

Dossier d'Habilitation à diriger les Recherches

Spécialité : Informatique

Donner accès au contenu des documents textuels
Acquisition de connaissances et analyse de corpus spécialisés

Adeline NAZARENKO

Laboratoire d'Informatique de Paris-Nord
Université Paris-Nord

Présentée le 10 décembre 2004 devant le jury composé de :

Madame Maria Tereza PAZIENZA (rapporteur)

Monsieur Benoît HABERT (président)

Monsieur Daniel KAYSER (examineur)

Monsieur François LEVY (examineur)

Monsieur Jean-Marie PIERREL (rapporteur)

Monsieur Bernard VICTORRI (rapporteur)

Adeline Nazarenko
Maître de Conférences au département d'informatique de l'Université Paris 13

LIPN – Institut Galilée

Av. J.-B. Clément 93430 Villetaneuse
Tél. 01 49 40 40 89 adeline.nazarenko@lipn.univ-paris13.fr

DOSSIER D'HABILITATION A DIRIGER LES RECHERCHES

En complément du mémoire d'Habilitation à Diriger les Recherches, ce dossier comporte les pièces suivantes :

- Un Curriculum Vitae détaillant les responsabilités scientifiques, pédagogiques et administratives assumées
 - Une liste de publications
 - Les 4 articles de recherche ci-dessous
1. Adeline Nazarenko, Pierre Zweigenbaum, Benoît Habert and Jacques Bouaud. "Corpus-based Extension of a Terminological Semantic Lexicon", In *Recent Advances in Computational Terminology* (D. Bourigault, C. Jacquemin and M.-C. L'Homme, eds.). John Benjamins, 2001
 2. Thierry Hamon and Adeline Nazarenko. "Detection of synonymy links between terms: experiment and results", *Recent Advances in Computational Terminology* (D. Bourigault, C. Jacquemin and M.-C. L'Homme, eds.). John Benjamins, 2001.
 3. Touria Aït El Mekki, Adeline Nazarenko. « L'index de fin de livre, une forme de résumé indicatif ? », *Traitement Automatique des Langues (TAL)*, à paraître.
 4. Erick Alphonse, Sophie Aubin, Philippe Bessières, Gilles Bisson, Thierry Hamon, Sandrine Lagarrigue, Adeline Nazarenko, Alain-Pierre Manine, Claire Nédellec, Mohamed Ould Abdel Vetah, Thierry Poibeau, Davy Weissenbacher, "Event-based Information Extraction for the biomedical domain: the Caderige project". *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and Its Applications (COLING'2004)*, Genève.

CURRICULUM VITAE DETAILLE

Curriculum Vitae.....	4
1. Etat Civil.....	4
2. Formation.....	4
3. Carrière professionnelle	4
4. Distinction.....	4
Responsabilités scientifiques	5
1. Encadrement	5
1.1 Travaux de thèses	5
1.2 Travaux de DEA.....	5
1.3 Autres travaux.....	6
2. Contrats	7
2.1 Contrats terminés.....	7
2.2 Contrats en cours	8
2.3 Contrats soumis	8
3. Animation de la recherche.....	9
3.1 Animation et groupes de recherche.....	9
3.2 Organisation de séminaires et colloques.....	10
3.3 Evaluations et expertises	10
Activités et Charges pédagogiques.....	12
4. Liste de cours	12
4.1 Informatique générale.....	12
4.2 Intelligence Artificielle.....	13
4.3 Traitement automatique des langues et des données textuelles	13
4.4 Divers	14
5. Charges pédagogiques.....	14
5.1 Responsabilités pédagogiques.....	14
5.2 Définition de programmes.....	15
Autres activités administratives	16
Liste de Publications	17
1. Ouvrages, chapitres de livres	17
2. Revues Internationales avec Comité de Sélection.....	18
3. Conférences ou Workshops Internationaux avec Comité de Sélection	18
4. Revues Nationales avec Comité de Sélection.....	19
5. Conférences Nationales avec Comité de Sélection	20
6. Autres Revues.....	21
7. Autres Conférences.....	21
8. Mémoires	22
Articles de recherche.....	23

CURRICULUM VITAE

1. Etat Civil

Adeline NAZARENKO née PERRIN, 39 ans, française

Née le 29 sept. 1965, à Paris 15^{ème} Mariée, 3 enfants

2. Formation

- 1983 : Baccalauréat, **série C** à Lyon
- 1983-85 : Hypokhâgne et khâgne, série S (Lettres et Sciences Sociales), au Lycée du Parc à Lyon
Littérature, philosophie, histoire, langues, sciences sociales et mathématiques
- 1985-89 : Scolarité à l'**Ecole Normale Supérieure**
- 1985-86 : **Licence de Linguistique Informatique** à l'Université Paris VII
Licence de Lettres Modernes à l'Université Paris IV
- 1986-87 : **Maîtrise de Linguistique Informatique** à l'Université Paris VII
- 1988 : **Agrégation de Lettres Modernes**
- 1988-89 : **DEA d'Intelligence Artificielle** à l'Université Paris XIII (Directeur : D. Kayser, mention B)
- 1989-90 : **Séminaires de Linguistique** à l'Université de Stanford, Californie
- 1990-93 : **Thèse d'Informatique** de l'Université Paris XIII (Directeur : D. Kayser, soutenue le 31/01/94, mention très honorable avec félicitations du jury). Jury : G. Bès (rapp), M. Borillo, J. Fargues, J. Jayez, D. Kayser (directeur), G. Ligozat (rapp)
- 2004 : Thèse d'Habilitation à Diriger les recherches (Université Paris XIII) : Jury : D. Kayser, F. Lévy, M.T. Paziienza (rapp), J.-M. Pierrel (rapp), B. Victorri (rapp), Soutenance prévue le 10/12/04.

3. Carrière professionnelle

- 1985-89 : **Fonctionnaire stagiaire à l'Ecole Normale Supérieure**
- Juil-août 1987: **Stage de recherche** à l'Institut de la Communication Parlée (Grenoble), dirigé par G. Bailly
- Mars-août 1989 : **Stage de DEA** : recherche et développement au Centre Scientifique d'IBM, stage dirigé par J. Fargues
- 1989-90 : **Professeur de langue et stylistique française** à l'Université de Stanford, Californie
Stage de recherche au CSLI (Center for the Study of Language and Information, Université de Stanford), dirigé par S. Peters
- 1990-93 : **Stage de thèse** : recherche et développement au Centre Scientifique d'IBM, stage dirigé par J. Fargues
- 1994-96 : **ATER en Informatique** à l'Ecole Normale Supérieure de Fontenay-St Cloud
- Depuis sept. 1996 : **Maître de Conférence en Informatique** à l'Université Paris 13

4. Distinction

Médaille de bronze du CNRS, commission interdisciplinaire 45 (« Cognition, langage, traitement de l'information : systèmes naturels et artificiels ») en 2004.

RESPONSABILITES SCIENTIFIQUES

Les références mentionnées ici entrez crochets renvoient à la liste de publications figurant à la fin du présent document.

1. Encadrement

1.1 Travaux de thèses

- Thèse d’informatique de Thierry Hamon (1997-2000, soutenue le 20/12/2000). Ce travail porte sur la détection de liens de synonymie entre les termes d’un corpus. Titre : « Variation sémantique en corpus spécialisés : Acquisition de relations de synonymie à partir de ressources lexicales ».

J’ai assuré l’encadrement scientifique sous la direction de Christophe Fouqueré.

Jury : H. Boccon-Gibod (Examineur), B. Daille (Présidente), C. Fouqueré (Directeur), B. Habert (Rapporteur), A. Nazarenko (Encadrante), J. Veronis (Rapporteur)

- Thèse d’informatique de Thierry Poibeau (1999-2002, soutenue le 8/3/2002). Cette thèse porte sur le développement d’une méthode et d’outils génériques mais adaptables pour l’extraction d’information à partir de textes. Titre : « Extraction d’information à base de connaissances hybrides ».

Cette thèse a fait l’objet d’une bourse Cifre entre Thalès et l’Université Paris 13. J’ai assuré l’encadrement scientifique sous la direction de Daniel Kayser.

Jury : C. Fluhr (Président), C. Jacquemin (Rapporteur), D. Kayser (Directeur), A. Nazarenko (Encadrante), P. Saint-Dizier (Rapporteur), C. Sedogbo (Examineur) et Y. Wilks (Rapporteur)

- Thèse d’informatique de Touria Aït El Mekki (2000-2004, soutenance prévue le 21 déc. 2004). Cette thèse porte sur le développement d’une méthode et d’un outil pour l’indexation fine de documents. Titre : « Construction semi-automatique d’index de fin de livres ».

J’assure l’encadrement scientifique sous la direction de Daniel Kayser.

Jury : B. Daille (Rapporteur), P. Desclés (Rapporteur), D. Kayser (Directeur), A. Nazarenko (Encadrante), J. Véronis (Examineur).

- Thèse d’informatique de Davy Weissenbacher (2003- ?). Cette thèse porte sur le développement de méthodes de normalisation de corpus combinant des techniques de TAL et d’apprentissage automatique.

J’assure l’encadrement scientifique sous la direction de Christophe Fouqueré.

1.2 Travaux de DEA

Figurent en italiques les travaux que j’ai plus supervisés que véritablement encadrés.

Dans ce qui suit, « DEA d’IA&OC » fait référence au DEA d’Intelligence Artificielle et Optimisation Combinatoire de l’Université Paris 13.

- Stage de DEA d’IA&OC de Thierry Hamon (mention Bien, 1997). Ce stage a été effectué à la Direction des Études et Recherches de EDF-GDF. Titre : « Exploitation d’informations de la langue générale pour l’aide à la structuration de terminologie ».
- *Stage de DEA d’IA&OC de Hacène Cherfi (mention Bien, 1999). Ce stage a été effectué à l’Action AID – INRIA Sofia-Antipolis. Il a été co-encadré avec Brigitte Trousse (INRIA) qui a principalement dirigé le travail.*
- *Stage de DEA d’IA&OC d’Odile Cadassou (2000). Ce stage est effectué à Paris V. Il a été principalement encadré par Gabriel Michel (MC, Paris V).*
- Stage de DEA d’IA&OC de Touria Aït El Mekki (2000), effectué au LIPN. Titre : « Acquisition de connaissances pour l’aide au support technique : analyse de corpus et réflexion méthodologique ».
- Stage de DEA d’IA&OC de Manal El Zant (2002), effectué au LIPN. Titre : « Normalisation de textes techniques pour l’extraction d’information ».
- *Stage de DEA d’IA&OC d’Izoumi Lhérieau (2001), effectué dans la société Sinequa. Il a été principalement encadré par Claude de Loupy (Sinequa). Titre : « Analyse des questions et recherche sur Internet pour un système de questions-réponses ».*
- Stage de DEA d’IA&OC de Zhyiu Qian (mention AB, 2003), effectué au LIPN. Titre : « Etude de la contribution de la terminologie au filtrage de fragments de texte ».
- Stage de DEA d’IA&OC de Davy Weissenbacher (mention TB, 2003), effectué au LIPN. Titre : « Etude et reconnaissance automatique des relations de synonymie et de renommage dans les textes de génomique ».
- Stage de DEA de Linguistique théorique et formelle de Goritsa Ninova (2004, Université Paris 7). Ce stage a été effectué au LIPN. Il a été co-encadré par Sylvie Szulman. Titre : « Mesures de l’adéquation des ressources lexicales et ontologiques en corpus ».
- Stage de DEA d’IA&OC de Frederik Cailliau (mention AB, 2004), effectué dans la société Sinequa et co-encadré par Claude de Loupy (Sinequa). Titre : « Vers une gestion intégrée des ressources linguistiques ».
- Stage de DEA d’IA&OC d’Oliver Hamon (2004), effectué dans la société ELDA et co-encadré par Sylvain Surcin (ELDA). Titre : « Evaluation automatisée et méta-évaluation des systèmes de traduction automatique ».

1.3 Autres travaux

- Stage de DESS d’Ingénierie multilingue de Christine Reynaud (1995, 6 mois), effectué à l’ENS de Fontenay. Titre : « Etude statistique d’un corpus de réponses libres : des formes graphiques aux unités sémantiques ».
- Stage de fin d’études de Danièle Cohen (ingénieur CNAM (2001, 9 mois), effectué au LIPN. Titre : « Analyse linguistique et terminologique de corpus techniques pour l’acquisition de connaissances ».
- Stage PostDoctoral de Thierry Charnois (1999-2000, 18 mois), effectué au LIPN dans le cadre du projet ASTUXE.

- Stage de maîtrise de Touria Aït El Mekki (Université Cadi Ayyad, Marrakech, 1999, 2 mois) effectué au LIPN. Ce stage a été co-encadré par Thierry Hamon (LIPN). Titre : « Installation et évaluation d'un analyseur syntaxique ».
- Stage de maîtrise de Lise Pâris (Université Paris 7, 2003, 2 mois) effectué au LIPN. Ce stage a été co-encadré par Touria Aït El Mekki (LIPN). Titre : « Construction automatique d'index : validation de résultats et évaluation ».
- Stage post-doctoral de Guillaume Vauvert (2004, 6 mois), effectué au LIPN dans le cadre du projet Alvis.

2. Contrats

Depuis que je suis Maître de Conférences, une part non négligeable de mon travail de recherche a été consacrée au montage et à la réalisation de contrats de recherche. Ceux-ci présentent un triple intérêt à mes yeux :

- Les collaborations avec des industriels ou d'autres types d'utilisateurs définissent des contextes applicatifs (y compris les jeux de données et la définition des besoins) dans lequel les méthodes et outils développés en laboratoire peuvent être testés et mis en œuvre.
- Ces contrats sont l'occasion d'établir des collaborations sur la durée avec d'autres équipes et de travailler à plusieurs sur les problèmes de recherche qui ont été identifiés, en particulier quand ils font appel à des compétences mixtes.
- Les budgets associés à ces contrats permettent de financer des stages d'étudiants mais aussi des PostDoctorants et ingénieurs contractuels qui sont précieux pour mener à bien des tâches lourdes sur le plan expérimental ou en termes de développement.

A deux exceptions près (en italique dans la liste ci-dessous), j'ai assuré le montage et la responsabilité scientifique de ces projets pour la partie LIPN.

2.1 Contrats terminés

- *Action de Recherche Concertée (ARC 4, 1995-97) de l'AUPELF-UREF pour l'évaluation des systèmes de compréhension de textes, sous la direction de Paul Sabatier (CNRS, Laboratoire d'Informatique de Marseille). Cette étude avait pour but d'étudier les conditions d'une évaluation de tels systèmes et de jeter les bases d'un protocole d'évaluation. La contribution du LIPN a porté sur les aspects proprement sémantiques et pragmatiques de la compréhension de textes [9,58].*
- Contrat de d'Études de Recherche et Développement entre la DER-EDF et l'Université Paris 13 pour l'aide à la structuration de terminologies. Ce contrat de 40 KF a duré un an (1998). Le stage de DEA de Thierry Hamon a été effectué dans ce cadre. Cette collaboration a permis de conduire des expériences intéressantes concernant l'étude de la synonymie entre termes [15,44].
- *Plan Pluriformations de l'Université Paris 13 pour le plan quadriennal 1997-2000 en collaboration avec le Laboratoire de Linguistique Informatique de l'Université Paris 13. (le LLI). Le LIPN est coordinateur. Ce projet a permis de rapprocher les deux laboratoires et d'initier des travaux communs sur l'analyse de vastes corpus textuels.*

- Bourse Cifre avec Thalès Recherche & Technologie (anciennement Thomson-LCR) : 1998-2001.
- Projet RNRT ASTUXE, associant Spartacom (PME, éditeur de logiciel), le Thalès Recherche & Technologie (anciennement Thomson-LCR), le LRI et le LIPN. Ce projet précompétitif vise à développer un module d'analyse automatique des mails de demande d'intervention adressés à un centre de support technique. Le financement prévu pour le LIPN s'élevait à 1 125 KF pour 27 mois (2000-02). Ce projet a connu des difficultés suite à la défection et au remplacement du coordinateur ainsi qu'au changement d'objectifs qui en a découlé. Il a été interrompu au bout de 15 mois. En dépit de cet échec, ce projet a permis d'établir une collaboration durable avec Claire Nédellec (LRI puis MIG-INRA). Il a été relancé sur de nouvelles bases dans le cadre du projet ExtraPloDocs.
- Projet de bioinformatique IMPG inter-EPST CADERIGE 1 et 2, associant différents partenaires dont le LIPN (Financement global pour le LIPN : environ 15 Keuros). Ce projet a pour objectif de concevoir des bases de connaissances d'interactions géniques grâce à l'automatisation de l'extraction d'information dans la bibliographie de MedLine (novembre 2001 à octobre 2003). Il est à l'origine de mon travail sur la biologie [12,53,54].

2.2 Contrats en cours

- Projet Cederillic (2003-04), associant l'ERSS, l'unité STIM de l'AP-HP, FranceTélécom et le LIPN pour la construction d'index pour des ouvrages au format électronique et papier. Ce projet, modestement financé par FranceTélécom, a débouché sur la publication d'un ouvrage scientifique doté d'un riche index. Cet ouvrage a été publié à la fois sur CD-rom et sous forme papier [34].
- Projet RNTL ExtaPloDocs (30 mois, 2002-2005), associant le LIPN qui en assure la coordination, l'Unité Mathématique Informatique et Génome de l'INRA et la société Isoft (PME, éditeur de logiciel de datamining). Ce projet exploratoire vise au développement d'un outil d'extraction d'information destiné à faciliter les recherches bibliographiques effectuées par les biologistes en génomique. Le financement pour le LIPN s'élève à 170 Keuros.
- Projet STREP Alvis (6^{ème} PCRD, 36 mois, 2004-2006), associant différents partenaires européens universitaires et PME. Ce projet vise au développement d'un moteur de recherche sémantique et distribué. Les aspects sémantiques assurent une meilleure adéquation des résultats aux requêtes formulées par les utilisateurs. La recherche distribuée repose sur la coopération de différents serveurs spécialisés dans des domaines différents. Le financement pour le LIPN s'élève à 210 Keuros.
- Plan Pluriformations pour le plan quadriennal 2005-2008 en collaboration avec 2 autres laboratoires de l'Université Paris 13 (le LimBio et le LLI). Le LIPN est coordinateur. Ce projet a pour but de développer des outils facilitant l'accès à l'information médicale. Ce projet a été accepté. Il doit débiter en janvier 2005. Le financement pour le LIPN doit s'élever à 70 Keuros sur toute la durée du projet.

2.3 Contrats soumis

- Coopération franco-québécoise entre le LIPN et l'École de bibliothéconomie et des sciences de l'information de l'Université de Montréal (2005-2006). Ce projet vise à faire tester et évaluer le prototype d'indexation fine de documents développé au LIPN par des professionnels des sciences de l'information. Il devrait déboucher sur des propositions d'amélioration. Ce projet a été déposé en juin 2004. Il est en cours d'évaluation.

- Bourse Cifre avec la société Sinequa pour la thèse de Frederik Cailliau, dans le prolongement de son stage de DEA. Le dossier est en cours d'évaluation.
- Bourse Cifre avec la société ELDA pour la thèse d'Olivier Hamon, dans le prolongement de son stage de DEA. Le dossier est en cours de constitution.

3. Animation de la recherche

3.1 Animation et groupes de recherche

J'ai pris la responsabilité de l'équipe « Représentation des Connaissances et Langage Naturel » du LIPN en décembre 2002. Cette équipe compte aujourd'hui 2 Professeurs, 8 Maîtres de Conférence et 1 chargé de Recherche 1^{ère} classe du CNRS. S'y ajoute actuellement 5 doctorants et deux ingénieurs contractuels.

En collaboration avec Claire Nédellec (LRI puis INRA), j'ai lancé et animé le groupe de travail A3CTE (Applications, acquisition et apprentissage de connaissances à partir de textes). Ce groupe avait pour but de rapprocher les communautés du TAL, de l'apprentissage automatique et de l'ingénierie des connaissances. Ce groupe a fonctionné jusqu'en juin 2003, à raison de 4 à 5 réunions de travail thématiques par an, réunissant 20 à 40 participants à chaque fois. Ce groupe a reçu le soutien de l'AFIA, du GRD I3 et du Réseau des Sciences Cognitives Ile de France.

Sans y assurer de responsabilité particulière, j'ai également participé aux travaux de différents groupes de travail :

- Le groupe ESPOIR autour de Jacques Bouaud et Pierre Zweigenbaum (AP-HP), Benoît Habert (LIMSI Univ. Paris X) et moi. Nous avons cherché à mesurer l'apport des méthodes d'exploration automatique ou semi-automatique de corpus de spécialité pour la construction d'ontologie. Ce groupe informel s'est réuni régulièrement entre 1996 et 2001 [6,8,10,18].
- Le groupe TIA (Terminologie et Intelligence artificielle, AFIA – GDR I3) que j'ai rejoint en 1998.
- L'Action Spécifique Assticcot du CNRS sur l'élaboration de ressources terminologiques à partir de corpus (2002-2003).
- Le bureau du réseau thématique pluridisciplinaire « Langage naturel, logique et philosophie du langage » (section 34 et 35 du CNRS, démarrage fin 2004), coordonné par Philip Miller (U. Lille 3 & Silex).

J'ai eu le souci faire reconnaître au niveau international les approches sémantiques à base de corpus que défendent les membres des groupes TIA et Assticcot. Cela s'est traduit par l'organisation de workshops internationaux mais aussi par une manifestation d'intérêt [66] que j'ai déposée au printemps 2002 dans le cadre du 6^{ème} PCRD et qui a compté un nombre important des participants (environ 80, de multiples pays).

3.2 Organisation de séminaires et colloques

a. Journées nationales

Organisation d'une journée d'étude de l'ATALA sur le principe de compositionnalité sémantique en janvier 1997. Coordination d'un numéro de la revue TAL (Traitement Automatique des Langues) sur ce même thème (1998).

Co-organisation d'une journée d'étude de l'ATALA sur les approches distributionnelles de l'analyse sémantique en janvier 1999 (en collaboration avec Benoît Habert, U. Paris X & Limsi).

Co-organisation, au LIPN et en 1999, d'une journée pluridisciplinaire sur la modélisation de la causalité, en collaboration avec Philippe Dague, François Lévy et Daniel Kayser (LIPN) et avec le soutien financier du pôle Cognicentre du RISC.

Co-organisation d'une journée d'étude de l'ATALA sur la variation en terminologie, en mars 2001 (en collaboration avec Thierry Hamon, LIPN). Coordination d'un numéro de la revue TAL (Traitement Automatique des Langues) sur ce même thème (2002).

Organisation d'une journée « Acquisition et apprentissage de connaissances à partir de textes » sur la Plate-forme AFIA 2001.

b. Colloques internationaux

Co-organisation d'un Workshop « Using Semantics for Information Retrieval and Filtering » associé à Langage Resources and Evaluation Conference (LREC'2002) en Espagne.

Co-organisation d'un Workshop « Machine Learning and Natural Language for Ontology Learning » associé à l'European Conference on Artificial Intelligence (ECAI'2002) à Lyon.

Co-organisation de « International Joint Workshop on Natural Language Processing in Biomedicine and Its Applications » associé à l'International Conference on Computational Linguistics (COLING'2004) à Genève.

3.3 Evaluations et expertises

J'ai participé à différents comités de programme ou de lecture, notamment pour les conférences TALN 2002 et 2003, TIA 2003, RFIA 2004, pour divers ateliers associés aux conférences TALN 1999, EKAW 2000, LREC 2002, ECAI 2002, LREC 2004, COLING 2004, ECAI 2004, et pour certains numéros spéciaux des revues TAL et Human Language Technology.

À plusieurs reprises, j'ai eu des évaluations de projets à faire : projets-contrats de la région de la Basse-Normandie, projets RNTL, projets de recherche de chercheurs canadiens.

Outre les thèses que j'ai encadrées, j'ai participé aux jurys des thèses suivantes :

- Examinatrice sur la thèse d'Elie Naulleau (Université Paris-13, 1998)
- Rapporteur (par dérogation) de la thèse de David Faure (Université Paris 11, 2000)
- Examinatrice sur la thèse d'Isabelle Debourges (Université Orléans, 2002)
- Examinatrice sur la thèse de Helka Folch (Université Paris-13, 2002)

- Examinatrice sur la thèse de Khalid Benabdeslem (Université Paris-13, 2004)

J'ai été nommée « référent scientifique » extérieur par le chef de département Biométrie et Intelligence Artificielle de l'INRA dans le cadre d'une procédure de titularisation de Nicolas Turenne comme chargé de recherche (CR2) : encadrement pendant 1 an et rapport d'évaluation.

ACTIVITES ET CHARGES PEDAGOGIQUES

J'ai enseigné comme ATER d'abord puis en tant que Maître de Conférence. À l'exception des périodes de congé de maternité, j'ai effectué un service complet de 192 h ou un peu plus.

4. Liste de cours

Mes enseignements s'organisent principalement autour de trois thèmes principaux : l'informatique générale, l'Intelligence Artificielle et le Traitement Automatique des Langues.

Tous les volumes horaires sont arrondis, ils sont donnés en heures équivalent TD et par année. Sauf mention contraire, ces cours ont eu lieu à l'Université Paris 13.

4.1 Informatique générale

Programmation impérative (DEUG MIAS puis Licence 1^{ère}, 1997-2005, 70h, cours)

Ce cours est destiné à des débutants en Informatique (1^{ère} année universitaire). Il s'agit d'une introduction à l'algorithmique et à la programmation (en Pascal au départ, en C désormais).

J'assure depuis 1997 l'un des deux cours d'amphi et la coordination de l'équipe pédagogique des chargés de TD, TP et projet (entre 8 à 14 groupes de TD selon les années). J'ai rédigé le support de cours (90 pages environ). Certaines années, j'ai également pris en charge un groupe de TD/TP sur un semestre.

Programme : Algorithmique (récursivité, types abstraits de données linéaires, recherche et tri). Les bases du C (structures de contrôles, tableaux et structures de données, fonctions, pointeurs, fichiers). Quelques éléments de systèmes d'exploitation, de codage, de compilation.

Structures de Programmes et Structures de Données (Licence d'informatique, 1996-99, 40h, TD)

Programme : Structures de données arborescentes (tas, arbres généraux et binaires, arbres de recherche, arbres équilibrés) et algorithmes associés (recherche, mise à jour,...) ; dérécursivation et tri externe.

C et Unix (Licence d'informatique, 1996-97, 20h, TD/TP)

Introduction à la programmation C et sous Unix : stage intensif de C, système de fichier, gestion des processus, outils de programmation sous Unix (grep, sed, make, haw...).

Théorie des langages et compilation (Option DEUG MIAS 2^{ème} année, 1996-97, 30h, TD/TP)

Programme : théorie des langages (problème de mots, types de langages et grammaires de réécriture ; automates) et compilation (analyse lexicale et syntaxique, génération de code)

LISP (Option info. du DEA de Linguistique Théorique et Formelle de l'Univ. Paris 7, 1993-96 30h, cours/TD/TP)

Programme : Programmation fonctionnelle et application au LISP

Bases de données relationnelles et SQL (IUT d'Orsay, 1993-94, 30h, TD/TP)

Programme : modèle relationnel, SQL.

4.2 Intelligence Artificielle

Ingénierie des connaissances (Maîtrise d'informatique, 2000-2005, 28h, cours/TD, en collaboration avec B. Biébow)

Programme : mise au jour, modélisation et formalisation de connaissances (UML, logique de description, OWL), résolution de problèmes et méthodes d'apprentissage.

Représentation des connaissances (Licence Pluridisciplinaire 2000-2005, 45h, cours et TD)

L'objectif de ce cours est d'amener les étudiants à analyser et formaliser les raisonnements ainsi qu'à structurer les connaissances. Ces étudiants se destinent à entrer en IUFM ou dans les métiers de la documentation.

Programme : logiques classiques, logique modale et temporelle, raisonnement causal, logiques de description, notion d'ontologie.

Représentation des connaissances (DESS GIGA, 1997-99, 25 h, cours/TD)

L'objectif de ce cours est d'initier les étudiants à la problématique de la représentation des connaissances en tant que telle. L'accent est mis sur les problèmes de modélisation et de représentation en liaison avec les approches reposant sur les bases de données ou la méthodologie orientée objet.

Programme logique classiques, logiques de description, graphes conceptuels, Assumption-based truth Maintenance Systems (ATMS).

L'héritage dans les modèles de représentation des connaissances (Option info. du DEA de Linguistique Théorique et Formelle de l'Univ. Paris 7, 1995-96, 20h, cours/TD/TP)

Programme : À travers une présentation comparée de différents formalismes de représentation des connaissances (Common LISP Object System, graphes conceptuels et logiques de description avec CLASSIC), les différentes facettes de l'héritage et son utilité pour le TAL.

Deux articles font le point sur cette expérience de l'enseignement de l'informatique dans une formation doctorale de linguistique [32,33].

4.3 Traitement automatique des langues et des données textuelles

Traitement Automatique des Langues (Formation d'ingénieur 3^{ème} année, option Communication Homme-Machine et Documents Electroniques, en collaboration avec D. Kayser, 2001-05, 15h, cours/TD)

Programme : Historique du TAL et de ses applications ; renouveau des approches à partir de corpus ; statistique textuelle (principes et notions de base) ; morphologie (connaissances et calcul morphologiques) ; syntaxe (connaissances et calcul syntaxiques) ; sémantique lexicale et grammaticale (D. Kayser) ; pragmatique (D. Kayser).

Traitement des données textuelles (DESS Exploration Informatique de Données, 2001-2005, 35h, cours/TD/TP)

Programme : Initiation à Perl, analyse des données textuelles (langagières ou autres), manipulation et exploration de données textuelles.

Option Ingénierie des connaissances textuelles (DEA d'Intelligence Artificielle et Optimisation Combinatoire, 2000-2005, en collaboration avec D. Kayser, S. Szulman et B. Biébow, 20h, cours/TD)

Programme : Analyse morphologique, syntaxique et sémantique, acquisition de connaissances à partir de données textuelles, construction d'ontologie.

Option Langage Naturel (DEA D'Intelligence Artificielle, 1996-1999, en collaboration avec D. Kayser, 20h, cours/TD)

Programme : Etiquetage morpho-syntaxique, modèles et analyse syntaxiques, analyse sémantique et pragmatique ; acquisition de connaissances à partir de données textuelles.

Option Langage Naturel (Licence et maîtrise d'informatique, en collaboration avec D. Kayser, (1996-00, 20h, cours/TD)

Programme : Panorama des applications du TALN et des techniques mises en oeuvre (traduction automatique, aide à la rédaction, documentation, acquisition de connaissances, interface, dictée automatique), formalismes grammaticaux et analyse syntaxique, traitement probabiliste de corpus, sémantique et pragmatique.

Représentation sémantique (DESS d'ingénierie multilingue de l'INaLCO, 1993-96, 20h, cours/TD)

Programme : Apport et limites des modèles sémantiques (Graphes Conceptuels, Théorie des Représentations Discusives...) pour le traitement de la détermination/quantification, du temps, des modalités, des ajouts (circonstanciels).

Introduction aux méthodes et outils statistiques (DESS d'ingénierie multilingue et de traductique de l'INaLCO, 1993-97, 10 h cours/TD)

Programme : méthodologie des études statistiques ; notions de base ; grandes lois de la statistique textuelle ; études de répartition, de distribution et de corrélation.

4.4 Divers

Informatique pour littéraires (Ecole Normale Supérieure de Fontenay, 1993-96)

À l'Ecole Normale Supérieure de Fontenay-St Cloud, j'ai assuré différentes formations, sous forme de stage, pour les étudiants et enseignants de l'Ecole, de formation littéraire pour la plupart.

Programme : bureautique (word, excel, messagerie électronique, navigation sur le web), initiation à Unix (systèmes de fichiers, notion de processus et commandes usuelles), gestion de données textuelles (encodage des caractères, balisage SGML)

Littérature et stylistique française (Univ. Stanford, 1989-1990, 150h)

5. Charges pédagogiques

5.1 Responsabilités pédagogiques

- Responsabilité de la maîtrise d'informatique durant 3 ans (2000-2003)

- Création et responsabilité de l'option Communication Homme-Machine et Documents Electroniques (environ 160 h) en 3^{ème} année de formation d'ingénieur (depuis 2001)

J'ai été membre du bureau de département d'informatique entre 2001 et 2002.

5.2 Définition de programmes

Dans le cadre de la réforme LMD, j'ai prévu une refonte des enseignements de TAL du département d'Informatique de l'Université Paris 13. Cette refonte prendra effet à la rentrée 2005. Le futur Master recherche en Informatique comportera un ensemble d'unités d'enseignement permettant à des informaticiens d'acquérir une solide formation en TAL (notions fondamentales de linguistique, traitements symboliques du TAL, approches statistiques et robustes) et dans des domaines connexes (méthodes d'apprentissage, ingénierie des connaissances).

Lors de la création de la formation d'ingénieur en Informatique de l'Institut Galilée, j'ai défini le programme de l'option de 3^{ème} année « Communication Homme-Machine et Documents Electroniques ». Cette option permet aux étudiants de découvrir de nouveaux champs de l'informatique : quelles voies ouvrent les nouvelles technologies de la langue, de l'image et de la parole pour la communication Homme-Machine ? En plus des cours fondamentaux, les étudiants ont l'occasion de découvrir des applications innovantes (par ex. les hypertextes adaptatifs) et de se former sur des technologies de base (programmation en Perl, XML).

Dans le cadre de mes fonctions d'ATER à l'ENS de Fontenay, j'ai participé en collaboration avec B. Habert à l'enseignement de l'informatique dans l'option Informatique du DEA de Linguistique théorique et formelle de l'Université Paris 7. La réflexion sur la définition du programme et les modalités d'enseignement de l'informatique dans une formation doctorale de linguistique a fait l'objet de deux articles [32][33].

Dans le même contexte, j'ai également été moteur du rapprochement entre l'ENS et les DESS gérés par le Centre de Recherche en Ingénierie Multilingue de l'INaLCO où j'avais demandé à exercer une partie de mon service d'enseignement. Ce rapprochement a débouché sur une collaboration de plus grande ampleur redéfinissant le programme des DESS en y incorporant une formation plus solide en informatique et centrée sur les traitements de corpus qui prenaient à l'époque une importance croissante [10]. J'ai participé avec B. Habert et M. Slodzian à la définition de ce programme, mais je n'ai pas assuré cet enseignement, ayant pris mon poste à l'Université Paris 13 en 1996.

AUTRES ACTIVITES ADMINISTRATIVES

Commissions de spécialistes

Je suis membre suppléant des commissions de spécialistes suivantes :

- Commission des spécialistes de 27^{ème} section à l'université Paris 13 ;
- Commission mixte des spécialistes (27^{ème} section) à l'université Paris 4-Sorbonne.

Participation à la vie du laboratoire

Je suis membre du conseil de laboratoire du LIPN depuis 2002. J'ai été responsable du site web du LIPN de 1998 à 2003.

Sociétés savantes

Je suis membre de l'AFIA (Association Française d'Intelligence Artificielle) et de l'ATALA (Association pour le Traitement Automatique des Langues) depuis 1996 et de l'Association pour la Recherche Cognitive.

Je suis membre du comité d'administration de l'ATALA depuis 1996. L'ATALA est une Association Loi 1901 regroupant environ 350 adhérents français et étrangers. Après avoir été trésorière de cette association (1996-2000), j'ai été responsable (2000-2004) de l'organisation des journées d'études (3 à 6 journées thématiques par an).

LISTE DE PUBLICATIONS

Mon travail de recherche s'est accompagné d'une activité continue de publication :

- 2 ouvrages en français dont 1 en collaboration
- 8 chapitres de livres dont 4 en anglais
- 1 article dans une revue internationale
- 12 conférences et workshops internationaux avec comité de sélection, dont 7 conférences majeures [16][18][19][21][22][23]
- 10 articles de revues francophones
- 15 conférences nationales avec comité de sélection
- Diverses autres publications

1. Ouvrages, chapitres de livres

- [1] **Adeline Nazarenko**. « Sur quelle sémantique reposent les méthodes automatiques d'accès au contenu textuel ? » . In *Sémantique et corpus* (A. Condamines, éd.), Lavoisier, Paris, à paraître en 2004.
- [2] **Adeline Nazarenko** et Thierry Poibeau. « Évaluation des systèmes d'analyse et de compréhension ». In *Evaluation des systèmes de traitement de l'information* (S. Chaudiron, éd.), Lavoisier, Paris, 2004.
- [3] Claire Nédellec and **Adeline Nazarenko**. "Ontology and Information Extraction". In *Handbook on Ontologies in Information Systems* (S. Staab & R. Studer eds.), Springer, 2003. Article non publié à ce jour, suite à une erreur des éditeurs.
- [4] Thierry Hamon and **Adeline Nazarenko**. "Dictionary-based detection of term semantic variation". In *Exploring the Lexicon: Theory and Computation* (Alessandro Lenci & Vittorio Di Tomaso, eds.). Edizioni dell'Orso, 2002.
- [5] Thierry Hamon and **Adeline Nazarenko**. "Detection of synonymy links between terms: experiment and results". In *Recent Advances in Computational Terminology* (D. Bourigault, C. Jacquemin and M.-C. L'Homme, eds.). John Benjamins, 2001.
- [6] **Adeline Nazarenko**, Pierre Zweigenbaum, Benoît Habert and Jacques Bouaud. "Corpus-based Extension of a Terminological Semantic Lexicon". In *Recent Advances in Computational Terminology* (D. Bourigault, C. Jacquemin and M.-C. L'Homme, eds.). John Benjamins, 2001.

- [7] **Adeline Nazarenko**. *La cause et son expression en Français*. OPHRYS, coll. L'essentiel Français, 2000 (160 pages).
- [8] Jacques Bouaud, Benoît Habert, **Adeline Nazarenko** et Pierre Zweigenbaum. « Regroupements issus de dépendances syntaxiques en corpus : catégorisation et confrontation à deux modélisations conceptuelles ». In *Ingénierie des Connaissances, évolutions récentes et nouveaux défis* (J. Charlet, M. Zacklad, G. Kassel et D. Bourigault, Eds.). Eyrolles, 2000, pp. 275-290.
- [9] Philippe Blache, Jacques Guizol, François Lévy, **Adeline Nazarenko**, Sylvain N'Guema, Monique Rolbert, Robert Pasero et Paul Sabatier « Évaluer des systèmes de compréhension de textes ». In *Ressources et Evaluation en Ingénierie de la Langue* (K. Chibout et al. Eds.), Duculot / De Boeck-Université, 2000, pp. 265-275.
- [10] Benoît Habert, **Adeline Nazarenko** et André Salem. *Les linguistiques de corpus*. U Linguistique. Armand Colin/Masson, Paris, 1997 (250 pages).

2. Revues Internationales avec Comité de Sélection

- [11] **Adeline Nazarenko**, Touria Aït El Mekki. « Building back-of-the-book indexes », *Terminology*, special Issue on “Application-driven Terminology engineering”, à paraître en 2005, John Benjamins.

3. Conférences ou Workshops Internationaux avec Comité de Sélection

- [12] Erick Alphonse, Sophie Aubin, Philippe Bessières, Gilles Bisson, Thierry Hamon, Sandrine Lagarrigue, **Adeline Nazarenko**, Alain-Pierre Manine, Claire Nédellec, Mohamed Ould Abdel Vetah, Thierry Poibeau, Davy Weissenbacher, "Event-based Information Extraction for the biomedical domain: the Caderige project". *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and Its Applications (COLING'2004)*, Genève.
- [13] Benoît Habert, **Adeline Nazarenko**, Pierre Zweigenbaum, and Jacques Bouaud. "Extending an existing specialized semantic lexicon". *Proceedings of the First International Conference on Language Resources and Evaluation* (Antonio Rubio, Navidad Gallardo, Rosa Castro, and Antonio Tejada, eds.), pp. 663-668, Granada, 1998. (sélection sur résumé)
- [14] Thierry Hamon and **Adeline Nazarenko**. "Using general semantic information to help the terminology structuration". *Proceedings First International Conference on Language Resources and Evaluation* (Antonio Rubio, Navidad Gallardo, Rosa Castro, and Antonio Tejada, eds.), pp. 675-680, Granada, may 1998. (sélection sur résumé)
- [15] Thierry Hamon, **Adeline Nazarenko** and Cécile Gros. "A step towards the detection of semantic variants of terms in technical documents". *Proceedings of the Eighteenth International conference on Computational Linguistics (Coling'98)*, Montréal, août 1998.
- [16] Pierre Zweigenbaum, Benoît Habert, **Adeline Nazarenko** and Jacques Bouaud. "Tuning an existing nomenclature for specific domain corpora: a syntax-based similarity method". *Journal of the American Medical Informatics Association*, 5(suppl), 1998.
- [17] Benoît Habert, Suzanne Bertrand-Castaldy, **Adeline Nazarenko**, Fernande Dupuis, Elie Naulleau, Monique Lemieux et Cynthia Delisle. « Recyclage d'analyses syntaxiques automatiques pour le repérage de variantes de termes ». *Proceedings of RIAO'97 conference*, Montréal, june 1997.

- [18] **Adeline Nazarenko**, Pierre Zweigenbaum, Jacques Bouaud, and Benoît Habert. "Corpus-based identification and refinement of semantic classes". *Proceedings of the 1997 American Medical Informatics Association (AMIA) Annual Fall Symposium* (Daniel R. Marys, ed.), Nashville, Tennessee, october 1997.
- [19] Benoît Habert, Elie Naulleau and **Adeline Nazarenko**. "Symbolic word clustering for medium-size corpora". *16th International Conference on Computational Linguistics*, volume 1, pp. 490-495, Copenhagen, Denmark, 5-6 August 1996.
- [20] **Adeline Nazarenko**, Benoît Habert, and Christine Reynaud. "Open response surveys : from tagging to syntactic and semantic analysis". *Actes des Journées Internationales en Analyse des Données textuelles*, Rome, Décembre 1995.
- [21] **Adeline Nazarenko**. "Representing natural language causality in conceptual graphs: the higher order conceptual relation problem". In *Conceptual Graphs for Knowledge Representation (Proceedings of the First International Conference on Conceptual structures*, G. W. Mineau, B. Moulin and J. Sowa, eds.), number 699 in Lecture Notes in Artificial Intelligence, pages 205-222. Springer-Verlag, August 1993.
- [22] **Adeline Nazarenko-Perrin**. "Causal ambiguity in natural language: Conceptual representation of 'parce que/because' and 'puisque/since' ". *Proceedings of the Fourteen Conference on Computational Linguistics*, volume 3, pages 880-885, Nantes, août 1992.
- [23] Jean Fargues and **Adeline Nazarenko-Perrin**. "Synthesizing a large concept hierarchy from french hyperonyms". *Proceedings of the Thirteen Conference on Computational Linguistics* (H. Karlgren ed.), volume 2, pages 112-117, Helsinki, 1990.

4. Revues Nationales avec Comité de Sélection

- [24] Touria Aït El Mekki et **Adeline Nazarenko**. « L'index de fin de livre, une forme de résumé indicatif ? », *Traitement Automatique des Langues (TAL)*, 2004, pp. 121-150.
- [25] **Adeline Nazarenko**. « Primitives logiques et réseaux sémantiques : quel bilan pour la terminologie ? ». In *Revue d'Intelligence Artificielle*, Vol 18 n°1/2004 (Techniques informatiques et structuration de terminologies), 2004, pp. 41-56.
- [26] Gaston Gross et **Adeline Nazarenko**. « Quand la langue cause : paramètres d'une étude linguistique de la causalité ». In *Intellectica*, 38(1), 2004.
- [27] **Adeline Nazarenko** et Thierry Hamon. « Structuration de terminologie : quels outils pour quelles pratiques ? ». In *Traitement Automatique des Langues*, n° 43(1), 2002, pp. 7-18.
- [28] Thierry Hamon, Daniela Garcia et **Adeline Nazarenko**. « Détection de liens de synonymie : complémentarité des ressources générales et spécialisées ». *Terminologies nouvelles*, 1999, pp. 61-69.
- [29] Thierry Poibeau et **Adeline Nazarenko**. « L'extraction d'information, une nouvelle conception de la compréhension de texte ? » *Traitement Automatique des Langues*, 40(2), 1999, pp. 87-115.
- [30] **Adeline Nazarenko**. « Le principe de compositionnalité sémantique, un enjeu pour le traitement automatique des langues ». *Traitement Automatique des Langues*, 39(1), pages 3-8, juillet 1998. (numéro thématique "Compositionnalité", coordonné par A. Nazarenko).
- [31] **Adeline Nazarenko**. « Structurer les graphes pour comprendre la causalité dans les textes ». *Revue d'Intelligence Artificielle*, 10(1), pages 163-198, 1996.

- [32] Benoît Habert, Serge Heiden et **Adeline Nazarenko**. « L'informatique dans une formation doctorale de linguistique ». *Traitement Automatique des Langues*, 37(1), pages 29-40, 1996. Numéro spécial « Enseignement du T.A.L. » (dir. Anne Abeillé et Benoît Habert).
- [33] Benoît Habert, Serge Heiden et **Adeline Nazarenko**. « Quelle informatique pour des linguistes, futurs chercheurs en TALN ? » *Traitement Automatique des Langues*, 37(2):174-177, 1996. Numéro spécial « Enseignement du T.A.L. » (dir. Anne Abeillé et Benoît Habert). (Cet article est un complément au précédent)

5. Conférences Nationales avec Comité de Sélection

- [34] Jean Charlet, Touria Aït El Mekki, Didier Bourigault, **Adeline Nazarenko**, Régine Teulier ad Baruk Toledano, « CEDERILIC : constitution d'un livre et d'un index numériques », *Actes de la Conférence Internationale sur le Document Numérique*, 2004.
- [35] Erick Alphonse, Sophie Aubin, Philippe Bessières, Gilles Bisson, Thierry Hamon, Sandrine Lagarrigue, **Adeline Nazarenko**, Alain-Pierre Manine, Claire Nédellec, Mohamed Ould Abdel Vetah, Thierry Poibeau et Davy Weissenbacher. « Extraction d'information appliquée au domaine biomédical: apprentissage et traitement automatique de la langue ». *Actes de la Conférence Internationale sur la Fouille de textes (CIFT'04)*, La Rochelle, mai 2004.
- [36] Touria Aït El Mekki et **Adeline Nazarenko**. « Une mesure de pertinence pour le tri de l'information dans un index de 'fin de livre' », *Actes de la onzième conférence annuelle sur le Traitement Automatique du Langage Naturel (TALN'04)*, Fèz, 2004.
- [37] Touria Aït El Mekki et **Adeline Nazarenko**. « Une mesure de pertinence pour le tri de l'information dans un document », *Plate-forme AFIA Atelier Acquisition, apprentissage et exploitation de connaissances sémantiques pour l'accès au contenu textuel*, Laval, 2003.
- [38] Touria Aït El Mekki et **Adeline Nazarenko**. « Le réseau terminologique, un élément central pour la construction d'index de document ». *Actes des 5èmes rencontres terminologie et Intelligence Artificielle (TIA'03)*, Strasbourg, mars-avril 2003, pp. 1-10.
- [39] Touria Aït El Mekki et **Adeline Nazarenko**. « Comment aider l'auteur à construire l'index d'un ouvrage ? L'architecture du système IndDoc ». *Actes du Colloque International sur la Fouille de Textes (CIFT'02)*, Hammamet, Tunisie, octobre 2002, pp. 141-167.
- [40] Touria Aït El Mekki et **Adeline Nazarenko**. « Quel index pour le document électronique ? ». In *Actes du Colloque International sur le Document Electronique (CIDE'01)*, Toulouse, octobre 2001, pp. 147-161.
- [41] Philippe Bessières, **Adeline Nazarenko** et Claire Nédellec. « Apport de l'apprentissage à l'extraction d'information : le problème de l'identification des interactions géniques ». In *Actes du Colloque International sur le Document Electronique (CIDE'01)*, Toulouse, octobre 2001, pp. 165-183.
- [42] Thierry Hamon et **Adeline Nazarenko**. « Exploitation de l'expertise humaine dans un processus de constitution de terminologie ». In *Actes de la huitième conférence annuelle sur le Traitement Automatique du Langage Naturel (TALN'01)*, Tours, juillet 2001.
- [43] Thierry Hamon et **Adeline Nazarenko**. « La structuration de terminologies » : une nécessaire coopération. *Actes de Terminologie et Intelligence Artificielle*, Nancy, 2001.

- [44] Thierry Hamon, Daniela Garcia et **Adeline Nazarenko**. « Détection de liens de synonymie : complémentarité des ressources générales et spécialisées ». *Actes de Terminologie et Intelligence Artificielle*, pages 45-58, Nantes, 1999.
- [45] Jacques Bouaud, Benoît Habert, **Adeline Nazarenko** et Pierre Zweigenbaum. « Regroupements issus de dépendances syntaxiques en corpus : catégorisation et confrontation à deux modélisations conceptuelles ». *Actes Ingénierie des connaissances*, pages 207-223, Roscoff, 1997.
- [46] Benoît Habert et **Adeline Nazarenko**. « La syntaxe comme marche-pied de l'acquisition des connaissances : bilan critique d'une expérience ». *Journées sur l'acquisition des connaissances*, pages 137-142, Sète, mai 1996. AFIA.
- [47] Pierre Zweigenbaum, Jacques Bouaud, **Adeline Nazarenko** et Benoît Habert. « Coopération apprentissage en corpus et connaissances du domaine pour la construction d'ontologies ». *Actes des 1ères Journées scientifiques et techniques FRANCIL 97*, pp. 501-508, Avignon, avril 1997. AUPELF-UREF.
- [48] **Adeline Nazarenko**. « Une méthode d'étiquetage sémantique pour l'analyse des données textuelles ». In *Actes de la troisième conférence annuelle sur le Traitement Automatique du Langage Naturel (TALN'96)*, pages 228-237, Marseille, mai 1996.

6. Autres Revues

- [49] Gérard Bailly and **Adeline Perrin (Nazarenko)**. "Converging approaches in speech synthesis and automatic translation of text". *Bulletin de l'I.C.P.*, 1989.

7. Autres Conférences

- [50] Touria Aït Ek Mekki and **Adeline Nazarenko**. "Semi-automatic method for the construction of a document index", *EUROLAN-2003 workshop : "Applied Natural Language Processing - possible applications for the Semantic Web -"*.
- [51] Touria Aït Ek Mekki and **Adeline Nazarenko**. "L'index, une représentation synthétique de document". Journée d'étude de l'Association pour le Traitement Automatique des Langues sur le "Résumé automatique", 15 décembre 2002.
- [52] **Adeline Nazarenko**. "Acquisition de connaissances à partir de corpus : élaborer des ressources pour l'extraction d'information", Rencontres «Sémantique et corpus», Université Toulouse Le Mirail, 20 et 21 mars 2002.
- [53] Philippe Bessières, Giles Bisson, Mohamed Ould Abdel Vetah, **Adeline Nazarenko**, Claire Nédellec and Thierry Poibeau. « Ontology learning for information extraction in genomics bibliography - the Caderige project ». Proceedings of the *Workshop on "Ontologies in molecular biology and information extraction from texts"*, Action Informatique, Mathématique et Physique pour la Génomique, Grenoble, avril 2001.
- [54] Claire Nédellec et **Adeline Nazarenko**. « Extraction de connaissances à partir de documents textuels ». Journée d'étude du LIPN "Exploration de données issues d'Internet", 2 mars 2001.
- [55] **Adeline Nazarenko**. « Structuration de terminologie ». Journée d'étude de l'Association pour le Traitement Automatique des Langues sur la "Structuration de terminologies: méthodes et évaluation", 10 mars 2001.

- [56] Benoît Habert and **Adeline Nazarenko**. « émantique distributionnelle : histoire et géographies ». Journée d'étude de l'Association pour le Traitement Automatique des Langues sur les "Approches distributionnelles de l'analyse sémantique", 30 janvier 1999.
- [57] **Adeline Nazarenko**. « Acquisition de connaissances sémantiques ». Journée d'étude de l'Association pour le Traitement Automatique des Langues sur "l'acquisition de connaissances à partir de corpus", janvier 1998.
- [58] Paul Sabatier, Philippe Blache, Jacques Guizol, François Lévy, **Adeline Nazarenko**, Sylvain N'Guema, Robert Guizol, and Monique Rolbert. « Evaluer les systèmes de compréhension de textes ». In *Actes de Journées Scientifiques et Techniques (JST)*, Avignon, France, avril 1997. Francil, AUPELF-UREF.
- [59] **Adeline Nazarenko** et Benoît Habert. « Acquisition de connaissances à partir de corpus ». VIIèmes Journées du LIPN, septembre 1997.
- [60] Benoît Habert et **Adeline Nazarenko**. « Acquisition de classes sémantiques en langue spécialisée et sémantique différentielle ». In *Rencontres de Sémantique et Pragmatique*, Saint-Denis, 1996.
- [61] Benoît Habert, Elie Naulleau, **Adeline Nazarenko** et André Salem. « Construction symbolique puis quantitative des catégories sémantiques d'un corpus ». In *Journées de Sémantique Lexicale Bretoises*, Brest, 1996.
- [62] **Adeline Nazarenko**. « Simuler l'interprétation causale des textes ». In *Actes des 3èmes Journées Jeunes Chercheurs en Sciences Cognitives*, pages 141-152, Presqu'île de Giens, 1996.
- [63] **Adeline Nazarenko**. « Des graphes conceptuels structurés pour représenter la sémantique des phrases complexes ». In *Actes de la journée "Graphes Conceptuels" du P.R.C.-G.D.R. Intelligence Artificielle*, pages 112-142, Montpellier, 25 mars 1994.
- [64] **Adeline Nazarenko**. « Formalisation et extension des graphes conceptuels : les types de concepts spéciaux ». In *Actes des 4èmes journées nationales PRC-GDR IA*, pages 73-82, Marseille, 19-21 octobre 1992. Technea.

8. Mémoires

- [65] **Adeline Nazarenko**. *Donner accès au contenu des documents textuels : acquisition de connaissances et analyse de corpus spécialisés*. Mémoire d'habilitation à diriger les recherches. Université Paris-Nord XIII, 10 décembre 2004. Jury : B. Habert, D. Kayser, F. Lévy, J.M. Pierrel (rapp.), M.T. Paziienza (rapp.) et B. Victorri (rapp.).
- [66] **Adeline Nazarenko** (coord). « SemTech. Acquiring specific semantic knowledge for an access to textual content », Expression d'intérêt déposée dans la cadre du 6^{ème} PPCRD.
- [67] **Adeline Nazarenko**. *Compréhension du Langage Naturel : le problème de la causalité*. Thèse d'Informatique. Université de Paris-Nord XIII, 31 janvier 1994. Directeur: Daniel Kayser.
- [68] **Adeline Perrin (Nazarenko)**. *Elaboration d'un dictionnaire sémantique informatisé : un réseau d'hyperonymes pour la compréhension du langage naturel*. Mémoire de DEA. Centre Scientifique IBM de Paris, Université de Paris-Nord XIII, septembre 1989. Directeur: Daniel Kayser.
- [69] **Adeline Perrin (Nazarenko)**. *Report on fluent : general framework of development*. Technical report, June 1990.

ARTICLES DE RECHERCHE

1. Adeline Nazarenko, Pierre Zweigenbaum, Benoît Habert and Jacques Bouaud. "Corpus-based Extension of a Terminological Semantic Lexicon", In *Recent Advances in Computational Terminology* (D. Bourigault, C. Jacquemin and M.-C. L'Homme, eds.). John Benjamins, 2001
2. Thierry Hamon and Adeline Nazarenko. "Detection of synonymy links between terms: experiment and results", *Recent Advances in Computational Terminology* (D. Bourigault, C. Jacquemin and M.-C. L'Homme, eds.). John Benjamins, 2001.
3. Touria Aït El Mekki, Adeline Nazarenko. « L'index de fin de livre, une forme de résumé indicatif ? », *Traitement Automatique des Langues (TAL)*, à paraître.
4. Erick Alphonse, Sophie Aubin, Philippe Bessières, Gilles Bisson, Thierry Hamon, Sandrine Lagarrigue, Adeline Nazarenko, Alain-Pierre Manine, Claire Nédellec, Mohamed Ould Abdel Vetah, Thierry Poibeau, Davy Weissenbacher, "Event-based Information Extraction for the biomedical domain: the Caderige project". *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and Its Applications (COLING'2004)*, Genève.