

Claire NEDELLEC

Responsable de l'Équipe *Bibliome*
Unité Mathématique, Informatique et Génome
Institut National de la Recherche Agronomique

Née le 24 Juillet 1963 à Brest
Nationalité française
Épouse de *Sainte Marie*, 3 enfants

DIPLÔMES

- 1994 • **Doctorat** en Informatique de l'Université de Paris-Sud (XI), mention Très Honorable, avec Félicitations du Jury.
- 1990 • **DEA IARFA** de l'Université Paris VI, mention Très Bien.
• **Magistère** d'Informatique Appliquée de l'Île de France de l'EHEI (Paris V). (après un DUT d'informatique en 2 ans et magistère en 3 ans)
- 1984 • **Diplôme d'instituteur** à l'École Normale de Quimper, mention Très Bien.

EMPLOIS D'ENSEIGNEMENT ET DE RECHERCHE

- 2002- • **Chercheur 1ère classe**, département de Mathématique et Informatique Appliquées (MIA), Unité Mathématique, Informatique et Génomique (MIG), Centre de Jouy-en-Josas, INRA.
- 2001-2002 • **Détachement à l'INRA**, dans l'unité MIG.
- 2000-2001 • **Délégation au CNRS** section 07 au Laboratoire de Recherche en Informatique (LRI) de l'Université Paris-Sud.
- 1994-2001 • **Maître de Conférences** au département informatique de l'Université Paris-Sud, Laboratoire de Recherche en Informatique (LRI). Stagiaire (1994-1995), 2^e classe (1995-1998), 1^e classe à partir de 1998.
- 1996 - 2001 • **Prime d'Encadrement** Doctoral et de Recherche.
- 1993 - 1994 • **ATER** au département d'informatique de l'IUT d'Orsay.
- 1990 - 1994 • **Étudiante en thèse** au LRI sous la direction de Yves Kodratoff, dans l'équipe Inférence et Apprentissage.
- 1990 - 1993 • **Responsable recherche et développement** de la société ISoft dans le cadre du projet Esprit II intitulé *Machine Learning Toolbox*.
- 1988- 1990 • **Spécification et coordination** du projet *CD ROM Littérature* pour les sociétés Nathan Logiciel et ACT Informatique. Conception et réalisation d'un **générateur de logiciels d'EAO** pour les sociétés Nathan Logiciel et Bordas.
- 1981 à 1985 • **Institutrice** stagiaire puis titulaire.

CHARGES COLLECTIVES

- **Participation aux jurys** de quatre concours INRA : CR2, IR2 et IE2 (2002, 2007 et 2009).
- **Participation aux jurys** de recrutement de maître de conférence Ensimag/LIG de l'Université Joseph Fourier de Grenoble (2010).

- **Membre de la Commission de Spécialistes** du Département Informatique de l'Université d'Orsay (1998-2001).
- **Membre de la Commission de Spécialistes** du Département Informatique de l'Université de Paris-Nord (2005-2008).
- **Membre du conseil de service** de l'unité MIG depuis 2012.
- **Membre suppléant du Conseil Scientifique** du Département INRA Mathématiques et Informatique appliqués (MIA) (2006-2011).
- **En charge du suivi du travail du Conseil Scientifique** du centre de recherche INRA de Jouy-en-Josas (2011-).
- **Membre nommé de la Commission Administrative Paritaire (CAP)** Chercheur de l'INRA (2008-).
- Editrice de la section texte et Connaissance du portail Math-Info de l'INRA depuis 2013.
- **Membre nommé de la Commission Administrative Paritaire (CAP)** Chercheur de l'INRA (2008-).
- **Co-animatrice du CATI** (Centre Automatisé du Traitement de l'Information) ICAT (Ingénierie des Connaissances et Analyse Textuelle) de l'INRA (2012-2015).
- Editrice en chef de la section *Texte et Connaissance* du [portail Math-Info](#) de l'INRA, depuis 2013.

RESPONSABILITES ADMINISTRATIVES D'ENSEIGNEMENT A LA FIIFO (Formation d'Ingénieur d'Ile de France à Orsay) (De 1994 à 2000)

- **Responsabilité du recrutement** niveau baccalauréat des élèves ingénieurs de la NFI informatique, Formation d'Ingénieur en Informatique de la Faculté d'Orsay (FIIFO). (600 dossiers environ, 3 étapes de sélection sur 9 mois, dont entretiens).
- **Responsabilité des études du 1er semestre Fiifo1** et coordination pédagogique avec le DEUG MIAS.
- Participation au **conseil de direction** de la FIIFO.

ENSEIGNEMENT (depuis 2001)

En extraction de connaissance et « text-mining » aux master recherche BIBS (Bioinformatique et Biostatistiques) de l'université Paris 11 et Egoist (BioInforMatique Etudes de Génomes : Outils Informatiques et Statistiques) de l'université de Rouen.

ENCADREMENTS EN RECHERCHE

Permanents de l'équipe Bibliome (unité MIG)

- **Encadrement** de *Nicolas Turenne* (CR2) de septembre 2003 à décembre 2005 sur l'extraction de connaissance temporelle. Logiciel Kaskad. Collaboration avec Marc Barbier (INRA SAD).
- **Encadrement** de *Alain Kotoujansky* (CR1) de mai 2005 à décembre 2009 sur la conception de corpus et l'acquisition de connaissance terminologique en biologie moléculaire des bactéries.
- **Encadrement** de *Robert Bossy* (IR2) depuis août 2007 sur la conception d'Alvis NLP/ML, une chaîne d'annotation sémantique de corpus basée sur l'apprentissage automatique.
- **Encadrement** de *Philippe Veber* (IR2) d'octobre 2009 à décembre 2010 sur l'apprentissage automatique appliqué à l'extraction d'information relationnelle (SPSK). Développement du moteur de recherche AlvisIR, de TyDI (interface d'acquisition termino-ontologique) et de l'Editeur d'annotation AlvisAE avec Frédéric Papazian.

Thèses au LRI

- **Co-encadrement en thèse** à 95 % avec Yves Kodratoff, de *Luc Poittevin* sous convention CIFRE LRI et SNCF, en apprentissage coopératif appliqué à la conception et maintenance d'un serveur minitel pour la Caisse d'Assurance Maladie et de Retraite de la SNCF, depuis octobre 1994. Date de soutenance : 11 Septembre 1998, mention très honorable. Ce travail a donné lieu à 5 publications dont 2 internationales.

- **Encadrement** de *David Faure* au LRI en tant qu'**ingénieur** de recherche et développement, sur l'application de l'apprentissage automatique au traitement du langage naturel, en partenariat avec la société Dassault Aviation, (Projet ESPRIT ILP 2) (Janvier 97 à Octobre 98). Puis **Co-encadrement de thèse** à 95 % avec Yves Kodratoff, sur l'apprentissage automatique d'ontologie de schémas de verbes et leur application à l'extraction de connaissances dans les texte, depuis octobre 98. Date de soutenance : 20 décembre 2000, mention très honorable. Ce travail a donné lieu à 9 publications dont 4 internationales.
- **Co-encadrement de thèse** à 95 % avec Yves Kodratoff (HDR), de *Mohammed Abdel Ould Vetah* sous convention CIFRE LRI et la société VALIGEN, en apprentissage appliqué à l'extraction de connaissance sur les interactions entre gènes à partir de texte, depuis septembre 2000. Date de soutenance : 21 septembre 2005, mention très honorable. Ce travail a donné lieu à 6 publications dont 2 internationales.

Thèses à MIG

- **Co-encadrement de thèse** avec Philippe Bessières (MIG), de *Alain-Pierre Manine* (ASC INRA bioinformatique), en apprentissage appliqué à l'annotation sémantique de texte en génomique, depuis novembre 2002. Date de soutenance 25 juillet 2006, mention très honorable. Ce travail a donné lieu à 4 publications dont 3 internationales.
- **Co-encadrement de thèse** avec Thierry Poibeau (Lattice, ENS Montrouge), de *Zorana Ratkovic*, sur l'acquisition de connaissance lexico-sémantiques par analyse distributionnelle, depuis septembre 2010. Ce travail a donné lieu à 3 publications internationales dont un journal (BioNLP, CICling, BMC Bioinfo).
- **Co-encadrement de thèse** avec Eric Gaussier (AMA, LIG, Université Joseph Fourier), de *Pierre Warnier*, sur l'apprentissage automatique de connaissances à partir de corpus, depuis décembre 2010. Ce travail a donné lieu à 3 publications internationales dont un journal (BioNLP, BMC Bioinfo).
- **Co-encadrement de la thèse** de *Dialekti Valsamou* avec Pierre Zweigenbaum (ILES, LIMSI, Université Paris-Sud), sur l'apprentissage automatique appliqué à l'extraction de relations à partir de corpus en biologie moléculaire, depuis janvier 2012.

Rapports de thèse

- **Rapporteur** de la thèse de *Lubomir Popelinsky* (Université Masaryk de Brno, République Tchèque) intitulée "Towards practical inductive logic programming", 2000.
- **Rapporteur** de la thèse de *Stefano Ferrili* (Université de Bari, Italie) intitulée "A Framework for Incremental Synthesis of Logic Theories: An Application to Document Processing." mars 2001.

Encadrement de post-doctorants

- **Encadrement** de *Lola Canamero* (1 an, de mars 2000 à mars 2001) et de *David Faure* (9 mois, de décembre 2000 à août 2001) dans le cadre du contrat RNRT Astuxe (voir ci-dessous) sur le thème de l'apprentissage et l'extraction de connaissance dans des courriers électroniques.
- **Encadrement** de *Erick Alphonse* (2 ans, de septembre 2003 à septembre 2005) dans le cadre du contrat RNTL ExtraPloDocs (voir ci-dessous) sur le thème de l'application de la PLI à l'EI.
- **Encadrement** de *Robert Bossy* (18 mois, de septembre 2005 à août 2007) dans le cadre des contrats européen Alvis et Epipagri (voir ci-dessous), sur le thème de l'annotation sémantique de documents.
- **Encadrement** de *Laurent Gillard* (1 an, de mai 2009 à mai 2010) dans le cadre du contrat Quaero, sur le thème de l'extraction de relation dans les documents.

Accueil de chercheur en détachement

- *Alain Lelu*, professeur au département informatique de l'université de Besançon sur le thème de la classification de documents. De septembre 2002 à août 2004.

Principaux stages et CDD

- **Encadrement** du CDD IE de *Paul Bui-Quang*, Ontologie et moteur de recherche sémantique sur l'implantation des cultures (fév 2012- août 2012, 7 mois). Projet GIS GC HP2E.

- **Encadrement** du CDD IE de *Dialekti Valsamou*, Extraction d'information sur les marqueurs génétiques du blé (déc 2010- nov 2011, 1 an). Projet FSOV SAM.
- **Co-Encadrement** du CDD IE de *Pierre Warnier* avec Robert Bossy, Classification de documents (juin 2010-, 7 mois). Projet Quaero.
- **Co-Encadrement** du CDD IE de *Julien Jourde* avec Philippe Veber, Reconnaissance d'entités nommées et renommage (sept 2008 - dec 2012). Une publication internationale (BioNLP'11). Projet Quaero.
- **Encadrement** du CDD IE de *Wiktorija Golik*, recherche et développement en traitement automatique de la langue et ingénierie des connaissances (sept 2009-juillet 2013). 5 publications internationales (EKAW, MTSR, BMC Bioinformatics, CiCLING, BioNLP) 3 publications nationales (TIA). Projet Quaero. Projet TriPHASE
- **Co-encadrement** du CDD IE de *Frédéric Papazian*, Reconnaissance d'entités nommées RenBio. Développement en IHM (terminologie TyDI et sémantique distributionnelle Asium (fév 2008-sept 2011). Avec Philippe Veber : Moteur de recherche AlvisIR. Editeur d'annotation. 3 publications internationales (LAW'11, BioCreative, MTSR) 2 nationales (TIA, *In_Ovive*) Projet Quaero.
- **Co-encadrement** avec Sophie Aubin du CDD IE de *Nicolas Rozanès*, Extraction d'entités nommées (juin 2009- août 2009, 4 mois). Projet Quaero.
- **Co-encadrement** avec Robert Bossy du CDD AI de *Ludovic Patey*, IHM pour la recherche documentaire (juin 2009- sept 2009, sept 2010 5 mois). Projet Quaero.
- **Co-encadrement** avec Robert Bossy du CDD AI de *Jasmeet Singh Jaggi*, Classification de documents (mai 2009- juillet 2009, 3 mois). Projet Quaero.
- **Co-encadrement** avec Sophie Aubin du CDD AI de *Sandra Makuntina*, acquisition de terminologie en agrobiotechnologie (3 mois été 2008). Projet Epipagri. Le stage se poursuit en 2008-2009 en parallèle avec le M1 (12 mois). Projet Quaero.
- **Encadrement** CDD IE2 de *Audrey Lemaçon*, modélisation d'ontologie en biologie (6 mois avril 2008 – sept 2008). Projet Quaero. Contrat d'apprentissage INRA pour les deux ans de formation en alternance du master bioinformatique Egoist de l'Université de Rouen (2 ans). Co-encadrement avec Philippe Bessières. Contribution au développement du logiciel CoCitation¹. Soutenance juillet 2010, mention Bien.
- **Encadrement** CDD IE2 de *Sophie Aubin*, recherche et développement en traitement automatique de la langue (sept 2002 – fév 2004, 17 mois, puis fév 2008 – sept 2009, 20 mois). 6 publications internationales (BioCreative, EKAW, SWAT4LS, IIA, RANLP, NLPBA), 3 publications nationales (CIFT, TALN, TIA). Projets ExtraPloDocs puis projet Quaero.
- **Co-encadrement** avec Robert Bossy du CDD IE2 de *Frédéric Papazian*, Reconnaissance d'entités nommées (fév 2008 – sept 2011, 40 mois). Projet Quaero.
- **Co-encadrement** avec Robert Bossy du CDD IE2 de *Sarhane Safraoui*, développement Web dans le domaine de la PI (7 mois en 2008). Projet Epipagri.
- **Co-encadrement** avec Robert Bossy du CDD IE2 de *Brice Durot*, sur la représentation des termes (6 mois en 2007). Projet ALVIS.
- **Encadrement** du CDD IR2 de *Aïcha Bourgognon*, sur l'acquisition de terminologie en microbiologie (6 mois en 2006-2007). Projet ALVIS.
- **Encadrement** du CDD IR2 de *Jérôme Azé*, sur l'apprentissage d'ontologie (10 mois en 2004-2005). Projet ALVIS.

¹ http://genome.jouy.inra.fr/migale/alemacon/public_html/cgi-bin/cocitations/

- **Encadrement** du CDD IR2 de *Julien Gobeill*, pour le challenge LLL (voir ci-dessous) (4 mois en 2005). Projet ALVIS.
- **Encadrement** du CDD IE2 de *Gaël Deral*, sur l'intégration de l'apprentissage dans une chaîne d'annotation sémantique (6 mois en 2004-2005). Projet ALVIS.
- **Encadrement** du CDD IE2 de *Gaëtan Lehman*, sur le filtrage de phrases pertinentes et l'apprentissage pour la reconnaissance d'entités nommées (6 mois en 2003). Projet ExtraPloDocs.
- **Encadrement** du stage de DEA de l'INALCO de *Sophie Aubin*, sur le choix et l'intégration d'un analyseur syntaxique pour l'apprentissage de connaissances sémantiques (2 mois en 2002). Projet ExtraPloDocs.
- **Encadrement** du stage de DEA I3 (Orsay) de *Zihyu Qian*, sur l'influence de la terminologie sur le filtrage de document et l'apprentissage d'ontologie (6 mois en 2002). Projet Caderige.
- **Encadrement** du stage de DEA I3 (Orsay) de *Mathieu Roche*, sur l'apprentissage d'ontologie en support technique logiciel (6 mois en 2001). Projet Astuxe.
- **Encadrement** du stage ingénieur de *Jérôme Maloberti* (FIIFO 5) au LRI, sur le développement d'une bibliothèque de formats d'entrée/sorties (FOL, GML, etc.) pour ASIUM (6 mois en 1999). Projet ILP II.
- **Co-encadrement** du stage ingénieur de *Mickaël Magnier* (FIIFO 5), LRI et Dassault Aviation, sur l'extraction d'ontologies et de schémas de verbe par apprentissage automatique à partir de textes de spécifications exprimés en langage naturel (5 mois en 1998). Projet ILP II.
- **Encadrement** du stage ingénieur de *François-David Collin* (Supelec, 3e année) au LRI, sur la désambiguïsation de l'analyse syntaxique de SYLEX (Ingénia) grâce aux schémas de sous-catégorisation représentés sous forme d'une base de données relationnelle (6 mois en 1998). Projet ILP II.
- **Encadrement** de *Thomas Jamet*, **stagiaire** du magistère MIAIF (1992), en **CDD ingénieur** chez Isoft (12 mois), puis **stagiaire** du DEA Paris VI (1993) sur une application de l'apprentissage automatique au domaine bancaire. Ce travail a donné lieu à une publication internationale (ECML'93). Projet MLT.

ANIMATION DE LA RECHERCHE ET COLLABORATIONS

Visites

- **Chercheur invité** au Knowledge System Laboratory (Stanford University, USA) sur l'apprentissage automatique de plans dans l'équipe du Pr. B. Hayes-Roth (1 mois et demi, en 1992).
- **Chercheur invité** au Computer Science Division, University of California à Berkeley (USA) dans l'équipe du Pr. Stuart Russell, sur l'identification et la modification automatique des connaissances de contrôle (ou biais), en apprentissage automatique (3 mois, en 1993).

Comité

- Comité du Ministère des Affaires Etrangères mixte franco-américain pour la science et la technologie, domaine bioinformatique sous la direction d'Alain Viari (depuis 2013).

Editeur de Journal International

- *BMC Bioinformatics*, BioMed Central (Pub.). Numéro spécial dédié à BioNLP Shared Task 2011, parution en juin 2012.

Participation au Comité d'Édition de Journaux Internationaux

- *Intelligent Data Analysis - An International Journal*, F. Famili (Ed.), Elsevier Science (Pub.). Depuis 1996. IOS Press Publisher depuis 1999. (1996-2007).
- *Bioinformatics Journal*, Elsevier Science (Pub.). Issue spéciale BioCreative III, 2010.

Participation au Comité d'Édition de Journaux Nationaux

- Revue TAL (Traitement Automatique des Langues). Numéro spécial sur la reconnaissance des entités nommées. <http://www.atala.org/> Editrices en chef, Sophia Ananiadou, Nathalie Friburger, Sophie Rosset, 2013.

Participation au Comité d'Édition de Journaux Électroniques

- *READ - Revue Électronique, l'Apprentissage par les Données*, G. Venturini (Ed.), co-responsable du thème "Apprentissage Automatique et Interactivité". 1996 - 2000.
- *The semantic web*, *Electronic Transaction on Artificial Intelligence*, G. Schreiber (Ed.). 2000 - 2001.

Co-présidence de Conférence Internationale

- *European Conference on Machine Learning*, ECML'98, co-présidée avec C. Rouveirol. 10^{ème} édition, à Chemnitz, (Allemagne) en Avril 1998. Principale conférence européenne d'apprentissage automatique. Actes publiés chez Springer Verlag (Lecture Notes).

Editeur (area chair) de Conférence Internationale

- *Joint 19th Conference on Machine Learning and 12th Conference on Principles and Practice of Knowledge Discovery in Databases (ECM/PKDD)*, Bled, Slovénie, septembre 2009.
- *Joint 18th Conference on Machine Learning and 11th Conference on Principles and Practice of Knowledge Discovery in Databases (ECM/PKDD)*, Warsaw, septembre 2007.

Organisation de challenges internationaux

- Présidence du challenge LLL, *Learning Language in Logics*, 2005. <http://genome.jouy.inra.fr/texte/LLLchallenge/>
- Membre du Scientific Advisory Committee du challenge *BioNLP 2011 shared task*, 2011. <https://sites.google.com/site/bionlpst/>.
- Présidence *BioNLP 2013 shared task*, 2013. <https://sites.google.com/site/bionlpst/>.

Présidence et Co-présidence d'Ateliers Internationaux

- Atelier intitulé "*Integration of Machine Learning and Knowledge Acquisition*", associé à la European Conference on Artificial Intelligence (ECAI), co-organisé avec Y. Kodratoff, Amsterdam, (Pays-Bas), août 1994.
- Atelier intitulé "*Knowledge Level Modeling of Machine Learning*", European Conference on Machine Learning (ECML), co-organisé avec D. Fensel et al., Heraklion, (Grèce), avril 1995.
- Atelier intitulé "*Machine Learning and Comprehensibility*", associé à la International Joint Conference on Artificial Intelligence (IJCAI), co-organisé avec Y. Kodratoff, Montréal, (Canada), août 1995.
- Atelier intitulé "*Ontologie Learning*", associé à la European Conference on Artificial Intelligence (ECAI), co-organisé avec A. Maedche, S. Staab, et P. Wiemer-Hastings, Berlin, août 2000.
- Atelier intitulé "*Learning Language in Logic*" 2nd édition, associé à la *Internationale Conference on Grammatical Inference (ICGI)* et la *Conference on Natural Language and Logic (CoNLL)*, Actes publiés chez Omni Press. Lisbonne, septembre 2000.
- Atelier intitulé "*Ontology Learning*" 2nd édition, associé à la Internationale Joint Conference on Artificial Intelligence (IJCAI-2001), co-organisé avec S. Staab, A. Maedche, et E. Hovy, Vancouver, Canada, août 2001.
- Atelier intitulé "*LLL, Learning Language in Logics*", associé à ICML, International Conference on Machine Learning, co-organisé avec James Cussens, Bonn, août 2005.

- Atelier intitulé "*Data and text mining in integrative biology*", associé à ECML/PKDD, European Conference on Machine Learning/European Conference on Principles and Practice of Knowledge Discovery, co-organisé avec Mélanie Hilario, Berlin, septembre 2006.

Participation au comité d'organisation scientifique d'Ateliers Internationaux

- Atelier intitulé "*Ontology engineering*", associé à la European Conference on Artificial Intelligence (ECAI), présidé par N. Aussenac et A. Maedche, Lyon, août 2002.

Co-présidence d'Ateliers Français

- Atelier intitulé "*Apprentissage Statistique et Symbolique pour la Recherche d'Information et la Fouille de Textes*", co-organisé avec P. Gallinari et C. Jacquemin sur la plate-forme AFIA'99 Ecole Polytechnique, Palaiseau, Juin 1999. (Voir <http://www.lri.fr/~cn/JASSRIFT>).
- Atelier intitulé "*Journée A3CTE, Applications, Apprentissage et Acquisition de Connaissances à partir de Textes Electroniques*", co-organisé avec A. Nazarenko sur la plate-forme AFIA'2001, Grenoble, juin 2001. (Voir <http://www-leibniz.imag.fr/AFIA>).
- Atelier intitulé "Conférence Internationale en Fouille de Texte" (CIFT) co-organisé avec Y. Toussaint (LORIA) sur la plate-forme CIDE (Conférence Internationale sur le Document Electronique), Hammamet, Tunisie, septembre 2002.
- Atelier intitulé "In-Ovive : *Intégration de sources/masses de données hétérogènes et Ontologies, dans le domaine des sciences du VIVant et de l'Environnement*", associé à la Plateforme AFIA, Villeneuve d'Asq, 2 juillet 2013. https://colloque.inra.fr/in_ovive_2013

Participation aux Comités de Programme de Conférences Internationales

ECAI

- *20th European Conference on Artificial Intelligence (ECAI)*, Ed. Luc de Raedt, Montpellier, France, 27-31 Août 2012.

ECML, PKDD, ICML

- *European Conference on Machine Learning (ECML)* : 2000 et 2001.
- *European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD)* : 2003 et 2004.
- *Joint 17th Conference on Machine Learning and 10th Conference on Principles and Practice of Knowledge Discovery in Databases (ECM/PKDD)*, Berlin, septembre 2006.
- *Joint 19th Conference on Machine Learning and 12th Conference on Principles and Practice of Knowledge Discovery in Databases (ECM/PKDD)*, Anvers, Belgique, septembre 2008.
- *Joint 20th Conference on Machine Learning and 13th Conference on Principles and Practice of Knowledge Discovery in Databases (ECM/PKDD)*, Bled, Slovénie, septembre 2009.
- *Joint 21th Conference on Machine Learning and 14th Conference on Principles and Practice of Knowledge Discovery in Databases (ECM/PKDD)*, Barcelone, Espagne, septembre 2010.
- *19th International Conference on Machine Learning, (ICML'2002)*, présidée par C. Sammut, Sydney, juillet 2002.
- *22th International Conference on Machine Learning, (ICML'2005)*, présidée par L. De Raedt et S. Wrobel, Bonn, août 2005.
- *23rd International Conference on Machine Learning, (ICML'2006)*, présidée par W. Cohen, Pittsburgh, juin 2006.
- *27th International Conference on Machine Learning, (ICML'2010)*, présidée par T. Joachims et J. Fuernkranz, Haïfa, Israël, août 2010.

Data-mining

- *Second International Symposium on Intelligent Data Analysis (IDA'97)*, présidé par X. Liu et P. Cohen, Londres, août 1997.

- *The First International Conference on Advanced Data Mining and Applications* (ADMA 2005), Wuhan, China, juillet 22-24, 2005.
- *The Second International Conference on Advanced Data Mining and Applications* (ADMA 2006), Xi'An, China, August 14-16, 2006.
- *The Third International Conference on Advanced Data Mining and Applications* (ADMA 2009), Beijing, 2009, China, 17-19 août, 2009.
- *8th International Conference on Data Warehousing and Knowledge Discovery - DaWaK '06* Krakow, Poland, 4 - 8 septembre 2006.
- *First edition of the International Conference on Pattern Recognition, Applications and Methods*, ICPRAM. 2012, présidée par Salvador Sánchez, Jaume et Pedro Latorre Carmona, 6-8 février 2012, Vilamoura, Portugal.
- *Second edition of the International Conference on Pattern Recognition, Applications and Methods*, ICPRAM. 2013, présidée par Ana Fred, 15-18 février 2013, Barcelone, Espagne.
- *Third edition of the International Conference on Pattern Recognition, Applications and Methods*, ICPRAM. 2014, présidée par Ana Fred, 6-8 mars 2014, Angers, France.

Semantic Web

- *International Semantic Web Conference*, D. Fensel (Ed), octobre 2003, Etats-Unis.
- *Artificial Intelligence: Methodology, Systems, and Applications, 11th International Conference*, (AIMSA 2004), C. Bussler, D. Fensel (Eds.), Varna, Bulgarie, septembre 2-4, 2004.
- *Semantic Web Track of the 16th International World Wide Web Conference (WWW2007)*, R. Studer, R. Mizoguchi (Eds), Banff, Canada, 8-12 mai, 2007.
- *Semantic / Data Web Track of the 17th International World Wide Web Conference (WWW2008)*, R. Guha, R. Mizoguchi (Eds), Pékin, Chine, 22 avril, 2008.

Bioinformatique

- *European Conference on Computational Biology (ECCB08)*, Cagliari, Italie, 22-26 septembre 2008.

Apprentissage homme-machine

- *Third International Conference on Human-System Learning (CAPS'3, Learning WWW)* présidée par K. Zreik, Paris, décembre 2000.
- *Fourth International Conference on Human-System Learning (CAPS'4, Learning WWW)*. présidée par R. Scott, et L. McKechnie (Eds) Glasgow, UK, 2003.
- *International Conference on Human System Learning (ICHSL'05)*, Marrakech, Maroc, 22-25 novembre 2005.

Traitement Automatique de la Langue Naturelle

- *11th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2006)*, Trento, Italie, 3-7 avril, 2006

Conférences Nationales

- *8th Ibero-American Conference on Artificial Intelligence IBERAMIA'02*, Seville, Espagne, novembre 2002.

Participation à l'organisation scientifique de Conférences Françaises

- *Journées Francophones d'Apprentissage* présidées par J. Nicolas, Lens, mai 1998.
- Participation au comité éditorial IA de la conférence *Recherche des Formes et Intelligence Artificielle (RFIA)*, présidée par R. Deriche et M.C. Rousset, Paris, février 2000.
- Participation au comité d'organisation de la *plateforme AFIA 2001, L'IA, l'Informatique de la Connaissance*, coordonnée par N. Balacheff, Grenoble, juin 2001.

Participation aux Comités de Programme de Conférences Françaises

- Journée commune des *Journées Francophones d'Apprentissage, de Validation et d'Acquisition (JAVA)* présidée par C. Rouveirol, Grenoble, avril 1995.
- *Journées Explications*, présidées par M. Joab et A. Giboin, Juan Les Pins, juin 1996.
- Co-responsable du thème *Ingénierie des Connaissances et Apprentissage Automatique* des Journées Ingénierie de Connaissances (IC, ex-JAC), présidées par M. Zacklad, Roscoff, mai 1997.
- Journées AFIA, *Jeunes Chercheurs en Intelligence Artificielle (RJCIA)*, présidées par P. Marquis, Toulouse, Septembre 1998.
- Conférence *Recherche des Formes et Intelligence Artificielle (RFIA)*, depuis 2002 (4 éditions).
- Conférence *Web'2002, (Conférence Francophone sur la Société de l'Information et de la Connaissance)*, présidé par M. Quafafou, Villetaneuse, Mars 2002.
- *Colloque International sur la Fouille de Texte (CIFT'04)* présidé par M.-H. Antoni et F. Yvon, associé à la SDN'0, La Rochelle, juin 2004.
- *Conférence Francophone d'Apprentissage Automatique (CAP)*, de 2001 à 2006 et 2010 (7 éditions).
- Journées *Ingénierie de Connaissances (IC)* de 1998 à 2000 (3 éditions).
- *XVIèmes rencontres de la Société Francophone de Classification*, Grenoble, 2 au 4 septembre 2009.
- *XVIIèmes rencontres de la Société Francophone de Classification*, Orléans, 28 au 30 septembre 2011.

Participation à des Comités de Programme d'Ateliers Internationaux

Apprentissage automatique, data-mining

- *Declarative Bias*", de la European Conference on Machine Learning (ECML), présidé par C. Rouveirol, Catane, (Italie), avril 1994.

Apprentissage et texte

- *Text Mining*", associé à European Conference on Machine Learning (ECML), présidé par Y. Kodratoff, Chemnitz, (Allemagne), avril 1998.
- *Learning Language in Logic*", associé à International Conference on Machine Learning (ICML), présidé par J. Cussens, Bled, (Slovénie), juin 1999.
- *Learning Language in Logic (LLL)*, associé à Inductive Logic Programming Conference (ILP) et à European Conference on Machine Learning (ECML), présidé par L. Popelinsky, Fribourg, (Allemagne), septembre 2001.
- *High-level Information Extraction*, associé à European Conference on Machine Learning (ECML), présidé par S. Blohm et al., Anvers, (Belgique), 29 septembre 2008.

Apprentissage et conception d'ontologie

- *Ontology Engineering*, présidé par N. Aussenac et A. Maedche, associé à ECAI, Lyon, août 2002.
- *Ontology Learning and Population - Towards Evaluation of Text-based Methods in the Semantic Web and Knowledge Discovery Life Cycle*, associé à European Conference on Artificial Intelligence (ECAI), présidé par Paul Buitelaar et al., Valence, Espagne, 22-23 juillet 2004.
- *Knowledge Discovery and Ontologies Workshop (KDO 2004)* associé à ECML/PKDD, 2004.
- *Learning and extending lexical ontologies by using machine learning methods*, (OL'2005) associé à ICML 2005.
- *2nd Workshop on Ontology Learning and Population (OLP), Bridging the Gap between Text and Knowledge*, Workshop at COLING/ACL, Sydney, Australie, 22 juillet 2006.

Data-mining and ontology

- *Information Integration and Mining in Databases and on the Web (DBFusion)*, présidé par Gerd Stumme, Karlsruhe, 4-5 juillet 2002.
- *VLDB Workshop on Ontologies-based techniques for DataBases and Information Systems (ODBIS)* joint avec VLDB, Trondheim, Norvège, 2-3 septembre, 2005.
- *VLDB Workshop on Ontologies-based techniques for DataBases and Information Systems (ODBIS)* joint avec VLDB, Seoul, Corée, 11 septembre, 2006.

- *4th ODBIS Workshop on Ontologies-based Techniques for DataBases in Information Systems and Knowledge Systems Co-located with VLDB 2008*, Auckland, Nouvelle Zélande, août 2008.

Web

- *Second International Workshop on the Semantic Web - SemWeb'2001*, associé à la WWW10, présidé par Decker S., Fensel D., Sheth A. et Staab S., Hongkong, mai 2001.
- *Semantic web mining*, associé à European Conference on Machine Learning (ECML) et Conference on Practical Knowledge Discovery in Data (PKDD), présidée par Gerd Stumme, Freiburg, septembre 2001.
- *Semantic web mining*, associé à European Conference on Machine Learning (ECML), présidé par Gerd Stumme, Helsinki, août 2002.
- *1st European Web Mining Forum* associé à European Conference on Machine Learning (ECML) et Conference on Practical Knowledge Discovery in Data (PKDD), présidée par A. Hotho, Dubrovnik, Croatia, 2004.
- *Semantic Web personalization* associé à AAAI, San Jose, Etats-Unis, 25-26 juillet 2004.
- *International Workshop on Intelligent Information Access (IIIA) workshop*, Helsinki, Finlande, 6-8 juillet 2006. <http://cosco.hiit.fi/search/IIIA2006/>
- *5th Workshop on Intelligent Techniques for Web Personalization (ITWP)*, joint to AAAI-07, Vancouver, Canada, 23 juillet 2007.
- *6th Workshop on Intelligent Techniques for Web Personalization (ITWP)*, joint to AAAI-08, Chicago, Etat-Unis, juillet 2008.
- *1st International Workshop on Recommender Systems and Personalized Retrieval (RSPR 2008)* associé à First IEEE International Conference on the Application of Digital Information and Web Technologies (ICADIWT 2008) présidée par G. Semeraro et al., Dubrovnik, Ostrava (République Tchèque), août 2008.
- *7th Workshop on Intelligent Techniques for Web Personalization (ITWP)*, joint to IJCAI-09, Chicago, Etat-Unis, juillet 2009.
- *8th Workshop on Intelligent Techniques for Web Personalization (ITWP)*, joint to UMAP, Hawaï, Etat-Unis, juillet 2010.
- *9th Workshop on Intelligent Techniques for Web Personalization* in conjunction with IJCAI'11, Barcelone, Espagne, 16 juillet 2011.

Apprentissage homme-machine

- *Machine Learning meets Human-Computer Interaction*", associé à International Conference on Machine Learning (ICML), présidé par V. Moustakis et J. Herrmann, Bari, (Italie), juillet 1996.

TALN

- *Joint Workshop on Natural Language Processing in Biomedicine and its Applications (NLPBA/BioNLP)*, joint avec Coling, Genève, Suisse, 28-29 août, 2004.

Bioinformatique

- *Second BioCreative challenge workshop*, Madrid, (Espagne), présidé par A. Valencia et al., mai 2007.
- *Third BioCreative challenge workshop*, Madrid, (Espagne), présidé par A. Valencia et al., octobre 2009.
- *Building and evaluating resources for biomedical text mining workshop* associé à International Conference on Language Resources and Evaluation (LREC), présidée par S. Ananiadou, Marrakech (Maroc), 26 mai 2008.
- *DTMBIO 2011 : ACM Fifth International Workshop on Data and Text Mining in Biomedical Informatics*. In conjunction with CIKM, (conference in information and knowledge management). Présidé par Doheon Lee et Sophia Ananiadou Glasgow, 24 oct 2011.
- *MIND 2011 : Workshop on Mining Complex Entities from Network and Biomedical Data*, présidé par Stefano Ferilli et al., Athens, Greece, 9th of September 2011.
- *DTMBIO 2012 : ACM Sixth International Workshop on Data and Text Mining in Biomedical Informatics*. In conjunction with CIKM, (conference in information and knowledge management). Présidé par Doheon Lee et Sophia Ananiadou, Hawaii, USA, 29 Octobre 2012.

- *BioNLP 2013 : BioNLP Shared Task 2013 Workshop*, Association for Computational Linguistics. Sofia, Bulgarie, Août 2013.
- *DTMBIO 2013 : ACM Seventh International Workshop on Data and Text Mining in Biomedical Informatics*. In conjunction with CIKM, (conference in information and knowledge management). Présidé par Doheon Lee et Atul Butte, San Francisco, USA, 1^{er} Novembre 2013.
- *BioASQ, Biomedical Question Answering Workshop*, Post-conference workshop après CLEF 2013, présidé par George Paliouras 27 septembre, Valence, Espagne.

Participation à des Comités de Programme d'Ateliers Nationaux

- *Journée informatique et transcriptome* jointe aux *Journées Post-Génomiques de la Doua (JPGD)*, Lyon, avril 2003.
- *Ontologies et Jeux de Données pour évaluer le web sémantique (OJD)* adossé à IC 2012, Paris, 26 juin 2012.
- Atelier intitulé " *Qualité et Robustesse pour le web de données*» (Q&R)", associé à la conférence IC (Ingénierie des Connaissances), 2 juillet 2013.

Principales relectures d'articles

- Depuis 1991, principales **conférences** nationales et internationales en Intelligence Artificielle et en Apprentissage : JAVA, IC, JFA, RFIA, JOBIM, ECML, EKAW, IDA, ECAI, ML, IJCAI, etc., et en traitement du document électronique, RIAO.
- Depuis 1994, **journaux** nationaux et internationaux :
 - *Revue d'Intelligence Artificielle (RIA)*, M. O. Cordier (Réd.), Hermès.
 - *Knowledge Engineering Review*, J. Fox (Ed), Cambridge University Press.
 - *Machine Learning Journal*, L. Saitta (Ed), Kluwer Academy.
 - *Intelligent Data Analysis - An International Journal*, F. Famili, (Ed), IOS Press.
 - *Journal Traitement Automatique des Langues*, C. Jacquemin, (Réd., n° spécial), ATALA/CNRS.
 - *ACM TOIT, special issue on Intelligent Techniques for Web Personalization*
 - *No spécial I3 : textes et RTO, Textes et ressources terminologiques et/ou ontologiques : évolution et maintenance* (2006). Numéro spécial "RFIA'08" (2008).
 - *Journal of Web Semantics* (Elsevier) (2008)
 - *AI Comm.*
 - *Journal of Bioscience and Bioengineering*, JBB.
 - *Data and Knowledge Engineering Journal*, DKE
 - *IEEE/ACM Transactions on Computational Biology and Bioinformatics(TCBB)* (2009)
 - *Briefings in Bioinformatics* (2010)
 - *Bioinformatics* (2010, 2011, 2012)
 - *Journal of Intelligent Information Systems*, JIIS (2012)
 - *TOIS, Transaction in Information Systems* (2012)
 - *PLOS One* (2013)

Animation de Groupe de Recherche

- Création et animation avec A. Nazarenko du groupe de recherche A3CTE, "*Application, Acquisition et Apprentissage de Connaissances dans les Textes spécialisés*", 13 réunions de octobre 1998 à mars 2003 (80 inscrits, 15 à 50 participants par réunion). Soutenu par le GDR-PRC I3 (thème Langue Ecrite et thème Apprentissage) et par le RISC (Réseau d'Ile de France en Science Cognitive). Groupe AFIA. Site <http://www-lipn.univ-paris13.fr/groupe/A3CTE>.
- Création et animation avec Odile Hologne (dir. INRA DV/IST) et Philippe Bessières (dépt INRA MICA) du réseau TC, "*Texte et Connaissance*". Organisation du premier séminaire les 3-4 juillet 2008 « méthodes, besoin, perspectives » au CR de Jouy-en-Josas (51 participants). Site <http://genome.jouy.inra.fr/bibliome/TC/>

Conférences invitées de colloques internationaux

- "Extraction of gene interaction networks from texts : the Caderige Project", présentation avec G. Bisson de l'INRIA, XRCE - INRIA workshop on Information Extraction in Genomics, Grenoble, juin 2000.
- "Knowledge extraction from text, a machine learning approach", CAPS'3, 3rd International Conference on Human-System Learning, , décembre 2000.
- «Machine Learning applied to Information Extraction in specific domains - an example, gene interaction extraction from bibliography in genomics », *Semantic Web Mining*, 2nd Workshop at ECML/PKDD, Helsinki, 20 août 2002.
- « Towards the automatization of knowledge extraction from text in genomics, strength and limits », NEMIS 2003 Launch Conference, 1st International Workshop on Text Mining & its Applications, Patras, Grèce, 5 avril 2003.
- « Ontology Learning and Populating in Genomics », *Ontology population*, ECAI-2004 Workshop, 2004
- "ML and NLP for Textual Information Annotation in Genomics", at *Dagstuhl Seminar on Machine Learning for the Semantic Web*, février 2005.
- "Machine Learning for semantic annotation of textual data in genomics", at *15th Amsterdam Colloquium*, décembre 2005.
- "Alvis Semantic Search Engine, Adaptation to Biology », Claire Nédellec, Robert Bossy, Alain Kotoujansky at *Machine Learning for Natural Language Processing workshop*, Université d'Amsterdam, mai 2007, Pays-Bas.

Conférences invitées de colloques nationaux

- "Une Application bancaire de l'Apprentissage Symbolique", *Quatrièmes Journées Francophones sur l'Induction Symbolique et Numérique*, Université Paris-Sud, France, mars 1994.
- "Le système d'apprentissage APT et la modélisation cognitives de phénomènes physiques", *journées de l'Institut des Sciences Cognitives et de la Communication*, Université Paris-Sud, janvier 1996.
- Avec Emmanuel Benoît. "Recherche d'information spécialisées, l'enjeu de l'automatisation", *Rencontres 2007 des professionnels de l'IST*, Nancy, 20 juin 2007.
- "L'apprentissage (semi)-automatique de connaissances à partir de corpus -- avec des ontologies, vers des ontologies --.", *Atelier OntoTexte* associé à TIA (Rencontres Terminologie et Intelligence Artificielle), Sophia-Antipolis, France, 10 octobre, 2007.
- "Apprentissage automatique pour l'analyse de contenu textuel en biologie", *Journée Apprentissage Artificiel: la carte, le territoire et l'horizon*, en l'honneur de Yves Kodratoff, INA-PG, Paris, 23 mai, 2008.
- "Ontologies et recherche d'informations dans le domaine scientifique", *Journée Communication technique de plein champ : recherche et innovation*, organisée par Patricia Minacori, Université Paris Diderot, Paris, 26 Mars 2011. <http://www.eila.univ-paris-diderot.fr/recherche/conf/ciel/communication-technique-plein-champ/2011/index>
- Avec Agnès Girard. "Traitement automatique du langage dans la gestion des données", *Journées Formation Frédoc de Renatis, (réseau national des professionnels de l'information scientifique du CNRS), Les données de la recherche : gestion et valorisation*, Aussois, 7-10 octobre 2013.

Tutoriels

- "Text Mining : Extraction de connaissances à partir de textes », aux *journées francophones d'extraction et de gestion des connaissances*, (EGC), Montpellier, janvier 2002.
- « Machine Learning and Natural Language Processing for Information Extraction in Genomics", *workshop text and literature mining of the Industry Programme of the European Bioinformatics Institute (EBI)*, janvier 2003.
- "Apprentissage de connaissances sémantiques spécialisées" *TCAN-CNRS tutorial serie, "Langue, connaissances, information"*, mai 2005.
- "Des outils d'analyse sémantique pour la recherche d'information" *formation INRA aux techniques d'analyse de l'information textuelle*, 20 mars 2007.
- "Texte et Connaissance - Quelques éléments méthodologiques", séminaire *Texte et Connaissance, méthodes, besoin, perspectives* Jouy-en-Josas, 3-4 juillet 2008. <http://genome.jouy.inra.fr/bibliome/TC/>.

- Organisation de la *training session text-mining* (1 semaine juin 2009) de l'action Marie Curie Transys. <http://genome.jouy.inra.fr/bibliome/Transys/> 3 cours.
- "Fouille de texte", séminaire *PEPI INRA Gestion de Données* Les-Loges-en-Josas, 6 mai 2011.
- "TyDI, acquisition de terminologie à partir de corpus", cours et TP sur PC, 10 scientifiques du projet *ATOL*, (dépt Phase, INRA), INRA Rennes, le 12 février 2012.
- "Thesaurus et recherche sémantique d'information", 7 documentalistes du projet *TriPhase* (dépt Phase, INRA), INRA Rennes, le 4 avril 2013.
- "Annotation sémantique avec *AlvisAE* pour la sélection du blé par marqueur", 13 sélectionneurs du projet *F5OV Sam Blé*, cours et TP sur PC, INRA Jouy-en-Josas, le 16 avril 2013.
- "TyDI, acquisition de terminologie à partir de corpus", cours et TP sur PC, 7 documentaliste du projet *TriPhase*, (dépt Phase, INRA), INRA Jouy-en-Josas, le 27 mai 2013. Durée 5 h.

Rapport d'expertise collective

- Contribution au rapport *d'autoévaluation de l'INRA pour l'AERES*, « Recherche en Sciences et Technologies de l'Information Scientifique et Technique et de la Communication ». 7 pages. novembre 2009.

Principaux séminaires invités depuis 1994

Groupe de recherche national (7)

- "Explications et Apprentissage Symbolique", *Groupe PRC-IA Explication*, Paris, avril 1994.
- "Apprentissage par *PLI* dans les Bases de Données Relationnelles", *Groupe AFCET-AFIA, GTRA*, Nantes, novembre 1997.
- "Acquisition des Connaissances et Apprentissage", *Journée Acquisition des Connaissances et Apprentissage pour l'interprétation des Signaux et des Images* organisée par l'Action *PRC 10.2*, Paris, novembre 1997.
- "Apprentissage de schémas de sous-catégorisation et d'ontologies", *Journée d'étude Extraction d'Information* (T. Poibeau), séminaire de *l'Association pour le Traitement Automatique des Langues (ATALA)*, Paris, mars 1999.
- "Classification conceptuelle pour la formation de classes sémantiques", *A3CTE*, Pitié-Salpêtrière, novembre 1999.
- "Extraction d'information à partir de notices sur les interactions géniques", *Séminaire IMPG, région Ouest*, IRISA, Rennes, avril 2001.
- « Extraction d'information appliquée au domaine biomédical – apprentissage et traitement automatique de la langue ». *Groupe Statistiques des Séquences Biologiques (SSB)*, janvier 2004.

Séminaire académique international (10)

- "Cooperative Machine Learning", au *Laboratoire d'Informatique de l'Université de Turin*, Italie, Juillet 1994.
- "Acquisition of semantic knowledge from specification texts by ML methods, the ASIUM system" au *Laboratoire d'Informatique de l'Université de Turin*, Italie, novembre 1997.
- "Application of Machine Learning to Knowledge Acquisition from Texts", *Colloquium serie*, Université de Karlsruhe, novembre 2000.
- "Apprentissage automatique de connaissance à partir de textes pour l'extraction d'information", séminaire de l'EPFL (*Ecole Polytechnique Fédérale de Lausanne*), mars 2001.
- « Apprentissage de connaissances pour l'extraction d'information textuelle en génomique" (*Machine Learning for Information Extraction in Genomics*), *CUI, Université de Genève*, décembre 2002.
- « *Some accomplishments and challenges in text-mining for biology*, Visit of *Greifswald Transcriptomics – Functional Genomics group*, Jouy-en-Josas, 26 January 2009.
- Philippe Bessières, Claire Nédellec *Bibliome group* « *Bibliome for genome annotation* », visit to *Mikrobiologie Institut für Mikrobiologie und Genetik*, University of Göttingen, 10 December 2009.
- « *Bibliome, text content analysis* », Présentation de Claire Nédellec à la délégation de l'université de Novosibirsk, *MIG*, 5 octobre 2009.

- Claire Nédellec and Bibliome group « Text Mining at Bibliome », visit to *DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH)*, novembre 2011, Braunschweig, Allemagne.
- Claire Nédellec, « Semantic search with the ATOL ontology » Présentation à la délégation du Leibniz Institute for Farm Animal Biology (FBN Dummerstorf), séminaire ATOL, Rennes, 5 mars 2013.

Séminaire international (1)

- "Semantic search of Animal Journal ", *Conseil d'administration de la revue Animal*, Edimbourg, 19 mai 2011.

Séminaire académique national (8)

- "Acquisition de connaissances sémantiques à partir de textes de spécifications par apprentissage automatique, le système ASIUM", *Xerox Research Center*, Grenoble, janvier 1998.
- "Apprentissage et traitement sémantique du langage naturel", *LIPN*, Villetaneuse, mai 1998.
- "Acquisition d'ontologies à partir de textes de spécialité, application aux brevets", *LORIA*, Nancy, octobre 1998.
- "Application de l'apprentissage automatique à l'acquisition de connaissances à partir de corpus", *LIMSI*, Orsay, décembre 1998.
- "Apprentissage automatique de connaissances à partir de documents textuels", *Journée du LRI*, Villemartin, juin 2000.
- "Apprentissage de connaissances et extraction d'information à partir de textes", *Séminaire du GREYC*, Université de Caen, décembre 2000.
- "Annotation sémantique de documents en biologie", *INIST*, Nancy, octobre 2006.
- "Acquisition de connaissances linguistiques par apprentissage automatique – application à la recherche documentaire", *Séminaire du LIG*, Université Joseph Fourier, Grenoble, 22 janvier 2010.

Journées thématiques (5)

- "Application de l'apprentissage à la recherche et à l'extraction d'information - Un exemple, le projet *Caderige* : identification d'interactions géniques.", présentation avec A. Nazarenko du LIPN, *Journée thématique Exploration de données issues d'Internet*, organisé par Y. Bennani, E. Janvier, R. Kanawati et S. Salotti, LIPN-CNRS, Institut Galilée, Université Paris 13, Villetaneuse, mars 2001.
- Nédellec C. et Ould Abdel Vetah M., Bessières P., Brun C. et Jacq B., "Text filtering for information extraction in genomics, a classification problem". *Journées IMPG Ontologie et Extraction d'Information en Génomique*, Grenoble, mai 2001.
- Bessières P., Bisson G., Nazarenko A., Nédellec C., Ould Abdel Vetah M. et Poibeau T., "Ontology learning for information extraction in genomics bibliography - the Caderige project", *Journées IMPG Ontologie et Extraction d'Information en Génomique*, Grenoble, mai 2001.
- Nédellec C. et Ould Abdel Vetah M., "Modélisation des interactions géniques à partir de textes.", *Journée Post-Génomique de la Doua (JPGD)*, Lyon, avril 2001.
- « Le projet Quaero », réunion du réseau IST des organismes de recherche nationaux, 27 avril 2006.

Séminaires INRA (11)

- « Acquisition automatique d'ontologie pour l'annotation sémantique de texte en biologie » séminaire de l'unité INRA, MetaRisk, février 2006.
- « Recherche sémantique d'information textuelle », Claire Nédellec, Pascale Sébillot, Philippe Bessières, Vincent Claveau, Journées INRA-INRIA, 30 mai 2007.
- "Annotation sémantique de documents scientifique" séminaire *Texte et Connaissance, méthodes, besoin, perspectives* Jouy-en-Josas, 3-4 juillet 2008. <http://genome.jouy.inra.fr/bibliome/TC/>.
- "Modélisation de connaissance termino-ontologique application à la recherche documentaire", *séminaire modélisation du département Phase*, Physiologie Animale, Paris, 9 mars 2010.

- « Gestion des connaissances - Ontologies », *Séminaire stratégique CEPIA (Département Caractérisation et Elaboration des Produits issus de l'agriculture)*, Paris, 17 mars 2010. https://intranet.inra.fr/cepia/vie_collective/seminaires_du_departement/seminaire_cepia_2010.
- « Conception d'ontologie et recherche sémantique dans le domaine scientifique », *Séminaire du groupe de travail ATO-Phase (Département Physiologie Animale)*, Rennes, 3 novembre 2010.
- « Extraction de régulations géniques à partir d'articles », *Séminaire du réseau MIA Inférence de réseau*, Paris, 9 février 2012.
- Claire Nédellec, Maarten van de Guchte, Philippe Bessières, "Réseau OntoBiotope", journées Animation MEM Days, 31 janvier 2012.
- Claire Nédellec, "Activités INRA sur l'extraction d'information sur les biotopes", journées Réseau OntoBiotope, 31 novembre 2012.
- Claire Nédellec, "Extraction d'information à partir de texte", journées ModalTub, Réseau Elucider le devenir de l'aliment dans le Tube Digestif, 21 mars 2013.
- Claire Nédellec, "Des technologies sémantiques pour la construction et l'exploration de corpus documentaires", journées DEPE (Direction de l'Expertise, de la Prospectives et des Etudes), 28 juin 2013.
- Claire Nédellec, Robert Bossy, "La tâche Gene Regulation Network in Bacteria – BioNLP-ST'13", *Séminaire du réseau MIA Inférence de réseau*, Paris, 12 septembre 2013.

Principaux séminaires invités de projet transversal (6)

- Claire Nédellec, Robert Bossy, Alain Kotoujansky, Annick Lacombe, Philippe Bessières. « Alvis Semantic Search Engine – Adaptation to Patent Search ». Revue annuelle du projet Epipagri, septembre 2007.
- Présentation publique : "Demonstration of the EPIPAGRI database and website", *Epipagri Network introduction meeting*, 22 avril 2008.
- Claire Nédellec, Robert Bossy, Frédéric Papazian, Sarhane Safraoui, Sophie Aubin, Sandra Makuntima "Epipagri.org, Presentation of the final version of the Web site", *Epipagri Final Review*, Barcelone – septembre 2008.
- « *Bibliome content analysis technology for system biology* » Micalis (Microbiology IdF) project. Jouy-en-Josas, 16 février 2009.
- Présentation publique : Claire Nédellec, R. Bossy, B. Teyssendier, W. Golik, F. Papazian, P. Veber, A. Kammoun « Recherche d'information dans les brevets du domaine de l'ARP VegA », *Assemblée Générale de l'ARP (Atelier de Recherche Prospective) VegA*, 12 janvier 2010. http://www.inra.fr/arpvega/travaux_de_l_arp_vega/assemblees_generales/3eme_assemblee_generale.
- « *Semantic Text Analysis* » conseil d'administration de la revue Animal (CUP, INRA, BSAS, EAAP), Edimbourg, 2011.

Participation à des Groupes de Recherche (8)

- GRACQ : Groupe National de Recherche en Acquisition des Connaissances, AFCET-AFIA (de 1992 à 1999).
- *Groupe Explication*, PRC-IA, AFCET-AFIA (1993-1998).
- GTRA : Groupe National de Travail sur l'Apprentissage Automatique, AFCET-AFIA (de 1994 à 2000).
- *Thème 5.2 Apprentissage*, GDR-III (de 1998-2003)
- *Groupe A3CTE*, GDR-III, RISC, AFIA, animation avec A. Nazarenko, de 1998 à 2003.
- *Corpus et terminologie*, ASSTICOT, Action Spécifique 34 CNRS, 2001-2003.
- *Intégration de sources/masses de données hétérogènes et ontologies*, réseau du département INRA MIA, animé par Juliette Dibie-Bathélémy, Pascal Neveu et Patrice Buche, depuis 2011.
- *Inférence de réseau biologique*, NetBio, réseau du département INRA MIA, animé par Matthieu Vignes, Simon de Givry et Marie-Laure Martin-Magniette, depuis 2012.

Membre du comité de sélection 2007 du prix de la meilleure thèse de l'ASTI (Association Française des Sciences et Technologies de l'Information et de la Communication)

Participation à des Associations

- *Association AFIA*, Association Française d'Intelligence Artificielle.
- *Société SPECIF*, Société des Personnels Enseignants et Chercheurs en Informatique de France.
- *Association ATALA*, Association pour le Traitement Automatique des Langues.

Organisation des séminaires de l'équipe Inférence et Apprentissage du LRI de 1993 à 2000, (env. 60).

PARTENARIATS DE RECHERCHE ET INDUSTRIELS

Responsabilité et participation scientifique à des Contrats de Recherche Internationaux

- **Coordination** pour le LRI du contrat bilatéral CNRS/JNICT intitulé *Disciple, un système apprenant*, engageant le LRI et le Laboratoire d'Informatique de l'Université de Coimbra (Portugal) sous la direction de Y. Kodratoff et de E. Costa (6 personnes). Thème de recherche : *Application bancaire*. De 1990 à 1993.
- **Participation** au contrat Esprit II *Machine Learning Toolbox* n°2154 (115 hommes année) pour les partenaires LRI et ISoft en tant que **responsable de la coordination** de la tâche "algorithmes d'apprentissage" (8 partenaires, 43 hommes an) et responsable de la recherche et du développement pour le *système d'apprentissage coopératif APT*. Production de documents de synthèse et démonstrations à ECAI'90, Esprit Technical Week 90 et 92 (Bruxelles), Journées Systèmes Experts 91, (Avignon), et IJCAI'93. De 1990 à 1993.
- **Participation** au contrat Esprit III "Basic Research Action" n° 6020 intitulé *Inductive Logic Programming* sous la direction de C. Rouveirol (LRI), pour la tâche "Declarative Biases" (4 partenaires, 22 hommes an). Thème de recherche : *PLI et systèmes configurables*. Coordination de la rédaction du rapport final de la tâche. De 1992 à 1994.
- **Coordination** pour le LRI du contrat CNRS bilatéral entre le LRI et le Laboratoire d'Informatique et de Mathématique de l'Université de Bar Ilan (Israël) intitulé *Incremental Revision of Relational Models in Robotics and Vision*, sous la direction de Y. Kodratoff et de M. Golumbic (6 personnes). Thème de recherche : *Utilisation des biais pour la révision des connaissances*. De 1992 à 1994. Participation à la suite du contrat de 1995 à 1998 sur le thème *fouille de textes*.
- **Participation** au projet European Science Foundation intitulé *Learning in Human and Machine* (80 personnes) dans le groupe *Changing Representation* dirigé par Lorenza Saitta. Thème de recherche : *Application du système APT à l'étude des changements cognitifs de représentation chez des étudiants en physique*. De 1995 à 1997.
- **Participation** pour le LRI au contrat Esprit IV "Long Term Research" n° 20237 intitulé *Inductive Logic Programming 2* (18 hommes an). Thèmes de recherche : *Application de systèmes de PLI configurable au "Data Mining" - Application de la PLI au traitement du langage naturel - Révision en PLI*. De 1996 à juin 1999.
- **Montage et coordination** pour MIG d'ALVIS: *Superpeer Semantic Search Engine*. IST-1-002068-STP (2004-2007). Coordination de la tâche WP6 : *Linguistic Resource Learning*. (Subvention MIG-INRA : 260 Keuros).
<http://www.alvis.info/alvis/>
- **Montage et coordination** pour MIG de *Quaero – Automatic multimedia content processing*. Oseo. Pré-contrat de décembre 2007 à mai 2008. Contrat de mai 2008 à décembre 2013. Coordination du WP3 (12 partenaires) : *Natural Language Processing du Core Technology Cluster*. Et de deux tâches 3.3 : *Ontology* et 3.4 : *Semantic Annotation*

Participation aux projets *Corpus* et *TIAE* (Text and Image Annotation Engine) avec la société Jouve. Subvention Equipe Bibliome : 1,083 million euros, budget total 250 millions.

- Participation au projet européen Specific Support Action, Priority Food quality and safety *Epipagri: Towards European Collective Management of Public Intellectual Property for Agricultural Biotechnologies*. Coordination du WP1: joint intellectual property information system. Responsabilité de la réalisation et de la maintenance du service de gestion de brevet (www.epipagri.org). Subvention MIG : 9 homme-mois.
- Participation au projet européen *Transys A systems approach to defining membrane protein networks and applications*. Contrat européen Marie Curie ITN-FP7 (2008-2012) Participation à Research Topic 4. *Modelling interactions and networks*. Voir : http://www.euprojekt.su.se/index.php/kb_1/io_1892/io.html. Organisation de la *training session text-mining* (1 semaine juin 2009) : <http://genome.jouy.inra.fr/bibliome/Transys/>
- Membre du conseil scientifique (Advisory Board) du projet européen FP7 **BioASQ**, - *A Challenge on Large Scale Biomedical Semantic Indexing and Question Answering*. (2012-2014).
- **Coordination** du projet **TriPhase**, Terminologie pour le recherche d'information en physiologie animale et système d'élevage ». avec CUP (Cambridge University Press), CA de la revue *Animal*, Financement INRA département Phase : 6 homme-mois (2013).

Participation scientifique à des Réseaux d'Excellence Européens

- **Membre du bureau du MLNet** (Machine Learning Network), réseau d'excellence Esprit V en apprentissage automatique, en tant que représentante d'un noeud principal. juin 1999 à 2001.
- Membre de **KDNet**: *Knowledge Discovery Network of Excellence*. FP5-IST-2001-33086 (2002-2005). <http://www.kdnet.org/>
- Membre de **KDubiq**: *Ubiquitous Knowledge Discovery Network of Excellence*. FP6-IST -021321 FET Open (Future and Emerging Technologies).(2005-2008). <http://www.kdubiq.org/>
- **Montage et coordination** avec A. Nazarenko (LIPN) de **l'Expression d'Intérêt** du 6ème programme cadre de la communauté européenne sur *l'acquisition de connaissance sémantique pour l'accès à l'information textuelle dans des domaines spécifiques* (plus de 50 partenaires), juin 2002. A contribué à la priorité thématique 2.4.7 *Semantic-based Knowledge and Content Systems*.

Responsabilité et participation scientifique à des Contrats de Recherche Nationaux

- **Montage et coordination** pour le LRI du contrat RNRT (Réseau National de Recherche en Télécommunication) pré-compétitif financé par Ministère de l'Industrie, intitulé *ASTUXE, Amélioration des Services d'Assistance Technique en Ligne par l'Extraction d'Information des Messages de Demande d'Intervention* (16 hommes an dont 3,5 et 1MF pour le LRI). Thème de recherche du LRI : *Apprentissage de règles d'extraction de connaissance à partir de textes*. De Novembre 2000, à Mai 2002, 30 mois.
- **Montage et coordination** pour le LRI du contrat bio-informatique CNRS-INRA-INRIA-INSERM, intitulé *CADERIGE, Catégorisation Automatique de Documents pour l'Extraction de Réseaux d'Interaction Géniques*, avec les laboratoires mixtes CNRS LEIBNIZ (Grenoble), LIPN (Villetaneuse), le projet INRIA AIDA (Rennes), les laboratoires INRA GM et LRV (Jouy-en-Josas) et MIG (Versailles). Thème de recherche du LRI : *Apprentissage de connaissances linguistiques conceptuelles à partir de textes sur le génome*. De novembre 2000 à octobre 2001, 1an.
- **Montage et coordination** pour MIG du contrat bio-informatique inter-EPST, *CADERIGE 2*, , avec les laboratoires mixtes CNRS LEIBNIZ (Grenoble), LIPN (Villetaneuse) et LRI (Orsay), le projet INRIA AIDA (Rennes), ENSAR-INRA Rennes. Thème de recherche de MIG : *Apprentissage d'ontologie à partir de textes sur le génome*. De novembre 2001 à novembre 2003, 2 ans.

- **Montage et coordination** pour MIG du contrat RNTL (Réseau National de Recherche en Technologie Logicielle), *EXTRAction de connaissance pour l'EXPLoitation de la DOcumentation Scientifique, Extraplodocs*, avec le laboratoire mixte CNRS-Université Paris 13, LIPN (Laboratoire d'Informatique de Paris-Nord) et la société ISOFT. Responsable des tâches *Acquisition d'ontologie* et *Extraction d'information*. Subvention : 181 Keuros. De 2003 à 2005.
- **Participation** au projet Décrypton *Cancer du rein et phosphorylation oxydative* avec le CGMC (Université C. Bernard Lyon). Subvention : 3 homme-mois. De 2006-2008.
- **Coordination** pour MIG du contrat ANR ARP (atelier de recherche prospective) **VegA** : *Quelles biomasses pour le futur?* . février 2008 - janvier 2010. Objectif de MIG : réalisation d'un système d'information de recherche de brevet dans le domaine des biotechnos vertes.
http://www.international.inra.fr/partnerships/with_the_private_sector/live_from_the_labs/vega foresight workshop
<http://www.inra.fr/arpvega/presentation>
- **Montage et coordination** pour MIG du contrat FSOV SAM Blé : *Sélection du blé assistée par marqueur*. Membre du comité de pilotage. Partenaires : UFS (Union française des semenciers), Arvalis, UMR INRA Génétique, Diversité et Ecophysiologie des Céréales. Objectif de MIG : extraction des information de relation marqueur – variété – phénotype. 1 homme-an. déc 2010 – nov 2014.
- **Coordination** pour MIG du contrat 3BCAR (bioénergie, biomolécules et biomatériaux du carbone renouvelable) **Astrali** : *Analyse Stratégique sur les Liquides Ioniques* : De leur conception à leurs applications dans un contexte de développement durable et d'innovation. Objectif pour MIG : conception d'un système de recherche bibliographique. 2010-2011.
- **Montage et coordination** pour MIG du projet *GIS Grande culture à hautes performances économiques et environnementales (GC HP2E)*, animé par *Stéphanie Potok* (Secrétaire Générale) sur la conception d'un moteur de recherche d'information, ½ homme-an. Partenaires : GIS, INRA IST, MIG. 2011-2012.

Responsabilité et Participation scientifique à des Projets de Recherche de l'institut

- Participation au thème *Apprentissage Automatique et Traitement des Connaissances*, animé par J. S. Liénard du projet pluriformation du Département Informatique de l'Université Paris XI. Thèmes de recherche : *Application de l'apprentissage automatique au traitement du langage naturel à l'aide de bases de données relationnelles*. Durée 2 ans : 1996-1997.
- **Co-responsabilité** avec Christian Jacquemin, du projet Bonus Qualité Recherche (BQR), *Construction automatique de thésaurus et d'index à partir de corpus* du Département Informatique de l'Université Paris XI. Durée 1 an : 1999-2001. (Voir <http://www.limsi.fr/Individu/jacquemi/BQR99>)
- Participation à l'AIP (action incitative programmée) *Contractualisation de l'Activité Scientifique et Technique entre l'INRA et les partenaires du développement agricole et territorial*, AIP 258, INRA, 2002-2003.
- Participation à l'AIP (action incitative programmée) *Texte et Connaissance*, animée par le chef de département MIA, 2010 -2012.
- **Coordination** pour MIG du réseau **OntoBiotope** du métaprogramme INRA MEM *Meta-omics for microbial ecosystems*. Partenaire : Micalis. Subvention : 10 Keuros (2012-2013) https://intranet.inra.fr/mica/metaprogramme_mem/reseau_ontobiotope.

Evaluation de projets en réponse à des appels d'offre institutionnels (rapports confidentiels)

Programme interdisciplinaire au CNRS sur le thème "traitement des connaissances, apprentissage et NTIC" (TCAN) : 7 évaluations.

ACI Masses de Données : 4 évaluations.

ACI IMPBIO : 3 évaluations. 2004

RNTL : 1 évaluation.
ACI santé français : 1 évaluation.
ANR blanc : 1 évaluation
Projet région Picardie : 1, 2007
ANR Masse de données et connaissances : 1, 2007
ANR Genomic : 1, 2008
SSHRC : *Major Collaborative Research Initiatives Program* Canada : 1 évaluation 2008
Allocation doctorale région Bretagne : 1 évaluation 2010
ANR programme Blanc : 1, 2011.
ANR Jeune équipe : 1, 2013.

Organisation d'une manifestation Industrie-Recherche

- **Organisation scientifique** de la Première Rencontre FIIFO (Formation d'Ingénieur en Informatique de la Faculté d'Orsay) Recherche-Industrie sur le thème de la "*Gestion documentaire*", Orsay, 10 Novembre 1999.

Partenariats Industriels nationaux et internationaux hors contrats de recherche institutionnels

- Responsabilité de la collaboration avec le **service d'Assurance Maladie et de Retraite de la SNCF**, Marseille pour la réalisation et la maintenance de la base de connaissance d'un serveur Minitel par des méthodes d'apprentissage coopératives. Cette collaboration a donné lieu à la CIFRE de Luc Poittevin. Durée 4 ans : 1994-1998. Le serveur minitel et l'outil de maintenance sont en service depuis Mai 96.
- Responsabilité de la collaboration avec le département R&D de **Dassault Aviation** sur l'application de l'apprentissage automatique à l'analyse sémantique de textes techniques. Durée 2 ans : 1996-1998. Officialisation du partenariat à travers le End-User Club du projet Esprit ILP 2.
- Responsabilité de la collaboration avec la société **Valigen** sur l'application de méthodes d'apprentissage automatique et de traitement automatique de la langue appliquées à l'extraction de connaissances sur les interactions géniques dans les notices de PubMed. Cette collaboration a donné lieu à la CIFRE de Mohammed Ould Abdel Vetah. juillet 2000 – juillet 2002.
- Responsabilité de la collaboration avec le département Phase et Cambridge University Press, pour l'**indexation de la revue Animal** Evaluation de l'ontologie ATO (Animal Trait Ontology), extraction terminologique, conception d'ontologie. Printemps 2010 – 2013.
- Responsabilité du MOU entre l'INRA et le **DBCLS** (Database Center for Life Science (Tokyo, Japon), accord de coopération dans le domaine des STIC, depuis 2012.

Valorisation de la recherche

- **Dépôt** du logiciel *ASIUM* par le CNRS auprès du CNISF (n°15072), dans le cadre d'un **dossier de valorisation** (n° 61560) géré par FIST (France Innovation Scientifique et Transfert). Les co-inventeurs sont David Faure, Claire Nédellec et Céline Rouveirol. Depuis 1998.

Challenges

- Organisation du **challenge international LLL** (Learning Language in Logics) pour l'apprentissage de relations sémantiques en biologie (13 participants), présentation au workshop LLL. Soutenu par le Réseau d'Excellence Pascal. Mise à disposition des données *LLL* : ensemble d'exemples d'apprentissage annotés pour les interactions géniques incluant les données linguistiques (lemmes et dépendances syntaxiques validées). Module d'évaluation en ligne avec données de test. <http://genome.jouy.inra.fr/texte/LLLchallenge/>, 2005.

- Participation à l'organisation du *BioNLP 2011 shared tasks*. <https://sites.google.com/site/bionlpst/>. Définition de la tâche et coordination de la préparation des données pour les trois tâches *Rename*, *Genic Interaction* et *Bacteria Biotope*.
- Responsable de l'organisation du *BioNLP 2013 shared tasks*. <https://sites.google.com/site/bionlpst/>. Sélection des 6 tâches, coordination. Définition des tâches de la *Bacteria Track*. En collaboration avec le DBCLS et le NACTEM.

Services en ligne

- Coordination du dévt et de la maintenance du service **BioAlvis**. Recherche sémantique documentaire en microbiologie (400 000 références PubMed) 2007-2011.
- Coordination du dévt et de la maintenance du service **Epipagri**. Recherche sémantique de brevet en agrobiotechnologie en 2008. <https://www.epipagri.org/> 2007-2012.
- Coordination du dévt et de la maintenance du service **VegA'lvis**. Recherche sémantique de brevet dans le domaine de l'exploitation de la biomasse, en 2010. <http://bibliome.jouy.inra.fr/A01H>
- Coordination du dévt et de la maintenance du service **GIS HP2E**. Systèmes de production de grande culture à hautes performances économiques et environnementales (GC-HP2E), en 2010-. <http://bibliome.jouy.inra.fr/gisdemo>
- Coordination du dévt et de la maintenance du service **FSOV SAM Blé**. Sélection du blé assistée par marqueur. <http://bibliome.jouy.inra.fr/test/alvisir/FSOV>. 2011-.
- Coordination du dévt et de la maintenance du service **Animal**. Physiologie animale et système d'élevage. <http://bibliome.jouy.inra.fr/test/alvisir/Animal/>. 2011-.
- Coordination du dévt et de la maintenance du service **CoCitation**. Cocitation de noms de gènes dans les références de PubMed. <http://migale.jouy.inra.fr/cocitations/>. 2012-.