créer un problème parce que je fais une ontologie qui colle trop au texte et qui ne donnera pas nécessairement une structure d'ontologie optimale. » Sujet 1

Il serait tentant de conclure qu'une approche de construction ascendante ne serait finalement pas une solution aussi intéressante qu'il n'y parait pour la réalisation d'ontologies consensuelles. C'est pourtant, justement, cette prise de conscience spontanée chez les participants qui a entraîné chez eux la nécessité de restructurer les ontologies existantes en leur rajoutant des classes intermédiaires lorsque nécessaire (Figure 8-3, page 115). Et c'est notamment pour cette raison que le sujet 8 a investi autant de temps à réaliser la restructuration des classes existantes (Figure 8-10, p.125). Il parait donc évident que les personnes qui participent à un échange de contenus annotés auront ainsi tôt ou tard le réflexe de réagencer les classes d'ontologie au fur et à mesure qu'ils s'apercevront des problèmes de cohérence à l'intérieur de ces ontologies.

8.8 Discussion

Au départ de cette thèse, nous avions formulé l'hypothèse que le nombre de descriptions augmenterait au fur et à mesure de l'implication des nouveaux intervenants dans une même chaîne de partage. L'expérience que nous avons réalisée a confirmé cette hypothèse en démontrant un taux de réutilisation d'annotation de 88% et un taux de réutilisation de classes d'ontologie de 99%.

A première vue, il pourrait paraître surprenant de constater que des sujets qui n'avaient pas nécessairement des points de vue communs puissent arriver aussi facilement à s'entendre sur la définition des mêmes éléments d'ontologie. Nos sujets provenaient de différents domaines et possédaient des expériences de travail très différentes les uns des autres. [Fong, 2003][Bhatt, 2000] s'entendent pour dire des expériences différentes de travail entraînent des structures cognitives distinctes; ce qui conditionne du même coup des perspectives différentes entre les individus. Ces différences de point de vue ne sont pas nécessairement insurmontables. Les interactions entre les utilisateurs sont possibles

si chacun arrive à faire abstraction des frontières cognitives qui l'entraînent à percevoir des fonctionnalités différentes d'un même sujet [Carlile, 2002]. Un moyen concret de surmonter ces frontières cognitives consiste simplement à utiliser un langage formel pour réaliser l'échange de connaissances tacites entre individus.

Nous savons que lorsqu'un individu qui possède une connaissance tacite sur un sujet est invité à interpréter cette connaissance à des pairs au moyen d'un langage codifié, l'interprétation de ces connaissances par des pairs peut alors elle-même donner lieu à la construction de nouvelles connaissances [Brown, 2001]. Et c'est justement ce qui s'est produit avec notre expérience où les ontologies étaient utilisées comme un moyen pratique d'expliquer le sens formel d'une section de texte. Si le sujet suivant arrivait à saisir la logique de classement de l'annotation récupérée, alors il pouvait à son tour utiliser cette même logique de classement pour rajouter de nouvelles descriptions aux annotations existantes.

Ce phénomène de création de connaissances à partir des connaissances tacites exprimées par les autres a été déjà été décrit par [Nonaka, 1998] comme un principe d'interactions cycliques entre, d'une part, l'individu et le groupe, et d'autre part, le savoir tacite et le savoir explicite qui y sont échangé. La Figure 8-12 présente ce modèle d'interactions cycliques appelé « modèle SECI » (Socialisation, Externalisation, Combinaison, Intériorisation).

Le processus de création des connaissances est réalisé en quatre cycles successifs de socialisation, d'externalisation, de combinaison et d'intériorisation des connaissances :

- Socialisation (de la connaissance tacite à la connaissances tacite): processus de partage d'expériences qui permet de créer des connaissances tacites. Lorsque, par exemple, des personnes utilisent des démonstrations plutôt que des explications directes sur un sujet particulier.

- Externalisation (de la connaissance tacite à la connaissance explicite): processus d'articulation des connaissance tacites en des concepts explicites pouvant être formalisées sous une forme écrite.
- Combinaison (des connaissances explicites): processus de systématisation des concepts en un système de connaissances au travers d'échanges de toutes sortes entre les différents acteurs.
- Intériorisation (de la connaissance explicite à la connaissance tacite): processus d'incorporation des connaissances explicites vers des connaissances tacites. L'internalisation permet la génération de nouvelles idées et l'acquisition de nouvelles connaissances tacites pour supporter une nouvelle action des individus.

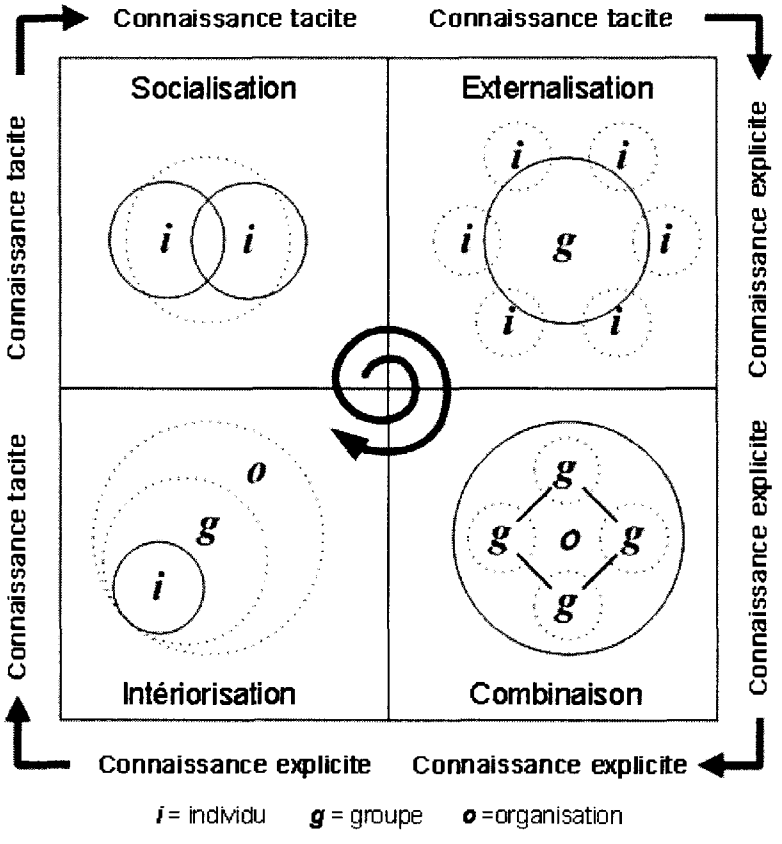


Figure 8-12 : La spirale de connaissance du modèle SECI [Nonaka, 1998].

Le modèle SECI démontre bien que la combinaison des connaissances explicites est une condition essentielle pour assurer le transfert des connaissances entre les différents intervenants en permettant notamment l'instauration d'un mouvement en spirale qui favorise la génération des connaissances. Les connaissances ne sont ainsi pas créées en ajoutant simplement de nouvelles informations, mais plutôt en publiant l'information tacite déjà existante à l'intérieur d'une organisation pour permettre à chacun de se l'approprier.

Et c'est justement ce qui s'est produit durant notre expérience lorsque les sujets ont rajouté de nouvelles annotations à leurs documents et qu'ils ont par la suite partagé ces annotations avec leurs pairs. Le nombre d'annotations n'a cessé de croître par la suite au point que certaines annotations ont fini par faire consensus et être systématiquement échangées entre les participants.

Il est important de se souvenir que le 12% de perte d'annotations observé entre les échanges s'applique ici dans un contexte où chaque sujet rajoutait toujours de nouveaux contenus annotés aux précédents. Cela signifie concrètement que les annotations du sujet 1 ont initialement subi une diminution de 12% mais que le risque de voir ces mêmes annotations être éliminées par la suite diminuait progressivement avec chaque nouvel échange réalisé entre les sujets. Si, par exemple, le sujet 2 rajoute de nouveaux contenus et que l'ensemble de ces annotations sont retranchées de 12% au moment de leur transfert vers le prochain sujet, alors le risque ponctuel de voir disparaître les annotations du premier sujet diminue proportionnellement au nombre de nouvelles annotations rajoutées par le second sujet. Le risque de voir disparaître les annotations du premier sujet devrait ainsi graduellement diminuer pour approcher le 0%. En d'autres termes, après plusieurs cycles d'emprunts, les annotations du premier sujet finiront par former un noyau d'éléments reflétant le consensus commun à l'intérieur de cette communauté d'échange.

Cette « sélection naturelle » des annotations a un impact direct sur la croissance des ontologies car ces annotations agissent comme les vecteurs de propagation permettant le transfert des ontologies entre sujets. L'utilisation continue des mêmes annotations signifie ainsi que les sujets finiront tous par posséder les mêmes ontologies. Il devrait donc y avoir ainsi un alignement spontané des ontologies entre les sujets. En d'autres termes, le fait de partager les mêmes annotations conditionne aussi le fait de partager les mêmes structures d'ontologie. C'est ce que nous avons appelé précédemment un « consensus ascendant ». Et c'est justement ce consensus ascendant qui permet de concevoir des ontologies sans avoir a recourir à des équipes d'experts, comme c'est actuellement le cas aujourd'hui.

Le domaine de connaissances et le niveau d'expertise des utilisateurs jouent probablement un rôle déterminant sur le taux de réutilisation observé durant notre expérimentation. Il est toutefois raisonnable de penser que l'effet de levier positif observé avec les emprunts d'annotations et les classes d'ontologies ne serait pas moins important pour d'autres types de contenus ou domaines. Le taux d'emprunt des annotations pourrait certainement varier sans toutefois compromettre pour autant l'effet de levier observé.

8.8.1 Qualité des ontologies produites

Il est important de se questionner sur la qualité des ontologies produites par les sujets. Certains sujets avaient, en effet, très tôt réalisé le danger de produire des ontologies dont la structure reprend trop fidèlement la structure même des contenus décrits.

Les ontologies produites par les différents sujets ne sont pas homogènes car elles contiennent à la fois :

- des éléments de table des matières organisés par sujets;
- des classes et sous-classes;
- des instances de classes.

Le mélange de ces trois types d'éléments à l'intérieur d'une même ontologie pose des difficultés importantes aux agents présents sur le Web sémantique en limitant leur capacité à exploiter ces informations pour réaliser un raisonnement logique. Des améliorations au prototype logiciel devraient être donc être entreprises avant de déployer une telle solution pour mieux encadrer les utilisateurs dans la réalisation d'ontologies.

Au départ de notre expérience, nous avions imaginé qu'il serait intéressant de comparer les ontologies produites avec des ontologies de haut niveau de manière à évaluer leurs lacunes et leurs faiblesses. Après mûres réflexions, nous avons abandonné l'idée d'une telle analyse parce :

1. Les ontologies produites sont des ontologies de bas niveau qui se distinguent justement des ontologies de haut niveau par l'adéquation de leur structure aux besoins particuliers d'une application.
2. Nous savons que l'expérience façonne le modèle cognitif des humains et cela se traduit concrètement par des points de vue différents pour un même sujet. Il sera donc toujours difficile de trancher sur la validité d'une ontologie de bas niveau parce ces différents points de vue entrent forcément en jeu dans la construction de ces ontologies (ce qui n'est pas nécessairement le cas des ontologies de haut niveau qui traitent de concepts très généraux).
3. Nous souscrivons à la thèse de [Cahier, 2005] qui stipule que la sémantique Web qui sera la plus utile et la plus efficace, sera assurément celle qui aura été créée au sein de communautés. Nous rajoutons toutefois que peu importe la taille ou l'importance des communautés, ce sera l'application qu'en feront ces communautés qui déterminera si ces ontologies seront utiles ou non.
4. Il n'existe pas de bonnes ou de mauvaise ontologies. Tout dépend finalement du contexte d'application auquel l'ontologie est destinée.

En définitive, la qualité d'une ontologie ne devrait jamais être évaluée en fonction d'un idéal absolu mais plutôt en fonction de l'application à laquelle elle est destinée. La qualité des ontologies actuellement produites avec notre prototype logiciel ne permettrait néanmoins pas de supporter le raisonnement logique d'un agent sur le Web sémantique et il faudrait encore améliorer ce prototype logiciel pour fournir un meilleur guidage dans la conception d'ontologie.

CHAPITRE 9 : CONCLUSION

Notre objectif était de concevoir une nouvelle méthodologie de conception d'ontologie. Une méthodologie se définit comme l'ensemble des méthodes, des règles et des techniques utilisées par un domaine particulier. Il est ainsi possible de représenter une méthodologie sous la forme d'une structure en étage où chaque niveau sert de base au niveau suivant.

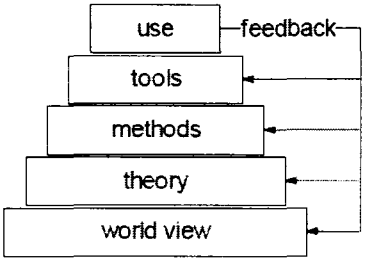


Figure 9-1 : Pyramide méthodologique [Schreiber, 2000].

Le premier niveau de cette « pyramide méthodologique » est occupé par un **constat du monde** qui met en lumière les principes et les constats qui appuient la méthodologie. Ce constat du monde supporte la **théorie** requise pour décrire les connaissances déclaratives propres à la méthodologie. Cette théorie permet à son tour de supporter les **méthodes** qui rendent son opération applicable. Ces méthodes supportent finalement l'utilisation des **outils** qui appuieront l'**usage** de la méthodologie proposée.

Nous avons à notre tour établi cette pyramide en faisant l'état de l'art de l'utilisation des repositoires de cours, de l'utilisation des métadonnées, de la conception d'annotation, de la conception de descriptions RDF ainsi que de la conception d'ontologies pour le Web sémantique. Nous avons présenté les principes théoriques concernant l'alignement des ontologies et la nécessité de recourir à des équipes spécialisées pour réaliser le consensus dans la définition des éléments d'ontologie. Nous avons élaboré notre propre hypothèse sur le comportement des concepteurs de cours qui partagent ensemble des connaissances formelles au moyen de contenus annotés. Nous avons élaboré une méthode pour la conception d'annotations robustes et instauré cinq principes pour

garantir l'alignement constant des annotations produites avec les ontologies correspondantes. Nous avons construit un outil qui exploite cette méthode pour réaliser la production de documents sur le Web sémantique. Nous avons par la suite exploité cet outil pour procéder à une expérimentation avec des sujets. Nous avons finalement construit un logiciel d'analyse qui produit des annogrammes pour mettre en évidence les stratégies d'emprunts entre les différents sujets. Nous avons ainsi réalisé la conception et la validation d'une nouvelle méthodologie pour la production ascendante et décentralisée d'ontologies sur le Web sémantique.

9.1 Contribution à l'avancement des connaissances

L'objectif de cette thèse était de démontrer qu'il est possible d'utiliser des contenus annotés et échangés entre partenaires pour supporter la conception d'ontologies consensuelles selon une approche ascendante. Nous avons démontré que le taux de réutilisation des annotations est de 88% alors que celui des classes échangées atteint 99%. Nous avons ainsi découvert l'existence d'un effet de levier important dans la conception de contenus annotés qui facilitera certainement la mise en place définitive du Web sémantique.

Il est toutefois important de souligner ici que les résultats obtenus ne proviennent que d'une seule expérimentation et que la généralisation des résultats reste en partie à tester. Il est toutefois raisonnable de penser que des résultats équivalents pourraient être obtenus en reprenant cette même expérimentation avec des contenus reliés à des domaines différents.

Les gains de cette découverte sont nombreux : notamment de ne plus être dépendant de la mise en place d'équipes spécialisées pour la production d'ontologies consensuelles, de réduire substantiellement la nécessité d'avoir à recourir à des techniques complexes d'alignement d'ontologies et de favoriser la capture des connaissances directement au niveau des concepteurs de contenu.

Nous avons conçu et validé une nouvelle méthodologie de conception de documents sémantiques en énonçant les principe nécessaires à la réalisation d'annotations robustes (page 89).

Nous avons construit un prototype logiciel qui supporte ces principes en facilitant la tâche de conception de pages Web et en automatisant la production de fichiers RDF destiné à réaliser la description de ces pages sur le Web sémantique (page 95).

Nous avons construit un système qui exploite ces descriptions RDF pour générer automatiquement un index de sites Web selon le principe d'une ontologie inversée (Figure 7-8). Cet index facilite le repérage des contenus annotés par les humains. Nous espérons que la présence d'un tel index motivera bientôt les concepteurs de cours à réaliser la description sémantique de leurs contenus.

Nous avons aussi jeté les bases d'un nouveau principe d'analyse des annotations appelé « annogramme » (page 127). Nous espérons que les annogrammes seront utilisés dans le futur pour mieux comprendre les stratégies mises en œuvre par les utilisateurs dans l'échange de contenus annotés sur le Web sémantique.

Finalement, pour reprendre les termes mêmes de Tim Berners-Lee (p.66), nous avons conçu et validé une nouvelle méthodologie pour la conception de descriptions RDF qui permet déjà de créer des relations nouvelles entre les objets d'apprentissage pour permettre "*a machine-readable content that will provide, say, automated translation between the output of a scientific device and the input of a datamining package used in some other discipline, or a self-evolving translator that allows one group of scientists to directly interact with the technical data produced by another. These new products will allow users to create relationships that allow communication when the commonality of concept has not (yet) led to a commonality of terms.*"

9.2 Suite des recherches

Le scénario de réutilisation de contenus par des concepteurs de cours peut être imaginé d'une multitude de manières différentes. Dans cette thèse, nous avons privilégié une chaîne d'échanges linéaire où chaque sujet travaillait à la suite d'un autre mais sans jamais toutefois directement interagir avec celui-ci. Les changements d'ontologies n'affectaient que le prochain sujet et le sujet précédent ne pouvait jamais directement bénéficier des changements apportés par ses pairs. Il serait toutefois intéressant de savoir comment cette connaissance des changements apportés par les pairs pourrait modifier le regard de chacun sur ses propres contenus. Comment se réalise la médiation des changements dans un contexte où chaque sujet peut interagir dynamiquement avec les changements apportés par les autres?

Notre expérience était limitée à la conception de classes d'ontologies seulement et les sujets n'avaient pas à définir de propriétés de classe. La définition des propriétés de classe demande un effort d'abstraction encore plus important et il se pose alors la question de savoir comment réussir à intégrer cette opération dans la tâche globale de conception de contenus sans nuire au volume ni à la qualité des annotations produites.

Le logiciel masque la complexité des ontologies pour l'utilisateur et il est raisonnable de penser que l'approche proposée ici, que nous avons testée avec des ontologies légères, fonctionnerait aussi bien avec des ontologies lourdes. L'interface du logiciel devrait toutefois être améliorée pour permettre à l'utilisateur de constater les contradictions entre les propriétés de classe et de les corriger si nécessaire.

Nous savons maintenant qu'il est possible de favoriser une construction ascendante des ontologies sur le Web sémantique en exploitant les contenus annotés échangés entre concepteurs de cours. Se pose maintenant la question de savoir comment réussir à harmoniser le développement des ontologies ascendantes avec les ontologies de haut niveau déjà présentes sur le Web? Comment réaliser l'alignement de ces ontologies?

Peut-on recourir à un système automatique pour guider « l'ascension » de ces ontologies?

Ce sont autant de questions qui demandent maintenant à être résolues. Nous espérons toutefois pouvoir poursuivre nos travaux pour bientôt répondre à plusieurs d'entre elles.

BIBLIOGRAPHIE

ABECKER, A., BERNARDI, A., HINKELMANN, K., KUHN, O., SINTEK M.. (1998) Toward a technology for organizational memories. In IEEE Intelligent Systems, May/June. pp. 41-48

ALBERT, L. K., (1993) YMIR: an ontology for engineering design, Ph.D. Thesis, University of Twente, Twente, The Netherlands.

ALEKSOVSKI, Z., TEN KATE, W., VAN HARMELEN, F. (2006). Exploiting the structure of background knowledge used in ontology matching. In Proc. of the Ontology Matching Workshop of the International Semantic Web Conference (ISWC), pp. 13-24

ANNETT, J., DUNCAN, K. D. (1967). Task analysis and training design. In Journal of Occupational Psychology, No.41, pp.211-221.

APPRIOU, A., AYOUN, A. ET AL. (2001). Fusion: General concepts and characteristics. In International Journal of Intelligent Systems, Vol.16, No.10, pp. 1107-1134.

ARROYO, S., DE BRUIJN, J., DING, Y. (2004) Ontology Mapping and Aligning, Deliverable 1.4 – Ontology Mapping and Aligning, University of Innsbruck, 21p

AUMUELLER, D., DO, H. H., MASSMAN, S., RAHM, E. (2005). Schema and ontology matching with COMA++. In Proc. of the 2005 ACM SIGMOD Conference , Baltimore, Maryland, pp. 906-908

AZOUAOU, F., CHEN, W., DESMOULINS C. (2004) Semantic Annotation Tools for Learning Material. In Workshop on Applications of Semantic Web Technologies for Adaptive Educational Hypermedia (SW-EL@AH'04), August 23-26, 2004, Eindhoven, The Netherlands.

BAADER, F., NUTT, W. (2003) Basic Description Logics. In the Description Logic Handbook, edited by F. Baader, D. Calvanese, D.L. McGuinness, D. Nardi, P.F. Patel-Schneider, Cambridge University Press, pp. 47-100

BARNARD, P.J. (1987). Cognitive Resources and the Learning of Human-Computer Dialogs. In: Carroll, J.M., (ed.) Interfacing Thought: Cognitive Aspects of Human-Computer Interaction, ch.6. MIT Press, Cambridge, MA,

BERLIN, J., MOTRO, A. (2002). Database Schema Matching Using Machine Learning with Feature Selection. In Proc. of the 14th Int. Conf. on Advanced Information Systems Eng. (CAiSE 02), LNCS 2348, Springer-Verlag, pp. 452-466.

BERLIN, J., MOTRO, A. (2001). Autoplex: Automated Discovery of Content for Virtual Databases. In Proc. of the Int. Conf. on Cooperative Information Systems (CoopIS), pp.108–122

BERNARAS, A., I. LARESGOITI, J. CORRERA, (1996) Building and reusing ontologies for electrical network applications. In Proc. of the 12th European conference on artificial intelligence (ECA196), Ed.John Wiley & Sons, Ltd., pp. 298-302

BERNERS-LEE T. (1998), Semantic Web Roadmap.
[http://www.w3.org/DesignIssues/Semantic.html]

BERNERS-LEE T., HENDLER J., LASSILA O. (2001). The Semantic Web. In Scientific American, (11 April 2001).
(http://www.sciam.com/2001/0501issue/0501berners-lee.html)

BERNERS-LEE T.; HENDLER J. (2001) Scientific publishing on the 'semantic web', In Nature, (Nature : Debates : technology developers : Future e-access to the primary literature) (http://www.nature.com/nature/debates/e-access/Articles/bernerslee.htm)

BERNERS-LEE, T. (1998), What the Semantic Web can represent. Parenthetical discussion to the Web Architecture at 50,000 feet and the Semantic Web roadmap. [http://www.w3.org/DesignIssues/RDFnot.html]

BERNHEIM BRUSH, A. J. (2002). Annotating Digital Documents for Asynchronous Collaboration, Technical Report UW-CSE-02-09-02, Computer Science & Engineering, University of Washington, 119 p.

BHATT, G. (2000). Information dynamics, learning and knowledge creation in organizations. In The Learning Organization, Vol.7, No.2, pp. 89-99

BOOCH, G., RUMBAUGH, J., JACOBSON, I. (1997). Unified Modeling Language User Guide, ISBN: 0-201-57168-4, Addison Wesley, est. publication

BRASE, J., NEJDL, W. (2003) Annotation for an open learning repository for computer science. In Annotation for the Semantic Web, IOS-press, Netherlands, pp. 212-227

BROWN, J. S., DUGUID, P. (2001). Knowledge and organization: A social-practice perspective. In Organization Science, Vol.12, No.2, pp. 198-213

CAHIER, J.-P (2005), Ontologies sémiotique pour le Web socio sémantique: étude de la gestion coopérative des connaissances avec des cartes hypertopiques, Thèse en informatique, Université de technologie de Troyes, France, 330 pp.

CAMPBELL, L., LITTLEJOHN, A., DUNCAN, C., (2001). Share and share alike: encouraging the reuse of academic resources through the Scottish electronic Staff Development Library. In ALT Journal, Taylor & Francis Group. Vol.9, No.2, pp. 28-38

CARD, S., MORAN, T., NEWELL, A. (1983). The Psychology of Human-Computer Interaction. Hillsdale, NJ: Erlbaum.

CARLILE, P. (2002). A pragmatic view of knowledge and boundaries: Boundary objects in new product development. In Organization Science, Vol.12, No.4, pp. 442-56

CASTANO, S., DE ANTONELLIS, V. (1999). A Schema Analysis and Reconciliation Tool Environment. In Proc. of the 1999 Int. Symposium on Database Engineering & Applications (IDEAS), pp. 53–62

CHANDRASEKARAN, B., JOSEPHSON, J.R., BENJAMINS, V.R. (1999). What are ontologies, and why do we need them. In IEEE Intelligent Systems, Vol.14, No.1, pp. 20-26

CHANG, C.C.K. & GARCIA-MOLINA, H. (1998). Conjunctive constraint mapping for data translation. Third ACM Conference on Digital Libraries, Pittsburgh, USA.

CHAWATHE, S., GARCIA-MOLINA, H., HAMMER, J., IRELAND, K., PAPAKONSTANTINOU, Y., ULLMAN, J., WIDOM, J. (1994). The TSIMMIS project: Integration of heterogeneous information sources. In Proc. of the 16th Meeting of the Information Processing Society of Japan (IPSJ), Tokyo, Japan, pp. 7-18

CIRAVEGNA, F., DINGLI, A., PETRELLI, D., WILKS Y. (2002) User-System Cooperation in Document Annotation based on Information Extraction. In Proc. of the 13th Int. Conf. on Knowledge Engineering and Knowledge Management (EKAW02), Ontologies and the Semantic Web, Sigüenza, Spain, Springer Verlag, pp. 122-137

CLIFTON, C., HOUSMAN,E., ROSENTHAL, A. (1997). Experience with a combined approach to attribute-matching across heterogeneous databases. In Proc. of the IFIP Working Conference on Data Semantics (DS-7), pp. 429-451.

CONRAD, R. (1951). Speed and load stress in a sensori-motor skill. In British Journal of Industrial Medicine, No.8, pp. 1-7

Conzilla, the Concept Browser. (http://www.conzilla.org/)

CROSSMAN, E. R. F. W. (1956). Perceptual activity in manual work. In Research, Vol.9, pp. 42-49

CRUZ, I. F., SUNNA, W., MAKAR, N., BATHALA, S. (2007). A visual tool for ontology alignment to enable geospatial interoperability. In Journal of Visual Languages and Computing & Computing, No.18, pp. 230–254

DO, H., AND RAHM, E. (2002). Coma: A system for fexible combination of schema matching approaches. In Proceedings of the 28th Conf. on Very Large Databases (VLDB).

DOAN, A., MADHAVAN, J., DHAMANKAR, R., DOMINGOS, P., HALEVY, A. (2003). Learning to Match Ontologies on the Semantic Web. In The Int. Journal on Very Large Data Bases (VLDB), Vol.12, No.4, 303–319.

DOAN, A.H., DOMINGOS, P., HALEVY, A. (2001). Reconciling Schemas of Disparate Data Sources: A Machine-Learning Approach. In Proc. of the ACM SIGMOD Conf. on Management of Data, pp. 509-520

Dublin Core Metadata Initiative. (http://dublincore.org/)

DZBOR, M., DOMINGUE, J., MOTTA, E. (2003) Magpie – towards a semantic web browser. In Proc. of the 2nd International Semantic Web Conference (ISWC2003), Sanibel Island, Florida, USA, pp. 690-705

Edutella Project Homepage. (http://edutella.jxta.org/)

ERIKSSON, H. (2007). The semantic-document approach to combining documents and ontologies. In the Int. Journal of Human-Computer Studies, Vol.65, No.7, pp. 624-639

EU-NSF Working Group on Metadata (1999) Metadata for Digital Libraries: a Research Agenda. (http://www.ercim.org/publication/ws-proceedings/EU-NSF/metadata.html)

EU-NSF Working Group on Metadata (1999). Digital Libraries Initiative Phase 2. (http://www.dli2.nsf.gov/internationalprojects/eu_d.html)

EUZENAT, J., BACH, T. L., BARRASA, J., BOUQUET, P., MAYNARD, D., STAMOU, G., STUCKENSCHMIDT, H., ZAIHRAYEU, H., HAUSWIRTH, M., EHRIG, M., JARRAR, M., SHVAIKO, P., DIENG-KUNTZ, R., HERNÁNDEZ, R. L., TESSARIS, S., ACKER, S. V. (2004). State of the art on current alignment techniques, Knowledge Web. [http://knowledgeweb.semanticweb.org]

FERNÁNDEZ-LÓPEZ, M., GÓMEZ-PÉREZ, A. (2002). A Survey on Methodologies for Developing, Maintaining, Evaluating and Reengineering Ontologies, Deliverable 1.4, OntoWeb Consortium, 56 p.

FONG, P. (2003). Knowledge creation in multidisciplinary project teams: an empirical study of the processes and their dynamic interrelationships. In Int. Journal of Project Management, Vol.21, No.7, pp. 479-86

FRIESEN, N. (2004). International LOM survey: Report. ISO/IEC JTCI/SC36 subcommittee [http://mdlet.jtc1sc36.org/doc/SC36_WG4_N0109.pdf]

GARGANTILLA, J. A. R., GÓMEZ-PÉREZ, A. (2004) A survey on ontology-based applications. e-commerce, knowledge management, multimedia, information sharing and educational applications will deserve special attention. Deliverable 1.6. OntoWeb: Ontology-based Information Exchange for Knowledge Management and Electronic Commerce, (OntoWeb), 32p.

GARWOOD, K.L., LORD, P.W., PARKINSON, H., PATON, N.W., GOBLE, C.A., Pedro (2005) Ontology Services: A Framework for Rapid Ontology Markup. In Proc. of the 2nd European Semantic Web Conference. Springer Verlag, A. Gomez-Perez, J. Euzenat (eds), pp. 578-591

GIBSON, C. B. (2001). From knowledge accumulation to accommodation: cycles of collective cognition in work groups. In Journal of Organizational Behavior, Vol.22, No.2, pp. 121-34

GILBRETH, F. (1909). Bricklaying System, NY and Chicago, The Myron C. Clark Publishing Co., [Easton, PA, Hive Publishing (reprint), 1974.], 338 p.

GLIGOROV, R., Aleksovski, Z., ten Kate, W., van Harmelen, F. (2007). Using google distance to weight approximate ontology matches. In Proc. of the International Semantic Web Conference (ISWC), pp. 767-775

GODBY, C.J. (2004). What do Application Profiles Reveal About the Learning Object Metadata Standard? ARIADNE, Vol.41, pp. 1-15

GÓMEZ-PÉREZ, A. (2002) A survey on ontology tools. Deliverable 1.3. OntoWeb: Ontology-based Information Exchange for Knowledge Management and Electronic Commerce, OntoWeb Consortium, 119p.

GRUBER, T. R. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. In International Journal Human-Computer Studies, No.43, pp.907-928.

GRUBER, T. R. (1994) Toward Principles for the Design of Ontologies Used for Knowledge Sharing. In International Journal of Human-Computer Studies, Vol.43, No.5/6, pp.907-928

GRUNINGER, M., FOX, M.S. (1995). Methodology for the Design and Evaluation of Ontologies. In Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal.

GUARINO, N. (1994). The Ontological Level. In R. Casati, B. Smith and G. White (eds.), Philosophy and the Cognitive Science. Hölder-Pichler-Tempsky, Vienna: 443-456.

GUARINO, N. (1997) Understanding, Building, and Using Ontologies, In International Journal of Human-Computer Studies, 46(2/3), pp. 293--310

GUARINO, N. (1998) Formal Ontology and Information Systems. In N. Guarino, (Ed.) Formal Ontology in Information Systems. IOS Press, Amsterdam, Netherlands. pp. 3-15

GUARINO, N. AND GIARETTA, P. (1995) Ontologies and Knowledge Bases: Towards a Terminological Clarification. In N. Mars (ed.) Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing 1995. IOS Press, Amsterdam: pp.25-32.

GUHA, R.V. (1991). Contexts: A formalization and some applications. Ph.D. Thesis, Stanford University, (Technical Report STAN-CS-91-1399-Thesis et MCC ACT-CYC-423-91).

SCHREIBER, G. AKKERMANS, H., ANJEWIERDEN, A., DE HOOG, R., SHADBOLT, N., VAN DE VELDE, W., WIELINGA, B. (1999). Knowledge Engineering and Management: The CommonKADS Methodology. MIT-Press, Cambridge, MA,.

HANDSCHUH, S., STAAB, S. (2002) Authoring and annotation of web pages in CREAM. In Proc. of the 11th Int. WWW Conference, Honolulu, Hawaii, USA, ACM Press, pp. 462-473

HARRIS, M.C., THOM, J.A. (2006.). Challenges Facing the Retrieval and Reuse of Learning Objects. In ECDL Workshop on Learning Object Repositories as Digital Libraries: Current challenges, Alicante, Spain.

HEERY, R. (2005). Digital Repositories Review. Permalink, UKOLN, University of Bath and Sheila Anderson, Arts and Humanities Data Service. [http://www.ukoln.ac.uk/repositories/publications/review-200502/]

HORI, M., ABE, M., ONO, K. (2003) Extensible Framework of Authoring Tools for Web Document Annotation. In Proc. of the International Workshop on Semantic Web Foundations and Application Technologies (SWFAT), Nara, Japan.

IEEE (2002) IEEE Learning Object Metadata. Publication of IEEE, 1484.12.1-2002

IEEE Learning Technology Standards Committee. (http://ltsc.ieee.org/)

IMS Content Packaging Specification. (http://www.imsglobal.org/content/)

IMS Meta-data Specification. (http://www.imsglobal.org/metadata/)

ISO (1998) International Standard 9241-11: Ergonomic requirements for office work with visual display terminal, Part 11: Guidance on usability. International Organisation for Standards, Geneva.

ISO/IEC JTC1/SC36 - Information Technology for Learning, Education, and Training. (http://jtc1sc36.org/)

JENA (2003). Jena 2 - A Semantic Web Framework, Hewlett-Packard, [http://www.hpl.hp.com/semweb/jena.htm]

JOHNSON, P. (1992). Human Computer Interaction. London: McGraw-Hill

JOHNSON, P., DIAPER, D., LONG, J. (1985). Task Analysis in Interactive Systems Design and Evaluation. In Analysis, Design and Evaluation of Man-Machine Systems. (Eds. Johannsen, G., Mancini, G. and Martensson, L.), Pergamon Press, pp.123-127.

KAHAN J., KOIVUNEN M.-R. (2001) Annotea: an open RDF infrastructure for shared Web annotations. WWW10 conference, pp. 623-632 (http://www10.org/cdrom/papers/488/)

KAHAN, J., KOIVUNEN, M.-R., (2001) Annotea: an open RDF infrastructure for shared Web annotations. In Proc. of the 10th International WWW Conference, pp. 623-632

KALYANPUR, A., HENDLER, J., PARSIA, B., GOLBECK, J. (2003) SMORE - Semantic Markup, Ontology, and RDF Editor. [http://www.mindswap.org/papers/SMORE]

KASHYAP V., SHETH A. (1996) Semantic Heterogeneity in Global Information Systems: The Role of Metadata, Context and Ontologies. In M. Papzoglou, and G. Schlageter, (eds.), Cooperative Information Systems: Current Trends and Directions, Academic (1998), pp. 139-178,(http://lsdis.cs.uga.edu/lib/download/KS97.ps)

KETTLER, B., STARZ, J., MILLER, W., HAGLICH, P. (2005) A Template-based Markup Tool for Semantic Web Content. In Proc. of the 4th Int. Semantic Web Conf. (ISWC 2005), Galway, Ireland, pp. 446-460

KIRWAN, B., AINSWORTH, L. K. (1992). A guide to task analysis. London: Taylor & Francis, UK, 427 p.

KIU, C. C., LEE, C. S. (2006). Ontology Mapping and Merging through OntoDNA for Learning Object Reusability. In Educational Technology & Society Journal, Vol.9, No.3, pp. 27-42

KLEIN, M. (2001). Combining and Relating Ontologies: An Analysis of Problems and Solutions. In Workshop on Ontologies and Information Sharing (IJCAI-2001), Seattle, USA, pp. 309-327

KOPPI, T., LAVITT, N. (2003). Institutional Use of Learning Objects Three Years on: Lessons Learned and Future Directions. In P. Kommers & G. Richards (Eds.), Proc. of World Conference on Educational Multimedia, Hypermedia and Telecommunications, Chesapeake, VA. pp. 644-648

KOTIS, K., VOUROS, G.A., STERGIOU, K. (2006). Towards automatic merging of domain ontologies: The HCONE-merge approach. In Elsevier's Journal of Web Semantics (JWS), Vol.4, No.1, pp. 60-79

LANFRANCHI, V., CIRAVEGNA, F., PETRELLI, D. (2005) Semantic Web-Based Document: Editing and Browsing in AktiveDoc. In Proc. of Second European Semantic Web Conference (ESWC 2005), Heraklion, Crete, Greece.

LASSILA, O., MCGUINNESS, D. (2001) The Role of Frame-Based Representation on the Semantic Web, Technical Report, KSL-01-02, Knowledge Systems Laboratory, Stanford, University, Stanford, California. 10 p.

LAWRENCE S., GILES C. L. (1999) Accessibility and Distribution of Information on the Web, Accessibility of information on the web. In Nature, Vol.400, (July 8), pp. 107-109.

LEWIS, J. R. (1995). IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. In Int. Journal of Human-Computer Interaction, Vol.7, No.1, pp. 57-78

LI, W.S., CLIFTON, C. (1994). Semantic Integration in Heterogeneous Databases Using Neural Networks. In Proc. of the 20th Int. Conf. on Very Large Data Bases (VLDB), pp.1-12

LI, W.S., CLIFTON, C. (2000). SemInt: A Tool for Identifying Attribute Correspondences in Heterogeneous Databases Using Neural Network. In Data and Knowledge Engineering, Vol.33, No.1, pp. 49–84

LINECAR P., SIVITER D., SIVITER P. (1994) Collaborative Courseware Development for the Analysis and Design of Software Systems Workshop and paper at the XXIX Annual Conference of the Association for Educational & Training Technology, Napier University, Edinburgh, Scotland (April 1994).

MADHAVAN, J., BERNSTEIN, P.A., RAHM, E. (2001). Generic schema matching using Cupid. In Proc. of the 27th International Conference on Very Large Data Bases, pp. 49-58

MARSHALL CATHERINE C. (1998). Toward an ecology of hypertext annotation. In Proc. of the ACM conf. on Hypertext and Hypermedia: Links, Objects, Time and Space, Pittsburgh (USA). ACM Press, pp. 40-49

MARSHALL CATHERINE C. (2000) The Future of Annotation in a Digital (Paper) World. Successes and Failures of Digital Libraries. H. a. Twidale Ed. Urbana-Champaign, University of Illinois (USA), pp. 97-117

MARSHALL CATHERINE C., PRICE MORGAN N., GOLOVCHINSKY GENE, SCHILIT BILL N. (1999). Introducing a digital library reading appliance into a reading group. In Proc. ACM conf. on Digital libraries, Berkeley (USA), ACM Press, pp. 77-84

MARSHALL, C. C. (1997) Annotation: from paper books to the digital library. In Proc. of the second ACM Int. Conf. on Digital libraries, Philadelphia, Pennsylvania, United States, pp. 131-140.

MCDOWELL, L., ETZIONI, O., GRIBBLE, S. D., et autres (2003). Evolving the Semantic Web with Mangrove. UW Tech Report, Department of Computer Science & Engineering, University of Washington, Seattle, WA, United-State, 12p.

MCNAUGHT, C. (2003) Identifying the complexity of factors in the sharing and reuse of resources. In Littlejohn, A. (Ed), Reusing online resources: a sustainable approach to e-learning, Kogan Page, London, UK, pp. 199-211

MELNIK, S., GARCIA-MOLINA, H., RAHM, E. (2002). Similarity Flooding: A Versatile Graph Matching Algorithm. In Proc. of the 18th Int. Conf. on Data Engineering (ICDE), San Jose, CA, USA, pp. 117-128

MENA, E., ILLARRAMENDI, A., KASHYAP, V., SHETH, A. P. (1996). OBSERVER: An approach for query processing in global information systems based on interoperation across pre-existing ontologies. In Proc. of the First IFCIS Int. Conf. on Cooperative Information Systems (CoopIS'96), Brussels, Belgium, pp. 14–25

MERRILL D. M. (1998) Instructional Transaction Theory (ITT): Instructional Design Based on Knowledge Objects. Chapter 17 in C. M. Reigeluth (ed.), Instructional-Design Theories and Models: A New Paradigm of Instructional Theory. Mahwah, NJ: Lawrence Erlbaum Associates. (http://itech1.coe.uga.edu/itforum/paper22/paper22a.html)

MILLE, D. (2005). Modèles et outils logiciels pour l'annotation sémantique de documents pédagogiques, Université Joseph Fourier Grenoble 1, U.F.R. Informatique et Mathématiques Appliquées, 176 pp.

MILLER, R.J. et al. (2001). The Clio Project: Managing Heterogeneity. In ACM SIGMOD Record, Vol.30, No.1, pp. 78–83

MILO, T., ZOHAR, S. (1998). Using Schema Matching to Simplify Heterogeneous Data Translation. In Proc. of the 24 Int. Conf. on Very Large Data Bases (VLDB), pp. 122–133

MINSKY, M. (1974) A Framework for Representing Knowledge, MIT-AI Laboratory Memo 306. In Cognitive Science, Collins, Allan and Edward E. Smith (eds.) Morgan-Kaufmann, 1992.

MIT Open Knowledge Initiative. (http://web.mit.edu/oki/).

MITCHEL, R. NICHOLAS, S. (2005). Knowledge creating mechanisms and competitive advantage: The value of cognitive diversity, transactive memory, openmindedness norms and trans-specialist knowledge. In Proc. of the 6th European Conference on Knowledge Management, Academic Conferences Limited, Reading, UK, pp. 326-333

MITRA, P., WIEDERHOLD, G., JANNINK, J. (1999). Semi-automatic Integration of Knowledge Sources. In Proc. of Fusion'99, Sunnyvale, California, USA.

Moran, T. P. (1978). Introduction to the command language grammar, Report No.55L-78.3. Palo Alto, California, Xerox Corporation.

MOSTOWFI, F., FATOUHI, F., ARISTAR, A. (2005) OntoGloss: An Ontology-based Annotation Tool. In Proc. of 2005 Electronic Metastructure for Endangered Languages Data (E-MELD) Workshop, Wayne State University, 11 p.

NARDI, D., BRACHMAN, R. J. (2002) An Introduction to Description Logics. In the Description Logic Handbook (Chap. 1), edited by F. Baader, D. Calvanese, D.L. McGuinness, D. Nardi, P.F. Patel-Schneider, Cambridge Univ. Press, pp. 5-44

NECHES, R. FIKES, R. FININ, T.GRUBER, T., PATIL, R., SENATOR, T. SWARTOUT, W. R. (1991) Enabling technology for knowledge sharing, AI Magazine 12 (3), pp. 36-56.

NEJDL W., WOLF B., STAAB S., TANE J. (2002) EDUTELLA: Searching and Annotating Resources within an RDF-based P2P Network. In Proc. of the Semantic Web Workshop (WWW2002), Honolulu, Hawaii (7 May 2002).

NILSSON M. (2000) The Conzilla Design - The definitive reference, CID/NADA/KTH. (http://www.conzilla.org/doc/conzilla-design/conzilla-design.html).

NILSSON M., PALMÉR M. (1999) Conzilla - Towards a Concept Browser. CID-53, KTH, Stockholm, Sweden (kmr.nada.kth.se/papers/ConceptualBrowsing/cid_53.pdf).

NONAKA, I., KONNO, N. (1998). The Concept of 'ba': Building a Foundation for Knowledge Creation. In California Management Review , Vol.40, No.3, pp. 40-54

NOY, N. F., MCGUINNESS, D. L. (2001) Ontology Development 101: A Guide to Creating Your First Ontology. Knowledge Systems Laboratory (SMI-2001-0880), 21p.

NOY, N. F., MUSEN, M. A. (2000). PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment. In Proc. of Seventeenth National Conference on Artificial Intelligence, Austin, Texas, USA, pp. 450-455

NOY, N.F., MUSEN, M.A. (2000). PROMPT: algorithm and tool for automated ontology merging and alignment. In Proc. of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence (AAAI), Austin, Texas, pp. 450-455

OBENDORF, H. (2003) Simplifying annotation support for real-world-settings: a comparative study of active reading. In Proc. of the fourteenth ACM conference on Hypertext and Hypermedia, Nottingham, UK, pp. 120-121

PALOPOLI, L., TERRACINA, G., URSINO, D. (2000). The System DIKE: towards the semi-automatic synthesis of Cooperative Information Systems and Data Warehouse. In Proc. Int. Symposium on Advances in Databases and Information Systems, Prague, pp. 108-117

PATERNO, F., MANCINI, C., MENICONI, S. (1997). ConcurTaskTrees: A diagrammatic Notation for Specifying Task Models. In Proceedings INTERACT97, Chapman & Hall, pp. 362-369

PHILIPPE KRUCHTEN (1998). The Rational Unified Process, Addison Wesley. 255p.

PIROLLI, P., CARD, S. K., VAN DER WEGE, M. M. (2003). The effects of information scent on visual search in the hyperbolic tree browser. In ACM Transactions on Computer-Human Interaction (TOCHI), ACM Press, New York, NY, USA, Vol.10, No.1, pp. 20-53

POPOV, B., KIRYAKOV, A., OGNYANOFF, D., MANOV, D. KIRILOV, A., GORANOV, M. (2003). KIM – Semantic Annotation Platform. In Proc. of the 2nd Int. Semantic Web Conference (ISWC2003), 20-23 October 2003, Florida, USA, pp. 484-499

PRIÉ, Y., GARLATTI, S. (2004) Méta-données et annotations dans le Web sémantique. In Information - Interaction - Intelligence, Une Revue en Sciences du Traitement de l'Information, Hors Série 2004 Web Sémantique. (http://www.revue-i3.org/hors_serie/annee2004/).

QUILLIAN, M. R. (1967) Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities. In Behavioral Science, No.12, pp. 410-430

RAHM, E., BERNSTEIN, P. A. (2001). A survey of approaches to automatic schema matching. In the Int. Journal on Very Large Data Bases, Vol.10, No.4, pp. 334-350

RECKER M., WILEY D. A. (2001) A non-authoritative educational metadata ontology for filtering and recommending learning objects. In Interactive Learning Environments, Vol.9, No.3, pp. 255-271

RISVIK K. M., MICHELSEN R. (2002) Search engines and Web dynamics, In Journal Computer Networks, Vol.39, No.3, pp. 289-302

ROSCHELLE J., TEASLEY S. (1995) The construction of shared knowledge in collaborative problem solving. In O'Malley, C.E., (ed.), Computer Supported Collaborative Learning. Springer-Verlag, Heidelberg, pp. 69-97.

SABOU, M., D'AQUIN, M., MOTTA, E. (2006). Using the Semantic Web as Background Knowledge for Ontology Mapping. In Proc. of the Int. Workshop on Ontology Matching (OM-2006), collocated with ISWC'06. Athens, USA, pp. 1-12.

SCAPIN, D., PIERRET-GOLBREICH, C. (1989). Une Méthode Analytique de Description des tâches. Colloque sur l'ingénierie des interfaces homme-machine, Cargèse.

SCHREIBER, G., AKKERMANS, H., ANJEWIERDEN, A., DE HOOG, R., SHADBOLT, N., VAN DE VELDE, W., & WIELINGA, B. (2000). Knowledge Engineering and Management - The CommonKADS Methodology. The MIT Press, Cambridge, Massachusetts; London, England, 476 p.

SCHREIBER, G., WIELINGA, B., AND JANSWEIJER, W. (1995) The KAKTUS View on the 'O' Word. In Proceedings of IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing. Montreal, Canada.Schreiber, G., Akkermans, H., Anjewierden, A., Shadbolt, N., Van de Velde, W., Wielinga, B. (2000). Knowledge Engineering and Management: The CommonKADS Methodology. MIT Press, Cambridge (USA), 455 p.

SEYMOUR, W.D. (1966). Industrial skills. Pitman & Sons Ed., New York, 401 p.

SHAH K., SHETH A. P. (1998) Logical Information Modeling of Web-Accessible Heterogeneous Digital Assets, ADL 1998, pp. 266-275

SHEPHERD, A.(1989). Analysis and training in information technology tasks. In Diaper, D (ED), Task Analysis for Human-Computer Interaction, Chichester, Ellis Horwood.

SHVAIKO, P., EUZENAT, J. (2005). A survey of schema-based matching approaches. In Journal on Data Semantics, Vol.4, pp. 146-171

SIVITER D. (1999) Objects in Education: from Courseware Widgets to Virtual Universities. In Educational Technology & Society, Vol.2, No.2.

STAAB, S., SCHNURR, H.-P., STUDER, R., SURE, Y. (2001). Knowledge processes and ontologies. In IEEE Intelligent Systems, Special Issue on Knowledge Management, 16 (1), pp. 26-34

STUDER, R., BENJAMINS, V. R., FENSEL, D. (1998) Knowledge engineering : principles and methods. In IEEE Transactions on Data and Knowledge Engineering, Vol.25, No.1-2, pp.161–197

STUMME, G., MAEDCHE, A. (2001). FCA-Merge: Bottom-up merging of ontologies. In Proc. of the 17th Int. Joint Conference Conf. on Artificial Intelligence (IJCAI '01), USA, pp. 225-230

SWARTOUT, B., R. PATIL, K. KNIGHT, T. RUSS. (1997). Toward distributed used of large-scale ontologies, Ontological engineering, AAAI-97 Spring symposium series, pp. 138-148.

T. R. GRUBER (1993) A translation approach to portable ontologies. In Knowledge Acquisition, Vol.5, No.2, pp. 199-220

TAYLOR, F. W. (1911). The Principles of Scientific Management. New York: Harper Bros Ed. [Published in Norton Library 1967], 148 p.

UREN, V., CIMIANO, P., IRIA, J., HANDSCHUH, S., VERA, V.M., MOTTA, E., CIRAVEGNA F. (2006) Semantic annotation for knowledge management: Requirements and a survey of the state of the art. In Journal of Web Semantics, Vol.4, No.1, pp. 14-28.

USCHOLD, M., GRUNINGER, M. (1996) Ontologies: Principles, Methods and Applications. Knowledge Engineering Review, 11 (2), 69 p.

VAN DER VEER, G.C., LENTING, B.F., BERGEVOET, B.A.J. (1996). GTA:Groupware Task Analysis - Modeling Complexity Acta Psychologica, No.91, pp. 297-322

VAN HEIJST, G., SCHREIBER, A. T., AND WIELINGA, B. J. (1996) Using Explicit Ontologies in KBS Development. In International Journal of Human and Computer Studies.

VARGAS-VERA, M., MOTTA, E., DOMINGUE, J., LANZONI, M., STUTT, A., CIRAVEGNA, F. (2002) MnM: Ontology Driven Semi-Automatic and Automatic Support for Semantic Markup. In Proc. of the 13th International Conference on Knowledge Engineering and Management (EKAW 2002), Springer Verlag, pp. 379-391

VISSER, P.R.S., CUI, Z. (1998). On accepting heterogeneous ontologies in distributed architectures. In. Proc. of the ECAI'98 workshop on applications of ontologies and problem-solving methods, Brighton, UK, pp. 112–119

VISSER, P.R.S., TAMMA, V.A.M. (1999). An experience with ontology clustering for information integration. In. Proc. of the IJCAI-99 Workshop on Intelligent Information Integration in conjunction with the Sixteenth International Joint Conference on Artificial Intelligence, Stockholm, Sweden.

WANG, T. D., PARSIA, B., HENDLER, J. (2006). A Survey of theWeb Ontology Landscape. In the 5th Int. Semantic Web Conf., Athens, GA, USA.

WEINSTEIN, P.C. & BIRMINGHAM, P. (1999). Comparing Concepts in Differentiated Ontologies. In Proc. of the Twelfth Workshop on Knowledge Acquisition, Modeling and Management (KAW99), Banff, Alberta, Canada.

WIELINGA, B. J. AND SCHREIBER, A. T. (1993) Reusable and sharable knowledge bases: a European perspective. In Proc. of First Int. Conf. on Building and Sharing of Very Large-Scaled Knowledge Bases. Tokyo, Japan Information Processing Development Center.

WILEY D. A. (2000). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley (ed.) The Instructional Use of Learning Objects. (http://reusability.org/read/chapters/wiley.doc)

WOLFE JOANNA L., NEUWIRTH CHRISTINE M. (2001). From the margins to the center: the future of annotation. Journal of Business and Technical Communication, Vol.15, pp. 333-371

YOUN, S., MCLEOD, D. (2006). Ontology Development Tools for Ontology-Based Knowledge Management. In Encyclopedia of E-Commerce, E-Government and Mobile Commerce. Idea Group Inc., 7p.

ZHANG S, BODENREIDER O. (2006). NLM anatomical ontology alignment system: Results of the 2006 ontology alignment contest. In Proc. of the Ontology Alignment Evaluation Initiative 2006 Campaign (OAEI 2006), pp. 145-156

ANNEXE B : FORMULAIRE DE CONSENTEMENT DES SUJETS

Nous vous invitons à bien lire ce formulaire et à poser des questions avant d'y apposer votre signature.

Titre de la recherche
Conception et validation d'une méthodologie pour la production ascendante et décentralisée d'ontologies sur le web sémantique.

Chercheur
Yan Bodain (yan.bodain@polymtl.ca)
Étudiant au doctorat
École Polytechnique de Montréal

Procédures
Nous vous demandons d'être disponible pour 2 séances d'une durée de 2 à 3 heures chacune.

Vous devrez concevoir une section de cours Web en :

1. Modifiant une section de cours pour la rendre conforme à votre compréhension du domaine.
2. Associer ces contenus de cours à des catégories (ontologies) distinctes.

Avantages et bénéfices

1. Pour votre participation, vous recevrez un montant de 15$/ heure.
2. Sur simple demande, nous vous transmettrons les résultats généraux de cette recherche, une fois l'étude terminée.

Risques et inconvénients
Chaque séance sera enregistrée. L'enregistrement vidéo permettra au chercheur d'analyser finement les détails de la tâche réalisée.

Clause de responsabilité
Le contenu des enregistrements vidéo restera confidentiel. Aucun extrait ne sera rendu public sans votre consentement préalable.

Votre identité sera aussi préservée au moment de la diffusion publique des résultats de recherche.

Confidentialité
Pour des fins de contrôle du projet de recherche, votre dossier pourra être consulté par une personne mandatée par le Comité d'éthique de l'École Polytechnique de Montréal ainsi que par des représentants du jury de thèse de doctorat. Tous adhèrent à une politique de stricte confidentialité.

Liberté de participation et liberté de retrait de l'étude :
Votre participation à cette étude est tout à fait volontaire. Vous êtes donc libre d'accepter ou de refuser d'y participer et vous pouvez vous retirer de l'étude en tout temps.

Personnes-ressources
Vous pouvez communiquer avec le Bureau de l'ombudsman de l'Université de Montréal pour obtenir des renseignements éthiques ou faire part d'un incident ou formuler des plaintes ou des commentaires au (514) 343-2100.

Adhésion au projet et signatures :
J'ai lu et compris le contenu du présent formulaire. Je certifie qu'on me l'a expliqué verbalement. J'ai eu l'occasion de poser toutes les questions concernant ce projet de recherche et on y a répondu à ma satisfaction.

Je certifie qu'on m'a laissé le temps voulu pour réfléchir et prendre ma décision. Je sais que je pourrai me retirer en tout temps.

Je soussigné(e) accepte de participer à cette étude.

Nom du participant	Signature du participant	Date

Je certifie a) avoir expliqué au signataire les termes du présent formulaire de consentement; b) lui avoir clairement indiqué qu'il reste à tout moment libre de mettre un terme à sa participation au présent projet et que je lui remettrai une copie signée du présent formulaire.

Nom du Chercheur	Signature du chercheur	Date

Informations de type administratif :

- Une copie signée sera remise au participant
- L'original du formulaire sera conservé au département MAGI de l'école Polytechnique de Montréal;
- Le projet de recherche et le présent formulaire de consentement ont été approuvés par le Comité d'éthique de la recherche avec des sujets humains (date à venir) :
  No de référence : CER-06/07-03
- Date de la version du présent formulaire : version 1 (17 octobre 2006).

ANNEXE C : QUESTIONNAIRE

Questionnaire sur l'utilisation d'un logiciel d'annotation pour le Web sémantique.

Sujet no.: _____ Date : _____

Sexe : **F** ☐ **H** ☐

Âge : **10-19** ☐ **20-29** ☐ **30-39** ☐ **40-49** ☐ **50-59** ☐ **60-69** ☐

Scolarité : **DEC** ☐ **BA** ☐ **M.A** ☐ **M. Ing** ☐ **M. Sc.A.** ☐ **PH.D** ☐ Autre : _____

Familiarité à l'environnement graphique Windows: --- ☐ -- ☐ - ☐ 0 ☐ + ☐ ++ ☐ +++ ☐

1. L'organisation de l'information est claire.	Pas du tout	--- ☐ -- ☐ - ☐ 0 ☐ + ☐ ++ ☐ +++ ☐	Absolument
2. L'information disponible contribue à soutenir ma tâche.	Pas du tout	--- ☐ -- ☐ - ☐ 0 ☐ + ☐ ++ ☐ +++ ☐	Absolument
3. Ce système est simple à utiliser.	Pas du tout	--- ☐ -- ☐ - ☐ 0 ☐ + ☐ ++ ☐ +++ ☐	Absolument
4. L'interface de ce système est agréable.	Pas du tout	--- ☐ -- ☐ - ☐ 0 ☐ + ☐ ++ ☐ +++ ☐	Absolument
5. Je suis satisfait de la facilité d'utilisation de ce système.	Pas du tout	--- ☐ -- ☐ - ☐ 0 ☐ + ☐ ++ ☐ +++ ☐	Absolument

		---	--	-	0	+	++	+++	
6. J'ai eu beaucoup de facilité à apprendre comment utiliser ce système.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
7. L'utilisation de ce système demande peu d'effort.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
8. J'ai été en mesure de produire facilement des pages Web.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
9. J'ai été en mesure d'assigner facilement des annotations aux pages Web.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
10. J'ai été en mesure de trouver facilement des ontologies sur le Web.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
11. J'ai été en mesure de récupérer facilement des ontologies sur le Web.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
12. J'ai été en mesure de modifier facilement la structure des ontologies.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
13. J'ai été en mesure de construire rapidement ma propre ontologie.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
14. Je me sens à l'aise avec ce système.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
15. Ce système est agréable à utiliser.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
16. J'ai été en mesure de compléter rapidement ma tâche avec ce système.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
17. Je crois que ce système améliore la qualité de mon travail.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument

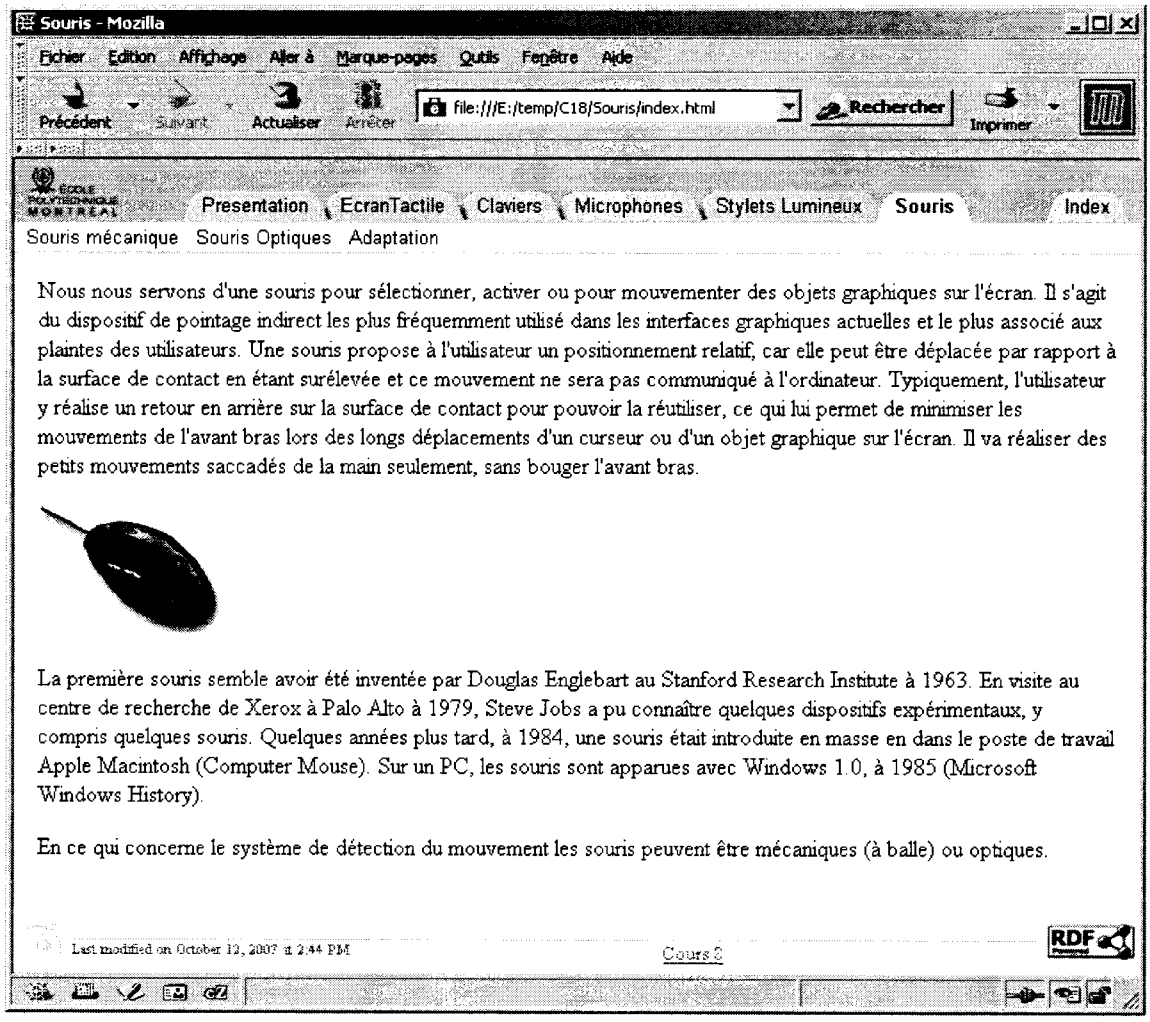
		---	--	-	0	+	++	+++	
18. Je crois que ce système améliore ma productivité (rapport en quantité produite et coût mis en œuvre).	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
19. Je crois que ce système améliore mon efficacité (qualité de produire des résultats utiles).	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument
20. Je crois que ce système simplifie mon travail.	Pas du tout	☐	☐	☐	☐	☐	☐	☐	Absolument

Suggestion : _____

ANNEXE D : CONTENU DE COURS

Les contenus de cours utilisés par les sujets se composaient de 53 pages différentes traitant sur le sujet des interfaces humain-ordinateur. Ce matériel de cours nous a été prêté par Walter Cybis et ce matériel ne peut être reproduit sans la permission explicite de l'auteur.

L'expérience que nous avons menée peut néanmoins être reproduite avec n'importe quel type de contenu. Par souci de transparence, nous présentons cinq pages écrans de ce matériel de cours.



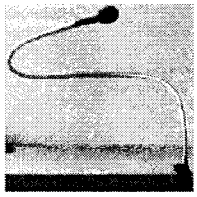
Adaptation à l'environnement - Mozilla

Fichier Edition Affichage Aller à Marque-pages Outils Fenêtre Aide

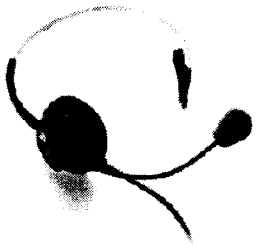
Précédent Suivant Actualiser Arrêter file:///E:/temp/C18/Microphones/Adaptation_a_l_environnem Rechercher Imprimer

L'utilisateur dans un environnement de travail informatique est souvent assis et a les mains occupé par des dispositifs d'entrée diverses, les plus souvent le clavier et la souris. Il a besoin d'avoir les mains libres et dans ce cas, il y a deux solutions ; les microphones avec support ou les microphones intégrés aux casques d'écoute.

Les microphones sur pied ou avec support sont souvent fournis avec les systèmes informatisés. Ces dispositifs sont faciles à utiliser et n'encombrent pas les utilisateurs. On remarque que du fait qu'ils sont fixes, la performance obtenue variera selon que l'utilisateur gardera la même orientation et maintiendra la même distance vis à vis le microphone. Des solutions alternatives sont des microphones plus sensibles et aux supports flexibles, comme ce da la figure en bas à droite. (36 Gooseneck microphone).



Due à leurs petites dimensions, les microphones intégrés à des casques d'écoute peuvent présenter des qualités acoustiques moindres. Cependant, comme l'orientation et la position de ce microphone par rapport à l'utilisateur est relativement fixe (en autant que le casque d'écoute soit porté de la même façon), ce type de dispositif permet généralement d'obtenir une meilleure performance en communication et reconnaissance vocale. Il faut noter que les utilisateurs sont parfois, rébarbatifs à l'emploi d'un casque d'écoute.



Des solutions sans fil, comme celle présenté ci-dessous (Sennheiser Wireless Headset Microphone), pourraient favoriser la mobilité de l'utilisateur dans son espace de travail déjà encombré par des câbles du clavier et souris.

Chargé

Caractéristiques générales - Mozilla

Fichier Edition Affichage Aller à Marque-pages Outils Fenêtre Aide

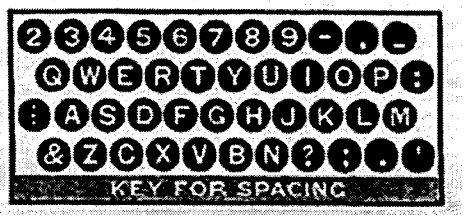
Précédent Suivant Actualiser Arrêter file:///E:/temp/C18/Claviers/Caracteristiques_ger Rechercher Imprimer

Presentation | EcranTactile | Claviers | Microphones | Stylets Lumineux | Souris | Index

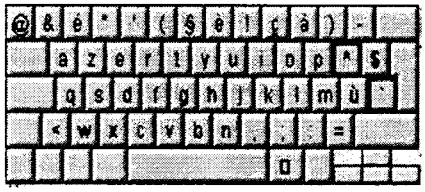
Caractéristiques générales Adéquation tâche Adaptation utilisateur Adaptation environnement

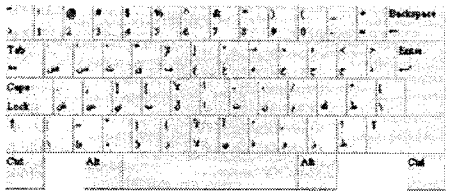
Claviers numériques Touches fonctions Déplacement curseur

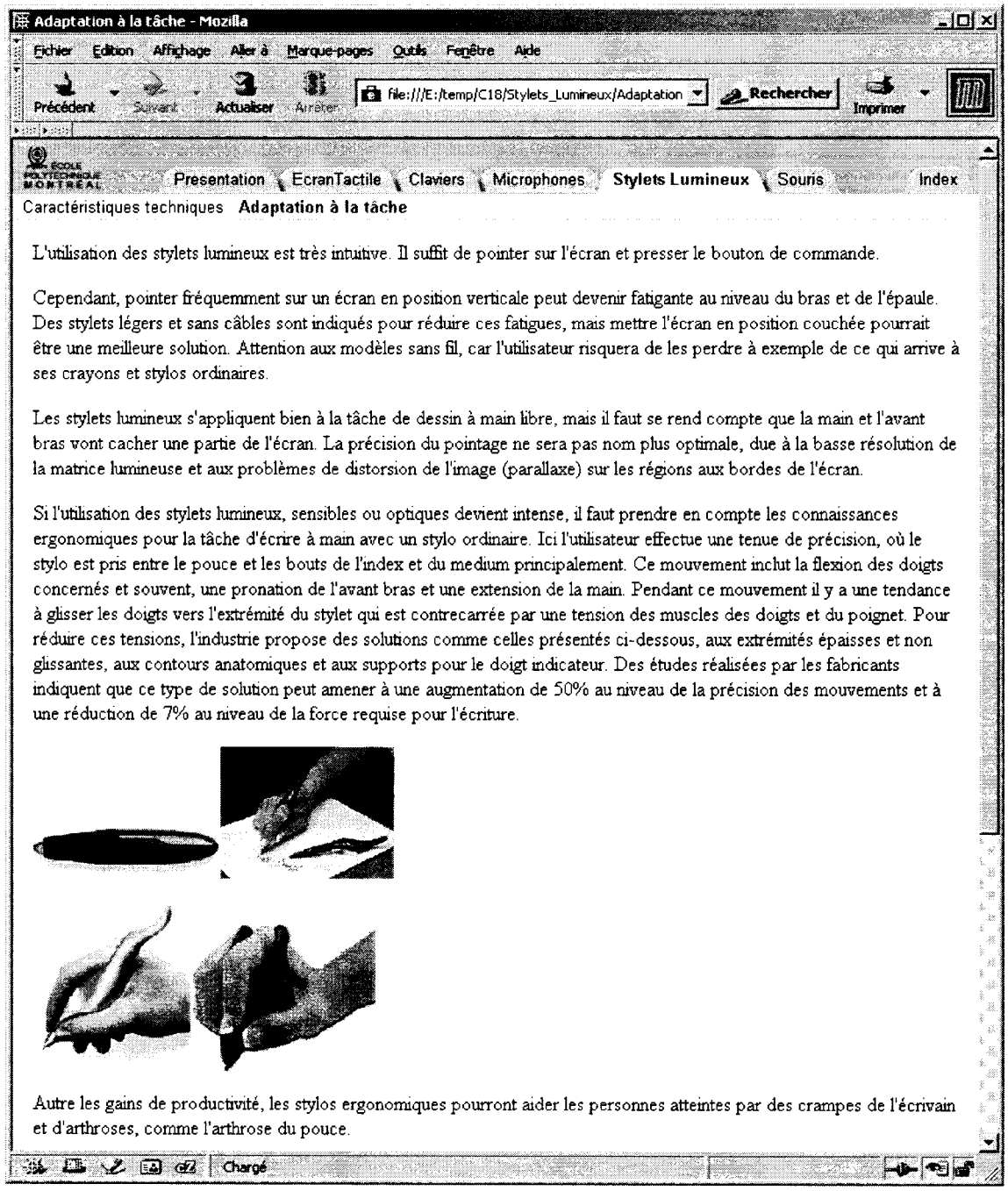
Le premier clavier commercial a été inventé par Christopher Latham Sholes en 1873, qui présentait déjà la disposition relative de touches QWERTY. Celle-la n'a pas été choisie en fonction de la commodité de l'utilisateur face à sa tâche, mais plutôt en raison du fonctionnement du système de barres de frappe dans les machines à écrire mécaniques. Des anecdotes associent l'origine de cette configuration au besoin de changer la disposition des touches des prototypes de façon à ralentir la vitesse de frappes des opérateurs, ce qui bloquait les barres mécaniques. Il y a aussi une version des faits, par laquelle cette disposition visait à favoriser l'écriture du mot TYPEWRITER, ce qui était important lors des démonstrations de vente de machine à taper. Voir un peut plus de ces histoires sur le site Web MYTHS about QWERTY.

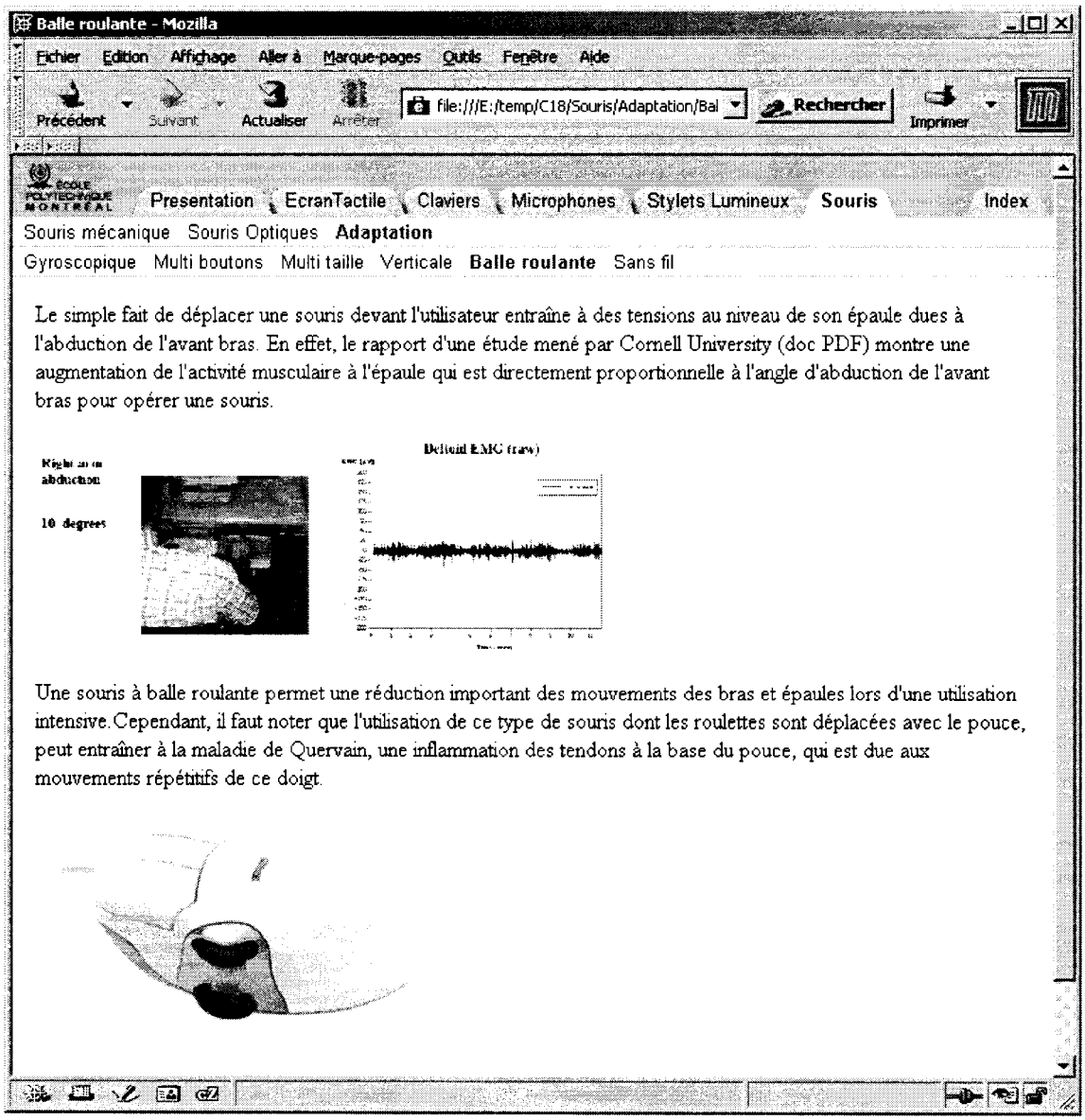


Le nom QWERTY désigne les lettres de la première rangée supérieure de touches alphabétiques, en partant de la gauche sur le clavier. Cette disposition est utilisé en Amérique du Nord et présente des déclinaisons régionales comme AZERTY (à droite), utilisée en France et QWERTZU, utilisée en Allemagne. Le site Web de WikiPedia présente d'autres dispositions de clavier utilisées dans le monde, comme le clavier arabe (à droite).






Adaptation à la tâche - Mozilla
Fichier Edition Affichage Aller à Marque-pages Outils Fenêtre Aide
Précédent Suivant Actualiser Arrêter
file:///E:/temp/C18/Stylets_Lumineux/Adaptation
Rechercher
Imprimer
Presentation EcranTactile Claviers Microphones Stylets Lumineux Souris Index
Caractéristiques techniques Adaptation à la tâche
L'utilisation des stylets lumineux est très intuitive. Il suffit de pointer sur l'écran et presser le bouton de commande.
Cependant, pointer fréquemment sur un écran en position verticale peut devenir fatigante au niveau du bras et de l'épaule. Des stylets légers et sans câbles sont indiqués pour réduire ces fatigues, mais mettre l'écran en position couchée pourrait être une meilleure solution. Attention aux modèles sans fil, car l'utilisateur risquera de les perdre à exemple de ce qui arrive à ses crayons et stylos ordinaires.
Les stylets lumineux s'appliquent bien à la tâche de dessin à main libre, mais il faut se rend compte que la main et l'avant bras vont cacher une partie de l'écran. La précision du pointage ne sera pas nom plus optimale, due à la basse résolution de la matrice lumineuse et aux problèmes de distorsion de l'image (parallaxe) sur les régions aux bordes de l'écran.
Si l'utilisation des stylets lumineux, sensibles ou optiques devient intense, il faut prendre en compte les connaissances ergonomiques pour la tâche d'écrire à main avec un stylo ordinaire. Ici l'utilisateur effectue une tenue de précision, où le stylo est pris entre le pouce et les bouts de l'index et du medium principalement. Ce mouvement inclut la flexion des doigts concernés et souvent, une pronation de l'avant bras et une extension de la main. Pendant ce mouvement il y a une tendance à glisser les doigts vers l'extrémité du stylet qui est contrecarrée par une tension des muscles des doigts et du poignet. Pour réduire ces tensions, l'industrie propose des solutions comme celles présentés ci-dessous, aux extrémités épaisses et non glissantes, aux contours anatomiques et aux supports pour le doigt indicateur. Des études réalisées par les fabricants indiquent que ce type de solution peut amener à une augmentation de 50% au niveau de la précision des mouvements et à une réduction de 7% au niveau de la force requise pour l'écriture.
Autre les gains de productivité, les stylos ergonomiques pourront aider les personnes atteintes par des crampes de l'écrivain et d'arthroses, comme l'arthrose du pouce.
Chargé


Balle roulante - Mozilla
Fichier Edition Affichage Aller à Marque-pages Outils Fenêtre Aide
Précédent Suivant Actualiser Arrêter
file:///E:/temp/C18/Souris/Adaptation/Bal
Rechercher
Imprimer
Presentation EcranTactile Claviers Microphones Stylets Lumineux Souris Index
Souris mécanique Souris Optiques Adaptation
Gyroscopique Multi boutons Multi taille Verticale Balle roulante Sans fil
Le simple fait de déplacer une souris devant l'utilisateur entraîne à des tensions au niveau de son épaule dues à l'abduction de l'avant bras. En effet, le rapport d'une étude mené par Cornell University (doc PDF) montre une augmentation de l'activité musculaire à l'épaule qui est directement proportionnelle à l'angle d'abduction de l'avant bras pour opérer une souris.
10 degrees
Une souris à balle roulante permet une réduction important des mouvements des bras et épaules lors d'une utilisation intensive. Cependant, il faut noter que l'utilisation de ce type de souris dont les roulettes sont déplacées avec le pouce, peut entraîner à la maladie de Quervain, une inflammation des tendons à la base du pouce, qui est due aux mouvements répétitifs de ce doigt.

ANNEXE E : ONTOLOGIES

Liste complète des noms de classes figurant dans les différentes ontologies.

Claviers.owl

- Thing
  - Alphabet
    - Braille
  - Appareil
    - Calculatrice
    - Telephone
    - Television
    - UrneÉlectronique
  - ApplicationLogicielle
    - ControleTraficAerien
      - Stars
    - UrneÉlectronique
  - ApplicationLogicielle
    - ControleTraficAerien
      - Stars
    - UrneÉlectronique
  - DispositifEntree
    - BatonCommande
    - BlocEffleurement
    - Clavier
      - Appareil
        - ANP
        - Ordinateur
          - OrdinateurPersonnel
          - OrdinateurPortable
        - Telephones
          - Portable
      - ClavierAMembrane
      - ClavierSansBlocNumerique
      - ClavierTouchesFonction
        - ReglesErgonomiques
        - ToucheFonction
          - ToucheFonctionProgrammable
      - ClavierVirtuel
        - AireSaisie
        - Performance
          - VitesseFrappe
      - DispositionClavier
        - AlphaNumerique
          - ABCDE
          - AZERTY
          - ClavierArabe

- QWERTY
- QWERTZU
- Numerique
  - Calculatrice
  - ClavierOrdinateur
  - ClavierTelephone
- Optimise
  - Accords
    - Avantages
  - DVORAK
    - Avantages
  - DataHandErgonomicKeyb
    - Avantages

- Genre
  - Accords
    - Avantages
  - Bouts_des_Doigts
  - Contours
  - Courbe
  - Surfaces_Sensibles
  - Vertical
- Histoire
  - MachineEcrire
  - Teletypes
- Illumine
- MicroClavier
  - AUneMain
  - EtuiClavier
  - Pliable
- Touches
  - Dimensions
  - Forme
  - Lettres
  - Retroaction
  - Texture
  - TouchesFonction
    - Deplacement
    - FonctionCritique
    - FonctionFrequente
- TouchesFonction
  - ReglesErgonomiques
  - ToucheFonction
    - ToucheFonctionProgrammable

- Souris
- Stylo

- DispositifSortie
  - Ecran
    - Classique
    - InterfaceUtilisateur
      - InterfaceGraphique
        - Bouton
        - ChampSaisie

        - Fenetre
        - Formulaire
    - Tactile
- Environnement
  - Contrainte
    - Bruit
    - Eclairage
    - Espace
    - Humide
    - Humidite
    - Poussiere
  - Types
    - Aeroports
    - SallesControle
    - Vehicule
- Evenement
  - SituationUrgence
    - IncidentCritique
- Langue
  - Anglaise
- Operations
  - Commande
  - Navigation
  - Selection
- Pays
  - Bresil
- Personnes
  - Aveugles
  - ChristohperLathamSholes
  - Concepteur
  - Utilisateur
    - ChargeTravail
      - Consequences
        - Deconfort
        - Troubles
          - SyndromeCanalCarpien
          - TensionsStatiques
      - Mouvements
        - DeviationPoignets
        - Extension
        - Inclinaison
        - Pronation
      - RepartitionDuTravail
        - Main
          - Droite
          - Gauche
    - Exclustechnologiques
    - Exclutechno
    - Expert
    - Formation
    - MemoireTravail
    - Novice

        - Qualification
            - Alphabetise
            - Dactylographie
        - UtilisateurExpert
        - UtilisateurNovice
- Reference
    - AdrianGerold
    - AviationToday
    - Buxton
    - CakirHartStewart
    - CardMackinlayRobertson
    - Card
    - DeafAndBlind
    - Dix
    - ErgonomicResources
    - Fingerworks
    - FolleyWallaceChan
    - Hinkey
    - JonMaxwell
    - LiebowitzMargolis
        - TypingErrors
    - Microsoft
    - Norme
        - ISO9241
    - Schneiderman
    - UseIt
        - AlertBox
    - VisorCentral
    - Wikipedia
- Tache
    - Mouvements
        - DeviationPoignets
        - Extension
        - Inclinaison
        - Pronation
    - Saisie
        - SaisieDigramme
        - SaisieDonnees
        - SaisieTexte
- Technologie
    - WebTv
- Utilisation
    - Performance
        - VitesseFrappe

Dispositif.owl

Thing
- User
- Interface
  - InterfaceHumainOrdinateur
    - Dispositif_entrees
      - Classification
        - Curseur
          - Correspondance
            - Complexe
            - Simple
          - Mouvement
            - Continu
            - Discret
          - Type
            - Ecran
            - Texte
        - Dimensionnalite
          - EspaceVolume
          - Lineaire
          - Lineaire
          - PlanAire
          - RotationTorque
        - Intermediaire
          - Direct
          - Indirect
        - Niveau_abstraction
          - Entree_donnees
          - Pointage
            - Positionnement
              - Absolu
              - Relatif
            - Systeme_coordonnees
        - Relatif_utilisateur
          - Intention
          - Sens
        - SensUtilisateur
      - Ecran_tactile
        - Adequation
          - Tache
          - Technologie
          - Utilisateur
            - Droitier
            - Gaucher
            - Taille_doigt
        - Domaine_Application
        - Domaines_Applications
        - Technologies
          - Accoustique
          - Capacitance

- Infrarouges
- Resistif

- Etudes
- Taxonomie
  - Dispositif_pointage
    - Pointage_Direct
    - Pointage_Indirect
      - Indirect_Mouvement
      - Indirect_Positionnement
    - Pointage_Semi_Direct

Souris.owl

- Thing
  - DispositifEntree
    - Souris
      - Action
      - Histoire
        - Inventeur
      - Interface
        - Adaptation
          - Balle_roulante
            - Inconvénient
          - Gyroscopique
          - Multibouton
          - Multitaille
          - Sans_fil
          - Verticale
        - Adapté
          - Balle_roulante
            - Inconvénient
          - Gyroscopique
          - Multibouton
          - Multitaille
          - Sans_fil
          - Verticale
        - AvecFil
          - BalleRoulante
            - Avantages
            - Inconvenient
          - Verticale
            - Avantages
        - SansFil
          - Avantages
          - Fabricant
          - Technologie
      - Proprietes
        - Configurable
        - DetectionMouvement
          - Mecanique
          - Optique
        - Pointage
        - Positionement
  - Etudes
  - Utilisateur
    - Caracterisques

Stylets_lumineux.owl

- Thing
  - SolutionsErgonomique
    - ExtrémitésNonGlissantes
  - FatigueBrasÉpaules
    - Arthrose
    - BasseResolution
    - CrampeDesÉrivains
      - MauvaisesHabitudesGestuelles
    - Extension
    - Flexion
    - Fluxion
    - Pronation
    - Solutions
      - SolutionsErgonomie
        - Avantages
        - ContoursAnatomiques
        - ExtrémitésÉpaisses
        - Support_Doigt
  - Mouvement
    - Actionner
    - Deplacer
    - Déplacer
    - Sélectionner
  - Technologies
    - Ecran
      - BasseResolution
      - BasseRésolution
      - EcranCache
      - EcranCouche
      - EcranCouché
      - EcranOrdinaire
      - EcranVertical
        - FatigueBrasEpaules
      - Grille
      - Parallaxe
      - ReperesInfrarouges
      - ÉcranCaché
      - ÉcranCouché
      - ÉcranOrdinaire
      - ÉcranVertical
        - FatigueBrasEpaules
    - Logiciel
      - LogicielOptiques
      - LogicielSensible
    - Stylets
      - BoutonDeCommande
      - Fabricants
      - Intuitivite
      - StyletsLegers
      - StyletsLumineux

        - CellulePhotosensible
    - StyletsLégers
    - StyletsOptiques
        - LecteursDeCodesBarre
        - Scanners
    - StyletsSansCables
    - StyletsSensibles

Microphone.owl

Thing
- Microphones
  - Adequation
    - Environnement
    - Tâche
    - Utilisateur
      - Intégrés
      - Sans_fil
      - À_support
    - Utilisateur
      - Intégrés
      - Sans_fil
      - À_support
    - Utilisateurs
  - Domaines_Application
    - Communication
    - Entrée_Commandes
    - Entrée_texte
    - Reconnaissance_paroles
  - Technologie
    - Fréquences
      - Anti-bruits
      - Distorsion
      - Intervalles_de_fonctionnement
    - Impédance
      - Faible
      - Haute
      - Moyenne
    - MIcros_a_Condensateurs
    - Micros_Dynamiques
    - Propriété_DIrectionnelles
      - Bidirectionnel
      - Omnidirectionnel
      - Unidirectionnel
    - Reponse_Fréquences
      - Anti-bruits
      - Distorsion
      - Intervalles_de_fonctionnement
  - Touches_controle
- Reference

ANNEXE F : ANNOGRAMME

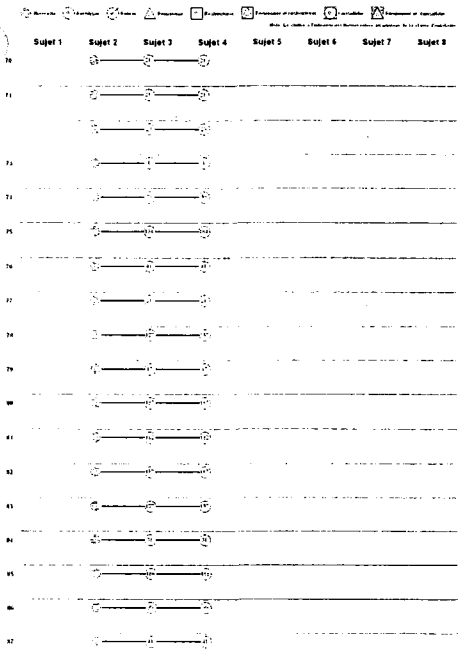
Table de correspondance des numéros de classes représentés sur l'annogramme.

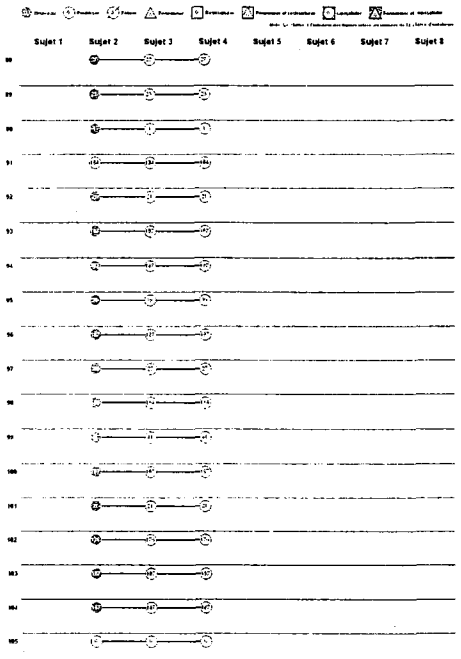
1	dispositif.owl#Tache
2	dispositif.owl#Taille_doigt
3	dispositif.owl#Environnement
4	dispositif.owl#Domaine
5	dispositif.owl#Entretien
6	dispositif.owl#Infrarouges
7	dispositif.owl#Positionnement
8	dispositif.owl#Complexe
9	dispositif.owl#Systeme_coordonnees
10	dispositif.owl#Classification
11	dispositif.owl#Pointage
12	dispositif.owl#Direct
13	dispositif.owl#Type
14	dispositif.owl#Mouvement
15	dispositif.owl#Pointage_Indirect
16	dispositif.owl#Espace
17	dispositif.owl#Correspondance
18	dispositif.owl#Intention
19	dispositif.owl#Continu
20	dispositif.owl#Texte
21	dispositif.owl#Dispositif_entrees
22	dispositif.owl#Intermediaire
23	dispositif.owl#Pointage_Direct
24	dispositif.owl#Technologies
25	dispositif.owl#Indirect
26	dispositif.owl#Ecran
27	dispositif.owl#Indirect_Positionnement
28	dispositif.owl#Tâche
29	dispositif.owl#Droitier
30	dispositif.owl#Simple
31	dispositif.owl#Capacitance
32	dispositif.owl#Entree_donnees
33	dispositif.owl#Acoustique
34	dispositif.owl#Discret
35	dispositif.owl#Absolu
36	dispositif.owl#Frequence
37	dispositif.owl#Gaucher
38	dispositif.owl#Dispositif_pointage
39	dispositif.owl#Pointage_Semi_Direct
40	dispositif.owl#Indirect_Mouvement
41	dispositif.owl#Taxonomie
42	dispositif.owl#Resistivite
43	dispositif.owl#SensUtilisateur
44	dispositif.owl#SensUtilisateur_1
45	dispositif.owl#Pointage
46	claviers.owl#Forme_1
47	dispositif.owl#Dispositif_entrees
48	claviers.owl#Droite
49	dispositif.owl#Droitier
50	claviers.owl#ChargeTravail
51	dispositif.owl#Indirect_Positionnement
52	dispositif.owl#Ecran_tactile
53	truc.owl#Truc
54	dispositif.owl#EspaceVolume_1
55	dispositif.owl#Dispositif_pointage
56	dispositif.owl#Entree_donnees
57	dispositif.owl#Ecran
58	dispositif.owl#Domaines_Applications
59	dispositif.owl#Dispositifs_1
60	dispositif.owl#Discret
61	dispositif.owl#Direct
62	dispositif.owl#Dimensionnalite
63	dispositif.owl#Curseur
64	dispositif.owl#Correspondance
65	dispositif.owl#Continu
66	dispositif.owl#Complexe
67	dispositif.owl#Classification
68	dispositif.owl#Capacitance
69	dispositif.owl#Adequation
70	dispositif.owl#Absolu
71	dispositif.owl#Etudes
72	dispositif.owl#Gaucher
73	dispositif.owl#Indirect
74	dispositif.owl#Indirect_Mouvement
75	dispositif.owl#Infrarouges
76	dispositif.owl#Intention
77	dispositif.owl#Intermediaire
78	dispositif.owl#Lineaire
79	dispositif.owl#Mouvement
80	dispositif.owl#Mouvement_1
81	dispositif.owl#Niveau_abstraction
82	dispositif.owl#PlanAire
83	dispositif.owl#Pointage_Indirect
84	dispositif.owl#Pointage_Direct
85	dispositif.owl#Pointage_Semi_Direct
86	dispositif.owl#Position
87	dispositif.owl#Positionnement
88	dispositif.owl#Pression
89	dispositif.owl#Relatif
90	dispositif.owl#Relatif_utilisateur
91	dispositif.owl#Resistif
92	dispositif.owl#RotationTorque
93	dispositif.owl#Sens
94	dispositif.owl#Simple
95	dispositif.owl#Systeme_coordonnees
96	dispositif.owl#Tache
97	dispositif.owl#Taille_doigt
98	dispositif.owl#Taxonomie
99	dispositif.owl#Technologie
100	dispositif.owl#Technologies
101	dispositif.owl#Texte
102	dispositif.owl#Type
103	dispositif.owl#Utilisateur
104	souris.owl#Adaptation
105	souris.owl#Inconvénient
106	truc.owl#Sue
107	souris.owl#Verticale
108	souris.owl#Fonctionnement
109	souris.owl#DispositifEntree
110	souris.owl#Avantages_1
111	stylets_lumineux.owl#Avantages
112	claviers.owl#Fabricant
113	claviers.owl#Microphone
114	microphones.owl#Sans_fil
115	dispositif.owl#User
116	dispositif.owl#Accoustique
117	claviers.owl#Consequences
118	claviers.owl#Droitier

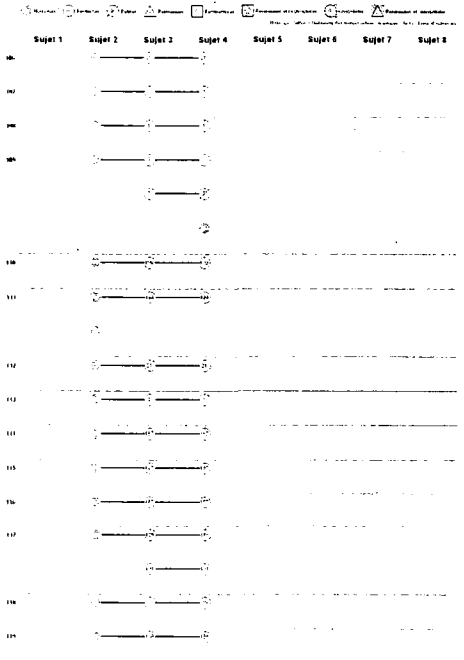
119 dispositif.owl#NatureDesEntrees
120 claviers.owl#Avantages
121 claviers.owl#CardMackinlayRobertson
122 claviers.owl#DVORAK
123 claviers.owl#Exclustechnologiques
124 claviers.owl#Espace
125 claviers.owl#Gauche
126 claviers.owl#Humidite
127 claviers.owl#Main
128 claviers.owl#Mouvements
129 claviers.owl#MemoireTravail
130 claviers.owl#Genre
131 claviers.owl#EtuiClavier
132 claviers.owl#Taille
133 claviers.owl#Contours
134 claviers.owl#EloTouchSystems
135 oijoji.owl#.lp.lp
136 SchemaDirectory
137 claviers.owl#ABCDE
138 claviers.owl#AZERTY
139 claviers.owl#AdrianGerold
140 claviers.owl#AlphaNumerique
141 claviers.owl#Appareil
142 claviers.owl#ApplicationLogicielle
143 claviers.owl#BatonCommande
144 claviers.owl#Bouton
145 claviers.owl#CakirHartStewart
146 claviers.owl#Calculatrice
147 claviers.owl#Calculatrice_1
148 claviers.owl#ChampSaisie
149 claviers.owl#ClavierArabe
150 claviers.owl#ClavierTelephone
151 claviers.owl#DispositifSortie
152 claviers.owl#Dispositif
153 claviers.owl#Eclairage
154 claviers.owl#Expert_1
155 claviers.owl#Fenetre
156 claviers.owl#Fonction
157 claviers.owl#Formulaire
158 claviers.owl#Histoire
159 claviers.owl#InterfaceGraphique
160 claviers.owl#InterfaceUtilisateur
161 claviers.owl#MachineEcrire
162 claviers.owl#Norme
163 claviers.owl#Numerique
164 claviers.owl#Ordinateur
165 claviers.owl#Personnes
166 claviers.owl#Portable
167 claviers.owl#QWERTY
168 claviers.owl#QWERTZU
169 claviers.owl#Telephones
170 claviers.owl#Teletypes
171 claviers.owl#ClavierTouchesFonction
172 claviers.owl#Novice
173 dispositif.owl#Utilisateur
174 dispositif.owl#Sens
175 dispositif.owl#Relatif_utilisateur
176 dispositif.owl#Niveau_abstraction
177 dispositif.owl#Écran_tactile
178 dispositif.owl#Curseur
179 dispositif.owl#Adéquation
180 dispositif.owl#Infrarouge
181 dispositif.owl#Domaine_Application
182 dispositif.owl#Accoustique
183 dispositif.owl#Relatif
184 dispositif.owl#Resistif
185 dispositif.owl#Dimensionnalite
186 dispositif.owl#Écran
187 dispositif.owl#Etudes
188 dispositif.owl#InterfaceHumainOrdinateur
189 dispositif.owl#Domaines_Applications
190 dispositif.owl#Ecran_tactile
191 dispositif.owl#Interface
192 claviers.owl#Clavier
193 claviers.owl#ClavierSansBlocNumerique
194 claviers.owl#Navigation
195 claviers.owl#Lettres
196 dispositif.owl#Sélectionner
197 claviers.owl#Selection
198 claviers.owl#Commande
199 claviers.owl#VitesseFrappe
200 claviers.owl#ClavierOrdinateur
201 claviers.owl#Deplacement
202 claviers.owl#FonctionCritique
203 claviers.owl#FonctionFrequente
204 claviers.owl#ToucheFonction
205 claviers.owl#TouchesFonction_1
206 claviers.owl#BlocEffleurement
207 claviers.owl#BlocEffleurement_1
208 claviers.owl#DispositifEntree
209 claviers.owl#Humide
210 claviers.owl#OrdinateurPersonnel
211 claviers.owl#AireSaisie
212 claviers.owl#Exclutechno
213 claviers.owl#SaisieDonnees
214 claviers.owl#Performance
215 claviers.owl#ToucheFonctionProgrammable
216 dispositif.owl#OndesAccoustiques
217 claviers.owl#ChristohperLathamSholes
218 claviers.owl#Alphabetise
219 claviers.owl#WebTv
220 claviers.owl#Utilisateur
221 claviers.owl#ControleTraficAerien
222 claviers.owl#Vehicule
223 claviers.owl#Poussiere
224 claviers.owl#VisorCentral
225 claviers.owl#Tache
226 claviers.owl#Telephone
227 claviers.owl#AviationToday
228 claviers.owl#Contrainte
229 claviers.owl#Stars
230 claviers.owl#SaisieTexte
231 claviers.owl#Concepteur
232 claviers.owl#SallesControle
233 claviers.owl#ClavierVirtuel
234 claviers.owl#ISO9241
235 claviers.owl#MicroClavier
236 claviers.owl#Formation
237 claviers.owl#Television
238 claviers.owl#Dactylographie
239 claviers.owl#Anglaise
240 claviers.owl#DispositionClavier
241 claviers.owl#IncidentCritique
242 claviers.owl#Aeroports
243 claviers.owl#Classique
244 claviers.owl#ANP
245 claviers.owl#Pliable
246 claviers.owl#TypingErrors
247 claviers.owl#UrneÉlectronique
248 claviers.owl#AlertBox
249 claviers.owl#SituationUrgence
250 claviers.owl#TouchesFonction

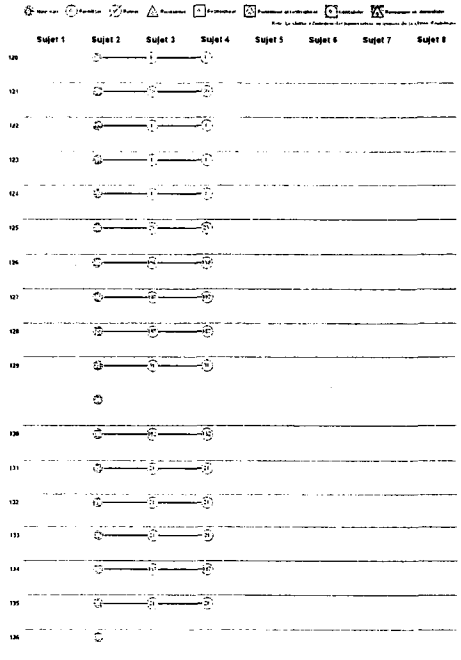
251 claviers.owl#Illumine
252 claviers.owl#ApplicationLogicielle_1
253 claviers.owl#Tactile
254 claviers.owl#Stylo
255 claviers.owl#AUneMain
256 claviers.owl#Touches
257 claviers.owl#UtilisateurExpert
258 claviers.owl#ClavierAMembrane
259 claviers.owl#Microsoft
260 claviers.owl#OrdinateurPortable
261 claviers.owl#Bresil
262 claviers.owl#UtilisateurNovice
263 claviers.owl#LiebowitzMargolis
264 claviers.owl#Pays
265 claviers.owl#ReglesErgonomiques
266 claviers.owl#Deconfort
267 claviers.owl#Inconfort
268 claviers.owl#Reference
269 claviers.owl#Accords
270 claviers.owl#Surfaces_Sensibles
271 claviers.owl#SaisieDigramme
272 claviers.owl#Bouts_des_Doigts
273 claviers.owl#Extension_1
274 claviers.owl#Pronation
275 claviers.owl#Déviation
276 claviers.owl#Courbe
277 claviers.owl#Tensions_1
278 claviers.owl#Inclinaison
279 claviers.owl#Épaules
280 claviers.owl#CanalCarpien
281 Schema
282 terrorOnt.owl#Sport
283 terrorOnt.owl#Male
284 terrorOnt.owl#Athlete
285 association.owl#PersonProjectAssociation
286 oijoji.owl#SchemaDirectory
287 owlweb.rdf#programmingLanguage
288 owlweb.rdf#objectOrientedLanguage
289 owlweb.rdf#softwareStatusType
290 dispositif.owl#EntréeDeDonnées
291 dispositif.owl#Capacitifs
292 dispositif.owl#OndesAccoustiques
293 claviers.owl#Buxton
294 dispositif.owl#Dispositif
295 dispositif.owl#Interface
296 dispositif.owl#InterfaceHumainOrdinateur
297 dispositif.owl#SystèmeInformations
298 dispositif.owl#PanneauDeVerre
299 dispositif.owl#Résistifs
300 claviers.owl#Hinkey
301 claviers.owl#Dix
302 dispositif.owl#Pointage_1
303 dispositif.owl#Environnement
304 dispositif.owl#ServicesMédicaux
305 dispositif.owl#Entretien
306 dispositif.owl#Adéquation
307 dispositif.owl#DomaineApplication
308 dispositif.owl#Domaines_Application
309 dispositif.owl#Applications
310 dispositif.owl#DomainesApplication
311 dispositif.owl#Microphones
312 dispositif.owl#Utilisate
313 dispositif.owl#Utilisateurs
314 microphones.owl#Entrée_texte
315 microphones.owl#Aetget_1
316 microphones.owl#Reponse_Fréquences
317 microphones.owl#À_support
318 microphones.owl#Intégrés
319 microphones.owl#Utilisateur
320 stylets_lumineux.owl#CellulePhotosensible
321 stylets_lumineux.owl#Scanners
322 stylets_lumineux.owl#ÉcranVertical
323 stylets_lumineux.owl#Flexion
324 stylets_lumineux.owl#LogicielLumineux
325 stylets_lumineux.owl#ÉcranCaché
326 stylets_lumineux.owl#Materiel
327 stylets_lumineux.owl#MaterielLumineux
328 stylets_lumineux.owl#Conséquences
329 stylets_lumineux.owl#ReperesInfrarouges
330 stylets_lumineux.owl#SolutionsErgonomie
331 stylets_lumineux.owl#Technologies
332 stylets_lumineux.owl#ÉcranOrdinaire
333 microphones.owl#Haute
334 microphones.owl#Impédance
335 microphones.owl#Reconnaissance_paroles
336 microphones.owl#Microphones
337 microphones.owl#Faible
338 microphones.owl#Entrée_Commandes
339 stylets_lumineux.owl#Grille_1
340 stylets_lumineux.owl#Grille
341 stylets_lumineux.owl#Logiciel
342 microphones.owl#Aaa
343 stylets_lumineux.owl#Stylets
344 stylets_lumineux.owl#StyletsLumineux
345 microphones.owl#Moyenne
346 stylets_lumineux.owl#BoutonDeCommande
347 claviers.owl#Ecran
348 dispositif.owl#Syst`mesMilitaires
349 dispositif.owl#Acoustique
350 microphones.owl#Micros_Dynamiques
351 stylets_lumineux.owl#LecteursDeCodesBarre
352 microphones.owl#Anti-bruits
353 microphones.owl#Omnidirectionnel
354 stylets_lumineux.owl#LogicielSensible
355 microphones.owl#MIcros_a_Condensateurs
356 stylets_lumineux.owl#StyletsSensibles
357 microphones.owl#Intervalles_de_fonctionnement
358 microphones.owl#Communication
359 stylets_lumineux.owl#StyletsOptiques
360 microphones.owl#Distorsion
361 microphones.owl#Unidirectionnel
362 stylets_lumineux.owl#Sélectionner
363 stylets_lumineux.owl#Ecran
364 stylets_lumineux.owl#Actionner
365 microphones.owl#Fréquences
366 microphones.owl#Bidirectionnel
367 stylets_lumineux.owl#Déplacer
368 microphones.owl#Propriété_DIrectionnelles
369 stylets_lumineux.owl#FatigueBrasÉpaules
370 stylets_lumineux.owl#Extension
371 stylets_lumineux.owl#Arthrose
372 stylets_lumineux.owl#BasseRésolution
373 stylets_lumineux.owl#StyletsLégers
374 stylets_lumineux.owl#Parallaxe
375 stylets_lumineux.owl#ExtrémitésÉpaisses
376 stylets_lumineux.owl#StyletsSansCables
377 claviers.owl#Vertical
378 stylets_lumineux.owl#Intuitivité
379 stylets_lumineux.owl#ExtrémitésNonGlissantes
380 stylets_lumineux.owl#CrampeDesÉrivains
381 stylets_lumineux.owl#ÉcranCouché
382 stylets_lumineux.owl#ContoursAnatomiques

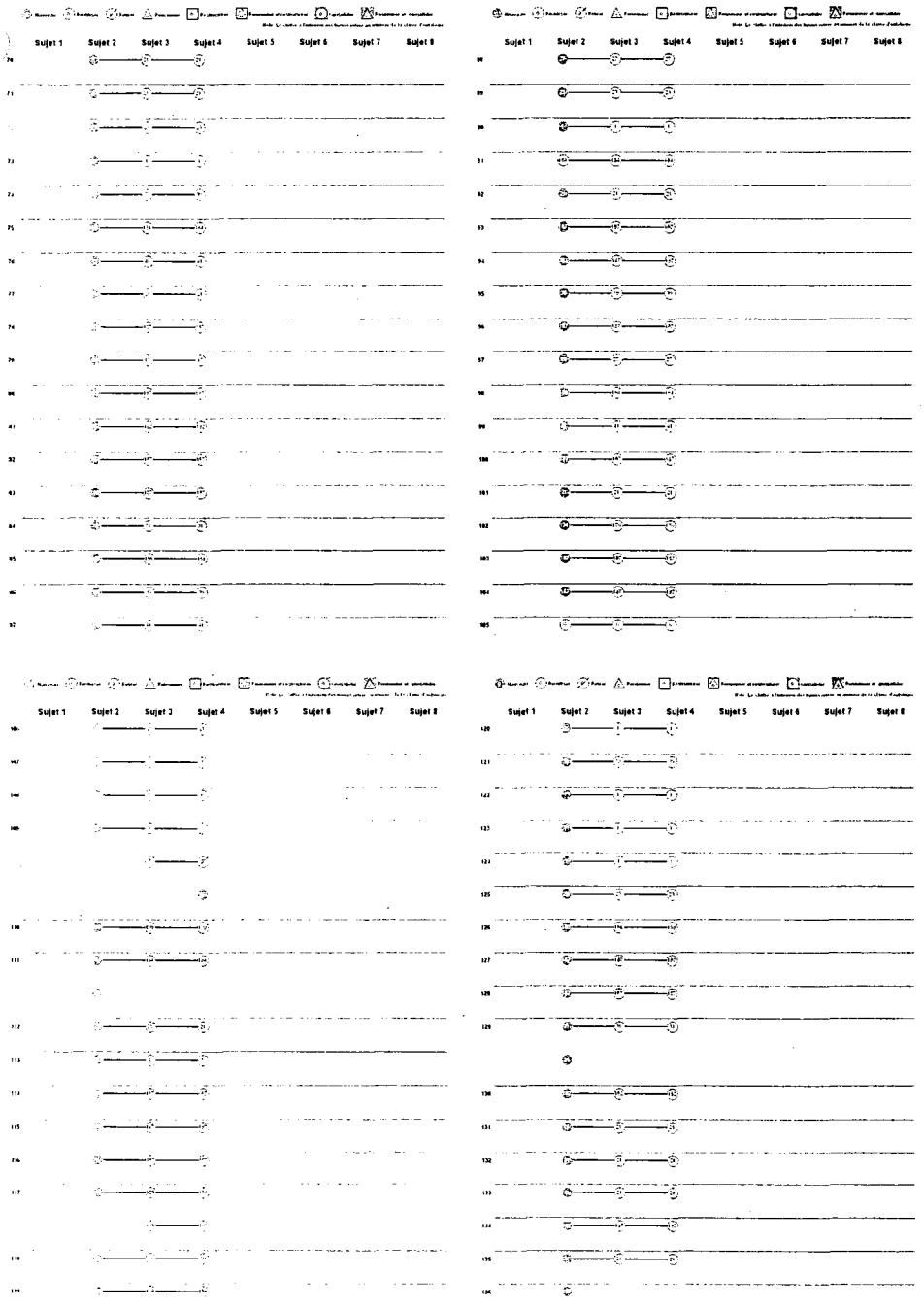
383 stylets_lumineux.owl#MauvaisesHabitudes Gestuelles
384 stylets_lumineux.owl#Pronation
385 dispositif.owl#Actionner
386 dispositif.owl#Déplacer
387 dispositif.owl#ConstructionCivile
388 dispositif.owl#PointsIsolants
389 dispositif.owl#FeuilleDEpolyester
390 microphones.owl#Aege
391 microphones.owl#Reconnaissance_Parole
392 claviers.owl#FolleyWallaceChan
393 claviers.owl#Schneiderman
394 dispositif.owl#Comportement
395 *dispositif.owl#Écrire*
396 dispositif.owl#Transactions
397 dispositif.owl#Lineaire_1
398 dispositif.owl#Informer
399 dispositif.owl#Principe
400 dispositif.owl#Positionner
401 dispositif.owl#Orienter
402 claviers.owl#Doigts
403 claviers.owl#Card_1
404 dispositif.owl#Dessiner
405 claviers.owl#DeafAndBlind
406 claviers.owl#TensionsStatiques
407 claviers.owl#Dimensions
408 claviers.owl#Fingerworks
409 claviers.owl#DeviationPoignets
410 claviers.owl#RepartitionDuTravail_1
411 claviers.owl#SyndromeCanalCarpien
412 claviers.owl#DataHandErgonomicKeyb
413 claviers.owl#JonMaxwell
414 claviers.owl#Extension
415 claviers.owl#Texture
416 claviers.owl#Braille
417 claviers.owl#Retroaction
418 claviers.owl#ErgonomicResources
419 claviers.owl#Aveugles
420 claviers.owl#Avantages_2
421 dispositif.owl#ActionsPossibles
422 claviers.owl#Souris
423 claviers.owl#Wikipedia
424 microphones.owl#Touches_controle
425 stylets_lumineux.owl#Fabricants
426 microphones.owl#Reference
427 stylets_lumineux.owl#Support_Doigt
428 souris.owl#Avantages
429 souris.owl#Multitaille
430 souris.owl#Caracterisques
431 souris.owl#Sans_fil
432 souris.owl#SansFil
433 souris.owl#Adapté
434 souris.owl#Optique
435 *souris.owl#Balle_roulante*
436 souris.owl#BalleRoulante
437 souris.owl#Action
438 souris.owl#Interface
439 souris.owl#DetectionMouvement
440 souris.owl#Technologie
441 souris.owl#Gyroscopique
442 souris.owl#Verticale_1
443 souris.owl#Avantages_2
444 souris.owl#Inventeur
445 souris.owl#Fabricant
446 souris.owl#Multibouton
447 souris.owl#Configurable
448 souris.owl#Etudes
449 souris.owl#Inconvenient
450 souris.owl#Mecanique
451 souris.owl#Pointage
452 souris.owl#Histoire
453 souris.owl#Souris
454 souris.owl#Positionement
455 claviers.owl#Operations
456 claviers.owl#Langue
457 claviers.owl#Troubles
458 claviers.owl#Avantages_1
459 claviers.owl#Qualification
460 claviers.owl#Bruit

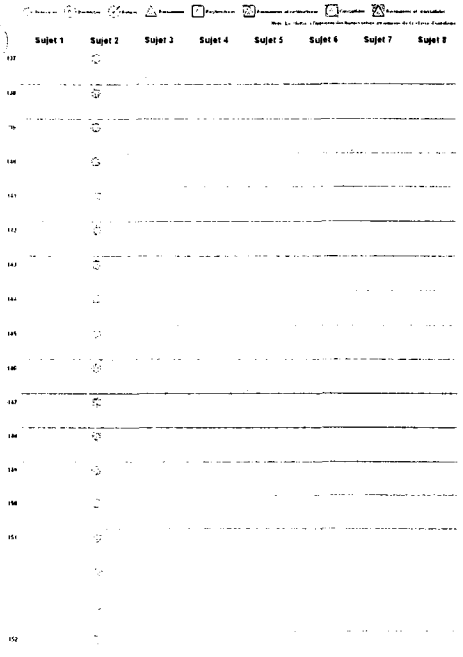

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

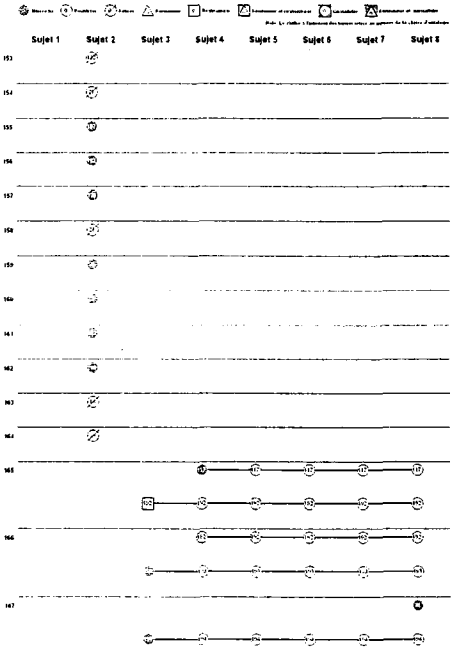

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

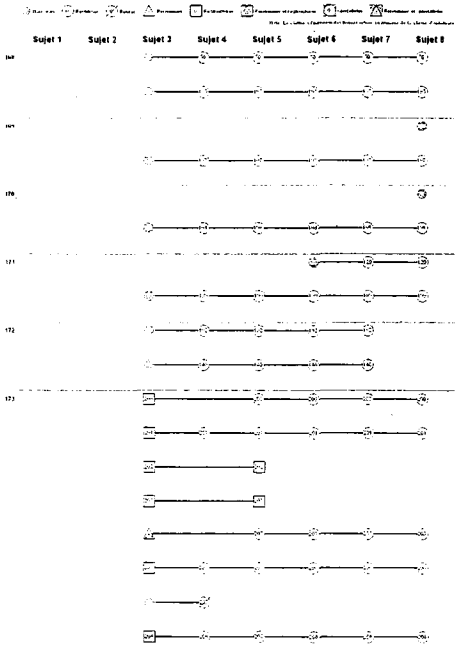

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

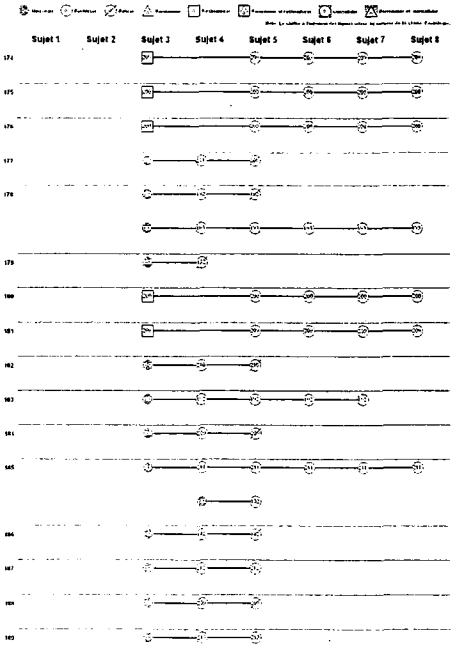

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

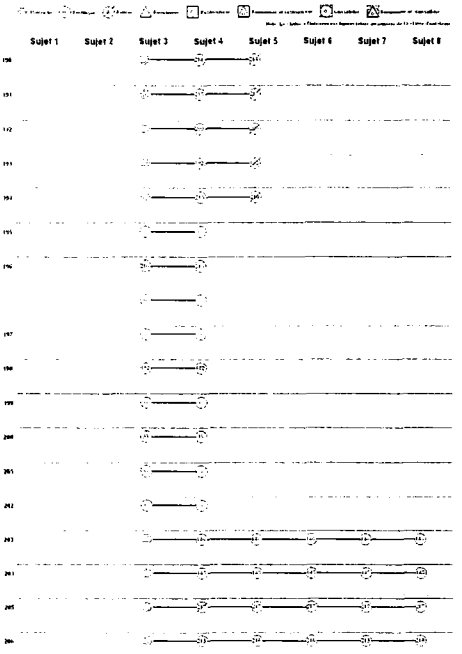

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

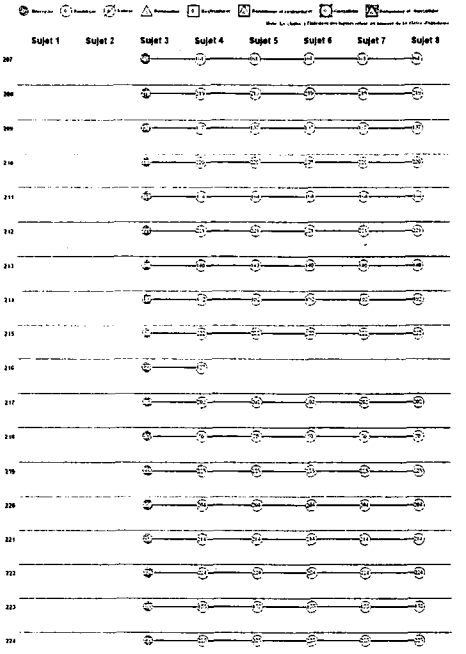

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

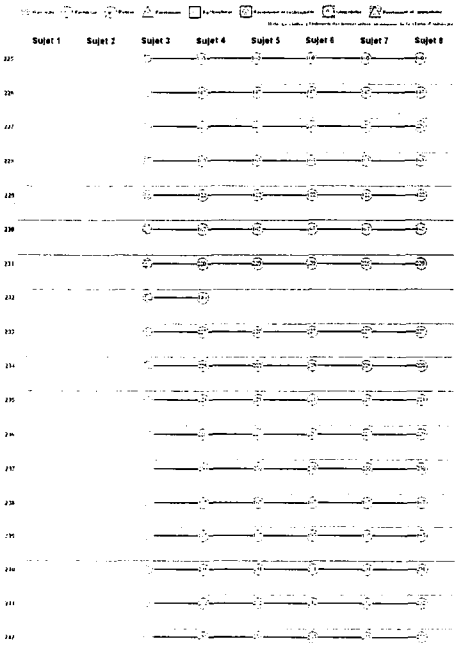

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

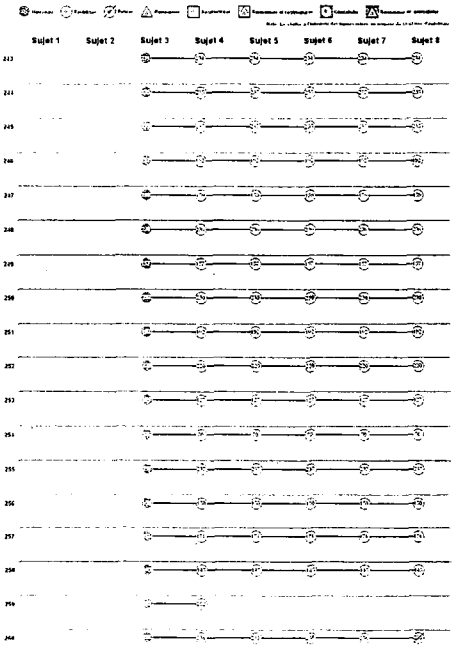

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

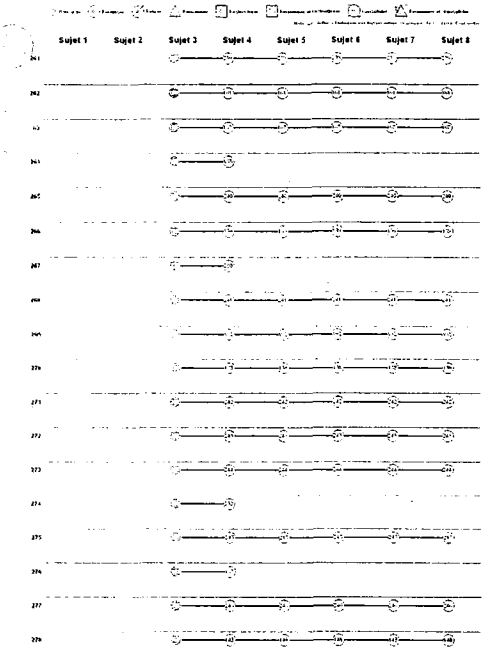

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

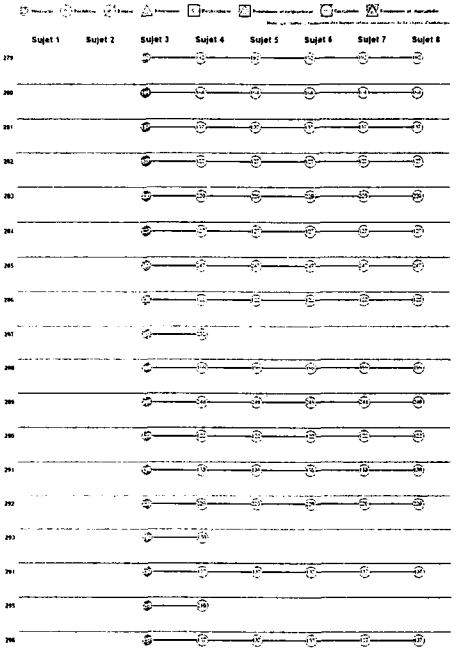

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

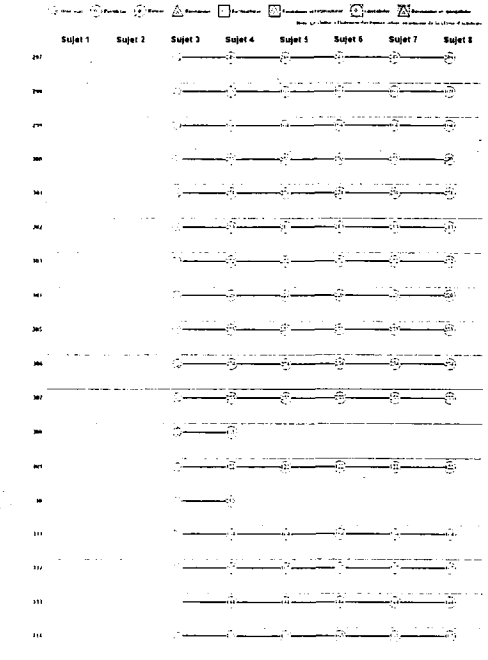

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

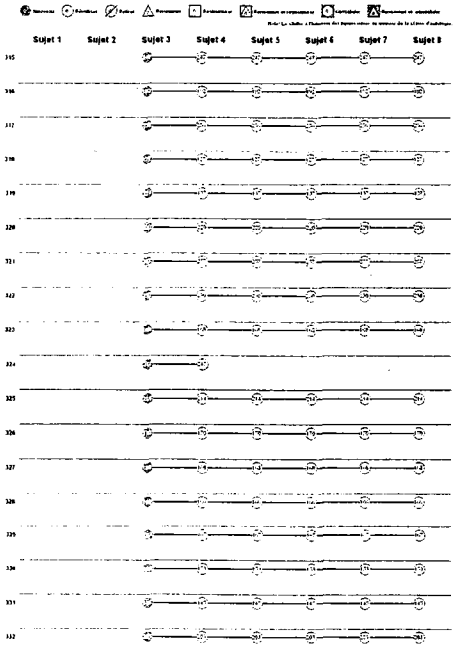

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

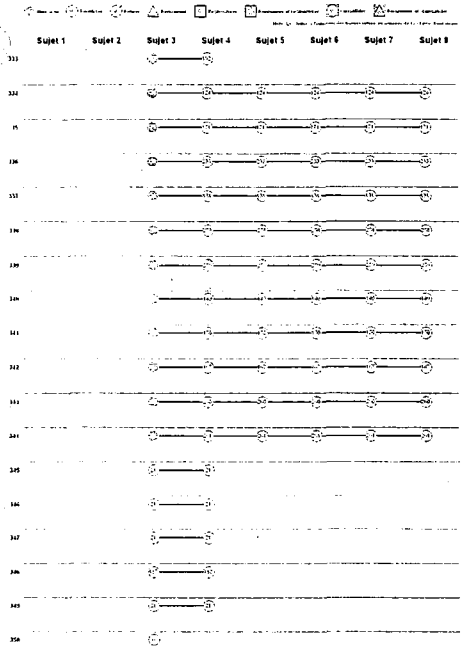

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

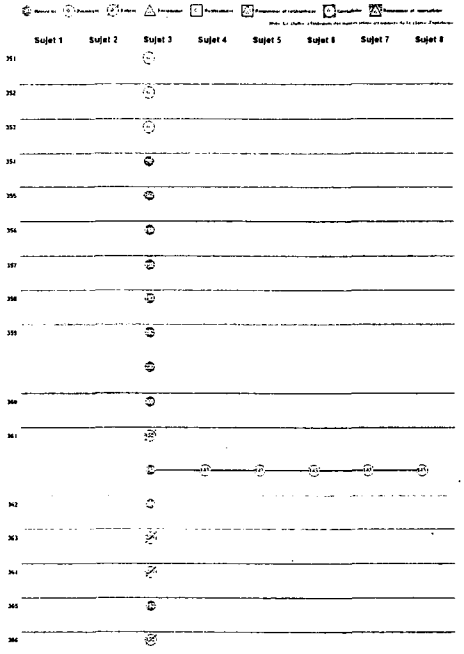

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

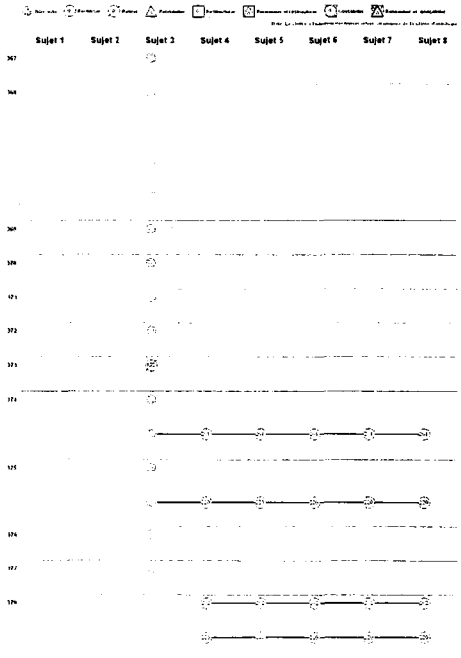

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

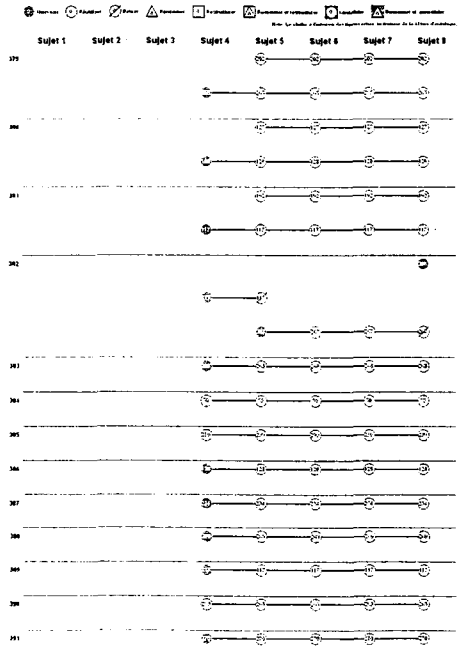

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

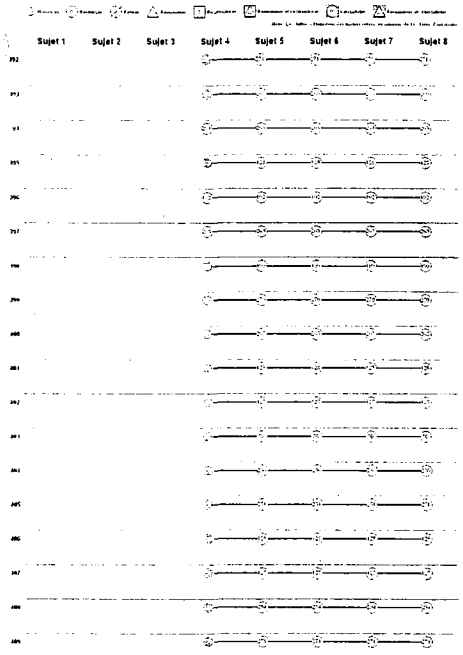

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

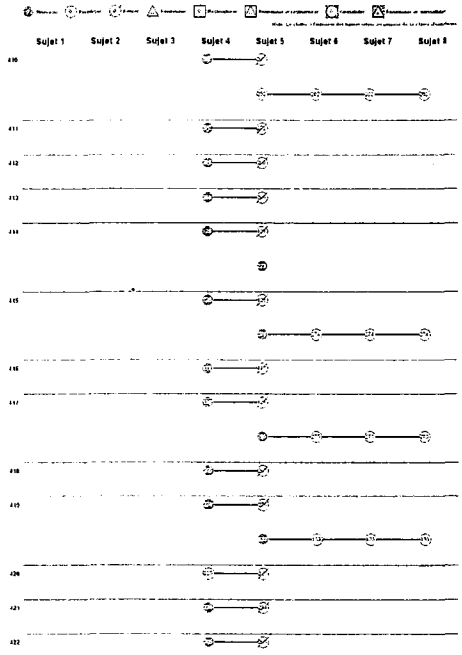

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

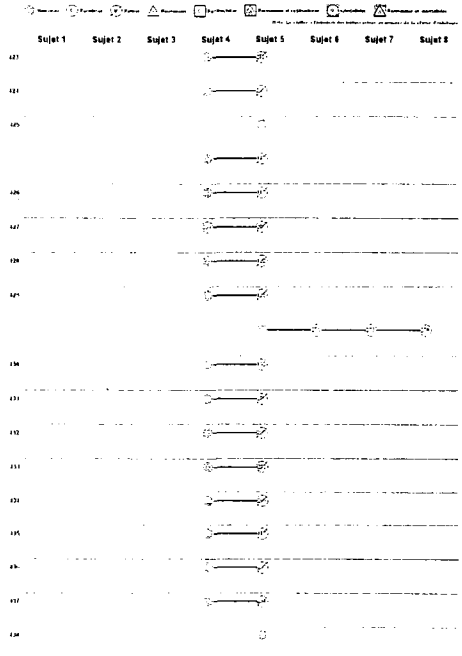

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

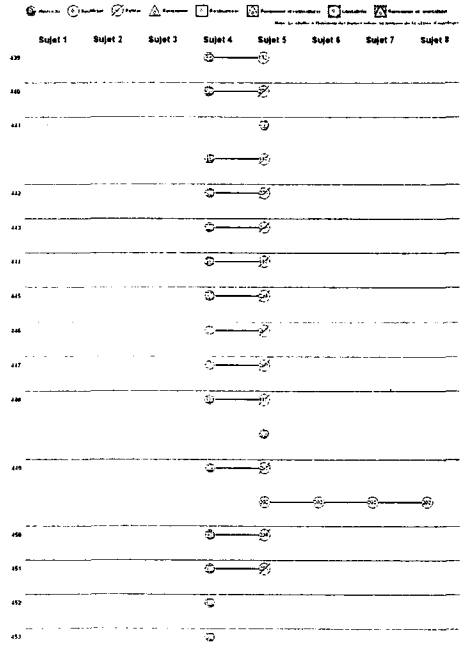

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

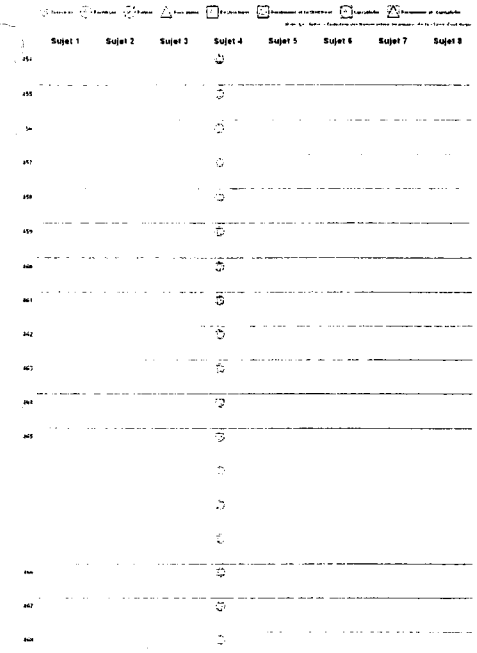

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

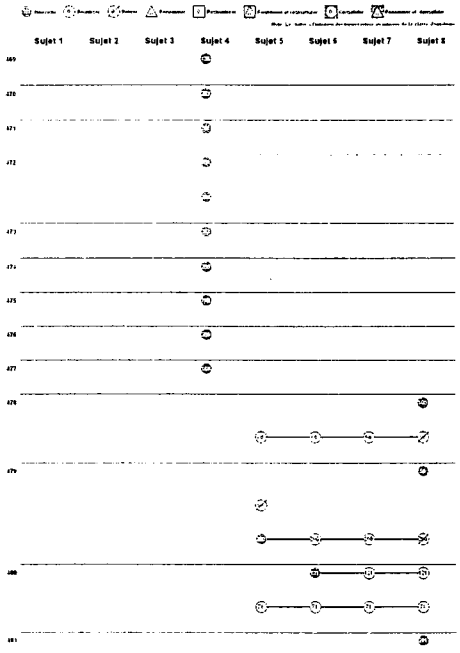

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

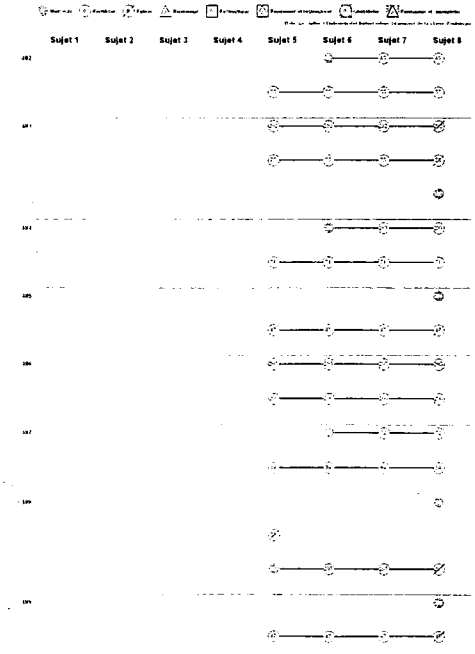

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

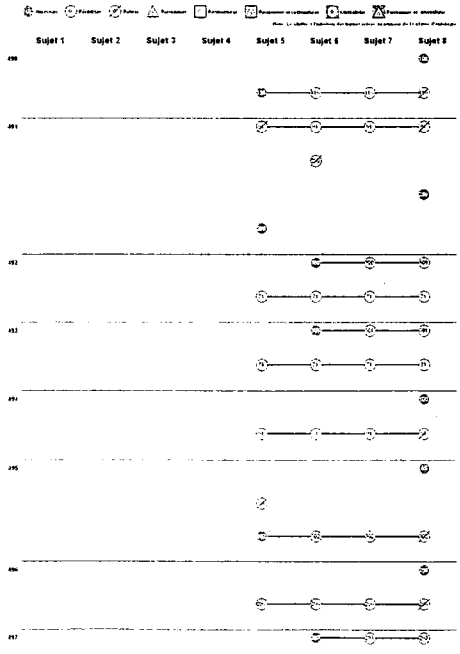

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

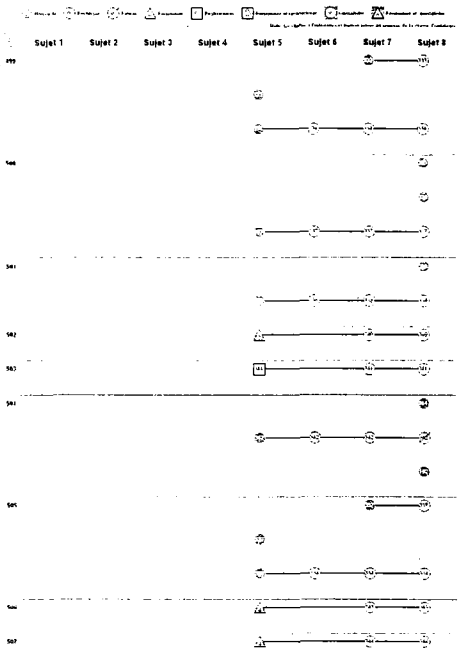

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

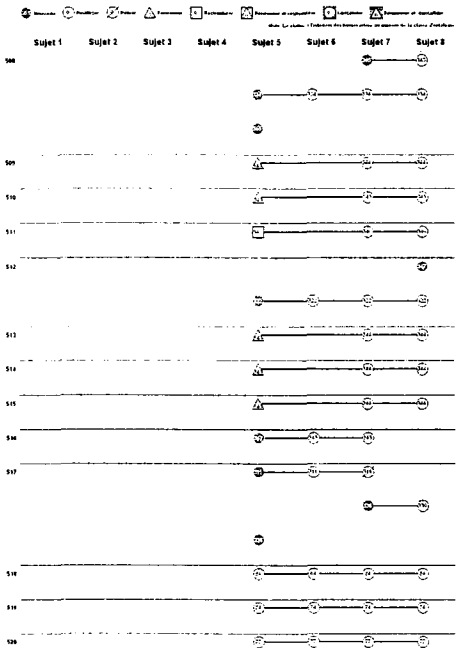

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

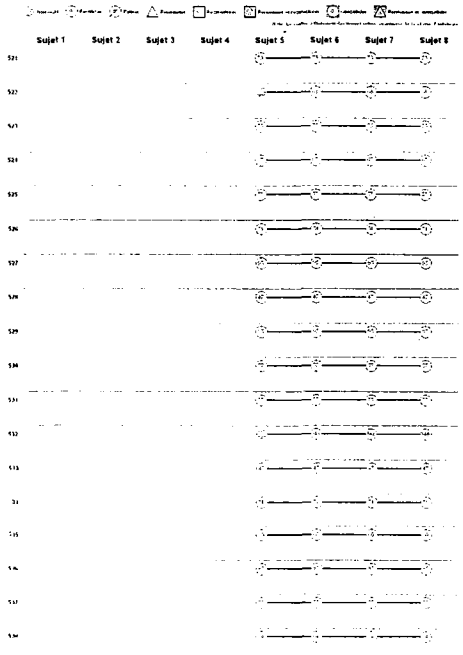

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

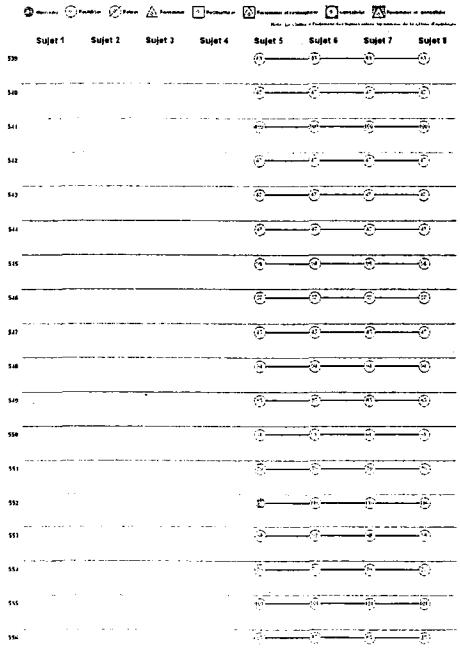

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

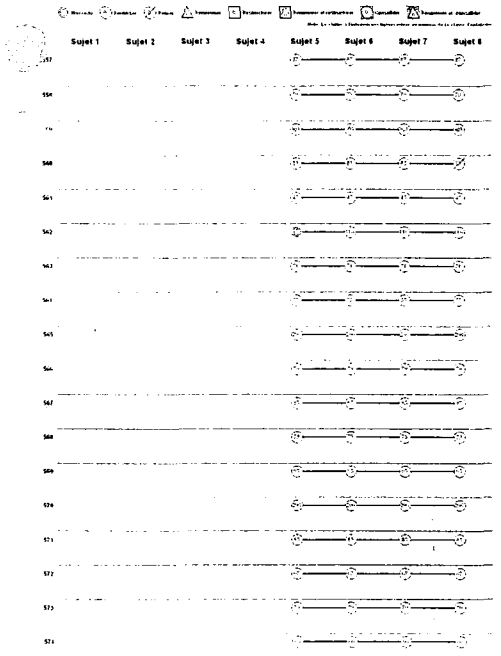

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

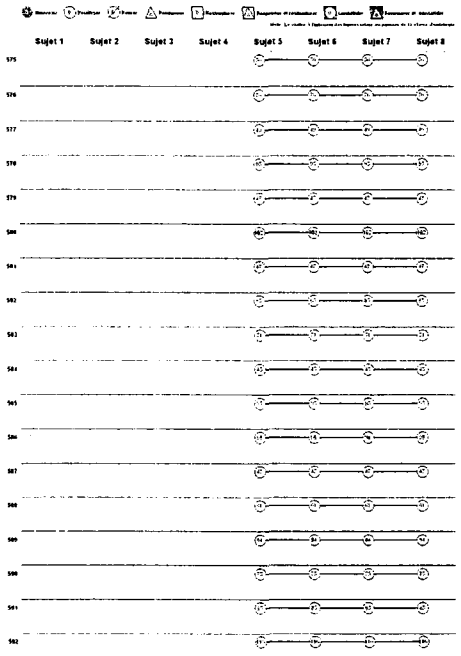

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

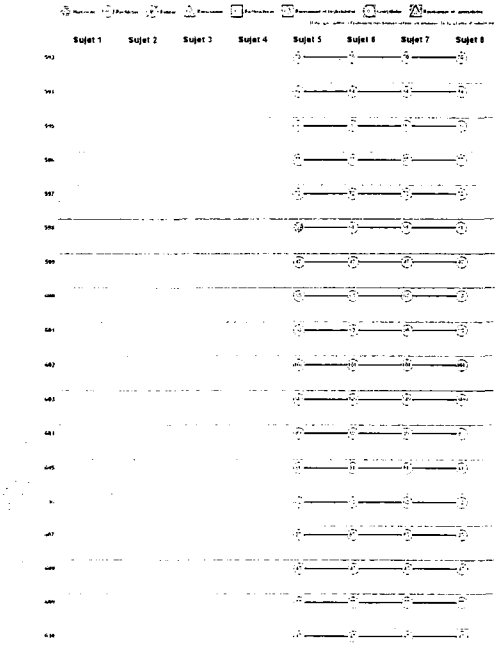

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

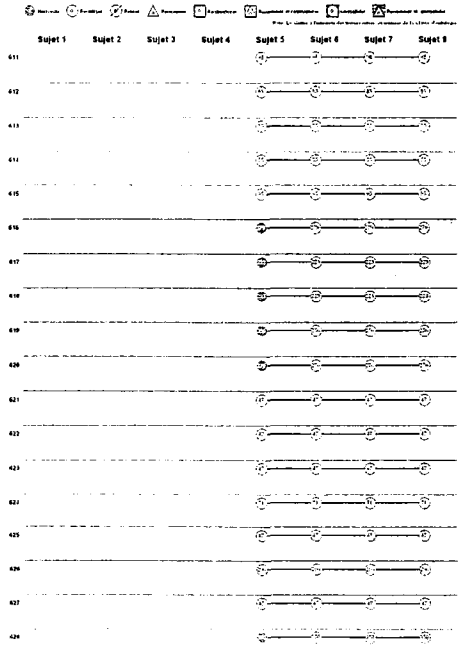

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

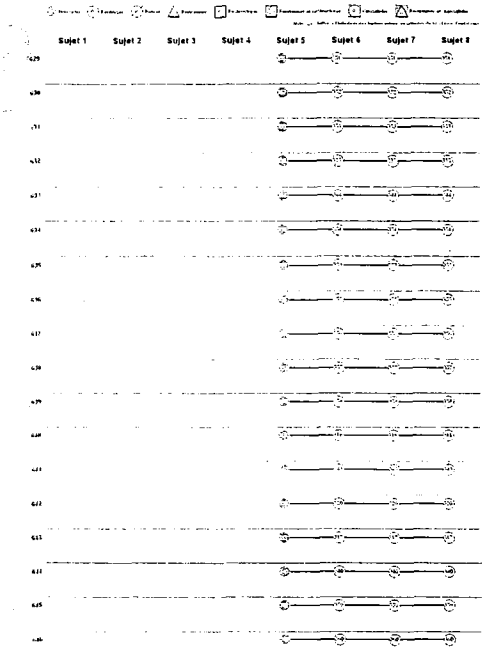

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

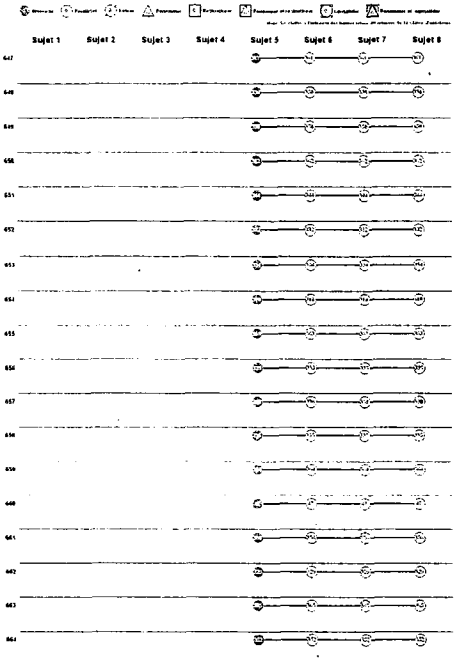

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

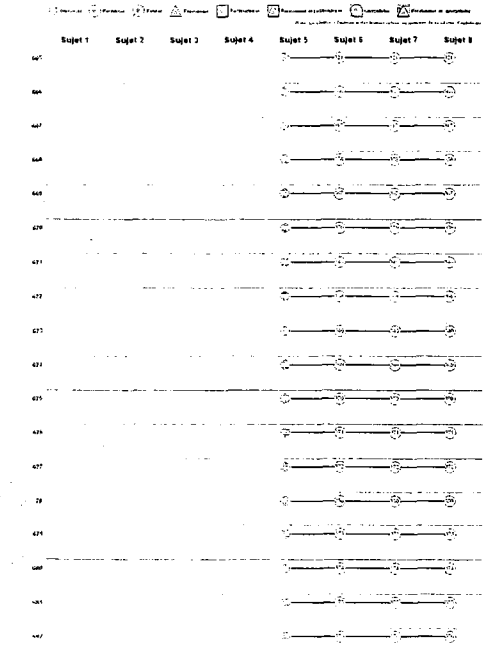

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

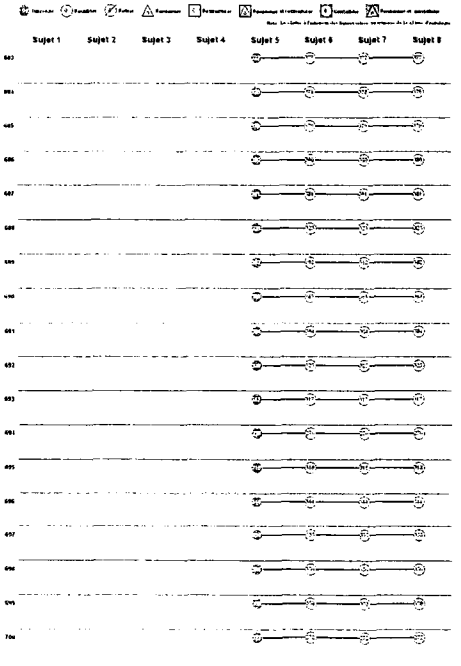

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

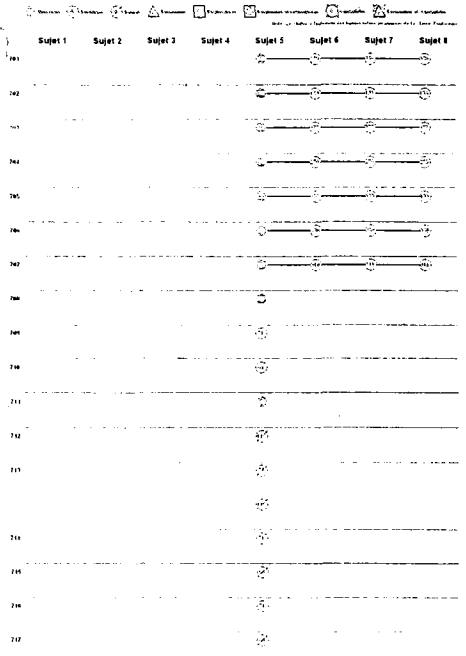

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

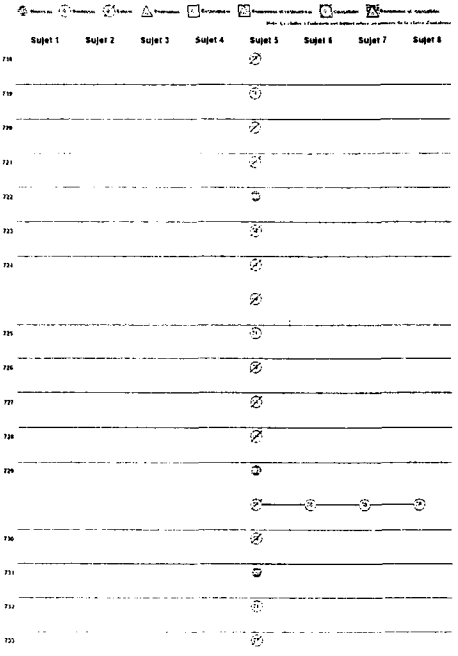

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

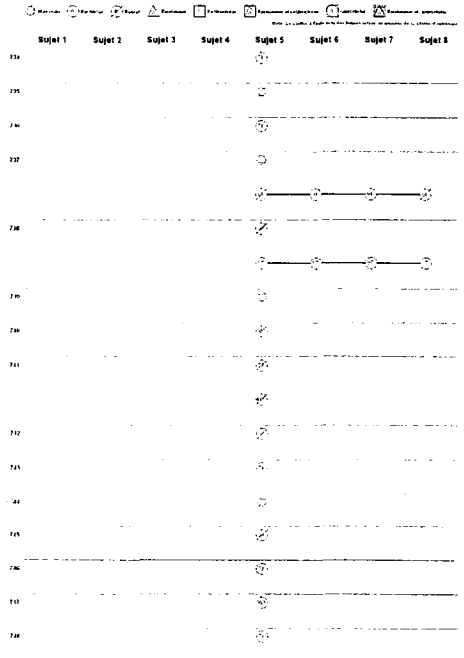

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

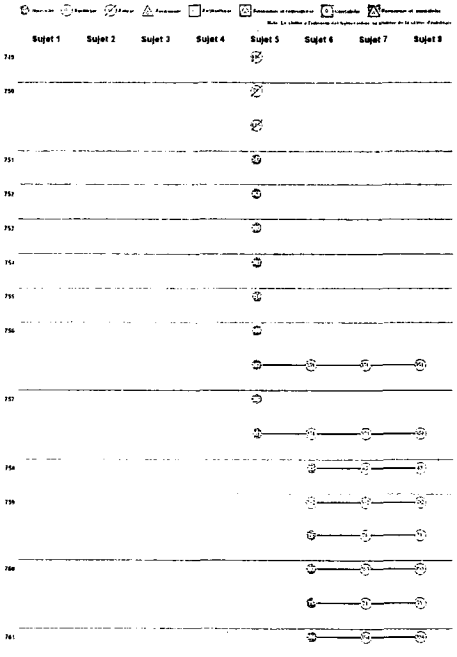

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

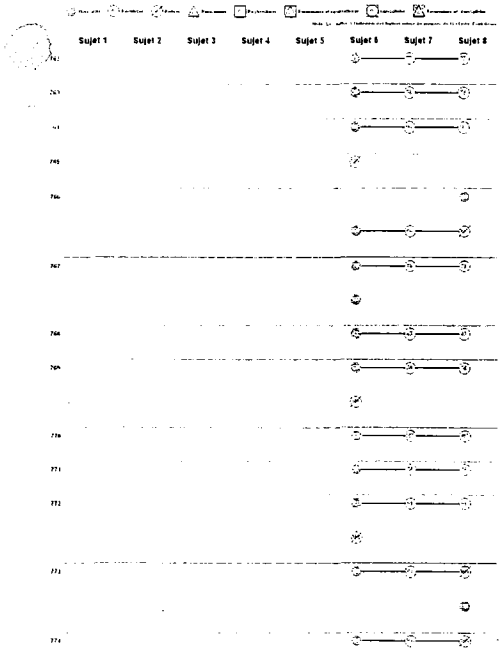

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

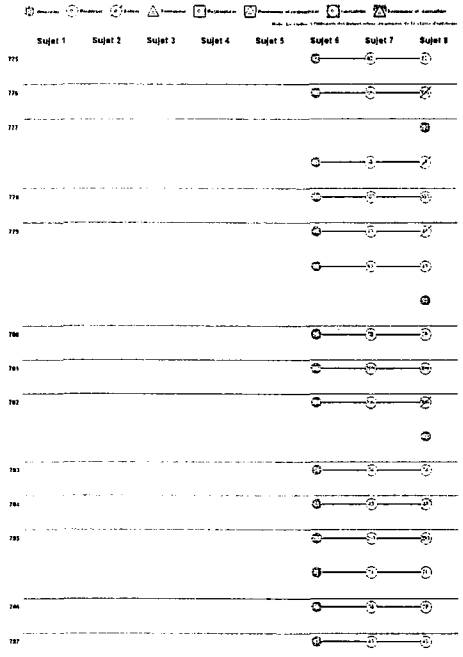

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

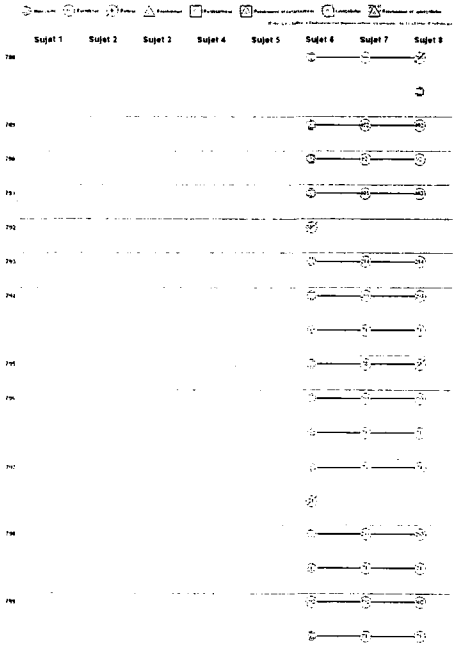

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

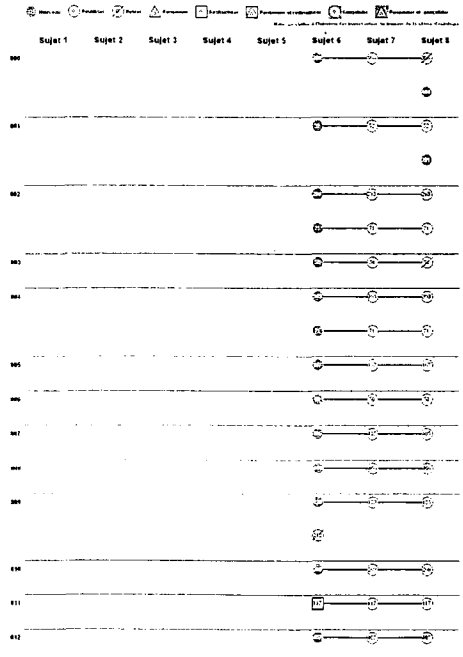

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

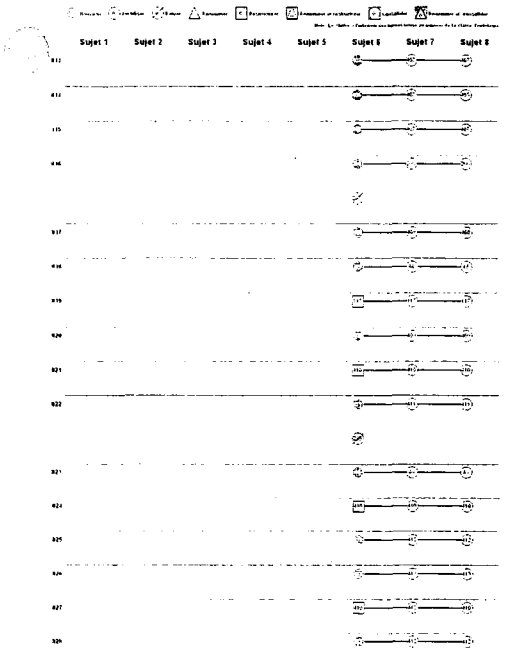

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

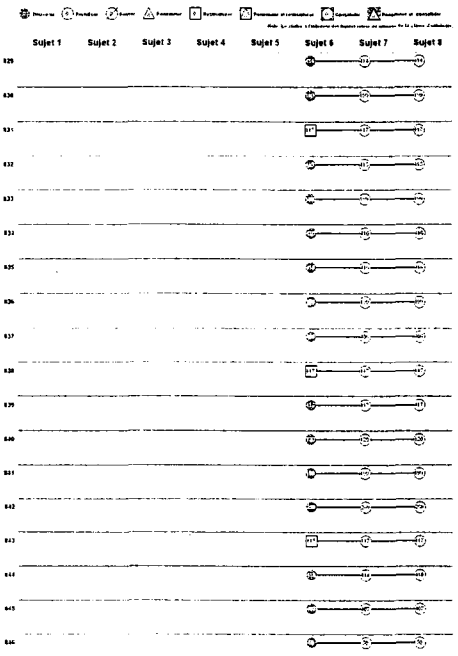

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

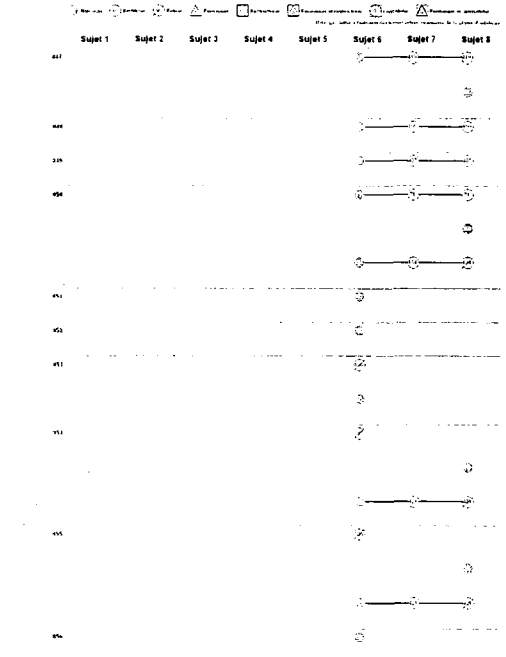

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

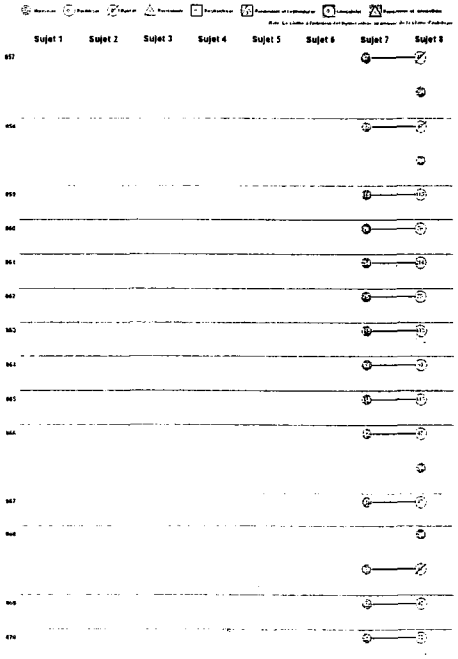

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8

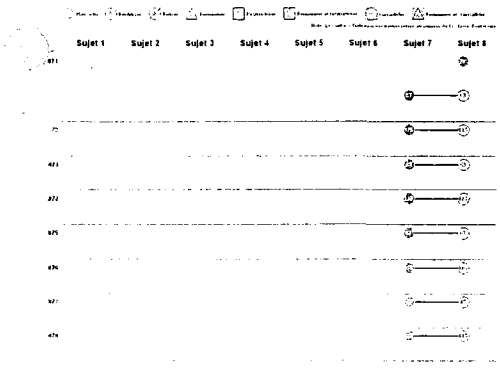

Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8


Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8


Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8


Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8


Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8


Sujet 1
Sujet 2
Sujet 3
Sujet 4
Sujet 5
Sujet 6
Sujet 7
Sujet 8