

Rapport 2007-2012

Projet 2014-2018

**Apprentissage Artificiel
et
Applications**

Section	Sub-sections	Page
Partie I	Le LIPN	page 5
Partie II	The LIPN at a glance	7
	Présentation générale	13
	Activités de recherche	17
	Production scientifique, valorisation, dissémination	31
	Rayonnement	41
	Formation à la recherche	47
	Publications	53
Publications antérieures à l'arrivée au LIPN	73	
Partie III	Acronymes	page 77

Partie I

Le LIPN

page 5

The LIPN at a glance

Scientific activities
Environment
Doctoral studies

7

Partie II

Equipe A3

page 11

Présentation générale

Overview
Membres de l'équipe

13

Activités de recherche

Bilan 2007-2012
Analyse AFOM/SWOT
Perspectives et projet

17

Production scientifique, valorisation, dissémination

Publications
Contrats de recherche
Valorisation de la recherche

31

Rayonnement

Comités d'organisation et d'édition
Expertises
Invitations

41

Formation à la recherche

Masters, Ecoles doctorale et post-doctorale
Thèses et Habilitations
Stages de master, post-doctorats

47

Publications

53

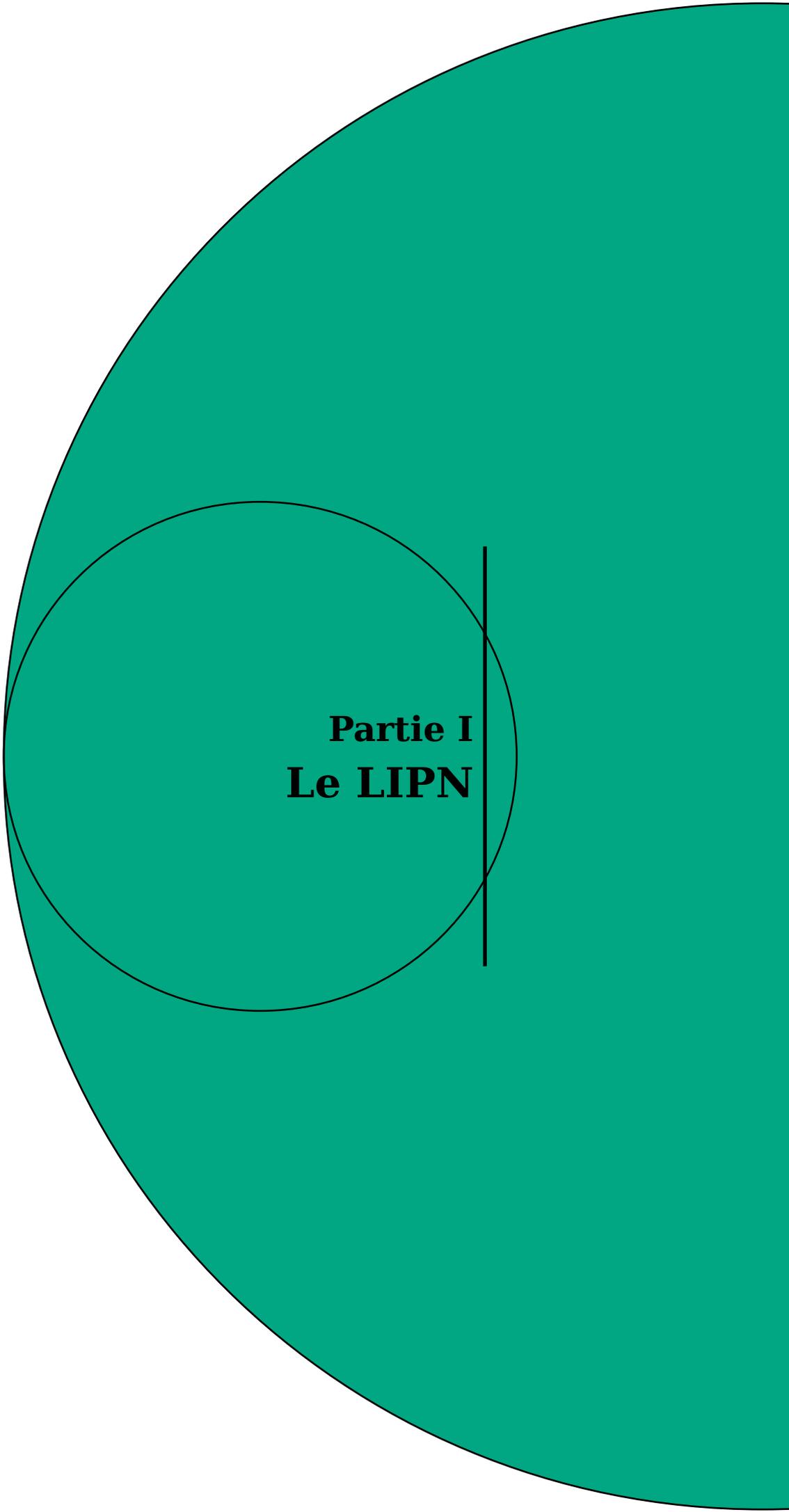
Publications antérieures à l'arrivée au LIPN

73

Partie III

Acronymes

page 77



Partie I
Le LIPN

Chapitre 1

The LIPN at a glance

The LIPN¹ (Laboratoire d'Informatique de Paris-Nord) was created in 1985, has been associated with CNRS since January 1992, before becoming a CNRS UMR in January 2001. The LIPN is the computer science laboratory of University Paris 13 which is part of the PRES Sorbonne Paris Cité.

The laboratory comprises 5 teams with 78 permanent staff (researchers or professors and associate professors), 8 technical and administrative staff, as well as PhD students and postdocs. It has overall more than 140 members. The LIPN has experienced an impressive growth during the last few years, which allowed for strengthening its major areas of expertise.

The scientific activities display many collaborations, be they national, international, or through research projects, in particular with the industry. LIPN members are extremely involved on the national scene, in national boards (presidency of the CNRS section, CNU — national universities council, etc.), and in institutional cooperation projects (ANR). On the international scene, members are widely present in editorial boards, programme committees, expert pools of different countries, and conduct collaborative research with colleagues worldwide.

1. <http://lipn.univ-paris13.fr>

1.1 Scientific activities

The laboratory is structured into 5 teams, focussing on the following main areas.

- **A³ (Apprentissage Artificiel et Applications): Machine Learning and Applications**

The A³ team tackles machine learning problems and covers a wide spectrum of issues, ranging from supervised and unsupervised learning to reinforcement learning. Its research is fed, coordinated and evaluated thanks to various applications in the field of pattern recognition and data mining. Research in A³ focuses on the following main topics: algebraic and logical models of learning, collaborative and transfer learning, cluster analysis, dimensionality reduction, link prediction in social networks, recommender systems, and learning structures from complex data.

- **AOC (Algorithmes et Optimisation Combinatoire): Algorithms and Combinatorial Optimisation**

The AOC team develops research in optimisation on graphs, mathematical programming, parallel and distributed computing. Optimisation on graphs is conducted with a particular focus on complexity, polyedral theory and approximation. The team has expertise in design and analysis of mathematical programming approaches and algorithms, and develops both exact and heuristic approaches for solving linear and non-linear problems. Many distributed environments issues are also considered: distributed middleware and architectures, distributed programming and distributed algorithms.

- **CALIN (Combinatoire, ALgorithmique et INteractions): Combinatorics, Algorithmics and Interactions**

The CALIN team brings together researchers having skills in a variety of aspects of combinatorics (analytic, algebraic and bijective combinatorics), interested in complexity of algorithms, trying to determine their behaviour in average or in distribution, in fine analysis of data structures, and also studying problems or using methods issued from physics. The team is organised in two axes: one focuses on the analysis of algorithms and combinatorial structures, the other is devoted to combinatorial physics.

- **LCR (Logique, Calcul et Raisonnement): Logic, Computation and Reasoning**

The LCR team tackles different aspects of computational and software models, from theory to applications. The specification and verification group addresses methodologies for the development of specifications, and verification through modular, compositional, distributed and parameterised approaches. The logic, theory of computation and programming languages group has expertise in computational issues of logical systems and formal semantics of programming languages: use of tools and methodologies issued from Linear Logic to analyse resource sensitive properties of programs, study graph formalisms for representing proofs ; algebraic and geometric foundations of computation.

- **RCLN (Représentation des Connaissances et Langage Naturel): Knowledge Representation and Natural Language**

The RCLN team is interested in natural language for its expressive capacity, and in knowledge representation, in particular for its connection to semantic analysis and text understanding. It tackles corpus analysis and semantic annotation of texts. Search and structuration of semantic web knowledge is achieved by the design of ontologies and knowledge patterns. Text understanding and knowledge acquisition are integrated in a virtuous circle between language, data and knowledge.

1.2 Environment

The LIPN has many research cooperations with its environment: other laboratories at University Paris 13, but also in the PRES Sorbonne Paris Cité, with industry, etc.

- **MathSTIC research federation, and University Paris 13 laboratories**

Research collaborations with other laboratories of University Paris 13 have been favoured over the years, in particular with LAGA² (UMR 7539, mathematics laboratory), L2TI³ (signal and image processing).

This led us to recently set up a new research federation in Mathematics and Information Technology (MathSTIC⁴), gathering members of LAGA, LIPN, and L2TI laboratories, in order to enhance cross-fertilisation in three of their major areas of expertise:

- Optimisation and learning applied to digital contents;
- High-performance computing, distributed systems;
- Mathematical physics, statistical physics, combinatorics.

- **Regional setting: PRES Sorbonne Paris Cité, and EFL LabEx**

The PRES Sorbonne Paris Cité comprises 8 universities and colleges. It includes University Paris 5, and University Paris 7, which both have computer science laboratories (LIPADE, LIAFA, and PPS). The PRES offers support for research projects between these laboratories.

The team RCLN at LIPN is a key actor in the EFL LabEx (Empirical Foundations of Linguistics). Its multidisciplinary nature is a key to evolution of research topics at the meeting point of linguistics and computer science.

- **International cooperations**

Members of the laboratory have many collaborations worldwide, with exchanges of researchers. Approximately 15 foreign researchers are invited for a month at our laboratory each year to carry collaborative research. Cooperation projects also involve laboratories in Germany, Italy, Norway, Poland, Tunisia, Canada, Brazil, Chile, Uruguay, Singapore, Vietnam.

- **Industrial collaborations**

Most research areas of the LIPN also participate in collaborations with the industry, mainly through research projects (ANR, FP7, FUI, FEDER) and CIFRE PhDs. Some large national or European collaborative projects involve both academic and industrial research laboratories. Some researchers of the LIPN have also created a start-up company, or benefit from a few months industrial experience. Moreover, the LIPN is involved in different boards of 3 business clusters (pôles de compétitivité) which favour technology transfer and research projects with companies at a regional level.

1.3 Doctoral studies

The doctoral studies are organised within the “Galilée” doctoral school, which addresses four major scientific areas:

1. mathematics, computer science, signal processing;
2. physics, materials, engineering sciences;

2. <http://www.math.univ-paris13.fr/laga/>

3. <http://www-l2ti.univ-paris13.fr/>

4. <http://mathstic.univ-paris13.fr/>

1.3 Doctoral studies

3. health sciences, medicine, human biology, chemistry;
4. ethology.

LIPN is part of the first group.

The doctoral school organises mandatory courses, and allocates grants to doctoral students. In 2012, six of them were allocated to the LIPN (out of 25 for the whole doctoral school).

• **PhDs and Habilitations**

Over the 2008–2012 period, 65 PhD defences took place. Half of the doctoral students were financed by a grant from the doctoral school, and the others through research projects, industry, or international support.

During the same period, 15 habilitations (Habilitations à diriger les Recherches) were delivered.

The laboratory organises an oral presentation halfway through the PhD (after 18 months), in front of a jury comprising an external member, who often is to become one of the reviewers. The aim is thus to assess the work that has already been done, and confirm the future directions of the doctoral work. Doctoral students benefit from a broader view of their work as well as advice from the jury.

• **Masters courses**

A masters programme in computer science is delivered at University Paris 13. It currently features two specialities:

- EID² (Exploration Informatique des Données et Décisionnel): Data Mining, Analytics, and Knowledge Discovery;
- PLS (Programmation et Logiciels Sûrs): Programming and Secure Software.

Both specialities lead to either industrial or research careers, depending on the master thesis.

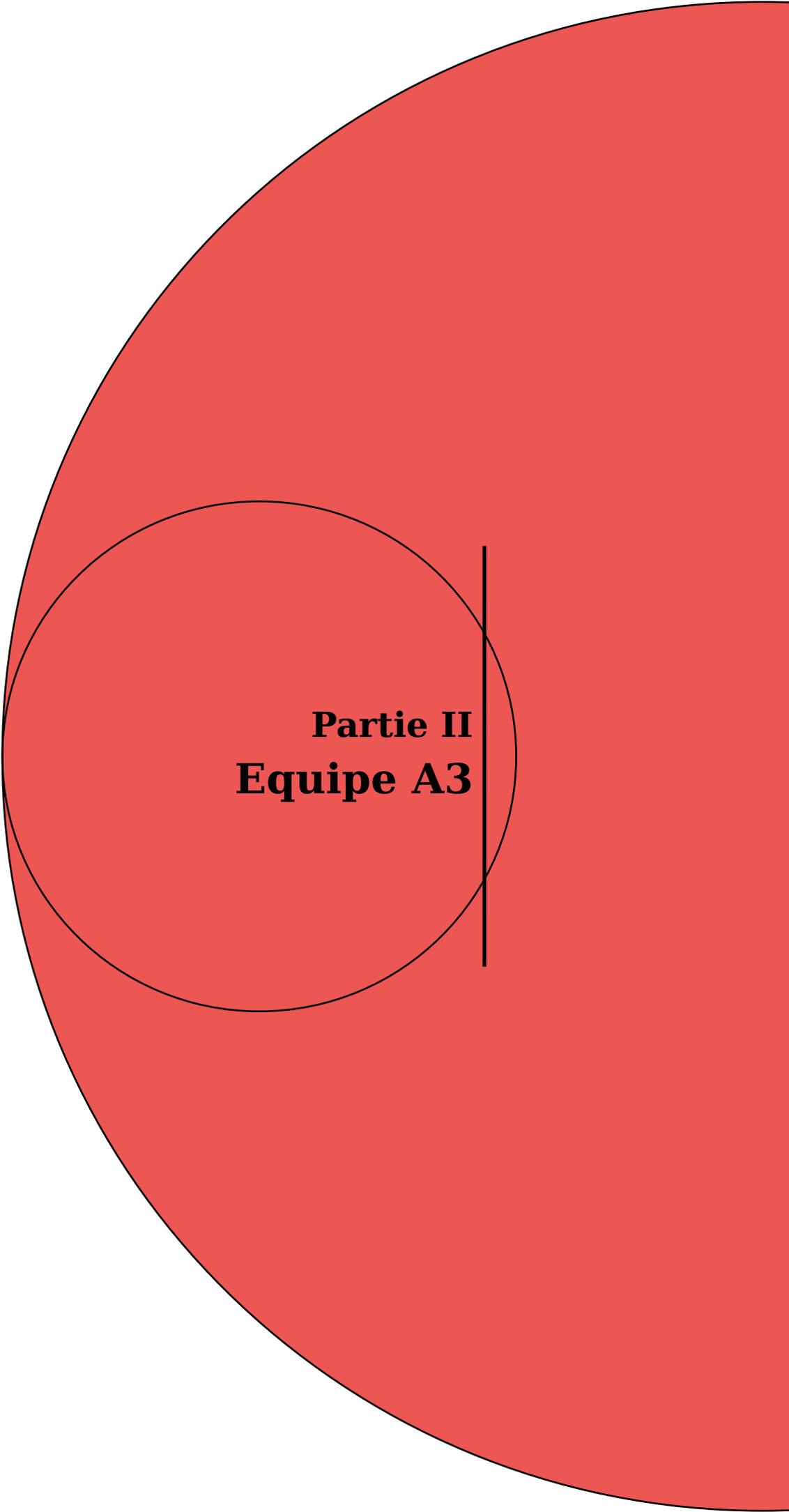
LIPN researchers also deliver courses in other masters programmes of universities in the parisian region.

Furthermore, they are involved in international cooperations for setting up masters programmes in foreign countries, e.g. Madagascar, Vietnam.

• **Summer schools and conferences**

LIPN supports the organisation and participation in thematic schools, by encouraging doctoral students attendance (which is recognised as mandatory courses), talks by members of the laboratory, as well as financial support to the organisation of such events.

In particular, ADAMA 2012, Logics and Interactions 2008 and 2012, Proof Theory 2012, Alea 2012, AAFD, JPOC/ISCO were co-organised by LIPN members. Some of these thematic schools take place in University Paris 13, while others are elsewhere in France or abroad.



Partie II
Equipe A3

Chapitre 1

Présentation générale

1.1 Overview

Machine Learning is a scientific discipline concerned with the design and development of algorithms that improve their behaviour from experience (observations, labelled or not). It borrows techniques and theoretical background from the field of Artificial Intelligence, Logics, Statistics, to form a highly challenging domain. Machine Learning is now a mature field of computer science, with a number of solid theoretical models and results, and a broad range of applications, both in industry and in multi-disciplinary research. Our group is one of the largest ones in France in Machine Learning. It covers a remarkably broad range of topics, from Statistical Learning to Inductive Logic Programming and Reinforcement Learning. This allows the group to be quite reactive to cope with new challenges raised by emerging applications of Machine Learning. It also makes it possible to study innovative combinations of learning methods for tackling complex problems.

The team A³ was created in 2005, with the goal of gathering Machine Learning related research of LIPN within a strongly coherent group, highly visible on the national and international research scenes. It has smoothly evolved since 2007, with four arrivals – among which one change of laboratory inside Université Paris 13 – and one departure, a promotion inside Université Paris 13.

1.2 Membres de l'équipe

The structure of the group has evolved consequently to these moves, and now consists of three research axes, for which the group is widely recognized :

- *Algebraic and logical models of learning*, which gathers research that deals with supervised learning in logical concept languages;
- *Collaborative and evolutive approaches of unsupervised learning* which studies non supervised learning in the context of distributed data, stream data and unbalanced distributions.
- *Structure mining from heterogeneous data*, motivated by data mining issues, with a strong focus on mining structures, in particular graphs and logic programs.

In each of these axes, theoretical issues are addressed, and algorithms and software are developed, in the context of collaborative projects, including both academic and industrial partners.

The research topics of the team for the next five years aim at investigating promising areas where the team could be leader. Our researchers can provide fruitful contributions thanks to their know-how and the cross fertilization of different axes. The promising areas we have selected are the following. *Acting and Learning*, in particular the interactions of learning and planning in the context of Reinforcement Learning for structured representations on one hand, and collective learning on the other hand, that studies which interactions between agents that learn and reason may result in better learning performances. Two research projects concern unsupervised learning issues, namely *Transfer learning and matrix factorisation* that studies under which conditions learning for a given task may be improved by learning for related auxiliary tasks, and aims at studying matrix factorisation techniques for developing appropriate representations to do so; another project will deal with developments for *Massive distribution of unsupervised learning*. Finally, the last research project will deal with the *Analysis and Mining of Heterogeneous and Evolving Networks*, that extends techniques developed in the group for link prediction and community detection to inter-related evolving large networks.

1.2 Membres de l'équipe

1.2.1 Membres permanents

(membres ou, en italique, anciens membres, au 01/10/2012)

	Nom	Prénom	Situation	Institution
	Alphonse	Erick	MCF	IG
	Azzag	Hanane	MCF	IUTV
	Bennani	Younès	PU	IG
	Bouthinon	Dominique	MCF	IUTV
	Champesme	Marc	MCF	IG
	Chevaleyre	Yann	PU	IG
	Gérard	Pierre	MCF	IUTV
	Grozavu	Nistor	MCF	IG

1.2 Membres de l'équipe

	Guérif	Sébastien	MCF	IG
	Kanawati	Rushed	MCF	IUTV
	Lebbah	Mustapha	MCF	IUTV
	Osmani	Aomar	MCF	IUTV
	Rouveïrol	Céline	PU	IG
	Santini	Guillaume	MCF	IUTV
	Soldano	Henry	MCF	IG
	Viennet	Emmanuel	MCF	IUTV

**Présentation
générale**

1.2.2 Doctorants au 1er oct. 2012

	<i>Nom</i>	<i>Prénom</i>	<i>Situation</i>	<i>Institution</i>
	Alizadeh	Pegah	AM	Université Paris 13
	Chaïbi	Amine	CIFRE	Anticipo
	Chebil	Ines	AM	INCa
	Nhat-Quang	Doan	AM	Bourse Vietnam
	Essaidi	Moez	CIFRE	Intelligence Power
	Ghassany	Mohamad	AM	Université Paris 13
	Hamdi	Fatma	AM	ANR E-FRAUD
	Jaziri	Rakia	CIFRE	INA
	Mouhoubi	Karima	ATER	Université Paris 13
	Pujari	Manisha	AM	Université Paris 13
	Redko	Ievgen	AM	Université Paris 13
	Rodrigues	Christophe	ATER	Université Paris 13
	Yakoubi	Zied	AM	FUI Projet UrbanD

1.2 Membres de l'équipe

1.2.3 Personnel temporaire du 1/9/2007 au 1/9/2012

<i>Nom</i>	<i>Prénom</i>	<i>Situation</i>	<i>Début</i>	<i>Fin</i>	<i>Financement</i>
Aseervatham	Sujeevan	AM	oct. 2004	déc 2007	Université Paris 13
		ATER	déc. 2007	août 2008	Université Paris 13
Benchettara Cabanes	Nesrine Guénaël	Doctorante	déc. 2007	déc. 2011	ANR CADI
		AM	oct. 2007	déc. 2010	Université Paris 13
Chamroukhi	Faicel	ATER	dec. 2010	août 2011	Université Paris 13
		Postdoc	09/2011	02/2012	ANR E-Fraud
		ATER	02/2011	09/2011	Université Paris 13
Elati	Mohamed	ATER	sept. 2007	août 2008	Institut Galilée
Grozavu	Nistor	Doctorant	nov. 2006	déc. 2009	Institut Galilée
		ATER	déc. 2009	août 2010	Institut Galilée
Labiod	Lazhar	Postdoc	mars 2009	août 2010	ANR CADI
Manine	Alain-Pierre	ATER	sept. 2008	août 2009	Université Paris 13
Rogovschi	Nicoleta	AM	nov. 2006	déc. 2009	ANR Septia
		ATER	sept. 2009	août 2010	Université Paris 13

Présentation
générale

1.2.4 Evolution des membres de l'équipe

Les effectifs de l'équipe A³ sont en légère progression depuis 2007. Il est à noter que deux de ces mouvements se font au sein de l'Université Paris-Nord : une arrivée est une mutation depuis le Lim&Bio de l'Université Paris-Nord - Bobigny et un départ est une promotion vers le L2TI. Un recrutement de MCF est interne, les deux recrutements les plus récents sont externes. Un fait remarquable concernant la composition de l'équipe est qu'elle accueille de nombreux doctorants.

Départs

- E. Viennet, MCF, promotion PU au L2TI, Univ. Paris 13, 2009

Arrivées

- M. Lebbah, MCF, IUTV, 2008, LIM&BIO → LIPN
- N. Grozavu, MCF, IG, 2009 (thèse Paris 13)
- Y. Chevaleyre, PU IG, 2010 (MCF Paris 9, LAMSADE)
- G. Santini, MCF IUTV, 2010 (post-Doc Paris 6)

Chapitre 2

Activités de recherche

2.1 Bilan 2007-2012

L'équipe A³ (Apprentissage Artificiel & Applications) a été créée en 2005 suite à la réorganisation des thèmes de recherche de l'équipe ADAge (Apprentissage, Diagnostic et Agents), qui était devenue nécessaire après le départ de plusieurs de ses membres en début de quadriennal 2004-2008. Cette restructuration a été effectuée avec la volonté de tirer parti de la complémentarité des recherches autour de l'apprentissage automatique et de mettre en œuvre leur coopération afin de rendre plus visible cet axe de recherche au niveau national et international.

L'équipe A³ était structurée en 2008 en trois axes de recherche, organisés par types de méthodes d'apprentissage : un axe *Apprentissage Numérique*, un axe *Apprentissage Symbolique*, et un axe *Approches transversales* rassemblant des approches hybrides.

On retrouve dans l'équipe actuelle trois axes, cette fois-ci organisés en terme de grandes problématiques de l'apprentissage. L'axe *Modèles algébriques et logiques de l'apprentissage* rassemble les recherches centrées sur l'apprentissage supervisé dans des langages de concepts logiques, l'axe *Apprentissage non supervisé collaboratif et évolutif* traite des problèmes liés à l'apprentissage à partir de données distribuées, de flux de données et en distributions déséquilibrées. Enfin, le troisième axe *Apprentis-*

sage de structures à partir de données hétérogènes, plus orienté fouille de données, regroupe des recherches motivées par l'extraction de structures dans des données de toutes sortes, structurées ou non. Chaque axe aborde à la fois des recherches à caractère fondamental ainsi que des recherches plus appliquées, le plus souvent soutenues par des projets collaboratifs académiques et industriels.

Du point de vue national, c'est l'une des plus grosses équipes académiques sur le thème de l'apprentissage automatique, et elle est remarquable par la largeur de son spectre de compétences, allant de l'apprentissage statistique à l'apprentissage relationnel et à l'apprentissage par renforcement. Cette caractéristique lui permet de pouvoir aborder des problèmes complexes avec plusieurs points de vue et outils, tant théoriques que méthodologiques et algorithmiques, et d'explorer des combinaisons innovantes de ces méthodes, comme l'apprentissage par renforcement relationnel.

2.1.1 Modèles algébriques et logiques de l'apprentissage

Cet axe s'intéresse aux aspects fondamentaux de l'apprentissage, et est essentiellement concerné par les représentations logiques. On trouvera d'abord ci-dessous des travaux concernant des aspects théoriques de l'apprentissage relationnel, puis des travaux concernant la dimension collective de l'apprentissage. Enfin une dernière série de travaux concerne le rôle des connaissances, ou plutôt d'un modèle *a priori* sur le monde, dans un processus d'apprentissage.

• Aspects théoriques

Apprentissage relationnel D'un point de vue théorique, nous avons poursuivi un travail entamé depuis le début des années 90 en satisfaisabilité et depuis le début des années 2000 en Apprentissage Artificiel sur l'étude de la complexité typique des algorithmes d'apprentissage. Ce travail a pour but d'évaluer les algorithmes par rapport à la difficulté réelle des problèmes traités et de rechercher les meilleures heuristiques en fonction de la position dans le paysage de la complexité des problèmes. D'un point de vue pratique, nous avons également continué notre travail sur l'algorithmique de l'apprentissage relationnel par Programmation Logique Inductive (PLI). *Résultats* :

- Étude du paysage de la complexité dans le cas de l'apprentissage relationnel, proposition des meilleurs paramètres pour caractériser les problèmes intrinsèquement difficiles [CI-13, CI-24, CI-38, RI-7].
- Proposition d'un opérateur de raffinement de clauses, basé sur une reformulation en Apprentissage Multi-Instances, permettant de traiter les variables numériques de manière efficace en PLI [CI-54].

• Dimension Collective

Apprentissage Collectif L'apprentissage *collectif* dans une société d'agents se distingue de l'*apprentissage artificiel distribué* en ce que nous donnons la priorité à la faculté pour un agent d'avoir un raisonnement, des comportements et des choix propres. Nous avons étudié un modèle d'apprentissage collectif dit *apprenant-critique* dans lequel une théorie courante est révisée en fonction des informations fournies par l'environnement ou les agents, de manière à satisfaire une certaine notion de *consistance*. Les agents prennent alors différents rôles : *apprenti* (l'agent doit réviser la théorie courante pour en maintenir la consistance), ou *critique* (une modification de la théorie courante, proposée par un agent apprenti, est contredite par une information que possède l'agent critique). *Résultats* :

- Protocole d'apprentissage collectif par maintien de consistance pour l'apprentissage supervisé pour des représentations attribut-valeur et pour des sociétés d'agents *non-individualistes* maintenant ensemble une seule théorie courante [CI-8].

- Adaptation à des sociétés d’agents *individualistes*, maintenant chacun leur propre théorie courante, ce qui permet de mettre en œuvre des modes de décision collective par vote [CI-37].
- Proposition de protocoles permettant de maintenir les propriétés de consistance lorsque les entités ne communiquent pas toutes directement mais via un *réseau de voisinage* fixe ou dynamique [CI-36].
- Proposition de protocoles pour le cas où l’hypothèse en révision est proposée en parallèle, plutôt que séquentiellement, vers les voisins de l’agent apprenant [CI-51].

Décision Collective Nous avons étudié certains aspects théoriques de la décision collective, un premier problème de placement de panneaux pour canaliser des mouvements de population, et un deuxième problème de vote.

Résultats :

- Modélisation des mouvements d’une population sous forme de chaînes de Markov. Une fois le problème de placement de panneaux démontré NP-complet, formulation et résolution sous la forme d’un problème MILP (*Mixed Integer Linear Programming*) [CO-84*].
- Étude, pour différentes règles de vote, des propriétés théoriques d’un problème de vote pour lequel le nombre de candidats à une élection varie au cours de l’élection (ajout ou défection). Étude des conditions pour lesquelles le gagnant d’une élection peut changer [CI-69, RI-27].

Activités de
recherche

• Apprentissage et Action

L’équipe a entamé depuis 2007 un nouvel axe de recherche, s’appuyant sur des compétences déjà présentes en apprentissage relationnel et apprentissage par renforcement. Nous avons étudié l’*apprentissage par renforcement relationnel* qui manipule des langages de concepts restrictions de la logique d’ordre un, plus expressives que les représentations attribut-valeur utilisées usuellement en apprentissage par renforcement. Nous nous plaçons dans le cadre de l’apprentissage par renforcement *indirect* pour lequel un agent apprend en ligne un modèle du monde dans lequel il évolue et utilise ce modèle pour choisir ses actions et explorer le monde. Dans le cadre de l’ANR JCJC HARRI, nous avons étudié l’apprentissage d’un modèle explicite des actions, dans lequel chaque action est décrite par un ensemble de règles (STRIPS conditionnel).

Résultats :

- Approche de révision relationnelle en ligne [CI-40] : les exemples sont observés un par un selon les actions effectuées par l’agent, et le modèle est révisé chaque fois qu’il est contredit par un nouvel exemple.
- Adjonction à l’algorithme précédent d’un mécanisme d’apprentissage actif, qui apporte une amélioration significative du temps de convergence vers un modèle de qualité pour un ensemble de problèmes classiques en planification [CI-75].

• Le rôle des Connaissances

Apprentissage et Incertitude Nous avons proposé différents cadres logiques d’apprentissage dans le cas où les données recèlent des formes variées d’incomplétude.

Résultats :

- Formalisation de l’apprentissage logique en présence d’incomplétude et proposition d’un algorithme d’apprentissage à partir d’exemples ambigus par règles [CI-35]. Cette approche présente un avantage clair dans le cas où les exemples présentent un fort taux d’ambiguïté. Un logiciel a été développé et est disponible [LO-5].

Apprentissage et Abstraction Nous avons défini dans des travaux antérieurs des treillis de Galois particuliers, les *treillis de Galois projetés* dont les éléments sont l’image, par une fonction ayant des propriétés particulières et appelée *projection*, des éléments

du treillis de Galois. Nous avons en particulier défini la projection *Alpha*, qui prend en compte une catégorisation, connue *a priori*, des instances et/ou des attributs.

Résultats :

- Construction incrémentale de ces nouveaux treillis et extraction des règles d'associations dites *règles Alpha* qui y sont associées [CI-39].
- Extraction et utilisation des règles d'association Alpha pour traiter des problèmes de recommandation [CI-58].
- Utilisation de la projection de treillis de Galois pour formaliser la notion d'*abstraction*, interprétation en logique modale de cette forme d'abstraction [CI-56].
- Généralisation des *treillis de concepts abstraits*, lui-même modal [CI-55].

2.1.2 Apprentissage non supervisé collaboratif et évolutif

La tâche d'apprentissage non supervisé consiste en la découverte de similarités entre les observations, dans le but de regrouper celles-ci en sous-ensembles, appelés *clusters* ou *classes*. Nous nous intéressons en particulier à la conception et la mise en œuvre de systèmes d'apprentissage qui font coopérer différents apprenants (modèles) entre eux. Ces systèmes constituent une autre génération des modèles d'apprentissage automatique et apparaissent comme une solution prometteuse pour la réalisation de tâches complexes et de grande taille. Les études fondamentales portent sur les algorithmes d'apprentissage pour ce type de système, la décomposition de tâches et la conception de chaînes de traitement optimales.

• Apprentissage non supervisé collaboratif à partir de données distribuées

Nous nous plaçons dans le cas où nous disposons d'une collection d'ensembles de données distribuées sur différents sites. Une approche collaborative permet de distribuer les classifications et procéder *a posteriori* à un enrichissement des différents résultats. Notre approche opère en deux phases : une phase *locale* qui consiste à appliquer un algorithme de classification localement et indépendamment sur chacune des bases de données et une phase de *collaboration*, qui consiste à faire collaborer chacune des bases de données avec toutes les classifications obtenues lors de la phase locale. On obtient ainsi comme résultat sur chacun des sites une classification proche de la classification qu'on aurait obtenue en faisant collaborer les bases de données elles-mêmes. A l'issue des deux phases, toutes les classifications sont enrichies. *Résultats :*

- Approche de classification collaborative permettant de tenir compte dans une classification non supervisée d'autres résultats de classifications sans avoir recours aux données utilisées par ces dernières [RI-19, CI-66]. Une reformulation mathématique de ces approches d'apprentissage collaboratif dans le cadre de l'inférence variationnelle est en cours.

• Apprentissage non supervisé évolutif à partir de flux de données

Les flux de données posent plusieurs problèmes qui rendent caduques les applications des techniques classiques d'analyse de données. En effet, ces bases de données sont perpétuellement en ligne et grossissent au fur et à mesure de l'arrivée de nouvelles informations. De plus, la distribution de probabilité associée à ces données peut changer au cours du temps (*dérive de concept*).

Apprentissage non supervisé du nombre de classes Ce thème concerne le développement d'approches pour la découverte et le suivi de structures de classes dans les données par apprentissage non supervisé. *Résultats :*

- Méthode de clustering à deux niveaux simultanés guidée par le voisinage et la densité *DS2L-SOM* [CI-20], qui se base sur l'estimation, à partir des données, de valeurs de connectivité et de densité des prototypes d'une carte SOM.

- Méthode de détection automatique du nombre de clusters [CL-11] sur la base des valeurs calculées ci-dessus. Cette nouvelle méthode permet aussi une description condensée de la distribution des données.

Mesure de (dis)similarité entre structures Les propriétés de l'algorithme précédent rendent possible l'analyse de grandes bases de données, y compris de grands flux de données car il permet d'obtenir des représentations synthétiques de la structure des données permettant un stockage efficace des informations du flux. Cette représentation couplée avec une mesure de (dis)similarité entre structures permet la détection de changements ou de dérive de concept [CI-34].

Résultats :

- Algorithme de suivi des données d'un flux permettant le stockage régulier de la structure des données, ainsi que la compression de ces informations au cours du temps [CI-87]. Les informations stockées peuvent ensuite être comparées entre elles pour l'analyse de l'évolution de la structure du flux de données. Ces approches ont été testées sur deux applications réelles pour le suivi d'individus dans un dispositif RFID (ANR Blanc Sillages [RI-21] et ANR CADI [CI-33]).

Activités de
recherche

Apprentissage non supervisé avec mémoire Dans le cadre de l'analyse de données évolutives, il est naturel d'utiliser l'historique des données afin de découvrir la structure des données avec une meilleure qualité.

Résultats :

- Nouvelle stratégie d'apprentissage pour les algorithmes de classification topographique basés sur le modèle *SOM* [CI-29], qui consiste à choisir le neurone le plus actif en tenant compte de son *historique* d'activation appris dans une matrice de vote à partir de l'ensemble des données. La performance observée de cette stratégie est effectivement meilleure (erreur topographique plus faible et pureté de la classification plus élevée).

Caractérisation des classes découvertes La fonction de base d'une méthode de sélection de variables (attributs/caractéristiques) est de choisir un sous-ensemble de variables pertinentes à partir de la représentation vectorielle des formes observées.

Résultats :

- Proposition d'approches qui consistent à pondérer les variables en fonction de leur participation à la qualité de la classification [CL-10, RE-10, RE-10]. Cette pondération permet d'estimer durant le processus de la classification la pertinence de chaque variable et ainsi de sélectionner des sous-ensembles de caractéristiques propres à chaque classe découverte.

• Apprentissage en distributions déséquilibrées

Plusieurs aspects peuvent influencer les systèmes d'apprentissage durant la phase de conception. Un de ces aspects est lié au déséquilibre des classes dans lequel l'effectif (nombre d'observations) d'une classe dépasse fortement celui des autres classes. Dans ce type de cas assez fréquent, le système d'apprentissage rencontre des difficultés liées au déséquilibre inter-classes. En effet, la plupart des algorithmes d'apprentissage sont basés sur l'hypothèse que les données d'apprentissage doivent être un échantillon *i.i.d.* (indépendant et identiquement distribué) représentatif de la population sur laquelle le modèle sera appliqué. Ces deux hypothèses ne sont pas respectées pour certains modèles quand ils sont construits à partir de données déséquilibrées.

Résultats :

- Approche d'apprentissage à partir d'une seule classe fondée sur un ensemble d'opérateurs de projection orthogonale et un double *bootstrap* [CI-65]. L'approche combine une technique de rééchantillonnage et l'idée d'apprentissage d'ensemble.

Grâce à son algorithme d'apprentissage en ligne, l'approche peut également suivre les changements dans les données au fil du temps.

- Méthode de sous-échantillonnage structurel adaptatif. Le processus procède par sous-échantillonnage des données majoritaires, guidé par les données minoritaires tout au long de la phase d'un apprentissage semi-supervisé. Cette approche permet de découvrir la structure des données avec une meilleure qualité (erreur topographique plus faible et pureté de la classification plus importante) [CI-46].

2.1.3 Apprentissage de structures à partir de données hétérogènes

Cet axe rassemble les travaux autour de la conception d'algorithmes d'apprentissage dédiés à certains types de données structurées. Nous déclinons ces travaux en les organisant par type de structures de complexité croissante : les séquences, les arbres/graphes, puis les programmes logiques.

• Apprentissage par modèles de mélanges à partir de données séquentielles

Dans le cadre d'un problème d'apprentissage non supervisé à partir de données séquentielles (*non i.i.d.*), nous avons étudié les modèles d'apprentissage qui consistent à découvrir un concept sous une forme géométrique et topologique. Nous avons proposé une approche d'apprentissage statistique, qui combine les points forts des cartes topologiques et des modèles de Markov cachés, pour construire des modèles de la dynamique d'un ensemble de séquences. Ces modèles généralisent les chaînes de Markov en introduisant une relation spatio-séquentielle entre les états cachés.

Résultat :

- Définition des Modèles de Markov Cachés Topologiques Auto-organisés et d'un algorithme d'apprentissage correspondant [CI-64, CI-63].

• Apprentissage de modèles topologiques et hiérarchiques

Nous avons combiné les cartes topologiques et un algorithme de classification hiérarchique biomimétique [CI-74, RI-23]. Cette approche originale a donné lieu à des contributions concernant le résumé de graphes et la classification hiérarchique incrémentale.

Résultats :

- Méthode de résumé de graphes, qui décompose le graphe d'origine de manière hiérarchique en plusieurs sous-arbres auto-organisés, formant ainsi une forêt d'arbres projetés sur une carte 2D. Cela fournit un espace de visualisation simple tout en proposant une représentation du graphe riche en information [CI-83].
- Modèle combiné, hiérarchique et incrémental de construction d'arbres s'inspirant du modèle *Neural Gas* [CI-82]. L'espace de sortie de l'algorithme est un ensemble d'arbres disposés selon une certaine topologie dans un espace bidimensionnel, fournissant à la fois des informations sur la structure des données et leur topologie.

• Analyse et fouille de graphes de terrain

L'étude et l'analyse de grands réseaux d'interactions, dits aussi *graphes de terrain* est un thème qui pose un ensemble des problèmes d'apprentissage difficiles. Ces graphes de terrain sont au centre d'un grand nombre d'applications, car ils permettent de modéliser naturellement des systèmes d'interactions complexes. Les graphes traités dans nos travaux sont des graphes biologiques (statiques), des réseaux sociaux vus à la fois sous leurs aspects statiques (apprentissage de communautés) et dynamiques (prévision de liens).

Réseaux d'interactions et données biologiques à grande échelle L'équipe conçoit depuis plusieurs années des algorithmes d'apprentissage appliqués à des problèmes variés de bioinformatique, apprenant ou manipulant des graphes d'interaction.

Résultats :

- Méthode de classification non supervisée de réseaux d'interactions locaux construits à partir de données à grande échelle [↑RI-2], produisant des modules de gènes co-régulés plus pertinents que des groupes de gènes co-exprimés [RE-4].
- Méthode de classification non supervisée de structures protéiques [RI-24*], qui s'appuie sur un pré-traitement original du graphe des similarités structurales entre paires de protéines et applique une contrainte ternaire de transitivité pour filtrer les arêtes du graphe.

Prévision de liens La *prévision de nouveaux liens* est un sous-problème de modélisation de l'évolution des graphes de terrain, qui a des applications concrètes importantes, notamment pour le calcul de recommandations. Nous nous sommes concentrés sur l'étude des graphes d'interaction bipartis dans le cadre d'une approche dite *topologique*, qui prévoit l'évolution d'un graphe en utilisant uniquement des indicateurs topologiques calculés à partir de son historique.

Résultats :

- Mesures topologiques dyadiques spécifiques pour exprimer similarité entre nœuds dans un graphe biparti et dans les graphes projetés associés [CI-52]. Ces nouvelles mesures permettent d'améliorer significativement la qualité du modèle de prévision de liens.
- Application de l'approche précédente sur des graphes tripartis modélisant l'évolution de *folksonomies* [CI-76]. Des expérimentations sur des données de folksonomies réelles montrent un net avantage de notre nouvelle approche par rapport au filtrage collaboratif en terme de précision des recommandations.
- Adaptation à un cadre supervisé de techniques classiques d'agrégation de préférences (non-supervisées) pour la prévision de liens [CI-77*]. La contribution de chaque attribut au modèle prédictif est apprise puis utilisée lors de l'agrégation des préférences.

Extraction de communautés Les approches de prévision de liens décrites précédemment étant coûteuses lorsqu'elles sont appliquées à des grands graphes, nous souhaitons utiliser la structure en *communautés* des graphes étudiés (une communauté est un sous-graphe composé de nœuds plus liés entre eux qu'avec le reste du graphe), afin d'obtenir plus efficacement des modèles de prévision de liens.

Résultats :

- Algorithme de détection de communautés éventuellement chevauchantes [CI-62] qui identifie des nœuds « *meneurs de communauté* », puis rattache les autres nœuds aux « meneurs » par des techniques d'agrégation de préférences. Le choix du rattachement à une communauté dépend des choix de ses voisins directs.
- En collaboration avec l'équipe AOC (une thèse co-encadrée), proposition d'un algorithme de recherche des régions denses en 1 maximale, éventuellement chevauchantes, dans des matrices binaires bruitées. Cet algorithme combine un algorithme de graphes (flot maximal/coupe minimale), à une étape combinatoire énumérant des « graines » de régions denses [CI-60]. Une extension récente de cette méthode met en place un mécanisme de semi-supervision par le biais d'une classification fournie *a priori* [CI-78].

• Apprentissage de programmes logiques

On s'intéresse ici à l'apprentissage dans les langages de concepts les plus complexes : programmes logiques multi-prédicats, éventuellement récursifs.

Activités de
recherche

2.2 Analyse AFOM/SWOT

Apprentissage pour la conception d'entrepôts de données Dans les applications d'ingénierie des entrepôts de données actuelles, des experts produisent des règles de transformation de modèles, par exemple d'un modèle UML en un modèle OLAP. Nos travaux ont pour but d'assister les experts dans la mise au point de ces règles de transformation, en utilisant les techniques d'apprentissage.

Résultats :

- Modélisation du problème de transformation de modèles, et caractérisation des problèmes où l'apprentissage peut apporter une plus-value [CI-30, CI-48, CL-12].
- Méthode d'apprentissage multi-prédicats issue de la Programmation Logique Inductive, qui exploite un graphe de dépendances entre concepts à apprendre [CI-68].

Construction de réseaux d'interactions géniques par extraction d'information à partir d'articles scientifiques La construction d'un réseau d'interactions géniques à partir de la littérature scientifique nécessite l'extraction de multiples interactions de différents types et potentiellement récursives.

Résultats :

- Modélisation de l'apprentissage des patrons d'extraction comme une tâche d'apprentissage d'ontologie. Application de techniques de Programmation Logique Inductive multi-prédicats à l'extraction de réseaux d'interactions sur un corpus bactérien, avec de très bonnes performances [CI-42].

2.2 Analyse AFOM/SWOT

• Atouts

- *Homogénéité et couverture scientifique de l'équipe*, lui assurant une bonne lisibilité extérieure sur ses thèmes de recherches ;
- *Attractivité du thème et de l'équipe dans le monde académique et industriel* : nombreux doctorants, bon taux d'insertion académique et industriel à l'issue de la thèse, nombreux partenariats académiques (ANR) et industriels (FUI) ;
- *Formation à la recherche* : adossement à un master très visible (*EID*²) ;
- *Nombreuses initiatives d'animation de la recherche* : organisation de conférences françaises, d'une école d'hiver, de sessions spéciales dans des conférences internationales ;
- *Valorisation de la recherche* : brevets, logiciels distribués, transfert technologique par création de start-up).

• Faiblesses

- Collaborations internationales à renforcer ;
- *Absence de chercheurs permanents* : la politique de recrutement de chercheurs CNRS doit être dynamisée, notamment à travers le recrutement de postdocs et un recours plus fréquent à des mois de chercheurs invités ;
- *Diffusion de logiciels à renforcer* : au-delà des brevets déposés, des logiciels sont développés dans l'équipe, mais leur diffusion est encore trop restreinte.

• Opportunités

- *Large couverture des thèmes de l'apprentissage*, ce qui est original dans le paysage français d'où la possibilité de recherches transversales en apprentissage ;
- *Insertion dans le PRES* : renforcement de collaborations notamment avec le LIPADE, sur des thèmes déjà existants dans l'équipe ;
- *Création du pôle Math-STIC*, qui permettra de donner un cadre pour formaliser des collaborations inter-équipes (équipes AOC et RCLN) et avec le L2TI.

- **Menaces**

- Certains axes sont fragiles, car majoritairement animés par des MCF HDR candidats à la promotion ou en délégation pour création entreprise (apprentissage relationnel/transition de phase, par exemple) ;
- *Surcharge administrative* : les enseignants-chercheurs de l'équipe assument des responsabilités pédagogiques/administratives lourdes (présidents des départements d'informatique à l'Institut Galilée et à l'IUT, directrice adjointe chargée de l'école d'ingénieurs).

2.3 Perspectives et projet

Au cours du dernier quadriennal, les effectifs de l'équipe ont connu une légère augmentation, chacun des trois axes de l'équipe ayant été renforcé équitablement. Nous nous sommes recentrés sur des axes forts de l'équipe, toujours avec l'optique d'allier recherche théorique et applications innovantes. Notre projet s'appuie sur des compétences déjà abordées dans le bilan et suit la structure des axes déjà exposés. Certains thèmes sont transversaux aux axes de l'équipe (représentation adaptative, apprentissage collectif), nous encouragerons les recherches croisées à travers un séminaire d'équipe très actif et des groupes de travail.

L'axe un est stable en effectif : départ en délégation pour création d'entreprise d'E. Alphonse, et arrivée récente de Y. Chevaleyre, qui renforce l'équipe avec son profil théorique, et ses compétences en particulier en apprentissage par renforcement et en décision collective. Suite à ces mouvements et ceux à prévoir (notamment A. Osmani, récemment HDR et en projet de délégation pour création d'entreprise), l'activité concernant la complexité moyenne de l'apprentissage relationnel va diminuer ; nous souhaitons orienter le projet de cet axe autour du thème *apprentissage et action* (section 2.3.1). Il permettra de mobiliser et de combiner les compétences de l'axe un autour de recherches originales dans le paysage de la recherche française et internationale. Nous le déclinons en deux sous-thèmes : apprentissage et planification, et apprentissage collectif.

L'axe deux est en légère progression suite à un départ en promotion PR (E. Viennet), et deux arrivées MCF (N. Grozavu et M. Lebbah), et projette de développer son activité scientifique vers le thème très porteur de l'apprentissage par transfert et de la factorisation matricielle (section 2.3.2). Ce projet prendra la suite des travaux déjà réalisés sur l'apprentissage collaboratif. Il introduit un nouveau formalisme basé sur la factorisation matricielle couramment utilisée pour la compréhension de la structure latente des données. Ce formalisme a suscité récemment une attention croissante en raison de son élégance mathématique et de ses résultats empiriques encourageants pour une variété d'applications. Ce formalisme sera utilisé pour trouver une représentation adéquate des données pour l'apprentissage par transfert. Il s'agit d'un défi pour ce type d'apprentissage pour lequel il faut développer des approches qui détectent et évitent le transfert négatif des connaissances utilisant très peu d'informations sur la tâche cible.

Un projet complémentaire s'appuyant sur les travaux autour des modèles de mélanges de l'axe deux est l'apprentissage non supervisé massivement distribué (section 2.3.3).

Enfin le projet fédérateur de l'axe trois, renforcé par l'arrivée de G. Santini, s'organisera autour de l'analyse et de la fouille de grands graphes (section 2.3.4), à la fois d'un point de vue statique et dynamique. Ce projet s'appuiera également sur les résultats obtenus sur ce thème pendant le dernier quadriennal et leur donnera une perspective plus large. Cet axe est privilégié pour mener des recherches transverses au sein du laboratoire et de l'Université Paris 13, en collaboration avec les équipes AOC et RCLN, et au sein du projet Math-STIC.

Par ailleurs, au delà des projets décrits ci-dessous, nous continuerons à travailler sur les thèmes de recherche pour lesquels notre expertise est déjà reconnue, et qui

2.3 Perspectives et projet

caractérisent notre équipe, en particulier, l'apprentissage relationnel, l'analyse formelle de concepts, les systèmes de recommandation, les problèmes de décision collective . . . Nous avons précisé le rattachement principal de chaque membre permanent de l'équipe à un projet.

2.3.1 Apprentissage et Action

Participants : D. Bouthinon, M. Champesme, Y. Chevaleyre, P. Gérard, A. Osmani, H. Soldano

- **Apprentissage et planification dans des environnements structurés**

La plupart des algorithmes d'apprentissage par renforcement font l'hypothèse que l'environnement peut être représenté comme un Processus de Décision de Markov (PDM) comportant un nombre raisonnable d'états. Malheureusement, cette hypothèse ne tient pas dans la plupart des applications du monde réel. Par exemple, pour le jeu vidéo *Tetris*, sur lequel de nombreux algorithmes d'apprentissage par renforcement ont été testés, l'espace d'états comporte environ 10^{61} états [12]. Dès lors, les algorithmes qui ne se fondent sur aucune autre hypothèse que l'existence d'un PDM sous-jacent obtiennent ici de piètres résultats. De nombreux autres aspects, partiellement ou totalement ignorés par la communauté, nous paraissent cruciaux et constituent notre projet.

Étude des langages de représentation pour les PDMs factorisés Jusqu'à présent, pour représenter les PDMs factorisés dont l'espace d'états est un vecteur de bits, on trouve principalement (voir par exemple [10]) des fonctions de transition représentées sous la forme d'un réseau bayésien dynamique, et des fonctions de récompense représentées comme des fonctions additivement décomposables. Se limiter arbitrairement à ce type de fonctions ne semble pas judicieux face à l'extrême diversité des applications de l'apprentissage par renforcement. Nous avons déjà entrepris d'étudier d'autres langages de représentations, en particulier ceux issus de la logique, pour représenter les diverses composantes d'un PDM. Nous chercherons donc pour divers *fragments* de la logique d'ordre 0 et 1, à vérifier si ces langages offrent des propriétés désirables pour l'apprentissage par renforcement.

Complexité algorithmique La complexité algorithmique des opérations usuelles (planification, évaluation de politique) est un aspect crucial associé au choix du langage de représentation. Il existe déjà de nombreux résultats concernant la complexité algorithmique pour les PDMs factorisés classiques [14], mais tout reste à démontrer en logique d'ordre 1.

Compacité de représentation des stratégies Ainsi que l'a montré Kearns [1], il se peut que dans un PDM représenté de façon compacte, la meilleure stratégie (qui est ce que l'on cherche à calculer) ne possède pas de représentation de taille polynomiale. Ce type de résultats constitue un obstacle théorique bien plus profond que la complexité algorithmique pour résoudre des PDMs complexes. Nous chercherons (et nous avons déjà des résultats) à caractériser la complexité de représentation des langages que nous aurons choisis.

Conception d'algorithmes d'apprentissage par renforcement Ce point constitue l'aboutissement de ce projet. Pour ce faire, nous explorerons deux pistes. D'abord, nous explorerons le *lien avec l'apprentissage supervisé*. Récemment, il a été montré (entre autres dans [7, 13]) que l'on pouvait construire un algorithme d'apprentissage par renforcement à partir d'un algorithme d'apprentissage supervisé. Ces recherches en sont à leurs prémises, et de très nombreuses questions fondamentales n'ont pas encore été

abordées, en particulier le problème de l'exploration : comment guider l'exploration de l'agent apprenant lorsque cet agent exploite un algorithme d'apprentissage par renforcement? Une thèse a débuté en octobre 2012 sur cette question. Enfin, un dernier aspect relatif à la conception d'algorithmes sera abordé dans ce projet : il s'agit de *l'étude de l'apprentissage par renforcement actif*. A notre connaissance, cette question n'a jamais été traitée autrement que dans les PDMs non structurés [5].

- **Apprentissage Collectif et Action**

Nous avons récemment entamé des travaux étudiant l'apprentissage dans une société d'agents autonomes et dotés de capacités de communication. Nous proposons ici d'étendre le modèle d'apprentissage collectif *apprenant-critique* proposé dans un cadre purement cognitif (les agents changent leurs connaissances mais pas l'état du monde), à un cadre où les agents effectuent également des actions et changent ainsi leur environnement. Les contraintes propres à l'apprentissage collectif, comme la nécessité d'avoir un apprentissage en ligne, sous la forme d'un mécanisme de révision incrémental, restent les mêmes, mais plusieurs questions prennent alors un relief nouveau.

Ainsi un exemple, dans le cas de la révision d'un modèle d'action, est un triplet état-courant/action/nouvel-état dans lequel les états (courant et nouveau) peuvent être partiellement observés. Nous retrouvons alors la problématique de l'apprentissage en situation d'ambiguïté sur laquelle nous avons récemment travaillé [CI-35] : une partie seulement des faits sont déterminés, et des connaissances sur les états possibles, sous la forme de contraintes, aident à la convergence de l'apprentissage. Nous nous proposons d'abord d'exploiter nos travaux actuels sur l'apprentissage à partir de représentations ambiguës pour concevoir un mécanisme de révision permettant de traiter les observations partielles. La révision collective soulève alors la question de la distribution parmi les agents des informations concernant un même événement observé. Des protocoles d'interaction plus fins que ceux proposés pour les observations indépendantes restent à trouver. On s'appuiera par exemple sur les modèles d'interactions abductives, pour lesquels les interactions sont plus riches que celles que nous avons utilisées dans le cas purement inductif [CI-51]. Nous souhaitons proposer, et caractériser les propriétés de tels protocoles.

Enfin nous nous intéressons également aux aspects temporels des interactions liées à l'apprentissage, en particulier au cas où plusieurs théories courantes se propagent simultanément dans une société d'agents structurée en un graphe de communication. Une question posée est alors celle du choix et de la fusion de ces « croyances ». C'est un point actuellement très étudié [9] mais peu dans une perspective d'apprentissage collectif. Notre but est de nous donner ainsi, à travers un travail expérimental, une compréhension de ces aspects dynamiques et d'avancer vers leur modélisation.

**Activités de
recherche**

2.3.2 Transformation de l'espace de description pour l'apprentissage par transfert

Participants : Y. Bennani, N. Grozavu

L'apprentissage par transfert est le processus par lequel un individu utilise un apprentissage acquis dans une situation pour l'appliquer à une autre situation [16]. L'apprentissage par transfert implique donc d'utiliser de la connaissance acquise sur un jeu de tâches afin d'améliorer les performances pour une autre tâche analogue. Par exemple, les athlètes se servent de l'apprentissage par transfert quand ils pratiquent des activités auxiliaires afin d'être plus compétitifs dans leur activité principale.

L'apprentissage par transfert d'une certaine tâche cible est donc très dépendant de l'apprentissage d'une ou plusieurs tâches auxiliaires et il peut dégrader les performances sur la tâche cible si les tâches auxiliaires sont trop dissemblables de cette

2.3 Perspectives et projet

dernière. Un défi pour l'apprentissage par transfert est donc de développer des approches qui détectent et évitent le transfert négatif des connaissances, en utilisant très peu d'informations sur la tâche cible.

Par ailleurs, un problème fondamental dans de nombreuses tâches en apprentissage artificiel est de trouver une représentation adéquate des données. Une représentation utile rend généralement la structure latente dans les données explicite, et réduit souvent la dimensionnalité des données afin que d'autres méthodes de calcul puissent être appliquées. La *factorisation matricielle* est une approche couramment utilisée pour la compréhension de la structure latente de matrices de données pour diverses applications. Ces méthodes matricielles ont suscité récemment une attention croissante en raison de leur élégance mathématique et de leurs résultats empiriques encourageants pour une variété d'applications. Il existe de nombreuses formes de factorisation matricielle : Nonnegative Matrix Factorization (NMF) [4], Maximum margin matrix factorization (MMMF) [15], Probabilistic NMF [17] . . . Des travaux antérieurs ont montré que si la positivité est respectée, les résultats de la factorisation seront plus faciles à interpréter tout en étant comparables, voire meilleurs que ceux obtenus par d'autres techniques.

L'objectif de ce projet de recherche est donc d'une part de développer et d'étudier des méthodes de factorisation matricielle pour trouver une représentation adéquate des données dans le cadre de l'apprentissage par transfert, d'autre part d'identifier quelle connaissance doit être transférée et comment, et enfin d'exhiber les avantages et les inconvénients de ce paradigme d'apprentissage automatique avec des illustrations sur des données réelles.

Les principaux axes de recherche à explorer dans cette étude sont les suivants :

- Trouver des représentations adéquates des données pour ce type d'apprentissage ;
- Développer des approches adaptatives qui détectent et évitent le transfert négatif des connaissances en utilisant très peu d'informations sur la tâche cible ;
- Développer des approches adaptatives de transfert des connaissances ;
- Proposer des méthodes permettant d'étudier les propriétés de ce type d'apprentissage.

Ce projet de recherche se situe dans la continuité d'une part de l'un des axes majeurs développé par l'équipe depuis de nombreuses années portant sur l'apprentissage non supervisé. Une thèse a récemment débuté sur ce sujet, qui est d'autre part dans la continuité de recherches récemment développées dans le groupe, liées à l'apprentissage collaboratif et distribué (projet ANR MN COCLICO 2012-2016) et l'apprentissage à partir de flux de données multi-sources (Projet FUI Hermès 2012-2015).

2.3.3 Apprentissage non supervisé massivement distribué

Participants : H. Azzag, M. Lebbah

Actuellement, les algorithmes d'apprentissage existants ne sont pas adaptés aux données massives (*BigData*) qui peuvent atteindre des tailles de l'ordre du petaoctet ou du zettaoctet [6, 8]. La plupart des méthodes d'apprentissage exploitent une part réduite des informations générées et récoltées dans les systèmes d'information. La parallélisation massive du processus d'apprentissage est donc une voie à explorer.

Un nouveau paradigme *Map-Reduce* a été récemment proposé pour surmonter cet obstacle [2, 18]. Les travaux actuels, même si on note un projet à Berkeley [19], donnent des performances faibles pour les algorithmes itératifs du domaine, ces algorithmes n'ayant pas été conçus dans un cadre massivement distribué.

Le centre de ce projet est de développer des modèles d'apprentissage non supervisés (*clustering* et *co-clustering*) qui fonctionnent efficacement en parallèle et particulièrement avec les nouveaux paradigmes de parallélisation massivement distribués. Des méthodes d'apprentissage non supervisé distribuées peuvent être mises en œuvre de plusieurs manières [11, 3] selon que les algorithmes et les données sont centralisés ou

non. Le développement des approches visées nécessite de relever différents défis, en particulier :

- Elles devront être robustes aux choix des modèles et aux données aberrantes : la décentralisation des algorithmes et des données ne doit pas causer une dégradation de l'estimation du modèle global ;
- Elles devront être automatiquement décentralisées, massivement distribuées et appliquées d'une manière distribuée sur des données multi-vues, ce découpage multi-vue pouvant être lié i) à la nature des données décrites par des variables de natures et de sources différentes ii) à la répartition des données dans un système distribué ;
- Elles devront respecter l'anonymat et la confidentialité des données privées.

Ce travail pourra se faire en collaboration avec l'équipe AOC (axe Algorithmes, logiciels et architectures distribués). En effet ce groupe a une expertise scientifique et technique sur les architectures massivement parallèles comme le Cloud computing – il a par exemple « cloudifié » le paradigme MapReduce dans le Cloud SlapOS qu'il utilise dans le cadre du projet Resilience. L'équipe CALIN pourra également apporter à ce projet ses compétences en cryptographie.

Activités de
recherche

2.3.4 Analyse et fouille de graphes de terrain

Participants : S. Guérif, R. Kanawati, C. Rouveirol, G. Santini

La multiplication des sources de données en évolution continue (notamment en *open access*) et la grande quantité de données disponibles ne posent pas que des défis techniques d'exploitation mais offrent aussi une source précieuse pour imaginer de *nouveaux outils d'analyse et à l'exploitation de ces données complexes et éventuellement évolutives*.

Dans cette partie, nous considérons que ces données complexes sont vues comme un ensemble de *graphes complexes et dynamiques organisés en couches interconnectées*. Par exemple, à partir d'une base bibliographique classique, différents réseaux peuvent être extraits, un *réseau de co-publications* un *réseau de citations*, et différents *réseaux de termes* apparaissant dans les textes des publications (réseaux de co-occurrences, ...). Nous appelons réseau hétérogène l'union de ces différents réseaux.

Nous souhaitons développer d'une part des approches *efficaces* pour l'analyse et la caractérisation de *l'évolution conjointe* des différents réseaux dans un réseau hétérogène. Notre objectif est de construire des *modèles explicatifs* (par opposition à des modèles numériques, ou à base de calcul tensoriel) de l'évolution d'un réseau en s'appuyant sur l'historique du réseau analysé et des autres réseaux reliés. Nous pensons plus particulièrement à des applications de *recommandation* et de *détection de nouveauté*.

Le développement des applications visées nécessite de proposer des solutions efficaces pour de nombreux problèmes difficiles, en particulier :

- la détection de communautés (recouvrantes) et de leur évolution à partir de multiples points de vues ;
- la fouille simultanée de différents graphes dynamiques interconnectés. Nous avons obtenu dans nos travaux précédents de bons résultats en utilisant une représentation propositionnelle d'un réseau simple (mono-relationnel). Nous souhaitons concevoir des approches de fouille de graphes (supervisées et non supervisées) afin de prédire l'évolution des réseaux hétérogènes (multi-relationnels) finement, c'est-à-dire au niveau de l'apparition/disparition d'un lien ou d'un groupe de liens. Ces approches étant coûteuses, nous utiliserons afin de passer à l'échelle les méthodes de détection de communautés ci-dessus pour limiter la portée des méthodes de fouille de graphes.

D'autre part, nous poursuivrons nos travaux sur l'inférence de réseaux biologiques à partir de données hétérogènes, en mettant en place des méthodes hybrides de clas-

sification supervisée/non supervisée, reposant sur des techniques robustes de création/sélection de variables dans des données graphes.

Références

- [1] E. Allender, S. Arora, M. Kearns, C. Moore, and A. Russell. A note on the representational incompatibility of function approximation and factored dynamics. *Advances in Neural Information Processing Systems*, pages 447–454, 2003.
- [2] J. Dean and S. Ghemawat. Mapreduce : Simplified data processing on large clusters. In *6th Symposium on Operating System Design and Implementation (OSDI 2004)*, San Francisco, California, USA, December 6-8, 2004, pages 137–150. USENIX Association, 2004.
- [3] A. El Attar, A. Pigeau, and M. Gelgon. A decentralized and robust approach to estimating a probabilistic mixture model for structuring distributed data. In *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Volume 01, WI-IAT '11*, pages 372–379, Washington, DC, USA, 2011. IEEE Computer Society.
- [4] L. Eldén. *Matrix Methods in Data Mining and Pattern Recognition*. Society for industrial and applied mathematics, Philadelphia, 2007.
- [5] A. Epshteyn, A. Vogel, and G. DeJong. Active reinforcement learning. *Proceedings of the 25th international conference on Machine learning*, pages 296–303, 2008.
- [6] U. Fayyad. A data miner's story - getting to know the grand challenges. In Invited Innovation Talk, KDD, 2007 : Slide 61. Available at : http://videolectures.net/kdd07_fayyad_dms/. 2007.
- [7] A. Fern, R. Givan, and S. Yoon. Approximate Policy Iteration with a Policy Language Bias : Solving Relational Markov Decision Processes. *arXiv.org*, cs.AI, Sept. 2011.
- [8] R. L. Ferreira Cordeiro, C. Traina, Junior, A. J. Machado Traina, J. López, U. Kang, and C. Faloutsos. Clustering very large multi-dimensional datasets with mapreduce. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '11)*, pages 690–698, New York, NY, USA, 2011. ACM.
- [9] D. Gabbay, G. Pigozzi, and O. Rodrigues. Common foundations for belief revision, belief merging and voting. In G. Bonanno, J. Delgrande, J. Lang, and H. Rott, editors, *Formal Models of Belief Change in Rational Agents*, number 07351 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2007. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.
- [10] C. Guestrin, D. Koller, R. Parr, and S. Venkataraman. Efficient solution algorithms for factored MDPs. *J. Artif. Intell. Res. (JAIR)*, 19 :399–468, 2003.
- [11] E. Januzaj, H.-P. Kriegel, and M. Pfeifle. DBDC : Density based distributed clustering. In *Proceedings of the 9th Int. Conf. on Extending Database Technology (EDBT 2004)*, volume 2992 of *Lecture Notes in Computer Science*, pages 88–105. Springer, 2004.
- [12] M. G. Lagoudakis and R. Parr. Least-squares policy iteration. *The Journal of Machine Learning Research*, 4, Dec. 2003.
- [13] A. Lazaric, M. Ghavamzadeh, and R. Munos. Analysis of a classification-based policy iteration algorithm. *Rapport INRIA : inria-00482065, version 2*, 2010.
- [14] M. L. Littman, J. Goldsmith, and M. Mundhenk. The computational complexity of probabilistic planning. *Journal of Artificial Intelligence Research*, 9(1), Aug. 1998.
- [15] J. D. M. R. Nathan Srebro and T. Jaakkola. Maximum-margin matrix factorization. In A. Press, editor, *Advances in Neural Information Processing Systems (NIPS 04)*, 2004.
- [16] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10) :1345–1359, 2010.
- [17] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In A. Press, editor, *Advances in Neural Information Processing Systems (NIPS 07)*, pages 1257–1264, 2008.
- [18] T. White. *Hadoop : The Definitive Guide*. O'Reilly Media, 2009.
- [19] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauley, M. J. Franklin, S. Shenker, and I. Stoica. Resilient distributed datasets : a fault-tolerant abstraction for in-memory cluster computing. In *Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation, NSDI'12*, pages 2–2, Berkeley, CA, USA, 2012. USENIX Association.

Chapitre 3

Production scientifique, valorisation, dissémination

L'équipe A³ a largement diffusé ses résultats, et ceci autant dans la communauté académique en terme de publications, qu'à destination de l'industrie (brevets, logiciels distribués, création de start-up). Les publications scientifiques de l'équipe sont en nette progression ; la part des publications de l'équipe dans des conférences et de revues de premier plan est en augmentation, même si elle est encore sujette à amélioration.

Les recherches de l'équipe ont également débouché sur des logiciels, qui ont été développés et testés sur des projets concrets. Un effort de distribution à plus grande échelle fait partie de nos projets pour la prochaine période, notamment pour les axes 1 et 3. L'équipe A³ a poursuivi son investissement à de nombreux projets, de types très diversifiés (ANR Blancs et thématiques, FUI, Institut National du Cancer, BQR). Ces projets permettent d'évaluer le bien-fondé des modèles d'apprentissage proposés, et soulèvent également de nouvelles questions scientifiques propres renouveler les thématiques de recherche.

3.1 Publications

3.1 Publications

	2007	2008	2009	2010	2011	2012	Total
Livres	0	0	0	0	0	0	0
Chapitres de livre	3	2	3	3	2	0	13
Revue internationale	4	3	6	9	1	7	30
Autres revues	3	4	0	0	6	0	13
Conférences internationales	13	11	14	16	20	13	87
Autres communications	14	14	13	21	17	16	95
Total	37	34	36	49	46	36	238

3.2 Contrats de recherche

Production scientifique, valorisation, dissémination

Type	Coordinateur Resp. scient.	Nom	Financier	Période	Durée (mois)	Budget global (K€)	Part LIPN (K€)
ANR							
ANR	Y. Bennani	SILLAGES	MENESR	2005/2008	36	375	113
ANR	E. Viennet et Y. Bennani	SEPTIA	MENESR	2007/2009	24	554	100
ANR	Y. Bennani	CADI	MENESR	2008/2010	24	1 159	200
ANR JCJC	P. Gérard	HARRI	MENESR	2008/2011	36	126,2	
(ANR)	Y. Chevalere	(LARDONS)	MENESR	2010/2014	48	-	-
(ANR)	Y. Chevalere	(COMSOC)	MENESR	2010/2012	24	-	-
ANR	Y. Bennani	E-FRAUD	MENESR	2010/2013	36	1 092	200
ANR	Y. Bennani	COCLICO	MENESR	2012/2016	48	1 015	231
FUI							
FUI-06	A. Nazarenko	Infom@gic Oséo	Region IdF	2006/2009	36	19 300	250
FUI	R. Kanawati	Urban-D	Cons. Gén. 93	2010/2012	18	1 179	80
FUI	Y. Bennani	Hermes	MENESR	2012/2015	36	1 907	141,5
National							
Rech. Bioméd.	C. Rouveïrol	INSIGHT	INCa	2010/2013	36	464	117,3
Local							
BQR	Y. Bennani	FACE	UP13	2006/2008	24	30	
LIPN	L. Létocart	MOTIFS	LIPN	2009/2010	12	10	
BQR	R. Kanawati	LIPS	UP13	2009/2010	12	25	
BQR	C. Rouveïrol	ONTO-APP	UP13	2012/2013	12	18,5	36

Les projets entre parenthèses sont des projets dans lesquels des membres du LIPN sont impliqués mais pour lesquels le LIPN ne fait pas partie des partenaires.

• ANR SILLAGES : Suivi Individuel par Label électronique Au sein de Groupes complexes (2006-2008)

Partenaires Académiques : LIPN-CNRS, LEEC-CNRS, ULB (Université Libre de Bruxelles).
Industriels : SPACECODE

Les structures sociales animales sont des systèmes dynamiques caractérisées par un grand nombre d'interactions entre individus. Malgré leur diversité, ces structures sont basées sur un nombre limité de règles génériques (amplifications, inhibitions, compétition pour les ressources). Les dynamiques du groupe et ses réponses résultent de la synergie entre ces interactions, des capacités individuelles à traiter l'information et de la diversité de ces réponses individuelles.

Les objectifs de ce projet sont :

- le développement de nouveaux systèmes autonomes de suivi de l'activité spatio-temporelle des groupes, systèmes donnant accès à des données inaccessibles actuellement ;
- le développement d'outils d'analyse automatique de ces données et de quantification des comportements individuels ou collectifs ;
- l'intégration de ces quantifications aux outils de modélisation qui permettent de valider le lien entre les niveaux individuels et collectifs ;
- ces systèmes autonomes de suivi pourront contrôler des dispositifs et modifier l'environnement en fonction de l'identité des individus. Ceci permettra de tester notamment des hypothèses portant sur la plasticité comportementale.

Ces objectifs font de SILLAGES une action interdisciplinaire à l'intersection des sciences du comportement et du complexe et des sciences informatiques et de l'ingénieur. Les publications de l'équipe associées à ce projet sont [CO-37, RI-21, CI-71].

● **ANR SEPTIA : Solution automatique pour le traitement des interactions clients (2007-2009)**

Partenaires *Académiques* : LIPN-CNRS, LIP6-CNRS. *Industriels* : KXEN, TEMIS, EP-TICA

Le projet Septia visait à produire un logiciel avancé de traitement des interactions clients dans le cadre d'un centre de contacts. Axé de façon prédominante sur le traitement des emails, le projet a étudié également l'apport des données clients, des données d'interactions téléphoniques et des données d'interactions inter-clients dans un portail collaboratif. Les tâches principales identifiées :

- Extraction automatisée des thèmes des emails et constitution des ontologies associées, en utilisant des techniques statistiques et sémantiques.
- Aide à la réponse (recherche automatique du « meilleur » interlocuteur.
- Supervision des activités (prévision de séries temporelles)
- Portail collaboratif : étendre les outils à un portail (forum, base client, chats, ...)

Le projet prévoit la construction d'un démonstrateur qui sera testé sur des données réelles de grande taille. Il sera l'occasion de développer des technologies avancées dans le domaine de la fouille de texte, de la fouille de données et de l'étude des « réseaux sociaux ». La principale publication de l'équipe associée à ce projet est [CI-11].

● **ANR CADI : Composants Avancés pour la Distribution (2008-2010).**

Partenaires *Académiques* : LIPN-CNRS, LIP6-CNRS, LITIS. *Industriels* : KXEN, Num-sight, SPACECODE, SAMSE - La Boîte à Outils, MONDOMIX.

Le projet CADI vise à produire un prototype de composants logiciels pour la Distribution, tant brick-and-mortar (magasins) que vente en ligne (site de e-commerce). Dans le modèle économique de la distribution dit de la « Longue Traîne », on voit émerger d'un catalogue potentiellement illimité des multiplicités de marchés de micro-niches que le site commerçant peut espérer développer en prenant en compte toutes les sources d'information concernant les clients et les produits (communautés qui se créent sur les blogs, forums autour de ses produits, traces de navigations des clients sur le site, etc.). Visant un marché à l'intersection de quatre domaines en forte croissance (e-commerce, data mining, réseaux sociaux et RFID), le LIPN au sein du projet CADI a conçu un moteur de recommandation à base de règles d'associations abstraites et a montré que cette approche se comparait favorablement aux autres approches développées par les autres partenaires du projet [CI-58], conçu et expérimenté un algorithme de prévision de liens sur des données d'achat modélisées dans un graphe biparti [CO-59, CI-52]. Parmi les autres publications annexes : [RI-22, CI-50, CI-33].

Production scientifique, valorisation, dissémination

3.2 Contrats de recherche

• ANR JCJC HARRI

Le projet HARRI a pour objectif de résoudre des problèmes d'apprentissage par renforcement et plus généralement d'adaptation d'un agent informatique dans des contextes où les représentations relationnelles sont plus naturelles (objets discrets liés par des relations plutôt que vecteurs propositionnels). C'est souvent le cas pour les agents simulés et moins pour les robots et autres agents physiques. Plus particulièrement, nous avons porté nos efforts sur la question de l'apprentissage incrémental d'un modèle relationnel d'action (règles d'ordre 1) pour permettre à l'agent d'anticiper les conséquences de ses actions et accélérer son adaptation en ligne à un environnement a priori inconnu. Le problème d'apprentissage incrémental est traité comme un problème de révisions successives de modèles [CI-40]. Les algorithmes correspondants ont la particularité d'être ascendants et guidés par les données. Ils convergent dans le cas déterministe et ont montré empiriquement leur efficacité, y compris dans le cas de données bruitées [CI-41]. Les modèles d'action appris peuvent être utilisés en apprentissage par renforcement indirect ou en planification. Des mécanismes originaux d'apprentissage actif ont été proposés, adaptés au caractère ascendant de notre approche [CI-75].

• Lardons : Learning and Reasoning for Deciding Optimally using Numerical and Symbolic information. (2010-2014)

Partenaires LIP6-CNRS, GREYC-CNRS, LAMSADE, INRA

LARDONS est un projet de recherche fondamentale s'intéressant aux capacités d'apprentissage, de raisonnement et de prise de décision d'un agent artificiel autonome. Les objectifs de ce projet sont :

- Proposer de nouveaux modèles tenant compte du fait qu'un agent dispose typiquement de connaissances et de perceptions à la fois symboliques et numériques.
- Concevoir de nouvelles approches et de nouveaux algorithmes pour les problèmes en question, qui soient à la fois capables de traiter, mais aussi de tirer parti de cette dualité symbolique/numérique.

• COMSOC : Choix Social Computationnel. (2010-2012).

Partenaires LAMSADE-CNRS, CRIL-CNRS, PREG-CNRS, CREM-CNRS

La théorie du choix social est une branche de l'économie qui s'intéresse à la prise de décision collective, par l'intermédiaire entre autres de procédures de vote, de négociation, d'enchères. Ce projet explore les aspects computationnels de la prise de décision collective.

• ANR E-FRAUD BOX : Détection et Investigation de la fraude à la carte bancaire sur Internet (2010-2013)

Partenaires *Académiques* : LIPN-CNRS, LIP6-CNRS. *Industriels* : ALTIC, KXEN, THALES Communications France (TCF), GIE Cartes Bancaires « CB », Gendarmerie Nationale, Police Nationale.

Le projet E-fraud Box vise à développer une boîte à outils intégrée, dédiée à la détection et à l'investigation de la fraude à la carte bancaire sur Internet, et basée sur :

- des algorithmes d'apprentissage capables de modéliser les comportements frauduleux, de suivre la dynamique de la fraude dans le temps et d'étudier les interactions entre les transactions.
- des moteurs de détection et d'investigation de la fraude combinant les résultats des modèles.

Les moteurs seront utilisés pour :

- détecter le plus rapidement possible les cartes utilisées frauduleusement,
- construire des faisceaux d'indices caractérisant des comportements frauduleux et des scénarios de fraude,

- identifier les nouveaux modes opératoires des fraudeurs.

Les travaux développés dans l'équipe dans le cadre de ce projet ont mené aux publications suivantes : [CI-46, CI-65, CO-89] .

- **ANR COCLICO : COllaboration, CClassification, Incrémentalité et COnnais-sances (2012-2016)**

Partenaires *Académiques* : LSIIT, LIPN, AgroParisTech/INRA. *Industriels* : LIVE, ES-PACE DEV/IRD

La fouille de données est un maillon important dans la chaîne de traitements des données vers les connaissances. Ainsi, par exemple, appréhender et comprendre les processus de fonctionnement et de développement des systèmes plus ou moins anthropisés à diverses échelles spatiales et temporelles (urbanisation et pression sur les terres, érosion de la biodiversité etc.) à partir de données satellites ou autres devient un élément majeur dans différents domaines telles l'étude de l'environnement ou les politiques publiques d'aménagement du territoire. Or les techniques d'analyses actuelles sont de plus en plus limitées face à l'avalanche actuelle de données hétérogènes souvent incomplètes voire imprécises et de plus en plus souvent fournies en continu au fil de l'eau. Coclico est un projet de recherche visant à étudier et proposer une méthode générique innovante permettant une analyse multi-échelle de grands volumes de données spatio-temporelles fournies en continu de qualité très variable, mettant en œuvre une approche multistratégie incrémentale dans laquelle la collaboration entre les différentes méthodes de fouille de données sera guidée par des connaissances à la fois du domaine thématique (Géosciences, Géographie) formalisées en ontologies et du domaine de l'analyse (connaissances sur les méthodes), et garantissant un objectif de qualité finale prenant en compte à la fois la qualité des données et celles des connaissances.

Production scientifique, valorisation, dissémination

- **FUI Infom@gic (2006-2009)**

Partenaires *Académiques* : Universités Paris 6, Paris 8, Marne La vallée, GET, Limsi-CNRS. *Industriels* : Thales (coordination), EADS, ONERA, et une vingtaine de PME, dont Bertin, Vecsys, Sinequa.

Ce projet consistait à mettre en place, sur une période de trois ans, un laboratoire industriel de sélection, de test, d'intégration et de validation sur des applications opérationnelles des meilleures technologies franciliennes dans le domaine du traitement de l'information. Ce laboratoire s'appuie sur une plateforme UIMA commune couvrant tous les grands domaines de l'analyse d'information (la recherche et l'indexation, l'extraction de connaissances et la fusion d'informations), quelles que soient les sources de données considérées (données structurées, texte, images et sons). Les travaux réalisés par l'équipe dans ce projet ont donné lieu aux principales publications suivantes :[CI-29, CI-28, CI-6, CI-16, CL-9].

- **FUI UrbanD**

Partenaires *Académiques* : Télécom ParisTech. *Industriels* : LAA, UfO, Thinkout, Preview.

Le projet UrbanD consiste à développer un système de notation et de suivi (scoring and monitoring) destiné à augmenter la qualité de vie au sein des territoires urbains en vérifiant l'amélioration conjointe de réalisation des objectifs du développement durable. L'objectif innovant d'UrbanD est de lier qualité d'usages et développement durable en éclairant la compréhension des interactions positives et négatives au sein des éco-systèmes urbains.

Dans le cadre de ce projet, le LIPN a travaillé plus particulièrement sur le développement des approches automatiques pour l'analyse des rythmes des territoires. L'approche proposée consiste à modéliser la dynamique des territoires par un réseau complexe dynamique. Des approches de détection automatique de communautés sont utili-

3.2 Contrats de recherche

sées afin de découvrir des rythmes dans des vastes territoires difficiles à analyser avec les approches manuelles habituellement employées par les urbanistes et les anthropologues [CI-62, CO-80].

• **FUI Hermès : Relation Client Personnalisée et Contextualisée (2012-2015)**

Partenaires *Académiques* : INRIA, LEM, LIFL, LIPN, LITIS. *Industriels* : Auchan, Blogbang, Brand Alley, Cylande, Keyneosoftware, Leroy Merlin, Norsys, NumSight.

Le projet Hermès (FUI) adresse le champ du marketing one-to-one cross-canal et veut développer autant que possible la pertinence de l'information communiquée au client, dans le strict respect des règles d'éthique et de privacy. A partir des multiples sources hétérogènes susceptibles de délivrer de très nombreuses données sur le client (canaux traditionnels, devices connectés, réseaux sociaux, ...), il s'agira d'en extraire les connaissances utiles au fil de l'eau. A partir de ces connaissances, les moteurs de recommandations permettront de déterminer le contexte d'achat dans lequel se situe le client (l'étape de son parcours d'achat, sa motivation, son projet, son état d'esprit, ses contraintes, son comportement d'achat, ...), les propositions à faire au client, le contexte dans lequel elles lui seront délivrées (sur quel canal, à quel moment, Idots). Hermès a pour objectif de construire une plate forme de marketing contextuel générique, modulaire, industrialisée et automatisée capable :

- d'extraire des données provenant de tous les canaux (caisse, web, email, centre d'appel, réseaux sociaux, devices ...) ;
- de déterminer les contextes d'achat du client ainsi que les canaux pertinents d'interaction avec ce client ;
- de communiquer avec les systèmes opérationnels sur tous les canaux (notamment avec les devices en magasin et les réseaux sociaux ...).

• **INSIGHT : Identification de cibles thérapeutiques par la recherche de réseaux spécifiquement altérés dans les cancers (2010-2013)**

Partenaires Institut Curie ; LIPN-CNRS ; Equipe Recherche translationnelle en oncogénèse urogénitale, INSERM, Créteil ; Institut de Biologie Systémique et Synthétique (iSSB), Université d'Evry.

À partir des données du transcriptome, de l'expression des miARNs et des protéines des cancers de vessie et d'urothélium normal à différents stades de prolifération et de différenciation, le but du projet est de construire des réseaux de régulation — facteurs de transcription, miARN et gènes cibles — observés dans le tissu normal et pathologique, et par l'analyse de ces réseaux, de suggérer des régulations significativement perturbées dans les réseaux actifs dans le tissu pathologique.

Le LIPN s'intéresse à des problèmes ouverts en inférence de réseaux. Il a notamment i) proposé des heuristiques robustes pour le problème NP-difficile d'inférence de graphes [CO-93*] ii) étudié d'intégration des sources de données hétérogènes dans le mécanisme d'inférence de réseaux. Nous avons développé une méthode de classification croisée [CO-93*] capable d'engendrer des régions éventuellement recouvrantes à partir de données expérimentales bruitées [CI-60], puis nous avons étendu cette méthode pour prendre en compte une classification préalablement connue pour guider la construction de ces régions [CI-78].

• **BQR FACE : Détection et reconnaissance des mouvements faciaux liés au genre chez l'Homme (2006-2008)**

Partenaires : LIPN, LEEC

Ce projet concerne l'étude de la signalisation et de la reconnaissance du genre à partir des mouvements faciaux. La prise en compte des mouvements faciaux est relativement récente et a été effective dans certains travaux portant sur l'expression des

émotions et dans la communication parlée. Curieusement malgré des études comparatives, déjà anciennes sur les communications non verbales, peu de recherches récentes concernent le rôle des mouvements faciaux dans la signalisation d'appartenance à un groupe et dans la reconnaissance sociale chez l'Homme. Les objectifs du projet sont :

- de détecter les mouvements faciaux émis par des hommes et des femmes, appartenant à différents groupes culturels, dans des situations standardisées en situation contrôlée de laboratoire,
- de vérifier les résultats d'une étude pilote indiquant l'existence de mouvements faciaux latéralisés indicateurs du genre,
- de définir si ces mouvements sont perçus par des sujets « juges » des deux sexes, indépendamment d'indices morphologiques du visage,
- d'explorer ces mêmes informations par des réseaux connexionnistes (supervisés et non-supervisés),
- d'interpréter les phénomènes de causalité à partir des résultats de l'exploration de séquences spatio-temporelles des visages au moyen de modèles graphiques de type réseaux bayésiens.

- **BQR BigClust : Partitionnement de grandes bases de données médicales à très large dimensionnalité : Utilisation des modèles bio-inspirés (2007-2008)**

Partenaires : Equipe A³ du LIPN

Nous avons développé un modèle bio-inspiré combinant différentes approches et dédié à la classification de données complexes. Nous avons appliqué les approches développées sur des données médicales [CI-22, CI-74, CO-18].

Production scientifique, valorisation, dissémination

- **MOTIFS (Algorithmes de graphes pour la recherche de motifs fréquents flous, 2009)**

Partenaires : Equipes A³ et AOC du LIPN (*Projet interne LIPN*)

En collaboration avec l'équipe AOC, nous avons proposé dans ce projet d'étudier la combinaison de méthodes de fouille de données et de méthodes d'optimisation combinatoire pour rechercher des régions denses efficacement.

- **BQR émergent LIPS : Prédiction de liens dans les réseaux sociaux et application à la gestion des colloques (2009-2010)**

Partenaires : Equipe A³ du LIPN

Le but de ce projet est de développer des approches de prédiction de liens dans différents réseaux de collaborations scientifiques (notamment les réseaux de co-publication) afin de *prédire*, à l'aide de méthodes d'apprentissage supervisé, de nouvelles collaborations, chacune représentée par une nouvelle publication. La deuxième étape consiste à étudier comment les règles produites peuvent être utilisées pour recommander, par exemple des contacts scientifiques à prendre dans une conférence, ou des regroupements d'articles en session. Principales publications : [CI-52, CO-59].

- **BQR pluri-disciplinaire ONTO-APP : Ontologie et Apprentissage pour un système de suggestions nutritionnelles personnalisées (2012-2013)**

Partenaires : Lim&Bio (porteur), équipe A³ (LIPN), équipe déterminants, UREN (Unité de Recherche en Epidémiologie Nutritionnelle, UMR U557 Inserm/U1125 Inra/Cnam)

Le contexte général de ce travail est le développement d'une plateforme d'interrogation à base ontologique dédiée à la nutrition. Parmi les fonctionnalités de cette plateforme figure la possibilité d'obtenir, à partir d'une base de recettes, un ensemble de suggestions de recettes adaptées à un individu. Une suggestion sera établie à partir du profil de l'individu, du bloc de description des recettes (titre, ingrédients, étapes de préparation, etc.) et de leurs caractéristiques nutritionnelles. La finalité de ce BQR est

3.3 Valorisation de la recherche

i) la construction d'une ontologie du domaine de la nutrition, ii) son exploitation par des techniques d'apprentissage automatique symbolique dans le but de construire des catégories de recettes, et les descriptions symboliques de ces catégories. Les catégories construites et leurs descriptions seront validées par des experts du domaine. Le LIPN intervient particulièrement sur la deuxième tâche.

3.3 Valorisation de la recherche

3.3.1 Contrats industriels

- **CIFRE INA (2009-2012)**

Apprentissage non supervisé de données structurées. Application à la structuration automatique des flux télévisuels par analyse des répétitions.

Le principal objet de ce projet est de s'appuyer sur la masse des données récoltées (informations de répétition) pour tenter de reconstruire la structure des programmes télévisuels. Les problématiques suivantes sont abordées :

- Typage : le typage consiste, pour une séquence répétée, à définir son type à choisir parmi (par exemple) publicité, générique début ou fin, bande-annonce, extrait, ...
- Structuration : la structuration consiste à reconstruire des séries (tunnel publicitaire, émission unique, émission à base d'extraits, journal télévisé, inter-programmes). Ceci permet d'explicitier la structuration des flux (programmation et composition).
- Provenance : la provenance consiste à déterminer, pour chaque série retrouvée plus haut, l'origine de tout ou partie des éléments qui le constituent. En particulier, il s'agira de tenter de reconstituer autant que possible les liens de filiation entre les programmes qui incorporent des extraits d'archives.

- **CIFRE IntelligencePower (2009-2012)**

Automatisation du processus de génération d'un entrepôt de données.

Le projet MDDW a pour ambition de proposer une solution intégrée pour l'automatisation du processus de production d'un entrepôt de données. Habituellement, le processus industriel de développement d'entrepôts de données nécessite l'intervention de plusieurs experts à la fois au niveau conceptuel (couches physique, logique, conceptuelle) et au niveau de l'expertise-métier. Ceci engendre à la fois des coûts et des délais de développement importants dus à la disponibilité des experts et une réutilisation partielle des règles métier puisque celles-ci se limitent au biais apporté par l'expert et n'utilisent pas toute la connaissance pouvant être extraite de l'historique des projets similaires. L'objectif fixé dans ce travail de thèse est de proposer une approche globale pour réduire l'intervention de l'expert à la fois au niveau conceptuel et au niveau de la réutilisation de l'expertise pour améliorer la qualité des transformations et pour réduire les coûts et les délais. L'utilisation de l'approche dirigée par les modèles permet une meilleure structuration des processus de transformation en définissant notamment les invariants de transformation au niveau des méta-modèles et l'utilisation des techniques d'apprentissage automatique permet d'automatiser la réutilisation des traces des projets existants. Les publications principales sur les travaux développés dans cette convention sont : [CI-30, CI-48, CL-12, CI-68].

- **CIFRE Anticipo (2010 - 2013)**

Approches prédictives d'apprentissage statistique.

Anticipeo est une jeune entreprise innovante qui a développé une offre service et logiciel qui permet de construire aisément des budgets opérationnels détaillés et fiables. Anticipeo souhaite fournir des éléments de conseil aux clients dans les opérations de transformation des projections en prévisions (segmentation produits et segmentation clients). Nous travaillons dans le cadre de ce projet de recherche sur les thèmes suivants : la co-classification incrémentale, la détection des variations accidentelles, la prise en compte de données manquantes. Les publications principales sur les travaux développés dans cette convention sont : [CI-86],[CI-85], [CO-94].

3.3.2 Réalisations informatiques (logiciels), Brevets, Start-up

3.3.3 Logiciels

- **SOFTWARE (DS2L-SOM) (N° IDN.FR.001.490019.000.S.P.2008.000.20000, APP)**

Auteurs : G. Cabanes, Y. Bennani.

Ce logiciel implémente une méthode de description des données à partir de prototypes enrichis puis segmentés à l'aide d'un algorithme de classification non supervisée à deux niveaux. Cette méthode détermine automatiquement le nombre de groupes dans une distribution de données. Le logiciel propose aussi un procédé de visualisation capable de mettre en valeur la structure intra et inter-groupes des données.

Production scientifique, valorisation, dissémination

3.3.4 Brevets

- **VISUAL INFORMATION RETRIEVAL SYSTEM (WO/2010/066774 - PCT/EP2009/066702)**

Auteurs : H. Benhadda, Y. Bennani, N. Grozavu, M. Lebbah.

L'objectif du procédé proposé est de pouvoir classifier et visualiser des données à très fortes dimensionnalités multi-modales dans un espace de faibles dimensionnalités « l'espace de partitionnement », sans aucune information *a priori* sur le nombre de groupes. La première contribution consiste à définir le problème de fusion comme un problème de méta-classification dans un espace de variables catégorielles par une technique de classification automatique (l'analyse relationnelle). La deuxième contribution consiste à traiter le problème de fusion d'une manière modulaire, coopérative et évolutive. En effet, ce procédé est évolutif par rapport aux données et par rapport aux utilisateurs ou experts. Un processus de retour en arrière « Backward » et de raffinement des résultats de la classification globale est introduit dans le processus de fusion modulaire. L'utilisation des méthodes connexionnistes non supervisées comme moyen de recodage des données (quantification) et de l'analyse relationnelle comme méthode de fusion permettent une visualisation hiérarchique des résultats de la classification avec plusieurs niveaux de détails. L'efficacité de ce procédé est illustrée sur un problème de recherche et d'accès rapide à l'information visuelle dans une base de données d'images décrites par un ensemble de descripteurs numériques (descripteurs couleur et descripteurs texture) et textuels (plusieurs milliers de mots extraits des pages web).

- **METHODS FOR UPDATING AND TRAINING FOR A SELF-ORGANISING MAP (WO/2009/081005 - PCT/FR2008/052288)**

Auteur : Y. Bennani

Cette méthode propose une nouvelle stratégie de mise à jour des paramètres d'une carte auto-organisatrice pendant l'apprentissage. Cette méthode non supervisée à deux niveaux permet de déterminer automatiquement le nombre optimal de groupes, pendant l'apprentissage. Elle propose d'associer à chaque connexion de voisinage une valeur

3.3 Valorisation de la recherche

réelle qui indique la pertinence des neurones connectés. Étant donné la contrainte d'organisation de la carte, les deux meilleurs représentants de chaque donnée doivent être reliés par une connexion topologique. Cette connexion sera « récompensée » par une augmentation de sa valeur, alors que toutes les autres connexions issues du meilleur représentant seront « punies » par une diminution de leurs valeurs. Ainsi, à la fin de l'apprentissage, un ensemble de prototypes interconnectés sera représentatif d'un sous-groupe pertinent de l'ensemble des données : un cluster. Le principal avantage du procédé proposé, comparé aux méthodes classiques de classification, est qu'il n'est pas limité à la détection de groupes convexes, mais est capable de détecter des groupes de formes arbitraires.

3.3.5 Start-up

Erick Alphonse est actuellement (et depuis 2010) en délégation pour création d'entreprise. L'origine de ce projet d'essaimage est la volonté d'Erick Alphonse de valoriser ses travaux de recherche sur la fouille de données relationnelles. L'objet même de la société (IDaaS), créée en décembre 2009, est de développer et commercialiser des solutions innovantes d'Extraction de Connaissances à partir de Données. Le projet est de développer des outils, suivant le modèle SaaS, de valorisation de masses de données, issus de la recherche en Intelligence Artificielle : Intelligent Data Analysis as a Service (IDaaS). L'utilisateur loue des capacités d'acquisition, de stockage et de sécurisation de données dans un SGBDR (Système de gestion de base de données relationnelles) inductif et de la puissance de traitement et calcul sur ces données. L'outil développé s'appuie sur les recherches des laboratoires internationaux en Apprentissage Automatique (Machine Learning), dont fait partie A³.

Signalons qu'A. Osmani a obtenu de l'Université Paris 13 une délégation pour création d'entreprise sur le thème de l'Internet des objets.

**Production
scientifique,
valorisation,
dissémination**

Chapitre 4

Rayonnement

L'équipe Apprentissage Artificiel et Applications (A³) s'implique fortement au niveau national dans l'animation de la communauté scientifique de l'apprentissage automatique et de la fouille de données et des statistiques. Les membres de l'équipe sont particulièrement actifs, ils ont présidé des comités de programmes de conférences francophones (CAp, la conférences MARAMI/JFGG'2012 tout récemment). Ils organisent tous les deux ans et depuis 6 ans à l'Université Paris 13 les Journées d'Apprentissage Automatique et de Fouille de Données (AAFD). Enfin, ils participent et animent très régulièrement de nombreux groupes de travail et ateliers, dans le cadre de sociétés savantes ou attachés à des conférences d'apprentissage et de fouille de données françaises.

L'activité au niveau international est certes moindre, mais reste très significative : organisation d'ateliers internationaux (DYNAK'2010) et de sessions spéciales dans des conférences internationales de premier plan (ICMLA, ICONIP, IJCNN, IV). Les membres de l'équipe participent de façon récurrente à de nombreux comités de programmes de conférences internationales majeures du domaine. Enfin, nous participons au comité éditorial de la revue internationale IDA et dans celui de plusieurs issues spéciales.

4.1 Comités d'organisation et d'édition

4.1.1 Présidence de comités de programme

- *Journées thématiques "Apprentissage Artificiel et Fouille de Données"* : AAFD 2008, 2010 et 2012, Villetaneuse : Y. Bennani
- *Conférence francophone sur l'apprentissage artificiel, CAp 2009*, Plateforme AFIA, Hammamet, Tunisie : Y. Bennani et C. Rouveirol

4.1.2 Principaux comités de programme de conférences internationales

- *20th European Conference on Artificial Intelligence, ECAI'12*, Montpellier, France : Y. Bennani, Y. Chevaleyre, C. Rouveirol
- *26th AAAI Conference on Artificial Intelligence (AAAI-12)* : Yann Chevaleyre
- *Int. workshop on Complex networks & their applications, 2012* : R. Kanawati
- *3rd Int. Conference on Multimedia Computing and Systems (ICMCS'12)* : R. Kanawati
- *European Intelligence & Security informatics conference (ESIC'12)* : R. Kanawati
- *Int. Symposium on Foundations of open source intelligence and security informatics, 2012* : R. Kanawati
- *22nd Joint Conference on Artificial Intelligence (IJCAI-11)* : Yann Chevaleyre
- *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), 2010* : Y. Bennani
- *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), 2007 - 2012* : C. Rouveirol
- *Int. Conference on Data Mining (ICDM), 2007 - 2011* : C. Rouveirol
- *Int. Conference on Knowledge Discovery in Databases (KDD), 2007 et 2011* : C. Rouveirol
- *Int. Conference on Inductive Logic Programming (ILP), 2007 -* : C. Rouveirol, E. Alphonse
- *Int. Conference on Machine Learning (ICML), 2011* : C. Rouveirol, E. Alphonse
- *4th international conference on Electronics, Computers and Artificial Intelligence, 2011* : R. Kanawati
- *Int. Conference on Tools in Artificial Intelligence (ICTAI), 2010-2011* : C. Rouveirol
- *Int. conference on expert systems (Dexa 2010)* : R. Kanawati
- *Int. conference on collaboration Technology, 2010* : R. Kanawati
- *Second Int. Workshop on Mining Complex Data - MCD'08 - in Conjunction with IEEE ICDM'08, Pisa Italy* : Y. Bennani
- *Int. conference on expert systems (Dexa 2009)* : R. Kanawati
- *KDD'08 Workshop DMBA (Data Mining for Business Applications), Las Vegas, USA* : Y. Bennani, C. Rouveirol
- *Second Int. Workshop on Mining Complex Data - MCD'08 - in Conjunction with IEEE ICDM'08, Pisa Italy* : Y. Bennani

4.1.3 Principaux comités de programme de conférences nationales

Les membres de l'équipe participent aux conférences françaises principales du domaine ainsi que des ateliers associés.

- *Journées de Statistique* : Y. Bennani
- *Extraction et Gestion de Connaissance (EGC)* : H. Azzag, Y. Bennani, M. Lebbah, C. Rouveirol

- *Conférence Francophone d'Apprentissage (CAp)* : E. Alphonse, Y. Bennani, C. Rouveirol, H. Soldano
- *Concept Lattices and Application (CLA, 2011,2012)* : H. Soldano
- *MARAMI : Conférence sur les Modèles et l'Analyse des Réseaux : Approches Mathématiques et Informatiques - 2010 - 2012* : R. Kanawati

4.1.4 Organisation de conférences, sessions de conférences et ateliers

- *Internationales*
 - Comité d'organisation de la session spéciale *Co-clustering of Large and High Dimensional Data*, International Conference on Neural Information Processing, ICONIP 2012, Doha, Qatar : Y. Bennani
 - Comité d'organisation de la session spéciale *Nonnegative Matrix factorization paradigm for unsupervised learning*, International Joint Conference on Neural Networks, IJCNN 2012, Brisbane, Australia : Y. Bennani
 - Comité d'organisation de la session spéciale *Combining Multiple Learners*, ICOP'11, 14-17 November 2011 - Shanghai, China : Y. Bennani
 - Comité d'organisation de la session spéciale *Symposium on Advances in Interactive and Visual Data Clustering*, 16th International Conference Information Visualisation (iV'12), 2012, Montpellier France, <http://www.graphicslink.co.uk/IV2012/AIVDAC.htm> : H. Azzag, M. Lebbah
 - Comité d'organisation de la session spéciale *Machine Learning Methods for Modeling Treatment Outcomes in Cancer and Radiation Therapy*, ICMLA 2009 http://www.icmla-conference.org/icmla09/CFP_SpecialSession1.html : M. Lebbah
 - Comité d'organisation de la session spéciale *Incremental Topological Learning Models and Dimensional Reduction*, ICNNAI'2010, Biélorussie : N. Grozavu, M. Lebbah, Y. Bennani
 - Comité d'organisation du workshop ECML PKDD 2011 *Dynamic Networks and Knowledge Discovery*, Barcelone, Espagne : R. Kanawati et C. Rouveirol
- *Nationales*
 - Co-organisation de la troisième *Journée EGC Fouille de Grands Graphes (JFGG'2012)* associé à la conférence *MARAMI'2012*, Université Paris 13, <http://lipn.fr/jfgg12/JFGG12/> : R. Kanawati et G. Santini,
 - Co-organisation du colloque *MARAMI'2012*, 2012, Université Paris 13, <http://lipn.fr/marami12/MARAMI2012/> : R. Kanawati, C. Rouveirol et G. Santini
 - Co-organisation des 7^e, 8^e et 9^e édition de l'atelier *Fouille de données complexes - complexité liée aux données multiples* associé à la conférence **EGC**, <https://sites.google.com/site/afdcegc12> : M. Lebbah
 - Co-organisation de l'atelier *Apprentissage et Graphes pour les Systèmes complexes* associé à la plateforme AFIA'2009 : R. Kanawati et C. Rouveirol
 - Co-organisation du colloque *Fouille de Données Complexes (GT :EGC-FDC) : Complexité liée aux données multiples*, 2009, 2010, 2011 <http://eric.univ-lyon2.fr/~gt-fdc/> : M. Lebbah
 - Co-organisation d'une session spéciale *Apprentissage et modèles de mélanges* dans le cadres des 43^{es} journées de statistique, 2011, <http://jds2011.tn.refer.org/> : Y. Bennani, M. Lebbah
 - Co-organisation du colloque *Fouille et Visualisation de Données Massives (Big Data Mining and Visualization) (EGC-AFHIM)*. 2012, Polytech Tours, <https://sites.google.com/site/gtegcafhlm> : H. Azzag, M. Lebbah, R. Kanawati

4.1.5 Comités éditoriaux

- **Membre de comités éditoriaux**

- *Intelligent Data Analysis*, IOS Press, depuis 2008 : C. Rouveirol
- International journal on Social Network Mining, InderScience Publisher, depuis 2011 : R. Kanawati
- Special issue of *Intelligent Data Analysis journal on Knowledge Discovery in dynamic networks* (2013) : R. Kanawati, C. Rouveirol
- Special issue of international journal on Web-based communities : *Communities in complex networks* (2012) : R. Kanawati
- *Raisonnement à partir de cas : Surveillance, diagnostic et maintenance*, Editions Hermès (2007) : R. Kanawati

- **Autres activités de relecture**

Les membres de l'équipe ont participé à des relectures, en particulier dans des numéros spéciaux, pour les revues suivantes : *Technique et Science Informatiques*, *Computational Statistics and Data Analysis*, *Intelligent Data Analysis*, *Machine Learning*, *Journal of Machine Learning Research (JMLR)*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *Briefings in Bioinformatics*, *Computational Intelligence*, *Neurocomputing journal*, *Neural Computing & Applications*, ...

- **Comités d'évaluation**

- Membre du jury d'attribution du prix de thèse *AFIA* (2012) : H. Soldano

4.2 Expertises

4.2.1 Expertises de projets

- Evaluation projets pour l'ANR (Blanc, JCJC) : H. Azzag, Y. Bennani, R. Kanawati, C. Rouveirol
- Evaluation projets, programme *CONTINT* : H. Soldano
- Evaluation projets ACI Masses de Données : Y. Bennani
- Evaluation dossiers masters pour l'AERES : Y. Bennani
- Expertise auprès du Pôle de Compétitivité Cap Digital : Y. Bennani
- Expertise auprès l'INRETS pour évaluation des chercheurs : Y. Bennani
- Evaluation de projets de thèses CIFRE : R. Kanawati, M. Lebbah, C. Rouveirol
- Evaluation de projet CapDigital : R. Kanawati

4.3 Invitations

4.3.1 Invitations des membres du LIPN

- **Professeurs Invités**

- Visiting Professor Fellow, du 22 Août au 2 Sept. 2011 puis 23 Juillet au 21 Août 2012, Laboratoire des Technologies Innovantes de l'ENSA à Tanger : Y. Bennani

- **Conférences invitées, tutoriels, animation de tables rondes**

- Conférence invitée "Advances in unsupervised dimensionality reduction through topological clustering and variable weighting", International Workshop "Theory and Application of High-dimensional Complex Data Analysis in Economics and Management Science", Beijing (China) 2011 : Younès Bennani, Guénael Cabanès

- Tutoriel invité aux *Journées d'Intelligence Artificielle Fondamentale (JIAF)*, Toulouse, 2012, *Programmation Logique Inductive* : C. Rouveirol, en collaboration avec Christel Vrain (Univ. Orléans)
- Présentation invitée au NII *International Symposium on Symbolic Systems Biology (ISSSB'11)*, nov. 2011, Shonan Village, Japon : C. Rouveirol

4.3.2 Invitations à Paris 13

- Attilio GIORDANA, Université du Piémont Oriental, Italie, 2 mois en 2009
- Lorenza SAITTA, Université du Piémont Oriental, Italie, 2 mois en 2009
- Stefan KRAMER, Université Technique de Munich, 1 mois en 2009
- Djamel BOUCHAFFRA, Grambling State University, 1 mois en 2010 et en 2011
- Abdelouahid, LYHYAOUI, Université Abdelmalek Essadi de Tanger, 1 mois en 2011 et en 2012

Chapitre 5

Formation à la recherche

L'équipe A³ a une politique très active de formation à la recherche. Ceci se traduit par i) la responsabilité du master et la création et la coordination d'une spécialité de master « Exploration Informatique des Données et Décisionnel » (EID²) par Y. Bennani ii) la formation à la recherche à direction des étudiants de l'école d'ingénieurs de Paris 13, Sup Galilée par le biais de stages de recherche en deuxième année iii) l'encadrement de nombreux stages de recherche (3 ou 4 en moyenne chaque année) et d'un nombre important de doctorants iv) l'organisation régulière de tutoriels et d'une école d'hiver d'apprentissage (e-EGC). Enfin, la plupart des membres de l'équipe assurent couramment des cours dans le master d'informatique de l'université Paris 13, notamment dans la spécialité MICR (2007– 09), mais également dans la spécialité EID, puis EID² (depuis 2009) et dans le tronc commun de la 1^{ère} et 2^{ème} année du master d'informatique.

5.1 Masters, Ecoles doctorale et post-doctorale

5.1.1 Cours de master

Le travail de recherche d'une grande partie des membres de l'équipe a débouché sur la création de nouveaux enseignements au niveau du Master d'Informatique en particulier dans la spécialité « Exploration Informatique des Données et Décisionnel » EID².

– *Factorisation matricielle*, master 1 Informatique, Paris 13 : Y. Bennani

5.2 Thèses et Habilitations

- *Résolution de problèmes et Intelligence Artificielle*, master 1 Informatique, Paris 13 : Y. Chevaleyre, C. Rouveïrol, H. Soldano
- *Traitement Numérique des Données*, master 2 Informatique, Paris 13 : Y. Bennani
- *Informatique Décisionnelle et Data Mining*, Paris 13 : Y. Bennani : Y. Bennani
- *Modélisation biomimétique*, master 2 Informatique, Université de Tours, 2007-2008 : H. Azzag
- *Apprentissage numérique*, master 2 Informatique, Paris 13, spécialité MICR, 2007-2008 : M. Lebbah
- *Apprentissage Symbolique*, master 2 Informatique, Paris 13, spécialité MICR, 2007-2008 : C. Rouveïrol, H. Soldano
- *Fouille de Données Relationnelle*, master 2 Informatique, Paris 13, spécialité MICR, 2007-2008 : C. Rouveïrol
- *Apprentissage connexionniste (option)*, master 2 Informatique, spécialité EID2, Paris 13 : Y. Bennani
- *Apprentissage symbolique (option)*, master 2 Informatique, spécialité EID2, Paris 13 : Y. Chevaleyre, C. Rouveïrol, H. Soldano
- *Traitement numérique des données (tronc commun)*, master 2 Informatique, Paris 13 : Y. Bennani
- *Aide à la décision (option)*, master 2 Informatique, Paris 13 2010-2011 : Y. Chevaleyre
- *Réseaux sociaux (option)*, master 2 Informatique, Paris 13 2011 : C. Rouveïrol, R. Kanawati

5.1.2 Autres cours

- Président du Comité Scientifique de l'école d'hiver é-EGC'10, "Apprentissage Statistique et Data Mining", du 1 au 5 Février 2010 à Hammamet (Tunisie) : Y. Bennani
- Tutoriel à la conférence de *Reconnaissances des Formes et Intelligence Artificielle (RFIA 2012)*, Lyon, 2012 *Prévision de liens dans des grands graphes de terrain dynamiques : Techniques & Applications* : R. Kanawati et C. Rouveïrol
- Tutoriel à la conférence d'*Extraction et Gestions de Connaissances (EGC)*, Hammamet (Tunisie), 2010, *Prédiction de liens dans les réseaux sociaux* : R. Kanawati et C. Rouveïrol
- Conférencier invité à l'*Ecole Thématique sur l'Informatique et les Technologies de l'Information*, 11 au 15 décembre 2011, ORAN, <https://sites.google.com/site/labolutio/iti2011> : M. Lebbah

5.1.3 Jurys de thèse et d'HDR

- Jurys de thèse : H. Azzag (3), Y. Bennani (30 dont 16 comme rapporteur), M. Lebbah (5 dont 2 comme rapporteur), C. Rouveïrol (15 dont 9 comme rapporteur),
- Jury d'HDR : Y. Bennani (4), R. Kanawati (1), C. Rouveïrol (2 dont 1 comme rapporteur).

5.2 Thèses et Habilitations

5.2.1 Thèses en cours

Nom Prénom Titre	1ère insc. Directeur	Financement Co-encadrant
ALIZADEH Pegah Réduction de l'apprentissage par renforcement à l'apprentissage supervisé en ligne	01/10/12 Y. Chevaleyre	Alloc (12-15)

5.2 Thèses et Habilitations

REDKO Ievgen <i>Factorisation matricielle adaptative pour l'apprentissage par transfert</i>	01/10/12 Y. Bennani	Alloc (12-15)
CHAIBI Amine <i>Approches prédictives d'apprentissage statistique</i>	01/10/10 M. Lebbah	Cifre Anticepo H. Azzag
YAKOUBI Zied <i>Système de raisonnement à partir de cas pour l'intelligence collective</i>	01/10/10 H. Soldano	Projet UrbanD R. Kanawati
CHEBIL Inès <i>Fouille de graphes dans les données biologiques</i>	01/10/10 C. Rouveïrol	Projet Inca
PUJARI Manisha <i>Etude de la dynamique des grands graphes de terrain</i>	01/10/10 C. Rouveïrol	Alloc (10-13) R. Kanawati
DOAN Nhat-Quang <i>Approche hiérarchique et topologique pour la classification et la visualisation</i>	01/10/10 M. Lebbah	Convention USTH H. Azzag
BANNOUR Sondes <i>Apprentissage interactif de règles pour l'extraction d'informations dans des documents</i>	01/10/10 H. Soldano	Programme Quaero L. Audibert
GHASSANY Mohamad <i>Classification non-supervisée collaborative sous contraintes</i>	10/01/10 Y. Bennani	Alloc (10-13)
JAZIRI Rakia <i>Apprentissage non supervisé de données structurées</i>	10/01/09 Y. Bennani	ANR CADI - Cifre INA M. Lebbah
HAMDI Fatma <i>Apprentissage en distributions déséquilibrées</i>	10/01/09 Y. Bennani	ANR E-Fraud
MOUHOUBI Karima <i>Recherche de motifs contraints dans les données bruitées</i>	10/07/09 C. Rouveïrol	Alloc (09-12) L. Létocart
ESSAIDI Moez <i>Outils d'apprentissage pour un ETL dans un entrepôt de données</i>	01/12/08 C. Rouveïrol	Cifre A. Osmani
RODRIGUES BENTO Christophe <i>Apprentissage par renforcement relationnel</i>	01/10/07 C. Rouveïrol	contrat P. Gérard

Formation à la recherche

5.2.2 Thèses soutenues

(depuis 2007)

Nom Prénom	Situation ultérieure	Soutenance	Mention
Titre		Directeur	Co-encadrant
Jury : Jury			

5.2 Thèses et Habilitations

BENCHETTARA-HANIMED Nasserin		19/12/11	
<i>Prédiction de nouveaux liens dans les réseaux d'interactions bipartis : application au calcul de recommandation</i>			
		C. Rouveïrol	R. Kanawati
<i>Jury : Eric Gaussier (pdt), Gilles Bisson, Christine LARGERON, Clémence Magnien</i>			
CABANES Guénaël	Post-doc Australie	03/12/10	
<i>Classification non supervisée à deux niveaux guidée par le voisinage et la densité</i>			
		Y. Bennani	
<i>Jury : A. Hardy (rapp), Y. Lechevallier (rapp), M. Aupetit, H. Azzag, G. Cleuziou, A. Cornuéjols, J.L. Deneubourg, D. Fresneau</i>			
GROZAVU Nistor	MCF Paris 13	08/12/09	
<i>Réduction de dimensions en apprentissage non supervisé</i>			
		Y. Bennani	
<i>Jury : Pascale Kuntz (rapp), Michel Verleysen (rapp), Alexandre Aussem, Mustapha Lebbah, Jean-François Marcotorchino, Jean-Luc Zarader, Jean-Daniel Zucker</i>			
ROGOVSKI Nicoleta	MCF Paris 5	04/12/09	
<i>Classification à base de modèles de mélanges topologiques des données catégorielles et continues</i>			
		Y. Bennani	M. Lebbah
<i>Jury : Djamel Bouchaffra (rapp), Mohamed Nadif (rapp), Frédéric Alexandre, Khalid Benabdeslem, Bruce Denby, Catherine Recanati</i>			
ELATI Mohamed	MCF Evry	18/12/07	
<i>Apprentissage de réseaux de régulation génétique à partir de données d'expression</i>			
		C. Rouveïrol	
<i>Jury : Jean-François Boulicaut (rapp), Jean-Daniel Zucker (rapp), Florence d'Alché-Buc, Younés Bennani, Amedeo Napoli, François Radványi (invité)</i>			
ASEERVATHAM Sujeevan	Ingénieur	12/12/07	THF
<i>Apprentissage à base de Noyaux Sémantiques pour le Traitement de données textuelles</i>			
		Y. Bennani	E. Viennet
<i>Jury : Massih-Reza Amini, Christophe Fouqueré (pdt), Cyril Goutte, Jean-François Marcotorchino, Alain Rakotomamonjy, Michèle Sebag</i>			
KAROUI Hager	MCF Kairouan (Tunisie)	11/12/07	TH
<i>Système coopératif de type égal-à-égal pour la recommandation : Application à la gestion et la recommandation de références bibliographiques</i>			
		L. Petrucci	R. Kanawati
<i>Jury : Jean-Paul Barthès, Catherine Berrut, Sylvie Després, Alain Mille, Enric Plaza, Laurence Vignollet</i>			

5.2.3 Habilitations à Diriger des Recherches soutenues

(depuis 2007)

Nom Prénom	Soutenance
Titre	
Jury :	
Osmani Aomar	04/06/12
<i>Modélisation et raisonnement sur des données relationnelles</i>	
<i>Jury : Yann Chevaleyre, Antoine Cornuéjols (rapp), Philippe Dague, Christophe Fouqueré, Gérard Ligozat, Debasis Mitra (rapp), Lorenza Saitta (rapp), Céline Rouveïrol, Christel Vrain (rapp)</i>	
Lebbah Mustapha	27/01/12
<i>Contributions en apprentissage non supervisé à partir de données complexes</i>	

5.3 Stages de master, post-doctorats

Jury : Djamel Boucheffra (rapp), Marc Gelgon (rapp), Gérard Govaert (rapp), Djamel Abdelkader Zighed (rapp), Fouad Badran, Younes Benanni, Gérard Duchamp, Gilles Venturini, Céline Rouveirol

Soldano Henry

07/12/09

Apprentissage : Paradigmes, Structures et Abstractions

Jury : Younès Bennani, Daniel Kayser, Céline Rouveirol, Pierre Marquis (rapp), Amedeo Napoli (rapp), Lorenza Saitta (rapp), Jean Sallantin (rapp), Joël Pothier

Viennet Emmanuel

18/12/07

Contributions aux méthodes d'apprentissage numérique pour la fouille de données structurées

Jury : Stéphane Canu (rapp), Pascale Kuntz (rapp), Gilbert Ritschard (rapp), Younès Bennani, Françoise Fogelman-Soulié, Christophe Fouqueré, Gilbert Saporta

5.3 Stages de master, post-doctorats

5.3.1 Postdoctorats

- Guénael Cabanes, doctorat P13 (2010-2011) : Dérive de concepts (ANR E-Fraud).
- Lazhar Labiod, doctorat P6 (2009-2010) : Analyse relationnelle et classification non supervisée (ANR E-Fraud).
- Hager Karoui, doctorat P13 (2008) : Apprentissage et systèmes de recommandation (ANR CADI)
- Hani Hamdan, doctorat UTC (2007) : Modèles de mélanges pour la classification.

5.3.2 Stages de master

- AL OTHMAN, Mohamed, U. Paris 13 M2 Informatique MICR, 2007, 6 mois : H. Azzag, M. Lebbah
- CABANES, Guénaël, U. Paris 13 M2 Informatique, MICR, 2007, 6 mois : Y. Bennani
- CHELLI, Massinissa, U. Paris 13 M2 Informatique, MICR, 2007, 6 mois : Y. Bennani
- FERRADJ, Nabil, U. UVSQ, M2 TRIED, 2007, 6 mois : M. Lebbah
- RODRIGUES, Christophe, U. Paris 13 M2 Informatique, MICR, 2007, 6 mois : P. Gérard
- CHEREL, Guillaume, U. Paris 6, M2 Informatique IAD, 2008, 6 mois : Y. Bennani
- CHIBANI, Abderrahim, U. Paris 13, M2 Informatique MICR, 2008, 5 mois : R. Kanawati
- IORDATII, Maia, U. Paris 13, M2 Informatique MICR, 52008, 5 mois : Y. Bennani, M. Lebbah
- LAJNEF, Mohamed Ali, U. Paris 13, M2 Informatique MICR, 2008, 5 mois : H. Azzag, M. Lebbah
- LEI, Jin, U. Paris 13, M2 Informatique Math-Info, 2008, 5mois : E. Viennet
- AOUICHA, Kahina, U. VERSAILLES ST-QUENT., M2 Informatique Informatique : des concepts aux systèmes, 2009, 6 mois : Y. Bennani
- ELLOUZE, Slim, U. Paris 13, M2 Informatique MICR, 2009, 6 mois : S. Guérif
- KOLEILAT, Ghada, U. Paris 6, M2 Informatique, 2009, 6 mois : H. Soldano
- LAOUISSET, Brahim, U. Paris 8, M2 Informatique, CDSI, 2009, 6 mois : M. Lebbah, Y. Bennani
- MEDJAHED; Lakhdar, U. Paris 13, M2 Informatique EID, 2009, 5 mois : R. Kanawati
- MERABEK, Salah, U. CLERMONT-FERRAND 2, M2 MSIR, 2009, 6 mois : H. Azzag, M. Lebbah

Formation à la recherche

5.3 Stages de master, post-doctorats

- MOUHOUBI, Karima, U. Paris 13, M2 MICR, 2009, 6 mois : C. Rouveïrol, L. Létocart
- RAKOTONIRINA, Onja, U. Paris 13, M2 EID, 2009, 5 mois : R. Kanawati
- ARFAOUI, Aymen, U. Paris-DAUPHINE, M2 ISI, 2010, 6 mois : H. Azzag, M. Lebbah
- DESOLLE, Antoine, U. Paris 6, M2 IAD, 2010, 5 mois : H. Soldano
- MOHAND OUSAID, Mohamed, U. VERSAILLES ST-QUENT, M2 Informatique : des concepts aux systèmes, 2010, 5 mois : R. Kanawati
- PUJARI, Manisha, U. Paris-DAUPHINE, M2 ISI, 2010, 5 mois : R. Kanawati
- YACOUBI, Zied, U. Paris-DAUPHINE, M2 ISI, 2010, 6 mois : Y. Bennani, M. Lebbah
- ALTISINIK, Nejdet, U. Paris 13, M2 PLS, 2011, 6 mois : Y. Bennani
- LIU, Yu Qiong, U. Paris 13, M2 EID, 2011, 5 mois : M. Lebbah
- YOUSSEF, Sarra, U. Paris 13, M2 PLS, 2011, 5 mois : Y. Bennani
- ZHANG, Yaoyu, U. Paris 13, M2 EID, 2011, 5 mois : H. Azzag
- BARTCUS, Marius, U. Paris 13, M2 EID, 2012, 6 mois : Y. Bennani
- BERZOI, Victor, U. Paris 6, M2 Réseaux, 2012, 6 mois : Y. Bennani, N. Grozavu
- SARAZIN, Tugdual, U. Paris 13, M2 EID, 2012, 6 mois : M. Lebbah, H. Azzag

5.3.3 Stages ingénieurs

- BARRY, Mamadou Bobo, U. Paris 13, Sup Galilée, 2010, 3 mois : P. Gérard
- DUMONT, Laura, ENSIIE EVRY, 2010, 4 mois : H. Soldano
- GUIGOURES, Romain, EISTI CERGY, 2009, 5 mois : R. Kanawati
- RENAUD, Marion, EISTI CERGY, 2010, 4 mois : C. Rouveïrol
- DEBOIN, Sybil, ENSIIE EVRY, 2011, 3 mois : H. Soldano
- FLORENT, Charly, ENSIIE EVRY, 2011, 3 mois : H. Soldano
- GRANGE, Louis, ENSIIE EVRY, 2011, 3 mois : Y. Chevaleyre
- MICHELET, Julien, ENSIIE EVRY, 2011, 3 mois : H. Soldano
- MIZUNO, Toru, EC. CENTRALE LILLE, 2012, 5 mois : H. Soldano
- LE GOFF, Leni, ENSIIE EVRY, 2012, 2 mois : H. Soldano
- SONKO, Youssoupha, U. Paris 13, Sup Galilée Informatique, 2012, 3 mois : S. Guérif

5.3.4 Autres stages et encadrement

- AUBERT, Julien, U. Paris 13, M1 Informatique, 2011, 3 mois : Y. Bennani
- LAZZARINO, Rémi, U. Paris 13, M1 Informatique, 2011, 3 mois : Y. Bennani
- LONGUEU-BROCHART, Florian, U. Paris 13, M1 Informatique, 2011, 3 mois : Y. Bennani
- SOUAM Fatiha, U. de Tizi-Ouzou, Algérie, stage de recherche, 2010-2011, 2 mois : M. Lebbah

Chapitre 6

Publications

Les références comportant « * » sont acceptées et à paraître.

Signification des soulignements :

- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Sommaire

Articles dans des revues avec comité de lecture et chapitres de livres	
(ACL)	54
Communications avec actes (ACT)	57
Logiciels et Brevets (LO)	70
Thèses et Habilitations (TH)	71
Divers (Div)	72

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

- 2012** [RI-30*] H. Azzag, C. Guinot et G. Venturini. An artificial ants model for fast construction and approximation of proximity graphs. *Adaptive Behavior - Animals, Animats, Software Agents, Robots, Adaptive Systems*, 2012. 17 pages.
- [RI-29] R. Bossy, J. Jourde, A.-P. Manine, P. Veber, E. Alphonse, M. van de Guchte, P. Bessières et C. Nédellec. Bionlp shared task - the bacteria track. *BMC Bioinformatics*, 13(Suppl 11) :S3, June 2012. Selected articles from the BioNLP Shared Task 2011, 15 pages.
- [RI-28] G. Cabanes, Y. Bennani et D. Fresneau. Enriched topological learning for cluster detection and visualization. *Neural Networks*, 32 :186–195, August 2012.
- [RI-27] Y. Chevaleyre, J. Lang, N. Maudet, J. Monnot et L. Xia. New candidates welcome ! possible winners with respect to the addition of new candidates. *Mathematical Social Sciences*, 64(1) :74–88, 2012.
- [RI-26*] N.-Q. Doan, H. Azzag et M. Lebbah. Growing self-organizing trees for autonomous hierarchical clustering. *Neural Networks, Special Issue on Autonomous Learning*, 2012. 32 pages.
- [RI-25*] M. Ghassany, N. Grozavu et Y. Bennani. Collaborative clustering using prototype based techniques. *International Journal of Computational Intelligence and Applications (IJCIA)*, 2012. (16 pages).
- [RI-24*] G. Santini, H. Soldano et J. Pothier. Automatic classification of protein structures relying on similarities between alignments. *BMC Bioinformatics*, 2012. 16 pages.
- 2011** [RI-23] H. Azzag et M. Lebbah. Self-organizing tree using artificial ants. *Journal of Information Technology Research (JITR)*, 4(2) :1–16, 2011. Special Issue on Applications of Nature Inspired Computing.
- 2010** [RI-22] G. Cabanes et Y. Bennani. Unsupervised topographic learning for spatiotemporal datamining. *Advances in Artificial Intelligence*, 2010, Article ID 832542, 12 pages, 2010.
- [RI-21] G. Cabanes, Y. Bennani et D. Fresneau. Mining RFID Behavior Data using Unsupervised Learning. *International Journal of Applied Logistics*, 1(1) :28–47, 2010.
- [RI-20] M. Essaidi et A. Osmani. Model driven data warehouse using mda and 2tup. *Journal of Computational Methods in Sciences and Engineering*, 10 :119–134, 2010.
- [RI-19] N. Grozavu et Y. Bennani. Topological collaborative clustering. *Australian Journal of Intelligent Information Processing Systems*, 12(3), 2010. Machine Learning Applications (Part I).
- [RI-18] L. Labioud, N. Grozavu et Y. Bennani. Relational topological map. *International Journal of Computational Intelligence and Applications*, 9(4) :353–370, 2010.
- [RI-17] M. Lebbah et K. Benabdeslem. Visualization and clustering of categorical data with probabilistic self-organizing map. *Neural Computing and Applications*, 19(3) :393–404, 2010.
- [RI-16] N. Rogovschi, M. Lebbah et Y. Bennani. Learning self-organizing mixture markov models. *Journal of Nonlinear Systems and Applications. ISSN 1918-3704.*, 1(1-2) :63–71, 2010.
- [RI-15] Y. Zafoune, A. Moukhtari et R. Kanawati. Mobile codes Localization in Ad. hoc. Networks : A comparative Study of Centralized and Distributed Approaches. *International Journal of Computer Networks and Communications*, 2(2) :164–177, 2010.
- [RI-14] F. Zehraoui, R. Kanawati et S. Salotti. Hybrid neural network and case based reasoning system for web user behavior clustering and classification. *International Journal of Hybrid Intelligent Systems*, 7(3) :171–186, 2010.
- 2009** [RI-13] S. Aseervatham. A concept vector space model for semantic kernels. *International Journal on Artificial Intelligence Tools*, 18(02) :239–272, 2009.
- [RI-12] S. Aseervatham et Y. Bennani. Semi-structured document categorization with a semantic kernel. *Pattern Recognition*, 42(9) :2067–2076, 2009.
- [RI-11] U. Galassi, G. Cabanes et D. Fresneau. Modelling Evolving Behaviours in Ant Colonies. *Journal of Intelligent Systems*, 18(4) :353–376, 2009.

- [RI-10] A.-P. Manine, E. Alphonse et P. Bessières. Learning ontological rules to extract multiple relations of genic interactions from text. *International Journal of Medical Informatics*, 78(12) :31–38, 2009.
- [RI-9] N. Pisanti, H. Soldano, M. Carpentier et J. Pothier. A relational extension of the notion of motifs : application to the common 3d protein substructures searching problem. *Journal of computational biology*, 16(12) :1635–1660, 2009.
- [RI-8] J. Uckelman, Y. Chevalere, U. Endriss et J. Lang. Representing utility functions via weighted goals. *Mathematical Logic Quarterly*, 55(4) :341–361, 2009.
- 2008** [RI-7] E. Alphonse et A. Osmani. On the connection between the phase transition of the covering test and the learning success rate in ILP. *Machine Learning*, 70(2-3) :135–150, 2008.
- [RI-6] H. Kadri-Dahmani, C. Bertelle, G. H. Duchamp et A. Osmani. Emergent property of consistent updated geographical database. *International Journal of Modelling, Identification and Control*, 3(1) :58–68, 2008.
- [RI-5] M. Lebbah, Y. Bennani et N. Rogovschi. A probabilistic self-organizing map for binary data topographic clustering. *International Journal of Computational Intelligence and Applications*, 7(4) :363–383, 2008.
- 2007** [RI-4] H. Azzag, G. Venturini, A. Oliver et C. Guinot. A hierarchical ant based clustering algorithm and its use in three real-world applications. *European Journal of Operational Research*, 179(3) :906–922, 2007.
- [RI-3] D. G. Gomes, N. Agoulmine, Y. Bennani et J. N. d. Souza. Predictive connectionist approach for vod bandwidth management. *Computer Communications*, 30(10) :2236–2247, 2007.
- [RI-2] F. Picarougne, H. Azzag, G. Venturini et C. Guinot. A new approach of data clustering using a flock of agents. *Evolutionary Computation*, 15(3) :345–367, 2007.
- [RI-1] A. Tenenhaus, A. Giron, E. Viennet, M. Béra, G. Saporta et B. Fertil. Kernel logistic PLS : a tool for supervised nonlinear dimensionality reduction and binary classification. *Computational Statistics & Data Analysis*, 51(9) :4083–4100, mai 2007.

Articles dans des revues nationales ou internationales

- 2011** [RE-13] R. Kanawati. Leaders driven community detection algorithms : an empirical comparative study. *International journal of complex systems in science (IJCSS)*, 1 :1–10, 2011.
- [RE-12] F. Krief, Y. Bennani et J. Neuman de Souza. LECSOM : A Low-Energy Routing Algorithm based on SOM Clustering for Static and Mobile Wireless Sensor Networks. *IRECAP International Journal on Communications Antenna and Propagation*, 2011. Invited paper.
- [RE-11] L. Labiod, N. Grozavu et Y. Bennani. Classification relationnelle topographique. *Revue d'Intelligence Artificielle*, 25(3) :393–410, 2011.
- [RE-10] L. Labiod, N. Grozavu et Y. Bennani. Simultaneous topological categorical data clustering and cluster characterization. *International Journal of Computing (IJC)*, 10(1) :9–23, 2011. Special issue : 5th International conference on Neural Networks and Artificial Intelligence (ICNNAI'2010).
- [RE-9] N. Rogovschi, M. Lebbah et Y. Bennani. Modèles de mélanges topologiques pour la classification de données catégorielles et mixtes. *Revue des Nouvelles Technologies de l'Information*, RNTI-E-21 :53–80, 2011. Numéro spécial : Fouille de données complexes - Complexité liées aux données multiples.
- [RE-8] N. Rogovschi, M. Lebbah et Y. Bennani. A self-organizing map for mixed continuous and categorical data. *International Journal of Computing (IJC)*, 10(1) :24–32, 2011. Special issue : 5th International conference on Neural Networks and Artificial Intelligence (ICNNAI'2010).
- 2008** [RE-7] S. Aseervatham et E. Viennet. Méthodes à noyaux appliquées aux textes structurés. *Revue des Nouvelles Technologies de l'Information*, RNTI-A2 :185–205, 2008. Numéro spécial : Journées Apprentissage Artificiel et Fouille de Données.
- [RE-6] H. Azzag, D. Da Costa, C. Guinot et G. Venturini. Un aperçu de la fouille visuelle de données. *Revue des Nouvelles Technologies de l'Information*, RNTI-A2 :1–14, 2008. Numéro spécial : Journées Apprentissage Artificiel et Fouille de Données.
- [RE-5] Y. Bennani, E. Viennet et S. Guérif. Réduction des dimensions de l'espace de description. *Revue des Nouvelles Technologies de l'Information*, RNTI-A2 :135–163, 2008. Numéro spécial : Journées Apprentissage Artificiel et Fouille de Données.

- [RE-4] E. Birmelé, M. Elati, C. Rouveirol et C. Ambroise. Identification of functional modules based on transcriptional regulation structure. *BMC Proceedings*, 2(Suppl 4) :S4, Dec 2008. Selected papers of Machine Learning in Systems Biology : MLSB 2007, 6 pages.
- 2007** [RE-3] K. Benabdeslem et Y. Bennani. Approche connexionniste pour l'analyse des données issues d'usage d'internet : Classification et visualisation. *Revue des Nouvelles Technologies de l'Information*, RNTI-W1 :23–36, 2007. Numéro spécial : Fouille du Web.
- [RE-2] M. Elati, F. Radvanyi et C. Rouveirol. Fouille de données pour l'extraction de grands réseaux de régulation génétique. *Technique et Science Informatiques*, 26(1-2) :173–196, 2007.
- [RE-1] E. Viennet et F. Fogelman Soulié. Le traitement des refusés dans le risque crédit. *Revue des Nouvelles Technologies de l'Information*, RNTI-A-1 :23–45, sep 2007.

Chapitres de livre

- 2011** [CL-13] M. Elati et C. Rouveirol. *Unsupervised Learning for Gene Regulation Network Inference from Expression Data : A Review*, M. Elloumi et A. Y. Zomaya, éditeurs, *Algorithms in Computational Molecular Biology : Techniques, Approaches and Applications*, Chapitre 41. Wiley, 2011.
- [CL-12] M. Essaidi et A. Osmani. *Business Intelligence-as-a-Service : Studying the Functional and the Technical Architectures.*, *Business Intelligence Applications and the Web : Models, Systems and Technologies*, pages 199–221. IGI Global, 2011, ISBN 9781613500385.
- 2010** [CL-11] G. Cabanes et Y. Bennani. Learning the number of clusters in Self Organizing Map. *Self Organizing Maps*, pages 15–28. IN-TECH Publisher, 2010. Invited book chapter.
- [CL-10] N. Grozavu, Y. Bennani et M. Lebbah. Cluster-dependent feature selection through a weighted learning paradigm. F. Guillet, G. Ritschard, D. Zighed et H. Briand, éditeurs, *Advances in Knowledge Discovery and Management*, volume 292 de *Studies in Computational Intelligence*, pages 133–147. Springer Berlin / Heidelberg, 2010.
- [CL-9] M. Lebbah, Y. Bennani, N. Grozavu et H. Benhadda. Relational analysis for clustering consensus. Y. Zhang, éditeur, *Machine Learning*. IN-TECH Publisher, 2010.
- 2009** [CL-8] H. Azzag, F. Picarougne, C. Guinot et G. Venturini. Vrminer : A tool for multimedia database mining with virtual reality. *Database Technologies : Concepts, Methodologies, Tools, and Applications*, pages 1151–1167. IGI Global, 2009.
- [CL-7] H. Kadri-Dahmani, C. Bertelle, G. H. Duchamp et A. Osmani. *The evolution process of geographical database within self-organized topological propagation area.*, *Complex systems and self-organization modelling*, pages 97–106. Berlin : Springer, 2009, ISBN 978-3-540-88072-1.
- [CL-6] H. Karoui. *Collaborative and Social Information Retrieval and Access : Techniques for Improved User Modeling*, Chapitre COBRAS : Cooperative CBR Bibliographic Recommender System, pages 184–202. IGI Global, 2009.
- 2008** [CL-5] S. Aseervatham et Y. Bennani. A semantic tree representation for document categorization with a composite kernel. 2008. Invited Book Chapter.
- [CL-4] F. Badran, M. Lebbah et S. Thiria. Cartes auto-organisatrices et classification automatique. livre apprentissage statistique. *Apprentissage statistique*. Eyrolles, 2008.
- 2007** [CL-3] R. Kanawati. *Agents RàPC coopératifs pour la recherche d'information*, J. Renaud, B. C. Morello, B. Fuchs et J. Lieber, éditeurs, *Raisonnement à partir de cas 2, surveillance, diagnostic et maintenance*, volume 2 de *Informatique et système d'information*, Chapitre 6, pages 177–195. Hermès, 2007.
- [CL-2] M. Malek et R. Kanawati. *Case-Based Reasoning in Knowledge Discovery and Data Mining*, Chapitre CBR-based. John Wiley & Sons, 2007.
- [CL-1] F. Zehraoui, R. Kanawati et S. Salotti. *CASEP2 : Système Hybride pour le traitement de séquences de navigation sur le Web*, J. Renaud, B. Cheble-Morello, B. Fuchs et J. Lieber, éditeurs, *Raisonnement à Partir de Cas 2 : surveillance, diagnostic et maintenance*, volume 2 de *Informatique et système d'information*, pages 151–176. Hermès, 2007.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

- 2012**
- [CI-87] G. Cabanes et Y. Bennani. Change detection in data streams through unsupervised learning. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2012)*, pages 2659–2664, Brisbane, Australia, Jun 2012. IEEE.
 - [CI-86] A. Chaibi, H. Azzag et M. Lebbah. Automatic Group-Outlier Detection. *Proceedings of the 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2012)*, pages 393–398, 2012. (poster).
 - [CI-85] A. Chaibi, M. Lebbah et H. Azzag. Detection using a new group outlier factor. *Proceedings of the 19th International Conference On Neural Information Processing (ICONIP 2012), Part III*, volume 7665 de *Lecture Notes in Computer Science*, pages 364–372, Doha, Qatar, nov 2012. Springer.
 - [CI-84*] Y. Chevalleyre et A. Machado Pamponet. Adaptive probabilistic policy reuse. *Proceedings of the 19th International Conference On Neural Information Processing (ICONIP 2012)*, Lecture Notes in Computer Science, Doha, Qatar, nov 2012. Springer. 8 pages.
 - [CI-83] N.-Q. Doan, H. Azzag et M. Lebbah. Graph decomposition using self-organizing trees. *Proceedings of the 16th International Conference on Information Visualisation (IV 2012)*, pages 246–251, Montpellier, France, July 2012. IEEE Computer Society.
 - [CI-82] N.-Q. Doan, H. Azzag et M. Lebbah. Growing self-organizing trees for knowledge discovery from data. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2012)*, pages 251–258, Brisbane, Australia, Jun 2012. IEEE.
 - [CI-81] N.-Q. Doan, H. Azzag et M. Lebbah. Self-organizing map and tree topology for graph summarization. *Proceedings of the 22st International Conference on Artificial Neural Networks (ICANN 2012), Part II*, volume 7553 de *Lecture Notes in Computer Science*, pages 363–370. Springer, 2012.
 - [CI-80] M. Ghassany, N. Grozavu et Y. Bennani. Collaborative generative topographic mapping. *Proceedings of the 19th International Conference On Neural Information Processing (ICONIP 2012), Part II*, volume 7664 de *Lecture Notes in Computer Science*, pages 591–598, Doha, Qatar, nov 2012. Springer.
 - [CI-79] N. Grozavu, Y. Bennani et L. Labiod. Feature space transformation for transfer learning. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2012)*, pages 2637–2642. IEEE, 2012.
 - [CI-78] K. Mouhoubi, L. Létocart et C. Rouveirol. A knowledge-driven bi-clustering method for mining noisy datasets. *Proceedings of the 19th International Conference On Neural Information Processing (ICONIP 2012), Part III*, volume 7665 de *Lecture Notes in Computer Science*, pages 585–593, Doha, Qatar, nov 2012. Springer.
 - [CI-77*] M. Pujari et R. Kanawati. Link prediction in complex networks by supervised rank aggregation. *Proceedings of the 24th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2012)*, Athene, Greece, November 2012. IEEE.
 - [CI-76] M. Pujari et R. Kanawati. Tag recommendation by link prediction based on supervised machine learning. *Sixth International AAI Conference on Weblogs and Social Media (ICWSM 2012)*, pages 547–550, Dublin, june 2012. AAI.
 - [CI-75] C. Rodrigues, P. Gérard, C. Rouveirol et H. Soldano. Active learning of relational action models. S. Muggleton, A. Tamaddoni-Nezhad et F. Lisi, éditeurs, *Proceedings of the 21st International Conference on Inductive Logic Programming (ILP 2011)*, volume 7207 de *Lecture Notes in Computer Science*, pages 302–316. Springer Berlin / Heidelberg, 2012. (Revised Selected Papers).
- 2011**
- [CI-74] H. Azzag et M. Lebbah. A new approach for auto-organizing a groups of artificial ants. I. K. George Kampis et E. Szathmáry, éditeurs, *Advances in Artificial Life. Darwin meets Von Neumann*, volume 5778 de *Lecture Notes in Computer Science*, pages 440–447. Springer Berlin / Heidelberg, 2011. (ECAL'09).
 - [CI-73] L. Boudjeloud et H. Azzag. A cooperative biomimetic approach for high dimensional data mining. *13th Annual Genetic and Evolutionary Computation Conference, (GECCO 2011), Companion Material Proceedings*, pages 233–234. ACM, 2011. (poster).

- [CI-72] G. Cabanes et Y. Bennani. Coupling Clustering and Visualization for Knowledge Discovery from Data. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2011)*, pages 2127–2134, San Jose, California, USA, Aug 2011. IEEE.
- [CI-71] G. Cabanes, Y. Bennani et D. Fresneau. A new simultaneous two-levels coclustering algorithm for behavioural data-mining. B.-L. Lu, L. Zhang et J. T. Kwok, éditeurs, *Proceedings of the 18th International Conference On Neural Information Processing (ICONIP 2011), Part II*, volume 7062 de *Lecture Notes in Computer Science*, pages 745–752, Shanghai, China, nov 2011. Springer.
- [CI-70] F. Chamroukhi, A. Samé, P. Akin et G. Govaert. Model-based clustering with hidden markov model regression for time series with regime changes. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2011)*, pages 2814–2821, San Jose, California, USA, Aug 2011. IEEE.
- [CI-69] Y. Chevalyere, J. Lang, N. Maudet et J. Monnot. Compilation and communication protocols for voting rules with a dynamic set of candidates. *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK XIII, pages 153–160, New York, NY, USA, 2011. ACM.
- [CI-68] M. Essaidi, A. Osmani et C. Rouveirol. Transformation learning in the context of model-driven data warehouse : An experimental design based on inductive logic programming. *Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2011)*, pages 693–700. IEEE, 2011.
- [CI-67] N. Grozavu et Y. Bennani. Simultaneous pattern and variable weighting during topological clustering. B.-L. Lu, L. Zhang et J. T. Kwok, éditeurs, *Proceedings of the 18th International Conference On Neural Information Processing (ICONIP 2011), Part I*, volume 7062 de *Lecture Notes in Computer Science*, pages 570–579, Shanghai, China, nov 2011. Springer.
- [CI-66] N. Grozavu, M. Ghassany et Y. Bennani. Learning confidence exchange in collaborative clustering. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2011)*, pages 872–879, San Jose, California, USA, Aug 2011. IEEE.
- [CI-65] F. Hamdi et Y. Bennani. Learning random subspace novelty detection filters. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2011)*, pages 2273–2280, San Jose, California, USA, Aug 2011. IEEE.
- [CI-64] R. Jaziri, M. Lebbah, Y. Bennani et J.-H. Chenot. SOS-HMM : Self-Organizing Structure of Hidden Markov Model. *Proceedings of the 21st International Conference on Artificial Neural Networks (ICANN 2011)*, Lecture Notes in Computer Science, pages 87–94. Springer, 2011. June 14-17th, Espoo, Finland.
- [CI-63] R. Jaziri, M. Lebbah, N. Rogovschi et Y. Bennani. Probabilistic self-organizing maps for multivariate sequences. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2011)*, pages 851–858, San Jose, California, USA, Aug 2011. IEEE.
- [CI-62] R. Kanawati. Licod : Leaders identification for community detection in complex networks. *PASSAT/SocialCom 2011, Privacy, Security, Risk and Trust (PASSAT), 2011 IEEE Third International Conference on and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, pages 577–582. IEEE, 2011.
- [CI-61] L. Labioud et Y. Bennani. A spectral based clustering algorithm for categorical data with maximum modularity. *Proceedings of the 19th European Symposium on Artificial Neural Networks (ESANN 2011)*, pages 393–398, 2011.
- [CI-60] K. Mouhoubi, L. Létocart et C. Rouveirol. Itemset mining in noisy contexts : A hybrid approach. *Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2011)*, pages 33–40. IEEE, 2011.
- [CI-59] A. Perolini et S. Guérif. Simultaneous feature and model selection for high-dimensional data. *Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2011)*, pages 47–50. IEEE, 2011.
- [CI-58] B. Pradel, S. Sean, J. Delporte, S. Guérif, C. Rouveirol, N. Usunier, F. Fogelman-Soulié et F. Dufau-Joël. A case study in a recommender system based on purchase data. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2011)*, pages 377–385. ACM, 2011.

- [CI-57] M. Pujari et R. Kanawati. A supervised machine learning link prediction approach for tag recommendation. *Proceedings Online Communities and Social Computing - 4th International Conference, OCSC 2011, Held as Part of HCI International 2011*, volume 6778 de *Lecture Notes in Computer Science*, pages 336–344. Springer, 2011.
- [CI-56] H. Soldano. A modal view on abstract learning and reasoning. *Ninth Symposium on Abstraction, Reformulation and Approximation (SARA 2011)*, pages 99–106. AAAI Press, 2011.
- [CI-55] H. Soldano et V. Ventos. Abstract Concept Lattices. *International Conference on Formal Concept Analysis (ICFCA)*, volume 6628 de *LNAI*, pages 235–250. Springer, Heidelberg, 2011.
- 2010** [CI-54] E. Alphonse, T. Girschick, F. Buchwald et S. Kramer. A numerical refinement operator based on multi-instance learning. *Proceedings of the 20th International Conference on Inductive Logic Programming (ILP 2010)*, volume 6489 de *Lecture Notes in Computer Science*, pages 14–21. Springer, 2010. Revised Papers.
- [CI-53] H. Azzag, M. Lebbah et A. Arfaoui. Map-treemaps : A new approach for hierarchical and topological clustering. S. Draghici, T. M. Khoshgoftaar, V. Palade, W. Pedrycz, M. A. Wani et X. Zhu, éditeurs, *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'10)*, pages 873–878, Washington, DC, USA, Dec 2010. IEEE Computer Society.
- [CI-52] N. Benchettara, R. Kanawati et C. Rouveirol. A supervised machine learning link prediction approach for academic collaboration recommendation. *Proceedings of the 2010 ACM Conference on Recommender Systems, (RecSys 2010)*, pages 253–256. ACM, 2010.
- [CI-51] G. Bourgne, H. Soldano et A. ElFallah-Seghrouchni. Learning better together. *Proceedings 19th European Conference on Artificial Intelligence (ECAI 2010)*, volume 215 de *Frontiers in Artificial Intelligence and Applications*, pages 85–90. IOS Press, 2010.
- [CI-50] G. Cabanes et Y. Bennani. Learning topological constraints in Self-Organizing Map. *Proceedings of the 17th International Conference On Neural Information Processing (ICONIP 2010), Part II - Models and Applications*, volume 6444 de *Lecture Notes in Computer Science*, pages 367–374, Sydney, Australia, nov 2010. Springer.
- [CI-49] H. Elghazel, K. Benabdeslem et F. Hamdi. Consensus clustering by graph based approach. *Proceedings of the 18th European Symposium on Artificial Neural Networks (ESANN 2010)*, pages 493–498, 2010.
- [CI-48] M. Essaidi et A. Osmani. A unified method for developing data warehouses. *Proceedings of the 3d. International Conference on Information Systems and Economic Intelligence (SIIE'2010)*, pages 1–8, 2010.
- [CI-47] N. Grozavu, L. Labiod et Y. Bennani. Autonomous clustering characterization for categorical data. S. Draghici, T. M. Khoshgoftaar, V. Palade, W. Pedrycz, M. A. Wani et X. Zhu, éditeurs, *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'10)*, pages 607–613, Washington, DC, USA, Dec 2010. IEEE Computer Society.
- [CI-46] F. Hamdi, M. Lebbah et Y. Bennani. Topographic under-sampling for unbalanced distributions. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2010)*, pages 18–23, Barcelona, Spain, Jul 2010. IEEE.
- [CI-45] L. Labiod, N. Grozavu et Y. Bennani. Clustering categorical data using an extended modularity measure. K. W. Wong, B. S. U. Mendis et A. Bouzerdoum, éditeurs, *Proceedings of the 17th International Conference On Neural Information Processing (ICONIP 2010), Part II - Models and Applications*, volume 6444 de *Lecture Notes in Computer Science*, pages 310–320, Sydney, Australia, nov 2010. Springer.
- [CI-44] L. Labiod, N. Grozavu et Y. Bennani. Relational topological clustering. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2010)*, pages 1–8, Barcelona, Spain, Jul 2010. IEEE.
- [CI-43] M. Lebbah et H. Azzag. Topological hierarchical tree using artificial ants. K. W. Wong, B. S. U. Mendis et A. Bouzerdoum, éditeurs, *Proceedings of the 17th International Conference On Neural Information Processing (ICONIP 2010), Part I - Theory and Algorithms*, volume 6443 de *Lecture Notes in Computer Science*, pages 652–659, Sydney, Australia, nov 2010. Springer.

- [CI-42] A.-P. Manine, E. Alphonse et P. Bessières. Extraction of genic interactions with the recursive logical theory of an ontology. *Proceedings 11th International Conference Computational Linguistics and Intelligent Text Processing (CICLing 2010)*, volume 6008 de *Lecture Notes in Computer Science*, pages 549–563. Springer, 2010.
- [CI-41] C. Rodrigues, P. Gérard et C. Rouveirol. Incremental learning of relational action models in noisy environments. *Proceedings of the 20th International Conference on Inductive Logic Programming (ILP 2010)*, volume 6489 de *Lecture Notes in Computer Science*, pages 206–213. Springer, 2010. Revised Papers.
- [CI-40] C. Rodrigues, P. Gérard, C. Rouveirol et H. Soldano. Incremental learning of relational action rules. S. Draghici, T. M. Khoshgoftaar, V. Palade, W. Pedrycz, M. A. Wani et X. Zhu, éditeurs, *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'10)*, pages 451–458, Washington, DC, USA, Dec 2010. IEEE Computer Society.
- [CI-39] H. Soldano, V. Ventos, M. Champesme et D. Forge. Incremental construction of alpha lattices and association rules. *Proceedings of the 14th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2010)*, volume 6277 de *Lecture Notes in Computer Science*, pages 351–360. Springer, 2010. Part II.
- 2009** [CI-38] E. Alphonse et A. Osmani. Empirical Study of Relational Learning Algorithms in the Phase Transition Framework. *Proceedings of the European Machine Learning and Knowledge Discovery in Databases (ECML-PKDD 2009)*, volume 5781 de *Lecture Notes in Computer Science*, pages 51–66. Springer, 2009.
- [CI-37] G. Bourgne, D. Bouthinon, A. ElFallah-Seghrouchni et H. Soldano. Collaborative concept learning : Non individualistic vs individualistic agents. *Proceedings of the 21st IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2009)*, pages 653–657. IEEE Computer Society, 2009.
- [CI-36] G. Bourgne, A. ElFallah-Seghrouchni et H. Soldano. Learning in a fixed or evolving network of agents. *IEEE / WIC / ACM International Conference on Intelligent Agent Technology (WI-IAT 2009)*, pages 549–556, Milano, Italy, 2009. IEEE Computer Society.
- [CI-35] D. Bouthinon, H. Soldano et V. Ventos. Concept learning from (very) ambiguous examples. *Proceedings of the 6th International Conference on Machine Learning and Data Mining in Pattern Recognition (MLDM 2009)*, volume 5632 de *Lecture Notes in Computer Science*, pages 465–478. Springer, 2009.
- [CI-34] G. Cabanes et Y. Bennani. Comparing large datasets structures through unsupervised learning. C.-S. Leung, M. Lee et J. H. Chan, éditeurs, *Proceedings of the 16th International Conference On Neural Information Processing (ICONIP'09), Part I*, volume 5863 de *Lecture Notes in Computer Science*, pages 546–553, Bangkok, Thailand, dec 2009. Springer.
- [CI-33] G. Cabanes, Y. Bennani et F. Dufau-Joël. Mining Customers' Spatio-temporal Behavior Data using Topographic Unsupervised Learning. *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'09)*, pages 372–377, Miami Beach, Florida, USA, Dec 2009. IEEE Computer Society.
- [CI-32] G. Cabanes, D. Fresnau, U. Galassi et A. Giordana. A HMM-Based Approach to Modeling Ant Behavior. *4th International Indian Conference on Artificial Intelligence (IICAI'09)*, pages 2127–2139, Tumkur, India, 2009.
- [CI-31] G. Cabanes, D. Fresnau, A. Giordana et U. Galassi. Modeling Ant Activity by means of Structured HMMs. *International Symposium on Methodologies for Intelligent Systems (ISMIS'09)*, pages 341–350, Prague, Czech Republic, 2009.
- [CI-30] M. Essaidi et A. Osmani. Data warehouse development using mda and 2tup. *18th International Conference on Software Engineering and Data Engineering (SEDE-2009)*, pages 138–143. ISCA, 2009.
- [CI-29] N. Grozavu et Y. Bennani. A new competitive strategy for self organizing map learning. *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'09)*, pages 689–692, Miami Beach, Florida, USA, Dec 2009. IEEE Computer Society.
- [CI-28] N. Grozavu, Y. Bennani et M. Lebbah. From variable weighting to cluster characterization in topographic unsupervised learning. *Proceedings of the International Joint Conference*

on *Neural Networks (IJCNN 2009)*, pages 1005–1010, Atlanta, Georgia, USA, Jun 2009. IEEE.

[CI-27] M. Mbaye, F. Krief et H. Soldano. Collaborative strategy learning for distributed network self-configuring. *1st International Conference on Communications and Networking (ComNet'09)*, pages 1–8, 2009.

[CI-26] C. Recanati, N. Rogovschi et Y. Bennani. Hybrid Unsupervised Learning to Uncover Discourse Structure. Z. Vetulani et H. Uszkoreit, éditeurs, *Human Language Technology. Challenges of the Information Society - Third Language and Technology Conference (LTC 2007)*, volume 5603/2009 de *Lecture Notes in Artificial Intelligence, Computer Science*, pages 258–269, Poznan, Pologne, août 2009. Springer. volume of the best papers of LTC'07.

[CI-25] Y. Zafoune, A. Mokhtari et R. Kanawati. Mobile-agent approach for mobile code localization in ad hoc networks. *Proceedings of the 33rd Annual IEEE International Computer Software and Applications Conference (COMPSAC 2009)*, pages 36–39. IEEE Computer Society, 2009.

2008

[CI-24] E. Alphonse et A. Osmani. A model to study phase transition and plateaus in relational learning. *Proceedings of the 18th International Conference on Inductive Logic Programming (ILP 2008)*, volume 5194 de *Lecture Notes in Computer Science*, pages 6–23. Springer, 2008.

[CI-23] S. Aseervatham. A local latent semantic analysis-based kernel for document similarities. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2008)*, pages 214–219, Hong Kong, China, Jun 2008. IEEE.

[CI-22] H. Azzag et M. Lebbah. Clustering of self-organizing map. *Proceedings of the 16th European Symposium on Artificial Neural Networks (ESANN 2008)*, pages 209–214, 2008. Bruges, Belgium, April 23-25.

[CI-21] G. Bourgne, A. ElFallah-Seghrouchni, N. Maudet et H. Soldano. Multiagent incremental learning in networks. *Pacific Rim International Conference on Multi-Agents (PRIMA)*, volume 5357 de *Lecture Notes in Computer Science*, pages 109–120. Springer, 2008.

[CI-20] G. Cabanes et Y. Bennani. A local density-based simultaneous two-level algorithm for topographic clustering. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2008)*, pages 1176–1182, Hong Kong, China, Jun 2008. IEEE.

[CI-19] R. Kanawati. Collaborative Bibliography. *Proceedings of the 17th International Conference on Computing (CIC'08)*, Mexico, 2008.

[CI-18] R. Kanawati. On using SNA techniques for enhancing performances of on-line help-desks. *Proceedings of IADIS International Conference on E-commerce (E-commerce'08)*, pages 286–291, Amsterdam, 2008.

[CI-17] H. Karoui, R. Kanawati et L. Petrucci. Cooperative CBR system for peer agent committee formation. *Agents and Peer-to-Peer Computing*, volume 4461 de *Lecture Notes in Artificial Intelligence*, pages 51–62. Springer, 2008.

[CI-16] M. Lebbah, Y. Bennani et H. Benhadda. Relational analysis for consensus clustering from multiple partitions. M. A. Wani, X.-w. Chen, D. Casasent, L. A. Kurgan, T. Hu et K. Hafeez, éditeurs, *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'08)*, pages 218–223, Cincinnati, Ohio, USA, Dec 2008. IEEE Computer Society.

[CI-15] A.-P. Manine, E. Alphonse et P. Bessières. Information extraction as an ontology population task and its application to genic interactions. *Proceedings of the 20th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2008)*, pages 74–81. IEEE Computer Society, 2008.

[CI-14] N. Rogovschi, M. Lebbah et Y. Bennani. Probabilistic mixed topological map for categorical and continuous data. M. A. Wani, X.-w. Chen, D. Casasent, L. A. Kurgan, T. Hu et K. Hafeez, éditeurs, *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'08)*, pages 224–231, Cincinnati, Ohio, USA, Dec 2008. IEEE Computer Society.

2007

[CI-13] E. Alphonse et A. Osmani. Phase transition and heuristic search in relational learning. *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'07)*, pages 112–117, Cincinnati, Ohio, USA, Dec 2007. IEEE Computer Society.

- [CI-12] E. Alphonse et C. Rouveirol. Extension of the top-down data-driven strategy to ilp. *Inductive Logic Programming, 16th International Conference, ILP 2006, Revised Selected Papers*, volume 4455 de *Lecture Notes in Computer Science*, pages 49–63. Springer, 2007.
- [CI-11] S. Aseervatham, E. Viennet et Y. Bennani. A semantic kernel for semi-structured documents. *Proceedings of the 7th IEEE International Conference on Data Mining (ICDM'07)*, pages 403–408. IEEE Computer Society, 2007.
- [CI-10] H. Azzag, J. Lavergne, C. Guinot et G. Venturini. On building graphs of documents with artificial ants. *Proceedings of the 16th International Conference on World Wide Web (WWW 2007)*, pages 1299–1300. ACM, 2007. (poster).
- [CI-9] K. Bouandas et A. Osmani. Mining association rules in temporal sequences. *IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2007, part of the IEEE Symposium Series on Computational Intelligence 2007*, pages 610–615, 2007.
- [CI-8] G. Bourgne, H. Soldano et A. El Fallah-Seghrouchni. SMILE : Sound Multi-agent Incremental LEarning. *International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007)*, pages 164–171, Honolulu, Hawaii, 2007. ACM.
- [CI-7] G. Cabanes et Y. Bennani. A simultaneous two-level clustering algorithm for automatic model selection. *Proceedings of the International Conference on Machine Learning and Applications (ICMLA'07)*, pages 316–321, Cincinnati, Ohio, USA, Dec 2007. IEEE Computer Society.
- [CI-6] S. Guérif et Y. Bennani. Dimensionality Reduction Through Unsupervised Features Selection. *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks (EANN 2007)*, pages 98–106, Thessaloniki, Greece, aug 2007. Publishing Centre Alexander T.E.I. of Thessaloniki.
- [CI-5] R. Kanawati et M. Malek. Computing Social Networks for Information Sharing : A Case-based Approach. S. Douglas, éditeur, *Second international conference on Online Communities and Social Computing*, volume 4564 de *Lecture Notes in Computer Science*, pages 86–95, Beijing, China, 2007. Springer.
- [CI-4] J. Lavergne, H. Azzag, C. Guinot et G. Venturini. Incremental construction of neighborhood graphs using the ants self-assembly behavior. *Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, pages 399–406. IEEE Computer Society, 2007. Volume 1.
- [CI-3] M. Lebbah, N. Rogovschi et Y. Bennani. BeSOM : Bernoulli on Self-Organizing Map. *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2007)*, pages 631–636, Orlando, Florida, USA, Aug 2007. IEEE.
- [CI-2] C. Recanati, N. Rogovschi et Y. Bennani. Sequencing of verbs - a study on tense and aspect using unsupervised learning. G. Angelova, K. Bontcheva, R. Mitkov, N. Nicolov et N. Nikolov, éditeurs, *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 07)*, pages 499–503, Borovets, Bulgarie, septembre 2007. Institute for Parallel Processing, Bulgarian Academy of Sciences.
- [CI-1] Y. Zafoune, R. Kanawati et A. Moukhtari. Mobile Agents Localization in ad. hoc. Networks : A Comparative Study of Centralized and Distributed Approaches. *Proceedings of the 5th International Conference on Information and Communication Technologies (ICICT 07)*, pages 269–275, Cairo, Egypte, 2007.

Workshops et autres conférences avec comité de lecture

- 2012** [CO-95] A. Bossard et C. Rodrigues. Combining a Multi-document Summarization System with a Genetic Algorithm. I. Hatzilygeroudis et J. Prentzas, éditeurs, *Proceedings of the 2nd International Workshop on Combinations of Intelligent Methods and Applications (CIMA 2010)*, volume 8 de *Smart Innovation, Systems and Technologies*, pages 71–87, Arras, France, 2012. Springer.
- [CO-94] A. Chaibi, H. Azzag et M. Lebbah. Détection de groupes outliers en classification non supervisée. Y. Lechevallier, G. Melançon et B. Pinaud, éditeurs, *Extraction et Gestion des Connaissances (EGC'2012)*, volume RNTI-E-23 de *Revue des Nouvelles Technologies de l'Information*, pages 119–125. Hermann, 2012.

- [CO-93*] I. Chebil, M. Elati, C. Rodrigues, R. Nicolle et C. Rouveirol. Licorn* : construction de réseaux de régulation chez l'homme. *Conférence francophone sur l'Apprentissage automatique (CAp'2012)*, page to appear, may 2012.
- [CO-92] Y. Chevaleyre, F. Koriche et J.-D. Zucker. Learning linear classifiers with ternary weights from metagenomic data. *Conférence francophone sur l'Apprentissage automatique (CAp'2012)*, May 2012. 16 pages.
- [CO-91] N.-Q. Doan, H. Azzag et M. Lebbah. Clustering multi-niveaux de graphes : hiérarchique et topologique. Y. Lechevallier, G. Melançon et B. Pinaud, éditeurs, *Extraction et Gestion des Connaissances (EGC'2012)*, volume RNTI-E-23 de *Revue des Nouvelles Technologies de l'Information*, pages 567-569. Hermann, 2012.
- [CO-90] N. Grozavu, Y. Bennani et L. Labiod. Transformation de l'espace de description pour l'apprentissage par transfert. Y. Lechevallier, G. Melançon et B. Pinaud, éditeurs, *Extraction et Gestion des Connaissances (EGC'2012)*, volume RNTI-E-23 de *Revue des Nouvelles Technologies de l'Information*, pages 35-40. Hermann, 2012.
- [CO-89] F. Hamdi et Y. Bennani. Apprentissage d'ensemble d'opérateurs de projection orthogonale pour la détection de nouveauté. Y. Lechevallier, G. Melançon et B. Pinaud, éditeurs, *Extraction et Gestion des Connaissances (EGC'2012)*, volume RNTI-E-23 de *Revue des Nouvelles Technologies de l'Information*, pages 41-52. Hermann, 2012.
- [CO-88] R. Jaziri, M. Lebbah et Y. Bennani. Classification probabiliste non supervisée et visualisation des données séquentielles. Y. Lechevallier, G. Melançon et B. Pinaud, éditeurs, *Extraction et Gestion des Connaissances (EGC'2012)*, volume RNTI-E-23 de *Revue des Nouvelles Technologies de l'Information*, pages 137-148. Hermann, 2012.
- [CO-87] R. Kanawati. CaSE : A cbr approach for re-ranking web search results. *International workshop on case-based reasoning (CBR-MD'12)*, Berlin, july 2012. IBAI.
- [CO-86] L. Labiod et Y. Bennani. Classification des données catégorielles via la maximisation spectral de la modularité. Y. Lechevallier, G. Melançon et B. Pinaud, éditeurs, *Extraction et Gestion des Connaissances (EGC'2012)*, volume RNTI-E-23 de *Revue des Nouvelles Technologies de l'Information*, pages 149-154. Hermann, 2012.
- [CO-85] K. Mouhoubi, L. Létocart et C. Rouveirol. Extraction de biclusters contraints dans des contextes bruités. *Actes de la Conférence Francophone d'Apprentissage (CAp 2012)*, Nancy, France, 2012.
- [CO-84*] N. A. Nguyen Thi, Y. Chevaleyre et J.-D. Zucker. Optimizing sign placements for crowd evacuation on road network in case of tsunami alert. *Proceedings of the 21st IEEE International WETICE conference, Track on Collaborative Technology for Coordinating Crisis Management*, 2012.
- [CO-83*] M. Pujari et R. Kanawati. Mining the dynamics of scientific publication networks for collaboration recommendation. *Second international workshop on mining communities and people recommender (COMMPER-ECML 2012)*, Bristol, United Kingdom, September 2012.
- [CO-82*] M. Pujari et R. Kanawati. Supervised rank aggregation approach for link prediction in complex networks. *3ème Conférence sur l'analyse et la modélisation des réseaux : approches mathématiques et informatiques (MARAMI 2012)*, Villetaneuse, October 2012.
- [CO-81] M. Pujari et R. Kanawati. Supervised rank aggregation approach for link prediction in complex networks. *Proceedings of the 21st international conference companion on World Wide Web (WWW '12 Companion)*, pages 1189-1196, Lyon, april 2012. ACM, New York, NY, USA. (MSND'12 workshop : Mining Social Networks Dynamics).
- [CO-80] Z. Yakoubi et R. Kanawati. Applying leaders driven community detection algorithms to data clustering. *The 36th Annual Conference of the German Classification Society on Data Analysis, Machine Learning and Knowledge Discovery (GfKI'12)*, 2012.
- 2011** [CO-79] H. Azzag et M. Lebbah. Une nouvelle approche visuelle pour la classification hiérarchique et topologique. A. Khenchaf et P. Poncet, éditeurs, *Extraction et Gestion des Connaissances (EGC'2011)*, volume RNTI-E-20 de *Revue des Nouvelles Technologies de l'Information*, pages 677-688. Hermann, 2011.
- [CO-78] G. Cabanes. Classification à deux niveaux selon le voisinage et la densité. *Actes de la 18ième Rencontre de la Société Francophone de Classification (SFC'2011)*, pages 107-110, Orléans, France, 2011.

- [CO-77] G. Cabanes et Y. Bennani. Apprendre les contraintes topologiques dans les cartes auto-organisatrices. A. Khenchaf et P. Poncelet, éditeurs, *Extraction et Gestion des Connaissances (EGC'2011)*, volume RNTI-E-20 de *Revue des Nouvelles Technologies de l'Information*, pages 137–148. Hermann, 2011.
- [CO-76] G. Cabanes et Y. Bennani. Visualisation de l'intra et inter structure des groupes en classification non supervisée. A. Khenchaf et P. Poncelet, éditeurs, *Extraction et Gestion des Connaissances (EGC'2011)*, volume RNTI-E-20 de *Revue des Nouvelles Technologies de l'Information*, pages 479–490. Hermann, 2011.
- [CO-75] N. Grozavu, M. Ghassany et Y. Bennani. Apprentissage de la confiance des échanges en classification collaborative non supervisée. *Actes de la Conférence Francophone d'Apprentissage (CAp 2011)*, pages 217–232, Chambéry, France, may 2011. Éditions Publibook, Presses Universitaires des Antilles et de la Guyane (PUAG). 7e Plateforme AFIA.
- [CO-74] R. Jaziri, M. Lebbah, Y. Bennani et J.-H. Chenot. Structuration automatique des flux télévisuels par apprentissage non supervisé des répétitions. A. Khenchaf et P. Poncelet, éditeurs, *Extraction et Gestion des Connaissances (EGC'2011)*, volume RNTI-E-20 de *Revue des Nouvelles Technologies de l'Information*, pages 311–312. Hermann, 2011.
- [CO-73] R. Kanawati. A new algorithm for community detection in academic collaboration networks. *Third International Workshop on Social Network Analysis (ARS'11)*, Naples, 2011.
- [CO-72] R. Kanawati. A new Leaders-followers algorithm for detecting overlapping communities in social networks. *Interdisciplinary Workshop on Information and Decision in Social Networks (WIDS)*, MIT, Camrdidge, 2011.
- [CO-71] R. Kanawati. Overlapping communities detection algorithm. *International Conference on Network Science (NETSCI 2011)*, Budapest, june 2011.
- [CO-70] R. Kanawati. Query graphs analyzing for query similarity evaluation. *Proceedings of the 11th Dutch-Belgian Information Retrieval Workshop (DIR-11)*, pages 48–51, Amsterdam, 2011.
- [CO-69] K. Mouhoubi, L. Létocart et C. Rouveirol. Extraction de motifs ensemblistes dans des contextes bruités. *Actes de la Conférence Francophone d'Apprentissage (CAp 2011)*, pages 311–326, Chambéry, France, may 2011. Éditions Publibook, Presses Universitaires des Antilles et de la Guyane (PUAG). 7e Plateforme AFIA.
- [CO-68] K. Mouhoubi, L. Létocart et C. Rouveirol. Heuristique pour l'extraction de motifs ensemblistes bruités. A. Khenchaf et P. Poncelet, éditeurs, *Extraction et Gestion des Connaissances (EGC'2011)*, volume RNTI-E-20 de *Revue des Nouvelles Technologies de l'Information*, pages 467–472. Hermann, 2011.
- [CO-67] K. Mouhoubi, L. Létocart et C. Rouveirol. Heuristique pour l'extraction de motifs ensemblistes bruités. *12ème congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF'2011)*, Saint-Étienne, France, 2011.
- [CO-66] K. Mouhoubi, L. Létocart et C. Rouveirol. Une approche heuristique hybride pour l'extraction de motifs ensemblistes dans des contextes bruités. *Journée Fouille de Grands Graphes (JFGG 2011)*, Grenoble, France, 2011.
- [CO-65] G. Ravilly-Abadie, Y. Chevaleyre, M. Lacroix, J. Lang et N. Maudet. Que sait-on du résultat d'une élection avant que tous les votants se soient exprimés? *12ème congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF'2011)*, Saint-Etienne, France, 2011.
- [CO-64] C. Rodrigues, P. Gérard, C. Rouveirol et H. Soldano. Relational action model learning and planning integration. *Actes de la Conférence Francophone d'Apprentissage (CAp 2011)*, pages 361–376, Chambéry, France, may 2011. Éditions Publibook, Presses Universitaires des Antilles et de la Guyane (PUAG). 7e Plateforme AFIA.
- [CO-63] N. Rogovschi, M. Lebbah et N. Grozavu. Pondération et classification simultanée de données binaires et continues. A. Khenchaf et P. Poncelet, éditeurs, *Extraction et Gestion des Connaissances (EGC'2011)*, volume RNTI-E-20 de *Revue des Nouvelles Technologies de l'Information*, pages 65–70. Hermann, 2011.

- 2010** [CO-62] H. Azzag et M. Lebbah. Auto-organisation topologique et hiérarchique des données. S. Ben Yahia et J.-M. Petit, éditeurs, *Extraction et Gestion des Connaissances (EGC'2010)*, volume RNTI-E-19 de *Revue des Nouvelles Technologies de l'Information*, pages 555–560. Cépaduès-Éditions, 2010.
- [CO-61] H. Azzag et M. Lebbah. A new method for topological and hierarchical clustering. *Proceedings of the 34th Annual Conference of the German Classification Society (GfKI)*, page 52, July 2010.
- [CO-60] N. Benchettara, R. Kanawati et C. Rouveirol. Apprentissage supervisé pour la prédiction de nouveaux liens dans des réseaux sociaux bipartite. *Actes de la 17ième Rencontre de la Société Francophone de Classification (SFC'2010)*, pages 63–66, St. Denis, La Réunion, juin 2010.
- [CO-59] N. Benchettara, R. Kanawati et C. Rouveirol. Supervised machine learning applied to link prediction in bipartite social networks. *International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2010)*, pages 326–330. IEEE Computer Society, 2010.
- [CO-58] L. Boudjeloud-Assala et H. Azzag. Approche biomimétique coopérative pour la visualisation de grands graphes multidimensionnels. S. Ben Yahia et J.-M. Petit, éditeurs, *Extraction et Gestion des Connaissances (EGC'2010)*, volume RNTI-E-19 de *Revue des Nouvelles Technologies de l'Information*, pages 667–668. Cépaduès-Éditions, 2010.
- [CO-57] G. Bourgne, A. El Fallah-Seghrouchni et H. Soldano. Learning better together with forgetness. E. Mephu, éditeur, *Actes de la 12e Conférence d'Apprentissage automatique (CAp'2010)*, pages 177–188, Clermont-Ferrand, France, May 2010.
- [CO-56] G. Cabanes, Y. Bennani et F. Dufau-Joël. Comparaisons structurelles de grandes bases de données par apprentissage non-supervisé. S. Ben Yahia et J.-M. Petit, éditeurs, *Extraction et Gestion des Connaissances (EGC'2010)*, volume RNTI-E-19 de *Revue des Nouvelles Technologies de l'Information*, pages 115–120. Cépaduès-Éditions, 2010.
- [CO-55] M. Essaidi. ODBIS : towards a platform for on-demand business intelligence services. *Proceedings of the 2010 EDBT/ICDT Workshops, EDBT'10*, pages 1–6, New York, NY, USA, 2010. ACM.
- [CO-54] M. Essaidi et A. Osmani. Towards model-driven data warehouse automation using machine learning. *ICEC 2010 - Proceedings of the International Conference on Evolutionary Computation*, pages 380–383. SciTePress, 2010.
- [CO-53] N. Grozavu et Y. Bennani. Classification collaborative non supervisée. E. Mephu, éditeur, *Actes de la 12e Conférence d'Apprentissage automatique (CAp'2010)*, pages 249–276, Clermont-Ferrand, France, May 2010.
- [CO-52] N. Grozavu, L. Labioud et Y. Bennani. Dimensionality reduction for binary data. *5th International Conference on Neural Networks and Artificial Intelligence (ICNNAI'10)*, pages 30–33, Brest, Belarus, 2010.
- [CO-51] R. Kanawati. LIPS : A SNA-based system for Intelligent Management of Academic Conferences. *Social networking in Cyberspace conference (SNIC'2010)*, 2010.
- [CO-50] L. Labioud, N. Grozavu et Y. Bennani. Classification relationnelle topographique. E. Mephu, éditeur, *Actes de la 12e Conférence d'Apprentissage automatique (CAp'2010)*, pages 277–288, Clermont-Ferrand, France, May 2010.
- [CO-49] L. Labioud, N. Grozavu et Y. Bennani. Relationship between the modularity criterion and the relational analysis. *Proceedings of the IEEE International Conference on Data Mining and Intelligent Information Technology Applications (ICMIA'10)*, pages 229–236, Seoul, Korea, 2010.
- [CO-48] L. Labioud, N. Grozavu et Y. Bennani. Topological organization for categorical data clustering. *5th International Conference on Neural Networks and Artificial Intelligence (ICNNAI'10)*, pages 34–41, Brest, Belarus, 2010.
- [CO-47] M. Lebbah et Y. Bennani. Sous-échantillonnage topographique par apprentissage semi-supervisé. S. Ben Yahia et J.-M. Petit, éditeurs, *Extraction et Gestion des Connaissances (EGC'2010)*, volume RNTI-E-19 de *Revue des Nouvelles Technologies de l'Information*, pages 121–126. Cépaduès-Éditions, 2010.
- [CO-46] M. M. Ousaid et R. Kanawati. Analyse des graphes sans échelle pour l'évaluation des similarités entre requêtes web. *Actes de la 1er journée de fouille de grands graphes (FGG'2010)*, 2010.

- [CO-45] M. Pujari et R. Kanawati. Link Prediction Approach for Tag Recommendation in Folksonomy. *Actes de la 1er journée de fouille de grands graphes (FGG'2010)*, 2010.
- [CO-44] C. Rodrigues, P. Gérard et C. Rouveirol. Apprentissage incrémental de règles d'actions relationnelles. *17ème congrès francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle (RFIA 2010)*, 2010.
- [CO-43] C. Rodrigues, P. Gérard et C. Rouveirol. Incremental learning of relational action models in noisy environments. *Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite des systèmes (JFPDA 2010)*, 2010.
- [CO-42] N. Rogovschi et N. Grozavu. A content-based image retrieval system based on unsupervised topological learning. *Proceedings of the IEEE International Conference on Data Mining and Intelligent Information Technology Applications (ICMIA'10)*, pages 398–394, Seoul, Korea, 2010.
- 2009**
- [CO-41] E. Alphonse. Étude empirique de la transition de phase en apprentissage relationnel. *Actes de la 11e Conférence d'Apprentissage, Plateforme AFIA (CAp 2009)*, pages 173–184, Hammamet, Tunisie, may 2009. Presses universitaires de Grenoble.
- [CO-40] E. Alphonse et A. Osmani. Analyse en moyenne de la complexité des algorithmes d'apprentissage relationnel. *Journées Nationales de l'IA Fondamentale (JIAF 2009)*, 2009. GDR I3.
- [CO-39] N. Benchettara, R. Kanawati et C. Rouveirol. Supervised Machine Learning for Link prediction in two Mode Social network. *International Conference on Two-Mode Social Network Analysis*, Amsterdam, 2009.
- [CO-38] D. Bouthinon, H. Soldano et V. Ventos. Apprentissage de concept à partir d'exemples (très ambigus). *Actes de la 11e Conférence d'Apprentissage, Plateforme AFIA (CAp 2009)*, pages 197–208, Hammamet, Tunisie, may 2009. Presses universitaires de Grenoble.
- [CO-37] G. Cabanes, Y. Bennani et D. Fresneau. Exploration de données de traçabilité issues de la RFID par apprentissage non-supervisé. J.-G. Ganascia et P. Gançarcki, éditeurs, *Extraction et Gestion des Connaissances (EGC'2009)*, volume RNTI-E-15 de *Revue des Nouvelles Technologies de l'Information*, pages 451–452. Cépaduès-Éditions, 2009.
- [CO-36] N. Grozavu et Y. Bennani. Apprentissage topographique non supervisé avec mémoire. *Actes de la 11e Conférence d'Apprentissage, Plateforme AFIA (CAp 2009)*, pages 351–354, Hammamet, Tunisie, may 2009. Presses universitaires de Grenoble.
- [CO-35] N. Grozavu et Y. Bennani. Voting memory based self-organizing map. *TopoLearn'09 : International Workshop on Topological Learning, in conjunction with ISMIS'09, International Symposium on Methodologies for Intelligent Systems*, Prague, Czech Republic, 2009.
- [CO-34] N. Grozavu, Y. Bennani et M. Lebbah. Caractérisation automatique des classes découvertes en classification non supervisée. J.-G. Ganascia et P. Gançarcki, éditeurs, *Extraction et Gestion des Connaissances (EGC'2009)*, volume RNTI-E-15 de *Revue des Nouvelles Technologies de l'Information*, pages 43–54. Cépaduès-Éditions, 2009.
- [CO-33] S. Guérif. Sous-échantillonnage et test de permutations pour la sélection de variables non supervisée. *Actes de la 11e Conférence d'Apprentissage, Plateforme AFIA (CAp 2009)*, pages 355–358, Hammamet, Tunisie, may 2009. Presses universitaires de Grenoble. (poster).
- [CO-32] R. Kanawati et H. Karoui. A P2P collaborative bibliography recommender system. *Proceedings of 4th International Conference on Internet and Web Applications and Services (ICIW '09)*, pages 90–96, 2009.
- [CO-31] R. Kanawati et C. Rouveirol. LIPS : A SNA-based System for intelligent Management of Academic Conferences. *4th International Conference on Applications of Social Network Analysis (ASNA'2009)*, Zurich, Zurich, 2009.
- [CO-30] N. Rogovschi, M. Lebbah et Y. Bennani. Un algorithme pour la classification topographique simultanée de données qualitatives et quantitatives. *Actes de la 11e Conférence d'Apprentissage, Plateforme AFIA (CAp 2009)*, pages 209–224, Hammamet, Tunisie, may 2009. Presses universitaires de Grenoble.

- [CO-29] Y. Zafoune, A. Moukhtari et R. Kanawati. Mobile agent approach for mobile code localization in ad hoc networks. *4th IEEE International Workshop on Engineering Semantic Agent Systems (ESAS 2009)*, Seattle, Washington, 2009.
- 2008**
- [CO-28] E. Alphonse. The phase transition of the bounded ILP consistency problem. F. zelezny et N. Lavrac, éditeurs, *18th International Conference on Inductive Logic Programming (ILP 2008)*, pages 1–6, 2008. Late Breaking Papers Track.
- [CO-27] E. Alphonse et A. Osmani. A problem generator to study heuristic search in relational learning. *Conférence d'Apprentissage (CAp'2008)*, pages 1–16, 2008.
- [CO-26] S. Aseervatham. Un modèle despace vectoriel de concepts pour noyaux sémantiques. F. Guillet et B. Trousse, éditeurs, *Extraction et Gestion des Connaissances (EGC'2008)*, volume RNTI-E-11 de *Revue des Nouvelles Technologies de l'Information*, pages 601–612. Cépaduès-Éditions, 2008.
- [CO-25] G. Cabanes et Y. Bennani. Classification topographique à deux niveaux simultanés à base de modes de densité. *Conférence d'Apprentissage (CAp'2008)*, pages 53–54, Porquerrolle, France, 2008.
- [CO-24] G. Cabanes et Y. Bennani. Un algorithme de classification topographique non supervisée à deux niveaux simultanés. F. Guillet et B. Trousse, éditeurs, *Extraction et Gestion des Connaissances (EGC'2008)*, volume RNTI-E-11 de *Revue des Nouvelles Technologies de l'Information*, pages 619–630. Cépaduès-Éditions, 2008.
- [CO-23] G. Cabanes, Y. Bennani, C. Chartagnat et D. Fresneau. Topographic connectionist unsupervised learning for RFID behavior data mining. *The Second International Workshop on RFID Technology (IWRIT), conjunction with ICEIS 2008*, pages 63–72, 2008.
- [CO-22] M. Elati et C. Rouveirol. Recherche adaptative de structures de régulation génétique. F. Guillet et B. Trousse, éditeurs, *Extraction et Gestion des Connaissances (EGC'2008)*, volume RNTI-E-11 de *Revue des Nouvelles Technologies de l'Information*, pages 427–432. Cépaduès-Éditions, 2008.
- [CO-21] N. Grozavu, Y. Bennani et M. Lebbah. Pondération locale des variables en apprentissage numérique non-supervisé. F. Guillet et B. Trousse, éditeurs, *Extraction et Gestion des Connaissances (EGC'2008)*, volume RNTI-E-11 de *Revue des Nouvelles Technologies de l'Information*, pages 321–330. Cépaduès-Éditions, 2008.
- [CO-20] S. Guérif. Unsupervised variable selection : when random rankings sound as irrelevancy. *Journal of Machine Learning Research - Proceedings Track*, 4 :163–177, 2008. (FSDM workshop at ECML-PKDD).
- [CO-19] R. Kanawati. A CBR framework for implementing community-aware web search engine. *proceedings of second international workshop on adaptive information retrieval (AIR'08)*, pages 32–38, London, UK, 2008.
- [CO-18] M. Lebbah et H. Azzag. Segmentation hiérarchique des cartes topologiques. F. Guillet et B. Trousse, éditeurs, *Extraction et Gestion des Connaissances (EGC'2008)*, volume RNTI-E-11 de *Revue des Nouvelles Technologies de l'Information*, pages 631–642. Cépaduès-Éditions, 2008.
- [CO-17] A.-P. Manine, E. Alphonse et P. Bessières. Genic interaction extraction by reasoning on an ontology. *Proceedings of the 3rd International Symposium on Semantic Mining in Biomedicine (SMBM 2008)*, pages 93–100, 2008.
- [CO-16] C. Rodrigues, P. Gérard et C. Rouveirol. On and off-policy relational reinforcement learning. F. zelezny et N. Lavrac, éditeurs, *18th International Conference on Inductive Logic Programming (ILP 2008)*, pages 99–104, 2008. Late Breaking Papers Track.
- [CO-15] C. Rodrigues, P. Gérard et C. Rouveirol. Relational TD Reinforcement Learning, 2008. 8th European Workshop on Reinforcement Learning (EWRL 2008).
- 2007**
- [CO-14] E. Alphonse et A. Osmani. On the connection between the phase transition of the covering test and the learning success rate. *Conférence d'Apprentissage (CAp'2007)*, pages 1–16, July 2007.
- [CO-13] E. Alphonse et C. Rouveirol. Extension of the Top-Down Data-Driven Strategy to ILP. *Conférence d'Apprentissage (CAp'2007)*, pages 17–18, July 2007.
- [CO-12] J. Azé, L. Gentils, C. Toffano-Nioche, V. Loux, J.-F. Gibrat, P. Bessières, C. Rouveirol, A. Poupon et C. Froidevaux. Towards a semi-automatic functional annotation tool based on decision tree techniques. *International Workshop on Machine Learning in Systems Biology (MLSB 2007)*, 2007.

- [CO-11] E. Birmelé, M. Elati, C. Rouveirol et C. Ambroise. Identification of functional modules based on transcriptional regulation structure. *International Workshop on Machine Learning in Systems Biology (MLSB 2007)*, 2007.
- [CO-10] M. Elati, P. Neuvial, C. Rouveirol, F. Radvanyi et E. Barillot. A methodology for discovering cooperative regulatory pattern from ternary gene expression data. *Conférence d'Apprentissage (CAp'2007)*, pages 91–106, Grenoble, France, July 2007.
- [CO-9] S. Guérif et Y. Bennani. Sélection de Variables en Apprentissage Numérique Non Supervisé. *Conférence d'Apprentissage (CAp'2007)*, pages 221–236, Grenoble, France, July 2007.
- [CO-8] R. Kanawati. A community aware web search engine. *In proceedings HCI International 2007*, Beijing, china, july 2007. poster.
- [CO-7] H. Karoui. Systèmes de recommandation de références bibliographiques et de formation de comités. *Actes des 3èmes Rencontre Inter-Associations (RIAs 07) : Recherche et Extraction d'Information*, Toulouse - IRIT, 2007.
- [CO-6] J. Lavergne, H. Azzag, C. Guinot et G. Venturini. Construction et exploration interactive d'un graphe de documents. *3èmes Rencontres Inter-Associations (RIA's 2007)*, Toulouse, France, 12-13 mars 2007. IRIT.
- [CO-5] J. Lavergne, H. Azzag, C. Guinot et G. Venturini. Une approche incrémentale d'une méthode de classification non supervisée par nuages d'insectes volants. *Actes de la 14ième Rencontre de la Société Francophone de Classification (SFC'2007)*, Paris, France, september 2007. ENST.
- [CO-4] J. Lavergne, H. Azzag, C. Guinot et G. Venturini. Construction incrémentale et visualisation de graphes de voisinage par des fourmis artificielles. M. Noirhomme-Fraiture et G. Venturini, éditeurs, *Actes des cinquièmes journées Extraction et Gestion des Connaissances (EGC'2007)*, volume RNTI-E-9 de *Revue des Nouvelles Technologies de l'Information*, pages 135–146. Cépaduès-Éditions, 2007.
- [CO-3] M. Lebbah, N. Rogovschi et Y. Bennani. BeSOM : Bernoulli on Self Organizing Map. *Conférence d'Apprentissage (CAp'2007)*, pages 283–297, July 2007.
- [CO-2] C. Recanati et N. Rogovschi. Enchaînements verbaux - étude sur le temps et l'aspect utilisant des techniques d'apprentissage non supervisé. *Actes de la 14ème conférence sur le Traitement Automatique des Langues Naturelles (TALN'07)*, pages 379–388, Toulouse, France, juin 2007. IRIT Press. 10 pages.
- [CO-1] C. Recanati, N. Rogovschi et Y. Bennani. The structure of verbal sequences analyzed with unsupervised learning techniques. *Proceedings of the 3rd Language and Technology Conference : Human and Language Technologies as a Challenge for Computer Science and Linguistics*, pages 325–329, Poznań, Poland, oct 2007.

Édition d'ouvrages collectifs

- 2012**
- [ED-15] E. Banissi, S. Bertschi, C. Forsell, J. Johansson, S. Kenderdine, F. T. Marchese, M. Sarfraz, L. J. Stuart, A. Ursyn, T. G. Wyeld, H. Azzag, M. Lebbah et G. Venturini, éditeurs. *16th International Conference on Information Visualisation, IV 2012, Montpellier, France, July 11-13, 2012*. IEEE Computer Society, 2012, ISBN 978-1-4673-2260-7.
- [ED-14] H. Cherifi et R. Kanawati, éditeurs. *Community Structure in Complex Networks*, volume 8 de *International Journal of Web-based Communities*. Inderscience, september 2012.
- [ED-13*] R. G. Pensa, F. Cordero, C. Rouveirol et R. Kanawati, éditeurs. *Special issue on Dynamic Networks and Knowledge Discovery*, volume 17 de *Intelligent Data Analysis*. IOS Press, 2012. Parution prévue en janvier 2013.
- 2011**
- [ED-12] B. Ben Yaghlane, G. Cleuziou, M. Lebbah et A. Martin, éditeurs. *Fouille de données complexes. Complexité liée aux données multiples*, volume RNTI-E-21 de *Revue des Nouvelles Technologies de l'Information*. Hermann, 2011.
- [ED-11] Y. Bennani et E. Viennet, éditeurs. *Apprentissage artificiel et fouille de données*, volume RNTI-A-5 de *Revue des Nouvelles Technologies de l'Information*. Hermann, september 2011. 4èmes journées thématiques apprentissage artificiel et fouille de données, 29-30 juin 2010.

- [ED-10] R. Kanawati et A. Termier, éditeurs. *Actes de la deuxième journée de Fouille de Grands Graphes*, Grenoble, 2011.
- 2010** [ED-9] H. Azzag, L. Boudjeloud, R. Kanawati et F. Picarougne, éditeurs. *Actes de la première journée de Fouille de Grands Graphes*, Toulouse, 2010.
- [ED-8] H. Azzag, M. Lebbah et G. Venturini, éditeurs. *Apprentissage et visualisation*, volume RNTI-A-4 de *Revue des Nouvelles Technologies de l'Information*. Hermann, 2010.
- [ED-7] Y. Bennani, éditeur. *Actes de la 1ère École thématique "Apprentissage Statistique et Data Mining"*. é-EGC, Presses universitaires de Tunis, feb 2010. 562 pages, Hammamet-Tunisie.
- [ED-6] R. G. Pensa, F. Cordero, C. Rouveirol, R. Kanawati, J. A. Troyano et P. Rosso, éditeurs. *Proceedings of the 1st Workshop on Dynamic Networks and Knowledge Discovery (DyNaK 2010), Barcelona, Spain, Oct 2010*. CEUR-WS, Vol-655, urn :nbn :de :0074-655-4.
- 2009** [ED-5] K. Benabdeslem, M. Elati, R. Kanawati, H. Kheddouci et C. Rouveirol, éditeurs. *Atelier Apprentissage et Graphes pour les Systèmes complexes (AGS'2009)*, May 2009. Plate-forme AFIA, 25 Mai, Hammamet-Tunisie (actes électroniques, 60 pages).
- [ED-4] Y. Bennani et C. Rouveirol, éditeurs. *Actes de la 11e Conférence d'Apprentissage, Plateforme AFIA (CAp 2009)*. Presses universitaires de Grenoble, may 2009. Plate-forme AFIA, Hammamet-Tunisie.
- [ED-3] Y. Bennani et E. Viennet, éditeurs. *Apprentissage artificiel et fouille de données*, volume RNTI-A3 de *Revue des Nouvelles Technologies de l'Information*. Cepadues, june 2009. 3èmes journées thématiques apprentissage artificiel et fouille de données, 8-9 avril 2008.
- 2008** [ED-2] Y. Bennani et E. Viennet, éditeurs. *Apprentissage artificiel et fouille de données*, volume RNTI-A2 de *Revue des Nouvelles Technologies de l'Information*. Cepadues, march 2008. 2èmes journées thématiques apprentissage artificiel et fouille de données, 27-28 avril 2006.
- 2007** [ED-1] H. Azzag et G. Venturini. *Fouille du Données et algorithmes biomimétiques*. *Atelier Fouille du Données et Algorithmes Biomimétiques aux 7ème EGC'07 Extraction et Gestion des Connaissances*, page 60, jan 2007.

Logiciels et Brevets (LO)

- 2012** [LO-5] D. Bouthinon et H. Soldano. Lear, 2012. Système pour apprendre un ensemble de clauses du premier ordre à partir d'exemples ambigus.
- 2010** [LO-4] H. Benhadda, Y. Bennani, M. Lebbah et N. Grozavu. System for searching visual information. WO/2010/066774. *European Patent Office (EPO)*. Brevet (THALES, Université de Paris 13), 2010.
- [LO-3] M. Champesme. Alphanet, 2010. Extension du logiciel Galicia pour la construction incrémentale et la fusion de treillis Alpha.
- 2009** [LO-2] Y. Bennani. Methods for updating and training a self-organising map (wo/2009/081005 - pct/fr2008/052288). Brevet Université Paris 13, 2009.
- 2008** [LO-1] G. Cabanes et Y. Bennani. Density-based Simultaneous 2-Level - Self-Organizing Map (DS2L-SOM). 2008. Enregistrement numéro IDN.FR.001.490019.000S.P.2008.000.20000 auprès de l'agence de protection logiciel.

Thèses et Habilitations (TH)

- 2012** [TH-10] M. Lebbah. Contributions en apprentissage non supervisé à partir de données complexes. HdR, Université Paris 13, 2012.
- [TH-9] A. Osmani. Modélisation et raisonnement sur des données relationnelles : apprentissage relationnel, diagnostic de systèmes dynamiques et raisonnement spatio-temporel. HdR, Université Paris 13, 2012.
- 2011** [TH-8] N. Benchettara. Prédiction de nouveaux liens dans les réseaux d'interactions bipartis : Application au calcul de recommandation. Thèse, Université Paris 13, 2011.
- 2010** [TH-7] G. Cabanes. Two-level Unsupervised Clustering driven by neighborhood and density. Thèse, Université Paris 13, 2010.
- 2009** [TH-6] N. Grozavu. Classification topologique pondérée : approches modulaires, hybrides et collaboratives. Thèse, Université Paris 13, 2009.
- [TH-5] N. Rogovschi. Classification à base de modèles de mélanges topologiques des données catégorielles et continues. Thèse, Université Paris 13, 2009.
- [TH-4] H. Soldano. Apprentissage : Paradigmes, structures et abstractions. HdR, Université Paris 13, 2009.
- 2007** [TH-3] S. Aseervatham. Apprentissage à base de noyaux sémantiques pour le traitement de données textuelles. Thèse, Université Paris 13, 2007.
- [TH-2] M. Elati. Apprentissage de relations de régulation complexes à partir de données d'expression. Thèse, Université Paris 13, 2007.
- [TH-1] E. Viennet. Contributions aux méthodes d'apprentissage numérique pour la fouille de données structurées. HdR, Université Paris 13, 2007.

Divers (Div)

- 2012** [Ra-1] K. Mouhoubi, R. Nicolle, L. Létocart, M. Elati, F. Radvanyi et C. Rouveirol. Discrete bi-clustering for mining noisy contexts. Rr, LIPN - CNRS UMR 7030 - Université Paris 13, 2012.

Chapitre 7

Publications antérieures à l'arrivée au LIPN

Les références comportent « ↑ » et correspondent à des publications antérieures à l'arrivée au LIPN, non référencées par ailleurs.

Signification des soulignements :

- *simple* : enseignant-chercheur ou chercheur titulaire,
- *double* : doctorant,
- *pointillé* : post-doc, ingénieur ou autre situation.

Sommaire

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)	74
Communications avec actes (ACT)	75

Articles dans des revues avec comité de lecture et chapitres de livres (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

- 2010 [↑RI-7] Y. Chevaleyre, U. Endriss et N. Maudet. Simple negotiation schemes for agents with simple preferences : Sufficiency, necessity and maximality. *Journal of Autonomous Agents and Multi-Agent Systems*, 20(2) :234–259, 2010.
- 2009 [↑RI-6] G. P. H. Santini, J. A. H. Cognet, D. Xu, K. K. Singarapu et C. Hervé du Penhoat. Nucleic acid folding determined by mesoscale modeling and NMR spectroscopy : solution structure of d(GCGAAAGC). *The Journal of Physical Chemistry B*, 113(19) :6881–93, 2009.
- 2008 [↑RI-5] Y. Chevaleyre, U. Endriss, S. Estivie et N. Maudet. Multiagent resource allocation in k-additive domains : Preference representation and complexity. *Annals of Operations Research*, 163(1) :49–62, 2008.
- [↑RI-4] P. E. Dunne et Y. Chevaleyre. The complexity of deciding reachability properties of distributed negotiation schemes. *Theoretical Computer Science*, 396(1) :113–144, 2008.
- [↑RI-3] J. Lang, U. Endriss, Y. Chevaleyre et N. Maudet. Preference handling in combinatorial domains : From ai to social choice. *AI Magazine*, 24(4) :37–46, 2008.
- 2007 [↑RI-2] M. Elati, P. Neuvial, M. Bolotin-Fukuhara, E. Barillot, F. Radvanyi et C. Rouveirol. Licorn : learning cooperative regulation networks from gene expression data. *Bioinformatics*, 23(18) :2407–2414, 2007.
- [↑RI-1] G. P. H. Santini, C. Pakleza, P. Auffinger, C. Moriou, A. Favre, P. Clivio et J. A. H. Cognet. Dinucleotide tpt and its 2'-O-Me analogue possess different backbone conformations and flexibilities but similar stacked geometries. *The Journal of Physical Chemistry B*, 111(31) :9400–9, 2007.

Articles dans des revues nationales ou internationales

- 2008 [↑RE-1] M. Lebbah, M. R. Temanni, C. Poitou-Bernert, K. Clément et J.-D. Zucker. Partitionnement des données pour les problèmes de classement difficiles : Combinaison des cartes topologiques mixtes et svm. *Revue des Nouvelles Technologies de l'Information*, RNTI-A2 :34–54, 2008. Numéro spécial : Journées Apprentissage Artificiel et Fouille de Données.

Chapitres de livre

- 2010 [↑CL-3] Y. Chevaleyre, F. Koriche, J. Lang, J. Mengin et B. Zanuttini. Learning ordinal preferences on multiattribute domains : the case of cp-nets. *Preference Learning*, pages 273–296. Springer-Verlag New York Inc, 2010.
- 2008 [↑CL-2] Y. Chevaleyre, U. Endriss et N. Maudet. Restricted classes of utility functions for simple negotiation schemes : Sufficiency, necessity, and maximality. V. T. Paschos, éditeur, *Combinatorial Optimization and Theoretical Computer Science*, pages 175–199. Wiley-ISTE, 2008.
- 2007 [↑CL-1] Y. Chevaleyre. The patrolling problem : Theoretical and experimental results. *Combinatorial Optimization and Theoretical Computer Science*, pages 161–174. Wiley Online Library, 2007.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

- 2010** [↑CI-10] R. Booth, Y. Chevaleyre, J. Lang, J. Mengin et C. Sombattheera. Learning conditionally lexicographic preference relations. *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI 2010)*, pages 269–274, 2010.
- [↑CI-9] Y. Chevaleyre, J. Lang, N. Maudet et J. Monnot. Possible winners when new candidates are added : the case of scoring rules. *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI 2010)*, July 2010.
- [↑CI-8] G. Santini, H. Soldano et J. Pothier. Use of ternary similarities in graph based clustering for protein structural family classification. *Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology (2010)*, BCB '10, pages 457–459, New York, NY, USA, 2010. ACM.
- 2009** [↑CI-7] N. Bredeche et Y. Chevaleyre. The robot swarm re-localization problem. *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO 2008)*, pages 1876–1881. IEEE, 2009.
- [↑CI-6] Y. Chevaleyre, J. Lang, N. Maudet et G. Ravilly-Abadie. Compiling the votes of a subelectorate. *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI 2009)*, pages 97–102, 2009.
- 2008** [↑CI-5] Y. Chevaleyre, U. Endriss et N. Maudet. Trajectories of goods in distributed allocation. *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, volume 2, pages 1111–1118, 2008.
- 2007** [↑CI-4] K. Benabdeslem et M. Lebbah. Feature selection for self-organizing map. *Proceedings of the 29th International Conference on Information Technology Interfaces (ITI 2007)*, pages 45–50, 2007.
- [↑CI-3] Y. Chevaleyre, U. Endriss, S. Estivie et N. Maudet. Reaching envy-free states in distributed negotiation settings. *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007)*, pages 1239–1244. AAAI Press, January 2007.
- [↑CI-2] Y. Chevaleyre, U. Endriss, J. Lang et N. Maudet. A short introduction to computational social choice. *Proceedings of the 33rd Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM 2007)*, volume 4362 de *Lecture Notes in Computer Science*, pages 51–69. Springer-Verlag, January 2007.
- [↑CI-1] Y. Chevaleyre, U. Endriss et N. Maudet. Allocating goods on a graph to eliminate envy. *Proceedings of the 22nd AAAI Conference on Artificial Intelligence (AAAI 2007)*, pages 700–705. AAAI Press, July 2007.



Partie III
Acronymes

A³	Apprentissage Artificiel et Applications
AM	allocataire moniteur
ANR	Agence National de la Recherche
AOC	Algorithmes et Optimisation Combinatoire
ATER	attaché temporaire d'enseignement et de recherche
BQR	Bonus Qualité-Recherche (université)
CIFRE	Conventions Industrielles de Formation par la Recherche
CNRS	Centre National de la Recherche Scientifique
EA	Equipe d'Accueil
EFL	Labex <i>Empirical Foundations of Linguistics</i>
FEDER	Fonds Européen de Développement Régional
FUI	Fonds Unique Interministériel
HDR	Habilitation à Diriger des Recherches
LabEx	Laboratoire d'Excellence (appel Grand Emprunt)
IG	Institut Galilée
IUTV	IUT de Villetaneuse
L2TI	Laboratoire de traitement et transport de l'information, université Paris 13 - EA 3043
LAGA	Laboratoire Analyse, Géométrie et Applications, université Paris 13 - UMR 7539
LAMSADE	Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision - UMR 7024
LCR	Logique, Calcul et Raisonnement
LIAFA	Laboratoire d'Informatique Algorithmique : Fondements et Applications
Lim&Bio	Laboratoire d'Informatique Médicale et Bio-Informatique, université Paris 13 - EA
LIPADE	Laboratoire d'Informatique Paris Descartes
LIPN	Laboratoire d'Informatique de Paris-Nord - UMR 7030
MCF	maître de conférences
MENESR	Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche
AOC	Algorithmes et Optimisation Combinatoire
CALIN	Combinatoire, algorithmique et interactions
PPS	Preuves Programmes et Systèmes - UMR 7126
PRES	Pôle de Recherche et d'Enseignement Supérieur
PU	Professeur des Universités
RCLN	Représentation des Connaissances et Langage Naturel
UMR	Unité Mixte de Recherche