

Curriculum Vitae

1. DONNEES PERSONNELLES

Nom : LAMIREL
Prénom : Jean-Charles François
Date et lieu de naissance : 21 Janvier 1960, Nancy (France)
Nationalité : Français
Situation familiale : Marié
Adresse personnelle : 60, rue des Ponts, 54000 NANCY
Téléphone : 33-6-24-36-54-91

Situation actuelle :

Maître de Conférences à l'IUT Robert Schuman département Sciences de l'Information et de la communication (Université de Strasbourg).

Activité de recherche au sein du projet INRIA CORTEX de 1996 à 2007. (Délégué de recherche à l'INRIA de 2001 à 2003).

Activité de recherche actuelle au sein du projet INRIA TALARIS.

Habilitation à Diriger des Recherches en cours de finalisation : soutenance Décembre 2010.

Titre de la thèse :

Application d'une approche symbolico-connexionniste pour la conception d'un système de recherche documentaire hautement interactif avec capacité d'apprentissage en ligne : le modèle NOMAD.

[PH95a] Rapport de recherche LORIA : [CRIN : 95-R-427](#)

Mots-clés : Analyse de données multi-vues, modèles neuronaux non supervisés, modèles multi-topographiques (*), évaluation de la qualité et optimisation du clustering, méthodes symboliques, réseaux bayésiens, fouille de données, règles d'association, visualisation hyperbolique, méthodes de visualisation de données, analyse d'étiquettes, webométrie, recherche documentaire, modèle utilisateur, multimédia, TICE.

Partenaires internationaux : NSC Taiwan, NISTAD (India), NIEHS (USA), WISELAB (China), KU Leuven (Belgique).

Langues :

- Anglais et Allemand : lus, parlés et écrits couramment
- Italien et Espagnol : lus et compris

Activités extra : Président de la Compagnie de danse contemporaine Ormone. Bénévole sur le Festival Musique Action (Scène nationale de Vandœuvre-lès-Nancy) depuis 1983. Photographie (niveau professionnel avec publications). Membre du bureau de l'association France-Palestine Solidarité, branche Est.

Voyages : Visite d'une grande partie des pays du monde, du Groenland aux îles Marquises, à l'exception (pour l'instant) de l'archipel Indonésien et de la Papouasie Nouvelle-Guinée.

2. ACTIVITE DE RECHERCHE : TRAVAUX PRELIMINAIRES

Le mode de communication usuel entre un utilisateur et un système de recherche et d'analyse de documents (SRAD) dans lequel l'utilisateur joue le rôle d'émetteur actif et le système celui de récepteur passif est sans doute à l'origine d'un grand nombre des problèmes posés par les SRAD traditionnels. Ce mode de communication, naturellement peu riche, ne permet pas de tenir correctement compte des diverses imperfections qui touchent aussi bien la définition du besoin de l'utilisateur que la connaissance du SRAD concernant le contenu du fonds documentaire incriminé. Ceci amène à penser que les processus de recherche et d'analyse de documents, pour être améliorés, doivent être envisagés comme des **actes de communication à double sens** entre l'utilisateur et le SRAD, permettant l'enrichissement mutuel des connaissances de l'un comme de l'autre. Cette dernière approche, qui s'inspire fortement de la **théorie systémique de la communication**, impose non seulement d'améliorer l'interactivité entre l'utilisateur et le SRAD, mais également d'équilibrer leurs rôles respectifs lors des opérations de recherche et d'analyse. Elle suggère aussi d'accorder au bruit, apparaissant inévitablement dans la communication entre l'utilisateur et le SRAD, un rôle constructif, aussi bien en phase opérationnelle qu'en phase d'apprentissage pour le SRAD.

D'un autre côté, le développement des réseaux mondiaux d'informations a déjà commencé à entraîner un glissement de la recherche et de l'analyse de documents traditionnelles, opérant sur une base unique, vers des formes fortement distribuées impliquant des fonds multiples. L'avenir de ces derniers types de recherche et d'analyse repose cependant sur la mise en place de nouvelles méthodes d'accès aux informations contenues dans les différents fonds qui, pour être efficaces, devront faire usage de structures fortement parallélisées. Mener à bien de tels types de recherche et d'analyse nécessitera également de pouvoir acquérir dynamiquement des connaissances concernant les différents fonds interrogés, ainsi que de pouvoir harmoniser l'interaction entre ces connaissances. Ceci implique bien évidemment que la connaissance sur un fonds, au cas où elle n'est pas directement disponible, puisse être construite de manière non supervisée et visualisée de manière adéquate. L'avènement de la recherche et de l'analyse distribuées encourage également à penser que la notion d'évolutivité d'un SRAD devra elle-même être remise en cause puisqu'elle devra entre autres prendre en compte la capacité de ce SRAD à intégrer à une échelle de temps proche de celle du temps réel de nouveaux fonds.

L'ensemble de ces constatations amène donc à revoir en profondeur l'architecture SRAD, ainsi que la forme des connaissances qu'ils sont amenés à manipuler. Notre travail de recherche dans ce domaine a débuté dans le cadre de notre fonction d'ingénieur d'étude du Département Recherche et Produits Nouveaux de l'INIST (CNRS). Celui-ci a consisté à mettre au point des outils de conception de réseaux de connaissances pour permettre l'analyse, aussi bien que la consultation associative, du contenu des bases documentaires. Il s'agissait notamment d'appliquer ces outils sur les bases PASCAL et FRANCIS. Les résultats encourageants de ce travail, qui ont été présentés dans le cadre du projet européen KWICK, et qui ont fait l'objet de plusieurs publications, ont suscité l'intérêt du CEDIAG (Département Intelligence Artificielle du groupe BULL), ainsi que celui de l'équipe EXPRIM du CRIN. Ceci nous a permis de démarrer un véritable projet de recherche afin d'approfondir notre approche.

Le modèle NOMAD-MULTISOM dont nous développons l'étude dans notre thèse, est à notre connaissance le premier SRAD à implanter une approche systémique. Il repose pour cela sur une modélisation symbolico-connexionniste inspirée des travaux de Kohonen et de Grossberg, et

utilise comme connaissance de base de multiples cartes topographiques auto-organisatrices (SOM) communiquant entre elles. Chaque carte est constituée d'une structure neuronale de faible dimensionnalité. Elle fournit un point de vue synthétique particulier sur le contenu de la base documentaire ainsi qu'une interprétation de la distance entre les documents de la base selon ce point de vue. Elle sert aussi bien d'outil de raisonnement au système que d'outil de navigation à l'utilisateur. Le modèle proposé gère l'interaction avec l'utilisateur (requêtes basées sur des mots-clés, bouclage de pertinence et/ou interrogation par l'exemple) grâce à un mécanisme de mémoire de session original basé sur la détection de nouveauté. Ce mécanisme permet à la fois de synthétiser le besoin de l'utilisateur, de caractériser le type de ce besoin (précis, thématique, exploratoire ou connotatif), de mettre en évidence de nouvelles alternatives de recherche, et d'évaluer la cohérence des décisions prises par l'utilisateur. Un apprentissage a posteriori permet la modification dynamique de la position des documents sur les cartes ainsi que la création ou la modification des liens entre ces documents. Il permet également de moduler l'effet des différents points de vue. Cet apprentissage est conditionné par le niveau de compétence estimé de l'utilisateur. Une gestion générale des stratégies de recherche est également assurée par le modèle. La stratégie de recherche peut changer en cours de session en fonction des résultats obtenus et du comportement de l'utilisateur. Une implémentation du modèle NOMAD-MULTISOM a été réalisée dans le cadre de la thèse. Ce travail a fait l'objet de plusieurs publications internationales. Le résumé de ce travail est décrit dans différents journaux internationaux [IC94a][IC94b][IC94d]. L'approche originale NOMAD-MultiSOM a été appliquée avec succès pour la mise en place d'interfaces de navigation intelligentes dans les collections muséologiques multimédia [IC00a][IC00e][IC01a][IC01b]. De même, pour l'analyse intelligente des bases de références bibliographiques et des bases de brevets [IJ01a][IC01c][IC01e].

3. CHAMPS DE RECHERCHE ACTUELS

3.1 Orientation principale

L'utilisation de modèles neuronaux présente l'avantage de procurer un parallélisme implicite au processus de recherche et d'analyse de documents. En suivant la même démarche que celle adoptée dans notre travail de thèse, nous continuons à explorer les possibilités d'intégration de tels modèles, spécialisés dans des tâches spécifiques, dans le processus de recherche et d'analyse de documents. Cette démarche implique également de développer la cohabitation entre le raisonnement neuronal et le raisonnement symbolique, ou entre des modèles de nature différente, de manière à couvrir l'ensemble des fonctions de la recherche et de l'analyse de l'information et à éliminer, sinon à réduire, les défauts inhérents à chacun des types d'approche.

Notre direction principale de recherche concerne cependant l'extension de l'approche systémique, initialement proposée dans le SRAD NOMAD, pour mettre en place un nouveau paradigme général d'analyse de données basé sur les points de vue multiples, paradigme que nous avons baptisé paradigme MVDA [PH06b]. Ce paradigme couvre à la fois le domaine de l'analyse de données et celui de la fouille de données. Selon ce paradigme, chaque analyse de données est considérée comme une vue différente sur les données. Le croisement entre les vues s'opère par l'intermédiaire d'un réseau bayésien construit, de manière non supervisée, à partir des données ou des propriétés partagées entre ces dernières. Le paradigme MDVA repose également sur l'exploitation de méthodes spécifiques de visualisation, comme la visualisation topographique ou la visualisation hyperbolique. La mise en place de nouveaux estimateurs de qualité de type

Rappel/Précision non supervisés basés sur l'analyse de la distribution des propriétés associées aux classes, et qui à la fois sont indépendants des méthodes de classification et des changements relatifs à leur mode opératoire (initialisation, distances utilisées ...), nous a permis de démontrer objectivement la supériorité de ce paradigme par rapport à l'approche globale, classique en analyse de données [IJ04a][IC03d]. Elle nous a également permis de comparer et d'intégrer dans le paradigme MVDA des méthodes de classification non supervisées (clustering) qui sont plus particulièrement adaptées à la gestion des données ultra-éparses et fortement multidimensionnelles, à l'image des données documentaires, ainsi que d'optimiser le mode opératoire de telles méthodes [IC04c][IC06h][IC06j]. Finalement, elle a servi de base au développement de méthodes d'extraction de règles d'association susceptibles de s'affranchir de la complexité de calcul et des problèmes de sélection d'information inhérents aux approches symboliques usuelles [IC05d][IC05e][IC06k]. Notre travail se concentre également sur le développement et sur la généralisation des mécanismes coeur du paradigme MVDA, à savoir, le mécanisme de communication entre les vues, le mécanisme de comparaison dynamique de vues et le mécanisme de synthèse en ligne de vues [IC04b][IC05c]. Il vise à illustrer comment ces mécanismes peuvent être employés pour contrôler les interactions, aussi bien locales que globales, entre les vues [IC04d], dans le but de mener à bien des tâches allant de l'inférence entre des propriétés de différentes natures, propres aux données d'analyse, à la comparaison de différents modèles de clustering menées sur ces mêmes données [IJ06a]. Un autre des buts importants de notre travail est celui de résoudre les défauts des méthodes classiques de visualisation, comme la surcharge cognitive produite par ces méthodes lorsqu'elle sont appliquées des données complexes, en développant des approches alternatives, comme celle basée sur les arbres hyperboliques [IC06e] permettant de visualiser l'interaction entre les données multidimensionnelles à différentes échelles et sans surcharge cognitive. La combinaison de telles approches globales avec des approches locales à base de modèle d'équilibre de forces, tel que les modèles de Spring, est également considérée [IC08d].

Une version préliminaire de l'approche MVDA a été appliquée avec succès à l'analyse élaborée et non supervisée des brevets d'invention [IC03c]. Cette approche a également été choisie comme l'une des deux approches de référence du projet européen IST-EISCTES [TR03a][TR03b], centré autour de l'analyse du Web. Une de ses versions plus récente a également ouvert de nouvelles perspectives pour l'analyse automatique de liens dans le domaine de la webométrie [IJ04b][IC04b][IJ06b][IP06a][IC06f], et dans le cadre plus général de l'analyse automatique de l'évolution de la recherche en scientométrie [IC10].

3.2 Travaux complémentaires

Etudes des interactions symboliques-numériques : Les limitations des méthodes de classification numériques sont liées aux erreurs d'interprétation qu'elles peuvent facilement produire quand elles sont employées sans précautions préliminaires, en particulier par des non spécialistes, pour mener à bien des analyses détaillées d'un domaine. De leur côté, les méthodes de classification symboliques présentent l'inconvénient de fournir des résultats souvent peu exploitables, car trop peu synthétiques, voire même de ne pouvoir surmonter les problèmes de complexité de calcul. Nous avons montré qu'il était possible d'établir une synergie entre ces deux types de méthodes en définissant des équivalences entre leurs résultats [IC03b][IC02a][IC01d], de même qu'en définissant des protocoles d'échange entre les classificateurs numériques et les classificateurs symboliques dans le cadre de modèles multi-agents [PH05a][IC03b][IC02a][IC01d]. Nous nous concentrons également sur l'élaboration des

méthodes symbolico-numériques originales d'évaluation/prédiction basées sur la projection et sur l'analyse d'étiquettes, en partant de l'hypothèse que ces dernières peuvent représenter à la fois des propriétés endogènes et des propriétés exogènes aux données. Nous avons récemment prouvé que ces nouvelles méthodes permettaient de faciliter l'interprétation des résultats, ainsi que d'optimiser l'apprentissage, des classificateurs numériques [IC06b][IC06c][IC08c]. Elles ont été expérimentées avec succès pour la prédiction de gisements miniers [PH05c]. Nous en réalisons actuellement de nouvelles expérimentations dans le domaine de la scientométrie, de manière à améliorer les résultats produits par les modèles de réseaux sociaux (SNA) [IP06b].

Filtrage intelligent d'information : Nous avons développé une nouvelle approche pour le filtrage des données documentaires basée sur le principe de détection de nouveauté. L'objectif principal de la détection de nouveauté est de souligner la nouveauté apparaissant dans des documents encore inconnus, en exploitant la connaissance extraite d'un corpus de documents de référence. L'idée fondamentale est donc d'apprendre un modèle des documents disponibles et de l'employer pour identifier des documents entrants différents (nouveauté) [IC05b][NC05a]. Dans le contexte de filtrage, le principe de détection de nouveauté est principalement appliqué d'une manière inversée, c.-à-d. que ce sont les documents qui sont semblables à un modèle appris par l'intermédiaire d'exemples positifs du besoin de l'utilisateur qui seront choisis. Le filtre spécifique de détecteur de nouveauté (IloNDF) que nous avons développé [IC06a][IC06d] est inspiré d'un modèle antérieur de détecteur de nouveauté proposé par Kohonen et de sa première adaptation dans le modèle de NOMAD-MultiSOM [IC94d]. Nos expériences récentes ont montré que ce nouveau modèle peut surpasser des approches de référence, comme celle basée sur les SVM, dans le cadre des tâches de filtrage et de catégorisation multi-classes, ou encore dans le cadre de l'apprentissage à partir d'exemples seuls, également nommé apprentissage à une seule classe. Par ailleurs, dans les tâches de filtrage proprement dites, l'un des avantages déterminants de notre approche est sa capacité d'apprentissage exacte du profil de l'utilisateur et d'évaluation, d'une manière conjointe, de la cohérence de son comportement, pendant son interaction avec le système [IC07a]. Grâce à son principe d'analyse en ligne, le filtre de nouveauté permet également de dépister les changements des centres d'intérêts de celui-ci au cours du temps.

Des expériences exhaustives ont été menées avec cette approche sur les collections-test de références telles que Reuters ou Oshumed (Medline). Elle a également été testée avec succès pour la distribution intelligente de sites Web dans le cadre du projet européen Sat-At-Once [TR05a][TR05b][TR05c]. De nouvelles expériences concernant sa combinaison avec l'approche de filtrage collaboratif sont actuellement étudiées.

Étiquetage des résultats de clustering : La mise au point de techniques d'étiquetage intelligent est nécessaire pour obtenir une vue d'ensemble des sujets principaux extraits par les méthodes de clustering, en particulier si celles-ci produisent originellement leurs résultats dans un espace fortement multidimensionnel. Ces techniques doivent permettre de fournir à la fois des informations synthétiques sur la nature des sujets couverts par chaque cluster, mais également de discriminer de manière optimale les informations fournies par les différents clusters. Dans un cadre de synthèses plus générales, elles doivent également pouvoir s'appliquer à l'étiquetage de regroupements hiérarchiques de clusters. Nous avons également récemment montré que les techniques d'étiquetage jouent un rôle déterminant dans l'analyse diachronique, ou encore dans l'analyse multi-source basées sur le clustering [IC09a] [IC10b]. Jusqu'ici, le développement de telles techniques demeurait un problème ouvert. Nous avons récemment proposé des méthodes

d'étiquetage originales basées sur l'utilisation des critères d'évaluation du contenu de clusters, comme la F-mesure probabiliste sur des propriétés des données associées aux clusters. Nous avons montré que ces méthodes, qui se rapprochent des méthodes de recherche de maximum de vraisemblance (EM), surpassaient, en terme de pertinence, les techniques d'étiquetage élémentaires, comme celles basées sur la fréquence des propriétés des données associées aux clusters, ou la dominance des composantes de leur profil, aussi bien que les méthodes statistiques usuelles, comme la méthode du χ^2 [IC08b]. Dans ce même cadre, nous avons également démontré l'efficacité de ces techniques pour l'étiquetage de structures complexes, telles que les arbres hyperboliques de clusters.

Détection de structures implicites des données par le clustering : Les réseaux neuronaux à apprentissage compétitif préservant la topologie représentent de puissants outils pour la visualisation des interactions entre les données basée sur leurs relations implicites. Dans de tels modèles de réseaux, ce type d'approche exige cependant la prise en compte des interactions entre les neurones, et non plus uniquement la considération individuelle de ces derniers en tant que clusters, ou groupes, de données. Nous étudions actuellement comment les connexions entre des neurones qu'il est possible de produire dans ces réseaux - particulièrement celle obtenues par un apprentissage Hebbien compétitif - peuvent être exploitées pour matérialiser différents niveaux d'abstraction et de structuration implicites dans l'organisation des données d'analyse. Le principe de nos méthodes est celui de définir des pôles attracteurs en calculant des scores de centralité basés sur le nombre et la stabilité des connexions originellement produites par l'apprentissage neuronal. Les pôles définissent eux-mêmes des groupes de clusters et seules les connexions entre les groupes sont conservées lors de la synthèse. Les expérimentations que nous avons menées sur l'analyse de distributions artificielles, telles que des mixtures de gaussiennes, ou d'autres distributions types, ont prouvé l'excellent pouvoir de synthèse et de discrimination de ces méthodes, en comparaison avec des méthodes de référence, comme celles basées sur les voisinages de Voronoi [IC08a].

Evaluation de la qualité du clustering : les méthodes d'analyse de la qualité du clustering sont généralement utilisées pour déterminer le nombre optimal de clusters dans un ensemble de données. A une échelle plus fine, elles peuvent également permettre de distinguer les clusters pertinents des clusters non pertinents. En dépit de leur succès, l'efficacité des indices basés sur la distance ou sur les moindres carrés diminue très rapidement lorsque les données à traiter deviennent fortement multidimensionnelles ou hétérogènes. Nous avons donc mis en place de nouveaux indices basés sur l'identification de formes-cœur, qui représentent elles-mêmes des groupes de propriétés fortement corrélés dans les données rattachées à chacun des clusters identifiés par une méthode de clustering. Cette technique se ramène à la caractérisation de sous-espace de description « locaux » dans un espace de description global de grande dimension. Des valeurs d'inertie basées sur le pourcentage de formes-cœurs dans chaque cluster (intra) et sur le recouvrement de formes-cœurs entre les clusters (inter) permettent à la fois d'estimer précisément la qualité locale de chaque cluster et d'évaluer la qualité globale d'un modèle de clustering. Des expériences étendues sur la collection Reuters-21578 nous ont permis de prouver l'efficacité de cette technique sur des données complexes, données sur lesquelles les méthodes classiques basées sur la distance, comme la méthode de Calinsky-Harabasz ou la méthode de Davies-Bouldin, s'avèrent totalement inopérantes [IC08b].

Dans ce dernier cadre, nous étudions également le comportement des méthodes d'analyse de la qualité du clustering pour discriminer les résultats de clustering homogènes des résultats de

clustering hétérogènes. Notre analyse a montré qu'aucun des indices existants, y compris ceux que nous avons définis jusqu'ici, ne permettaient de discriminer entre des résultats de clustering homogènes et des résultats hétérogènes. Leur défaut principal est donc celui de n'être pas assez sensibles à la présence d'un petit nombre de clusters hétérogènes de grande taille, particulièrement dans le cas de l'existence conjointe d'un grand nombre de clusters de petite taille, dans les résultats de clustering [IC10a]. Pour corriger cela, nous avons récemment proposé d'adapter notre approche basée sur les indices de Rappel/Précision non supervisés en définissant de nouveaux indices de Micro-Précision et de Micro-Rappel [IC10d] [IC10g], ainsi que des indices de Micro-F-mesure cumulée [IC10c], moyennés sur les propriétés des données associées aux clusters. Nos dernières expériences menées sur des collections test de notices bibliographiques de l'INIST nous ont permis de montrer que ces nouvelles mesures pouvaient effectivement être exploitées pour identifier de manière sûre les résultats de clustering hétérogènes [IC10e][IC10i][IC10j].

Clustering incrémental : le problème du clustering incrémental est un problème complexe pour lequel peu de solutions convaincantes ont été proposées jusqu'à aujourd'hui. Il s'agit en effet de pouvoir détecter de nouveaux sujets au fil de l'arrivée des nouvelles données, ceci avec une réactivité suffisamment bonne, ce qui revient à trouver un compromis optimal entre la plasticité et la stabilité du modèle de clustering lors de l'apprentissage. Dans ce but, nous étudions actuellement l'adaptation conjointe des solutions que nous avons proposées pour l'optimisation de la classification non supervisée, pour l'analyse multi-vues, pour la détection de nouveauté, et pour la validation fine du clustering [IJ07b][IC08a][IC08b]. Une première voie originale que nous explorons est celle qui consiste à ramener la comparaison entre données temporelles présentes à différents pas de temps à un problème de comparaison entre des vues constituées par des analyses de données associées à des macro-étapes de temps prédéfinies. Cette approche se base sur une adaptation du paradigme MVDA au raisonnement sur des groupes de données différenciées [IC09a][IC10b]. Elle produit d'ors et déjà d'excellents résultats, très supérieurs à ceux de l'état de l'art, et notamment à ceux fournis par les règles floues, ou ceux obtenus par la comparaison visuelle humaine des résultats de clustering, ceci sur des corpus de référence pour l'analyse diachronique comme le corpus "Optoélectronique" du projet européen PROMTECH [IC10]. Une seconde voie que nous explorons en parallèle est celle de l'adaptation d'algorithmes neuronaux existants, comme l'algorithme Growing Neural Gas, à un fonctionnement incrémental. Nous avons, pour cela, défini de nouvelles règles d'apprentissage locales qui remplacent les règles d'apprentissage globales usuellement employées dans ces méthodes. Ces règles ne reposant plus sur l'exploitation des distances classiques, celle sur des techniques de maximisation de la qualité de l'étiquetage des classes, elles mêmes basées sur le principe de maximisation de vraisemblance (EM) que nous avons préalablement exploité pour la caractérisation des étiquettes de classes. Nous avons notamment mis au point une adaptation de l'algorithme IGNG fonctionnant selon cette approche [IC10f] [IC10j][IC10k][IC10m]. Nous avons déjà pu montrer que ce nouvel algorithme fournissait des résultats supérieurs à tous ceux des algorithmes existants dans le cas du clustering de données hétérogènes. Ces travaux font l'objet d'une collaboration entre plusieurs équipes de recherche du LORIA et avec l'INIST (Projet CPER, démarré en 2007, et renouvelé pour 2 ans en 2009).

Wikis sémantiques évolutifs : Nous participons au projet WICRI dont le but est de développer le concept novateur de réseau auto-organisateur de wikis sémantiques. Notre rôle spécifique dans ce projet est celui de proposer des stratégies automatisées d'enrichissement du réseau de wikis par

l'intermédiaire de l'exploitation de données du Web. En effet, la fouille du Web représente un challenge important pour améliorer à la fois la réactivité, la flexibilité et la portée d'un tel type de réseau. D'un côté, ce processus est nécessaire pour assister les contributeurs potentiels lors de la phase de construction du réseau en leur fournissant des règles rédactionnelles fiables. D'un autre côté, il s'avère également décisif pour fournir aux utilisateurs finaux du réseau de l'information externe à celui-ci, avec comme valeur ajoutée de pouvoir la rapporter au contexte sémantique du réseau. Bien que le projet WICRI soit encore dans sa phase de démarrage, notre prototype de réseau de wikis est déjà opérationnel et peut ainsi être exploité comme une plate-forme de recherche et d'investigation collaborative disponible en ligne [IC10h].

4. RESPONSABILITES

4.1 Responsabilités scientifiques et institutionnelles internationales

2007-__ : Membre du comité de rédaction du nouveau journal international "**Collnet Journal of Scientometrics and Information Management**", Taru publications, New Delhi, Inde.

2005-__ : Membre de l'Ecole Doctorale Algéro-Française (EDAF). Cette école a pour vocation de former des formateurs algériens, ainsi que de monter des encadrements de recherche en co-tutelle d'étudiants algériens, dans les domaines des sciences du langage, de la littérature, de la didactique des TICE et de la fouille de textes.

2004-__ : Membre permanent du réseau de collaboration scientifique COLLNET. Ce réseau est un réseau interdisciplinaire de recherches visant à l'étude exhaustive des différents aspects des échanges et des collaborations dans les champs des sciences et de la technologie. Il dirige plus particulièrement des actions internationales de recherche dans le domaine de la scientométrie et dans celui de la webométrie. Organisateur de conférence internationale pour le réseau [CO06a].

2004-2007 : Membre du comité scientifique d'ARTIST. Ce réseau de recherche est spécialisé dans l'étude de l'appropriation par les chercheurs des nouvelles technologies de l'information scientifique et technique.

2001-2003 : Membre permanent du consortium européen DELOS : Réseau européen d'excellence dans le domaine des bibliothèques numériques. Ce consortium agit en tant que comité scientifique européen. Il dirige la plus grande partie des actions scientifiques européennes dans ce domaine, en particulier l'organisation des conférences et les ateliers thématiques, mais aussi les collaborations inter-universitaires.

4.2 Responsabilités institutionnelles nationales

2001-2008 : Membre titulaire de la commission de spécialistes UHP (Nancy) relative à la section CNU n° 71 : sciences de l'information.

2004-2008 : Membre suppléant de la commission de spécialistes UHP (Nancy) relative aux sections CNU n° 7, 9, 11, 12, 70, 71 : sciences de l'information, linguistique et sciences de l'éducation.

2001-2008 : Membre suppléant de la commission de spécialistes de l'université de Robert Schuman relative à la section CNU n° 71 : sciences de l'information.

4.3 Thèses présentées

1. [PH09a] Co-directeur de la thèse de **Randa Kassab**,
(taux d'encadrement : 90 %).
Titre : *Analyse des propriétés stationnaires et émergentes dans les flux d'information variant au cours du temps : application au filtrage et à l'analyse de l'information en provenance du Web.*
Soutenue le 11 Mai 2009, LORIA, Nancy, France.
Ce travail a mené à 1 publication de niveau national [NC05a], 5 publications de niveau international [IC05b][IC06d][IC07a][IC08a][IC08c], 3 rapports techniques de projet [TR05a][TR05b][TR05c] et 1 publication dans un journal international.
2. [PH06c] Participation à la direction de la thèse de **Mohammed Attik**,
(taux d'encadrement : 50 %).
Titre : *Traitement intelligent des données par réseaux de neurones artificiels.*
Soutenue le 13 Décembre 2006, LORIA, Nancy, France.
Ce travail a mené à 3 publications de niveau international [IC06b] [IC06c] [IC08b].
3. [PH06b] Co-directeur et membre du jury de la thèse de **Shadi Al Shehabi**,
(taux d'encadrement : 100 %).
Titre : *Modèles neuronaux topographiques pour l'analyse documentaire multimédia : application à l'analyse du Web.*
Soutenue le 28 Juin 2006, LORIA, Nancy, France.
Ce travail a mené à 18 publications de niveau international, 4 publications dans des journaux internationaux, et 3 rapports techniques de projet.
4. [PH06a] Co-directeur et membre du jury de la thèse de **Hager Kammoun**,
(taux d'encadrement : 100 %).
Titre : *Coopération de classifieurs pour la recherche d'information : le modèle SARCI.*
Soutenue le 20 Mars 2006, RIADI, Tunis, Tunisia.
Ce travail a mené à 1 publication [NC00b] de niveau national et à 6 publications de niveau international [IC00d][IC00e][IC01c][IC02c][IC05f][IC05g].
5. [PH01b] Participation à la direction et membre du jury de la thèse de **Mercedes Martinez**,
(taux d'encadrement : 30 %).
Titre : *Principles of Dynamic Management of the Relations between Documents in the Digital Libraries: Application to the Legal Domain.*
Soutenue le 24 April 2001, University of Valladolid, Spain.
Ce travail a mené à 1 publication de niveau national [NC00b] et à 1 publication de niveau international [IC00f].
6. [PH01a] Participation à la direction de la thèse de **Emmanuel Nauer**,
(taux d'encadrement : 30 %).
Titre : *Principes de conception de systèmes hypertextes pour la fouille de données multibases,*
Soutenue le 5 Janvier 2001, LORIA, Nancy, France.
Ce travail a mené à 1 publication de niveau national [NC95a] et à 2 publications de niveau international [IC97a][IC97b].

4.4 Thèses en cours

Participation à la direction de la thèse de **Mahmoud Zennaki**,
Sujet : **Image Classification using Semi-supervised Classification Methods**.

Ce travail a mené à 2 publications de niveau international [IC06i] [IC07e].

Co-encadrant français de la thèse de **Ghada Safi** menée à bien à **l'université d'Alep en Syrie**.
Sujet : **Extended communication and generalization models for the MVDA paradigm**.

Ce travail a mené à 2 publications de niveau international [IC10i][IC10].

Soutenance prévue : Janvier 2011.

Encadrant français dans le cadre de l'EDAF, de la thèse en co-tutelle de **Abelgammi Ammi**.

Sujet : **Analyse de l'usage des TICE dans l'enseignement supérieur en Algérie**.

(Débutée en 2009).

Rem : Encadrant et membre du jury du mémoire de Master du même étudiant, soutenu le 29 Mai 2008 à Médéa, Algérie.

Sujet : **Etude de l'acquisition d'un lexique spécifique par les TICE**.

Encadrant français dans le cadre de l'EDAF, de la thèse en co-tutelle de **Ghania Boudjelal**.

Sujet : **L'apport de l'enseignement des stratégies cognitives et des stratégies métacognitives dans l'apprentissage de la compréhension de l'écrit**.

(Débutée en 2009).

4.5 Stages de fin d'étude d'ingénieur, de Master et de recherche

2010

[Ma10a] Encadrement du stage de fin d'étude d'ingénieur ITT (Kampur, India) de **Navesh Priyankar**. Durée **3 mois** (en Internship INRIA).

Sujet : **Etude d'une approche d'analyse diachronique basée sur la combinaison de méthodes de clustering et de méthodes d'étiquetage des clusters dans le contexte du paradigme multi-vues MVDA**.

Ce travail a mené à 1 publication de niveau international [IC10].

[Ma10b] Encadrement du stage de fin d'étude d'ingénieur ITT (Hyderabad, India) de **Raghendra Mall**. Durée **5 mois** (en Internship INRIA).

Sujet : **Développement d'algorithmes neuronaux incrémentaux (2)**.

En cours.

2009

[Ma09c] Encadrement du stage de fin d'étude ENSI-TUNIS de Zied Boulila. Durée **5 mois**.

Sujet : **Développement d'algorithmes neuronaux incrémentaux (1)**.

Ce travail a mené à 1 publication de niveau national [IC10k] et a 1 publication de niveau international [IC10f].

[Ma09b] Encadrement du stage de Master de **Charif Haydar**. Durée **5 mois**.

Sujet : **Analyse multi-vues de l'évolution des thèmes de recherche exploitant la décomposition par pas de temps.**

Ce travail a mené à 1 publication de niveau international [IC09a].

[Ma09a] Co-encadrement du stage de Master de **Maha Ghribi**. Durée **5 mois**.

Sujet : **Comparaison des résultats des méthodes de clustering sur des corpus textuel de grande taille.**

2008

[Ma08a] Encadrant français dans le cadre de l'EDAF, du stage de master de **Abelgammi Ammi**.

Sujet : **Etude de l'acquisition d'un lexique informatique par l'usage des TICE en Algérie.**

2007

[Ma07b] Encadrement du stage de fin d'étude d'ingénieur IFI (Hanoï, Vietnam) de **Ta Anh Phuong**. Durée **7 mois** (en Internship INRIA).

Sujet : **Méthode avancées de visualisation et de sélection d'information pour la fouille et l'analyse de données.**

Ce travail a mené à 2 publications de niveau international [IC08d] [IC08b] et à 1 soumission dans un journal international [IJ09a].

[Ma07a] Co-encadrement du stage de fin d'étude IFI (Hanoï, Vietnam) de **Do Min Chau**. Durée **6 mois** (en Internship INRIA).

Sujet : **Collaboration entre filtrage basé sur le contenu et filtrage collaboratif pour l'analyse des usages.**

2006

[Re06a] Encadrement du stage de recherche doctorale de **Mahmoud Zennaki**. Durée **1 mois**.

Sujet : **Classification semi-supervisée d'images par machine à vecteurs supports.**

Voir la **section 4.4**, pour les résultats obtenus.

[Ma06a] Encadrement du stage de fin d'étude ENSEM (+ DEA) de **Nicolas Carrez**. Durée **5 mois**.

Sujet : **Extraction de connaissances à partir de modèle neuronaux non supervisés.**

Ce travail a mené à 1 publication de niveau international [IC06k].

2005

[Ma05a] Encadrement du stage de fin d'étude ISIMA de **Ronan Thomas**. Durée **3 mois**.

Sujet : **Mise en place d'une plate-forme générique d'expérimentation et d'évaluation des méthodes de clustering neuronal.**

2002-2003

[Ma02a] Encadrement du stage de fin d'étude ingénieur CNAM de **Martial Hoffmann**. Durée **18 mois**.

Sujet : **Mise en place opérationnelle de l'approche d'analyse de données multi-vues NOMAD-MultiSOM. Expérimentation sur les données Web institutionnelles.**

Travail contractuel dans le cadre du projet européen IST – EICSTES.

Ce travail a mené à 1 publication de niveau national, 2 publications de niveau international [IC03d][IC04a], 2 rapports techniques de projet [TR03a][TR03b] et 2 publications dans des journaux internationaux [IJ04a][IJ04b].

2000

[Ma00a] Co-encadrement du stage de DEA de **Matthieu d'Acquin**. Durée **6 mois**.

Sujet : **Combinaison de méthodes numériques et de méthodes symboliques pour l'analyse des données documentaires et multimédia**.

Ce travail a mené à 4 publications de niveau international [IC00g][IC01d][IC02a][IC03b].

1999

[Ma99a] Encadrement du stage de fin d'étude ingénieur CNAM de **Gérald Oster**. Durée **18 mois**.

Sujet : **Application de l'approche NOMAD à la découverte d'information dans les bases de données iconographiques**.

Ce travail a mené à 1 publication de niveau international [IC06k].

4.6 Gestion de projets

2010-_- : Direction scientifique, dans le cadre du projet QUAERO, de la tâche de gestion intelligente des données brevets. Le but de cette tâche est celui de mettre en place une plateforme permettant d'assister de manière efficace les ingénieurs experts en brevets dans le processus de contrôle de validation des dépôts. Atteindre un tel but implique de développer de nouvelles méthodes d'apprentissage semi-supervisées, en de modifier en profondeur les méthodes existantes pour permettre l'identification des relations pertinentes entre les classifications hiérarchiques de brevets et les références bibliographiques matérialisant les travaux de recherche se rapportant, directement ou indirectement, à ces différentes classes.

2007 -_- : Coordinateur côté français du projet de recherche international commun INRIA-NSC Taiwan IGAP-AREX. Le but du projet IGAP-AREX est de proposer un cadre fédérateur pour le traitement des données fournies par les puces à ADN. Il est basé sur le partage de compétences entre deux équipes de recherche et sur l'intégration des outils existants et des méthodologies originales dans le cadre d'un nouveau paradigme d'analyse de données, à savoir celui de l'analyse de données de multi-vues (MVDA). Un des objectifs à long terme du projet IGAP-AREX est de produire une plateforme intégrée pour l'annotation intelligente des gènes.

2007 -_- : Expert scientifique en matière de classification non supervisée dans le projet CPER-TALC classification incrémentale. Ce projet représente un projet innovant dans le domaine de l'analyse de données. Il vise à définir des méthodes de détection des sujets émergents apparaissant dans un flux continu de données documentaires. Beaucoup d'applications de portée internationale sont envisageables dans des domaines divers tels que l'évaluation des sciences, l'analyse du Web et la veille technologique.

2004-2005 : Expert scientifique en matière de filtrage d'information basé sur le contenu dans le cadre du projet Sat-At-Once, commandité par l'agence européenne des satellites (ESA). Ce projet avait pour but la mise en place d'un système de distribution personnalisée de sites Web. Notre contribution a concerné la conception d'un module de filtrage multilingue basé sur le contenu. Le projet a mené à un système de filtrage en ligne opérationnel.

2000-2004 : Expert scientifique en matière d'analyse de l'information et de visualisation dans le cadre du projet Européen IST-EISCTES. Ce projet se rattachait à l'installation d'outils et d'indicateurs de l'analyse de l'information institutionnelle en provenance du Web. L'approche NOMAD-MultiSOM y avait été choisie comme l'une de deux approches de référence du projet. Le projet a servi de base pour généraliser l'approche NOMAD-MultiSOM, et définir ainsi l'approche MVDA. Cette dernière a été alors mise en œuvre pour réaliser des expériences inédites de combinaison d'analyse de contenu et d'analyses de lien sur les grands ensembles de données Web.

2000-2003 : Directeur scientifique et administratif, pour le compte de l'INRIA, du projet d'Européen IST-SCHOLNET. Ce projet a visé à la conception d'une bibliothèque numérique universitaire distribuée qui devrait inclure de nombreuses fonctionnalités, parmi lesquelles : filtrage d'information, personnalisation, gestion du multilinguisme et gestion de l'information multimédia [TR01a]. Il a mené à une bibliothèque numérique en ligne distribuée opérationnelle.

1996 : Conseiller scientifique pour le projet SEDIME qui se rattachait à la réalisation d'un service Web européen pour la consultation des images médicales. Gestion, pour l'INED, du transfert des données bibliographiques du format UNIMARC vers le format GRISELI (SGML). Ce projet avait été commandité par le Ministère de l'Education Supérieure et de la Recherche. Gestion, pour la compagnie AEROSPATIALE, de la mise en place d'un système d'accès interactif aux images satellites, dans le cadre du projet COSME.

1993 : Directeur scientifique et administratif, pour le compte du département BULL-CEDIAG, du projet européen FAST (Financial Assistant for Servicing and Training). Le but de ce projet était de définir un environnement multimédia flexible adapté à la création de nouveaux services bancaires. Notre rôle dans le projet était de réaliser une étude technique des divers environnements hypermédia déjà disponibles sur le marché afin de définir l'environnement idéal pour le projet.

1992 : D'abord pour l'INIST, ensuite pour le département BULL-CEDIAG, nous avons assuré la responsabilité scientifique du projet européen KWICK (Knowledge Workers Intelligently Creating Collecting Consulting Knowledge) dont le but était créer un environnement multimédia destiné aux analystes de l'information. Ce projet s'est terminé avec succès. Il a permis de produire la plateforme d'infométrie SDOC [IC91a][IC91b][IC91c][IC93a], qui est encore opérationnelle à ce jour dans le cadre du portail INIST-STANALYST.

4.7 Propositions de projets en cours de validation

2009 : Le projet COLDLib vise à construire un espace numérique fédéré pour le réseau interdisciplinaire de recherches "collaboration en Science et en technologie" (COLLNET), afin de rassembler la connaissance existante relative au domaine des aspects quantitatifs de la science, et de favoriser ainsi la collaboration et la communication des chercheurs des pays émergents en analyse de la science et technologie et en politique de la science, par la combinaison et l'intégration d'approches qualitatives et quantitatives. Le but principal du projet est de à ces chercheurs fournir un accès global, efficace et fiable aux données de ce domaine de recherches, tout en accroissant l'accès à la connaissance à valeur ajoutée, plutôt que de leur laisser conserver une vue isolée et locale sur les données et les résultats de recherche. La proposition de projet est actuellement en cours d'achèvement.

2008 : Les administrations et les organismes publics échangent quotidiennement un grand volume d'informations écrites avec leurs administrés, leurs fournisseurs, d'autres administrations et d'autres partenaires (associations...). Les informations entrantes de ces organismes sont organisées en dossiers. Un dossier est un ensemble de documents différents mais cohérents et complémentaires qui permettent d'exprimer une "demande". Le but ambitieux du projet STRADA est de réaliser une nouvelle génération de système automatique et adaptatif de constitution et de traitement de dossiers consolidés. Ce système raisonne sur des informations multi-variées extraites par différentes techniques statistiques et linguistiques, et rendues cohérentes par des méthodes de clustering, en vue de produire des dossiers de synthèse personnalisés. Les dossiers finalisés sont ensuite obtenus par consolidation à partir d'informations métier.

2006 : Expert scientifique en matière de méthodes de classification non supervisées et d'analyse de données le projet national LIRENET (ANR). Ce projet vise à installer un environnement ouvert pour l'accès de la connaissance. Il est basé sur la classification dynamique d'information en provenance du Web en utilisant de la connaissance fédérée, comme la connaissance encyclopédique. Ce projet a été déjà accrédité par l'agence nationale française de recherches (ANR). Il est actuellement dans une phase de la redéfinition suite à des problèmes financiers rencontrés par le partenaire principal.

5. ENSEIGNEMENT

1997-2007 : Enseignant au département l'Information-Communication de l'IUT de Strasbourg-Sud, rattaché à l'Université Robert Schuman. Cours assurés : initiation à la communication numérique et multimédia, initiation à la veille technologique et à la recherche documentaire, services et normes du Web, théorie et pratique des bases de données, outils bureautiques et systèmes de gestion des données numériques.

Enseignant en Master-IST à l'UHP (Nancy) : Initiation aux outils et techniques pour la recherche et l'analyse de données documentaires et la veille technologique.

2002-2007 : Enseignant à l'école d'été nigériane des technologies de modélisation et de gestion de l'information (projet quadri-annuel reconduit, impliquant 4 universités Lagos (x2), Ibadan et Ife).

2005-2007 : Enseignement et conférences relatifs au domaine des nouvelles technologies de l'information et de la communication (TICE) et de la fouille de textes dans le cadre de l'EDAF.

2005 : Professeur invité pour un séminaire exceptionnel sur les techniques émergentes dans le domaine de la scientométrie et de la visualisation des données à l'Université Technologique de Dalian (DUT), Chine.

6. DEPOTS DE BREVETS

Le 8/11/2000, la version 1.0 du logiciel NOMAD-MultiSOM a fait l'objet, par l'INRIA, d'un dépôt à l'agence française pour la protection des programmes (APP). Numéro APP : **IDNN.FR.001.460018.00.R.P.2000.000.1000.**

Le 29/10/2004, la version 2.0 du logiciel NOMAD-MultiSOM a fait l'objet, par l'INRIA, d'un nouveau dépôt à l'agence française pour la protection des programmes (APP). Cette version incluait tous les développements réalisés dans le cadre du projet Européen IST-EISCTES. Même numéro APP que la version 1.

7. REFERENCES

7.1 Conférences invitées

[IP07b] Lamirel J.-C.,
The MVDA multi-viewpoint paradigm: a promising approach for automatizing Webometrics and Scientometrics studies,
Managerial Economics, Strategy and Innovation Workshops, Katholiek Universiteit (KU) Leuven, Leuven, Belgia (2007).

[IP07a] Lamirel J.-C.,
The central role of AI-based techniques for solving the new challenges in data management,
NISTAD Silver Jubilee Workshop on Data Gathering, Processing and Measurement of Science and Technology in the 21th Century, New Delhi, India (2007).

[IP06b] Lamirel J.-C.,
Highlighting characteristics of groups of collaboration using a new label analysis technique: a methodology proposal,
International Forum on Science Study and Scientometrics, Dalian, China (2006).

[IP06a] Lamirel J.-C.,
Network analysis using dynamic interaction between multiple sources,
Humboldt University, Berlin, Germany (2006).

[IP05c] Lamirel J.-C.,
Emerging techniques of hyperbolic visualization for Webometrics,
First DUT sessions on Scientometrics, Dalian, China (2005).

[IP05b] Lamirel J.-C.,
Link analysis based on unsupervised numerical knowledge extraction,
First DUT sessions on Scientometrics, Dalian, China (2005).

[IP05a] Lamirel J.-C.,
Evaluation of collaboration between European universities using dynamic interaction between multiple sources,
Proceedings of the 5th International ISSI Conference – 1st International Workshop on Webometrics, Informetrics and Scientometrics and 6th COLLNET meeting, Stockholm, Sweden (2005).

[IP04a] Lamirel J.-C.,
New neural methods for Webometrics,
Proceedings of the 4th International Conference on University Evaluation and Research Evaluation, Wuhan, China (2004).

[IP01a] Lamirel J.-C.,
Systemic approaches for data analysis and information retrieval,
University of Valladolid – Valladolid, Spain (2001).

7.2 Organisation de conférences et de sessions spéciales internationales

[CO10a] *Special Session on Incremental clustering and novelty detection techniques and their application to intelligent analysis of time varying information* in the framework of *IEA/IAE International Conference*,
Syracuse, NY, USA, June 2011.

[CO06a] *International Workshop on Webometrics, Informetrics and Scientometrics & Seventh COLLNET Meeting* in conjunction with the *Extra Session on Information Visualization for Webometrics, Informetrics and Scientometrics*,
Nancy, France, 10-12 May, 2006.

7.3 Organisation d'ateliers de travail internationaux

[WO11a] *Clustering incrémental et méthodes de détection de nouveauté et leur application à l'analyse intelligente d'information évoluant au cours du temps*,
EGC 2011 Workshop, Brest, France, January 2011.

[WO03b] *Economic Intelligence and Web Mining Techniques*,
SCI 2003, Orlando, FL, July 2003.

[WO03a] *EISCTES - Special Session on Information Visualization*,
INRIA, Nancy, France, February 2003.

[WO02a] *Information and Communication Technologies applied to Economic Intelligence (ICTEI'2002)* (co-organizer),
Ibadan University, Ibadan, Nigeria, 20-22 August 2002.

7.4 Comités de programme ou comités scientifiques dans des conférences nationales et internationales

[PC10b] *ICITST-2010: The 5th International Conference for Internet Technology and Secured Transactions*,
November 2010, London, UK.

[PC10a] *International Workshop on Webometrics, Informetrics and Scientometrics and COLLNET meeting (since 2005)*.

[PC02a] *IERA: economic intelligence*,
April 2003, INIST, Nancy, France.

[PC02a] *ICTEI: Information and Communication Technologies applied to Economic Intelligence*,
July 2002, Ibadan, Nigeria.

7.5 Présidence de session dans des conférences internationales

[CC10b] *Technological and Strategic Survey Conference – VSST 2010*,
Toulouse, France, October 2010.

[CC09b] **5st International Workshop on Webometrics, Informetrics and Scientometrics and 10th COLLNET meeting,**
Dalian, China, September 2009.

[CC09a] **Technological, Scientific and Strategic Survey Conference – VSST 2009,**
Nancy, France, March 2009.

[CC08a] **4st International Workshop on Webometrics, Informetrics and Scientometrics and 9th COLLNET meeting,**
Berlin, Germany, July 2008.

[CC07b] **International Conference on Tools with Artificial Intelligence (ICTAI),**
Patras, Grece, October 2007.

[CC07a] **3rd International Workshop on Webometrics, Informetrics and Scientometrics and 8th COLLNET meeting,**
New Delhi, India, March 2007.

[CC06a] **2nd International Workshop on Webometrics, Informetrics and Scientometrics and 7th COLLNET meeting,**
Nancy, France, May 2006.

[CC05a] **5th International ISSI Conference – 1st International Workshop on Webometrics, Informetrics and Scientometrics and 6th COLLNET meeting,**
Stockholm, Sweden, August 2005.

7.6 Journaux internationaux

[IJ09c] Lamirel J.-C., Phuong T.A., Al Shehabi S.,
Combination of hyperbolic visualization and novel labeling strategies for hierarchical representation of scientometrics study results,
Submitted to Journal of Classification, 2009.

[IC07b] Kassab R., Lamirel J.-C.,
Efficiently solving the one-class classification problem with enhanced novelty detection filtering,
Submitted to Machine Learning (under revision).

[IJ07a] Lamirel J.-C.,
Entropy-based methods for label analysis: highlighting relevant topics in large collections of information,
Accepted for publishing in COLLNET Journal of Scientometrics and Information Management, 2007.

[IJ06b] François C., Lamirel J.-C., Al Shehabi S.,
Combining Advanced Visualization and Automated Reasoning for Webometrics: A Test Study,
Journal of Information Management and Scientometrics (JIMS) **1(4)** 2006.

[IJ06a] Al Shehabi S., Lamirel J.-C.,
Evaluation of collaboration between European universities using dynamic interaction between multiple sources,
Journal of Information Management and Scientometrics (JIMS) **1(3)** 2006.

[IJ04b] Lamirel J.-C., Al Shehabi S., Francois C. , Polanco X., Hoffmann M.,
Using a compound approach based on elaborated neural network for Webometrics: an example issued from the EICSTES Project,
Scientometrics, 61(3) 2004.

[IJ04a] Lamirel J.-C., Al Shehabi S., François C., Hoffmann M.,
New classification quality estimators for analysis of documentary information: application to patent analysis and web mapping,
Scientometrics, 60(3) 2004.

[IJ01a] Polanco X., François C., Lamirel J.-C.,
Using artificial neural networks for mapping of science and technology: a MultiSOM approach,
Scientometrics, 51(1) 2001.

7.7 Edition d'actes de conférence

Lamirel J.-C., Kretschmer H., Francois C. and Ducloy J. (Eds.), *Proceedings of the 2nd International Workshop on Webometrics, Informetrics and Scientometrics and 7th COLLNET meeting*,
10-12 May 2006, Nancy, France. Nancy: SRDI-INIST-CNRS.

7.8 Revues de journaux internationaux et de conférences internationales

AICCSA Conference (2003), IJCAR Conference (2004), IJCAI Conference (2004), WISER IntJ (2006), COLLNET IntJ (2006-2009), Neural Networks IntJ (2007-2008), Neural Computing and Applications IntJ (2008-2009), Geographical Information Systems IntJ (2010).

7.9 Chapitres de livres

[BC10b] Lamirel J.-C.
A new multi-Viewpoint and multi-level clustering paradigm for efficient data mining tasks,
INTECH E-Books series, INTECH Open Access Publisher, Vienna, Austria, January 2011.

[BC10a] Lamirel J.-C., Cuxac P., Francois C.
Recherche des évolutions technologiques à partir de bases de données bibliographiques : apport de la classification incrémentale,
In: "Technologies de la Connaissance et Recherche d'Information en Contexte", STI Collection, Hermes Science Publishing Ltd, London, UK, January 2011.

[BC95a] Lamirel J.-C.,
Vers un approche neuromimétique à valeur ajoutée pour les systèmes documentaires,
In: Le neuromimétisme : Epistémologie, neurobiologie, informatique.
"Interdisciplinarité et Nouveaux outils" Collection, Hermes Ed., Paris, 1995.

7.9 Communications dans des conférences internationales avec comité de lecture

[IC10m] Lamirel J.-C, Mall R.,
Variations to incremental growing neural gas algorithm based on label maximization,
5th International ICST Conference on Performance Evaluation Methodologies and Tools, Paris
France, May 2011.

[IC10l] Lamirel J.-C, Safi G., Priyankar N., Cuxac P.,
Mining research topics evolving over time using a diachronic multi-source approach,
Proceedings of ICDM 2010 International Workshop on Mining Multiple Information Sources,
Sydney, Australia, December 2010.

[IC10k] Lamirel J.-C., Boulila Z., Ghribi M., Cuxac P., Francois C.,
***Un nouvel algorithme incrémental de gaz neuronal croissant basé sur des l'étiquetage des
clusters par maximisation de vraisemblance : application au clustering des gros corpus de
données textuelles hétérogènes,***
Proceedings of Technological and Strategic Survey Conference – VSST 2010, Toulouse, France,
October 2010.

[IC10j] Lamirel J.-C., Boulila Z., Ghribi M., Cuxac P., Francois C.,
***A new incremental neural clustering approach for performing reliable large scope
scientometrics analysis,***
Proceedings of 6th International Conference on Webometrics, Informetrics and Scientometrics
and 11th COLLNET Meeting, Mysore, India, October 2010.

[IC10i] Lamirel J.-C, Safi G.,
Use of distance-based indexes might well lead to misinterpretation of clustering quality results,
Proceedings of 1st International Workshop on Validation Statistics, Berlin, Germany, October
2010.

[IC10h] Ducloy J., Daunois T., Foulonneau M., Hermann A., Lamirel J.-C., Sire S.,
Vanoirbeek C.,
Metadata for Wicri, a network of semantic Wikis for communities in research and innovation,
Proceedings of the International Conference on Dublin Core and Metadata Application (DC-
2010), Pittsburgh, USA, October 2010.

[IC10g] Lamirel J.-C., Ghribi M., Cuxac P.,
***Unsupervised recall and precision measures: a step towards new efficient clustering quality
indexes,***
Proceedings of the 19th International Conference on Computational Statistics
(COMPSTAT'2010), Paris, France, August 2010.

[IC10f] Lamirel J.-C., Boulila Z., Ghribi M., Cuxac P.,
***A new incremental growing neural gas algorithm based on clusters labeling maximization:
application to clustering of heterogeneous textual data,***
Proceedings of the 23rd International Conference on Industrial, Engineering & Other Applications
of Applied Intelligent Systems (IEA-AIE 2010), Cordoba, Spain, June 2010.

[IC10e] Cuxac P., Lamirel J.-C., Ghribi M.,
***Les méthodes de classification non supervisées appliquées aux textes : mesure de la
performance des résultats de clustering de documents,***
Proceedings of ACSI 2010, Montreal, Canada, June 2010.

- [IC10d] Lamirel J.-C., Ghribi M., Cuxac P.,
Micro rappel précision non supervisés : vers de nouvelles mesures de qualité de clustering,
Proceedings of SFC'10 - Saint-Denis de La Réunion, France, June 2010.
- [IC10c] Lamirel J.-C., Ghribi M., Cuxac P.,
Exploitation of unsupervised cumulative precision measures for efficient clustering quality estimation,
Proceedings of the 42th Statistics Days (JDS 2010), Marseille, France, May 2010.
- [IC10b] Lamirel J.-C., Cuxac P., Haydar S., François C.,
A new general paradigm for mining research topics evolving over time,
Proceedings of the 3rd International Conference on Information Systems and Economic Intelligence (SIIE'2010), Sousse, Tunisia, February 2010.
- [IC10a] Ghribi M., Cuxac P., Lamirel J.-C., Lelu A.,
Mesures de qualité de clustering de documents : Prise en compte de la distribution des mots-clés,
Proceedings of the 10th International Francophone Conference on Knowledge Extraction and Management (EGC 2010), Hammamet, Tunisia, January 2010.
- [IC09b] Lamirel J.-C., Al Shehabi S.,
Intelligent patents analysis using a multiview neural-based approach,
Proceedings of Conference on Recent Trends in Patinformatics, Puna, India, December 2009.
- [IC09a] Lamirel J.-C., Haydar S.,
Analysis of research topics evolving over time using an unsupervised neural clustering in a multiview context,
Proceedings of the 5th International Conference on Webometrics, Informetrics and Scientometrics and 10th COLLNET Meeting, Dalian, China, September 2009.
- [IC08d] Lamirel J.-C., Ta A.P.,
Combination of hyperbolic visualization and graph-based approach for organizing data analysis results: an application to social network analysis,
Proceedings of the 4th International Conference on Webometrics, Informetrics and Scientometrics and 9th COLLNET Meeting, Berlin, Germany, August 2008.
- [IC08c] Kassab R., Lamirel J.-C.,
Feature-based cluster validation for high dimensional data,
Proceedings of the IASTED International Conference on Artificial Intelligence and Applications (AIA), Innsbruck, Austria, February 2008.
- [IC08b] Lamirel J.-C., Ta A.P., Attik M.,
Novel labeling strategies for hierarchical representation of multidimensional data analysis results,
Proceedings of the IASTED International Conference on Artificial Intelligence and Applications (AIA), Innsbruck, Austria, February 2008.
- [IC08a] Kassab R., Lamirel J.-C.,
Highlighting latent structures in clustering: a multilevel approach,
Proceedings of the IASTED International Conference on Artificial Intelligence and Applications (AIA), Innsbruck, Austria, February 2008.

- [IC07e] Zennaki M., Ech-Cherif A., Lamirel J.-C.,
Using reactive tabu search in semi-supervised classification,
Proceedings of the International Conference on Tools with Artificial Intelligence (ICTAI), Patras, Greece, October 2007.
- [IC07d] Lamirel J.-C.,
The MVDA approach: an emergent data analysis and data mining method for Toxicogenomics,
Proceedings of the 1st NIH-INRIA Workshop on Biomedical and Life Science Computing, Bethesda, Washington DC, USA, April 2007.
- [IC07c] Lamirel J.-C., Al Shehabi S.,
An extended neural gas model for efficient data mining tasks,
Proceedings of FLAIRS 2007, Key West, FL, USA, May 2007.
- [IC07b] Lamirel J.-C.,
Entropy-based methods for label analysis: an extended example of application to Scientometrics,
Proceedings of the 3rd International Conference on Webometrics, Informetrics, Scientometrics and 8th COLLNET Meeting, New Delhi, India, March 2007.
- [IC07a] Kassab R., Lamirel J.-C.,
Towards a synthetic analysis of user's information need for more effective personalized filtering services,
Proceedings of the 22th Annual ACM Symposium on Applied Computing (SAC-IAR 2007), Seoul, Korea., March 2007.
- [IC06k] Carrez N., Lamirel J.-C., Al Shehabi S.,
Extracting informative rules from high dimensional data using a numerical approach,
Proceedings of the IEEE International Conference on Data Mining (ICDM), Hong Kong, China, December 2006.
- [IC06j] Alusse A., Lamirel J.-C., Belaïd A.,
PFC : Un outil d'aide à la découverte des contenus des documents et à la création de dossiers,
Proceedings of the 8th International Colloquium on Electronic Document (CIDE.8), Fribourg, Switzerland, September 2006.
- [IC06i] Zennaki M., Ech-Cherif A., Lamirel J.-C.,
Tabu search meta-heuristic for image semi-supervised classification,
Proceedings of the International Conference on Graphical Models and Imaging (GMAI), London, UK, July 2006.
- [IC06h] Lamirel J.-C., Al Shehabi S.,
MultiSOM: a multi-view neural model for accurately analyzing and mining complex data,
Proceedings of the 4th International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV), London, UK, July 2006.
- [IC06g] Kammoun H., Lamirel J.-C., Ben Ahmed M.,
Apprentissage neuro-symbolique pour la RI coopérative, adaptative et évolutive : le modèle multi-agents SARCI,
Proceedings of INFORSID 2006, Hammamet, Tunisia, June 2006.

- [IC06f] François C., Lamirel J.-C., Al Shehabi S.,
Evaluation of collaboration between European universities using dynamic interaction between multiple sources,
Proceedings of the 2nd International Conference on Webometrics, Informetrics, Scientometrics and 7th COLLNET Meeting, Nancy, France, 2006.
- [IC06e] Al Shehabi S., Lamirel J.-C.,
A new hyperbolic visualization method for displaying the results of a neural gas model: application to webometrics,
Proceedings of the 14th European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium, April 2006.
- [IC06d] Kassab R., Lamirel J.-C.,
An innovative approach to intelligent information filtering,
Proceedings of the 21th Annual ACM Symposium on Applied Computing (SAC-IAR 2006), Dijon, April 2006.
- [IC06c] Attik M., Lamirel J.-C., Al Shehabi S.,
Clustering analysis for data with multiple labels,
Proceedings of the IASTED International Conference on Databases and Applications (DBA), Innsbruck, Austria, February 2006.
- [IC06b] Attik M., Al Shehabi S., Lamirel J.-C.,
Clustering quality measures for data samples with multiple labels,
Proceedings of the IASTED International Conference on Databases and Applications (DBA), Innsbruck, Austria, February 2006.
- [IC06a] Kassab R., Lamirel J.-C.,
A new approach to intelligent text filtering based on novelty detection,
Proceedings of the 17th Australasian Database Conference (ADC 2006), Hobart, Tasmania, AU, January 2006.
- [IC05g] Kammoun H., Lamirel J.-C., Ben Ahmed M.,
Machine-Learning applied to queries and documents for an adaptive and evolutive information retrieval,
Proceedings of the 2nd International Conference on Machine Intelligence (ACIDCA – ICMI), Tozeur, Tunisia, November 2005.
- [IC05f] Kammoun H., Lamirel J.-C., Ben Ahmed M.,
Neural-symbolic machine-learning for knowledge discovery and adaptive information retrieval,
Transaction on Engineering, Computing and Technology, Word Enformatika Society, Istanbul, Turkey, September 2005.
- [IC05e] Al Shehabi S., Lamirel J.-C.,
Knowledge extraction from unsupervised multi-topographic neural network models,
Proceedings of ICANN 2005, Warsaw, Poland, September 2005.
- [IC05d] Al Shehabi S., Lamirel J.-C.,
Efficient knowledge extraction using unsupervised neural network models,
Proceedings of the 5th Workshop on Self-Organizing Maps (WSOM 05), Paris, France, September 2005.

- [IC05c] Al Shehabi S., Lamirel J.-C.,
Multi-topographic neural network communication and generalization for multi-viewpoint analysis,
Proceedings of the International Joint Conference on Neural Networks – IJCNN'05, Montreal, Canada, August 2005.
- [IC05b] Kassab R., Lamirel J.-C., Nauer E.,
Novelty detection for modeling user's profile,
Proceedings of FLAIRS 2005, ClearWater, FL, USA, May 2005.
- [IC05a] Lamirel J.-C.,
Unsupervised rule extraction using neural numerical models: application to patent analysis and Web mining,
Proceedings of the 5th Sino-Franco Workshop on Information Technology, Tainan, Taiwan, March 2005.
- [IC04d] François C., Hoerlesberger M., Lamirel J.-C., Al Shehabi S.,
Visualization of Web data: complementarities between a graph and a multi-maps approach,
Proceedings of S&T 2004, Leiden, Holland, September 2004.
- [IC04c] Lamirel J.-C., Al Shehabi S.,
Comparison of unsupervised neural clustering methods for mining Web and textual data,
Proceedings of SCI 2004, Orlando, FL, USA, July 2004.
- [IC04b] Al Shehabi S., Lamirel J.-C.,
Inference bayesian network for multi-topographic neural network communication: a case study in documentary data,
Proceedings of ICTTA, Damas, Syria, April 2004.
- [IC04a] Lamirel J.-C., Al Shehabi S., Francois C., Polanco X., Hoffmann M.,
Using a compound approach based on elaborated neural network for Webometrics: an example issued from the EICSTES project,
Proceedings of the 5th COLLNET Conference. Roorkee, India, March 2004.
- [IC03d] Lamirel J.-C., Francois C., Al Shehabi S., Hoffmann M.,
New classification quality estimators for analysis of documentary information: application to patent analysis and Web mapping,
Proceedings of ISSI 2003, Beijing, China, August 2003.
- [IC03c] Lamirel J.-C., Al Shehabi S.,
Neural network driven patent analysis,
Proceedings of ACL 2003 Workshop on Patent Analysis, Sapporo, Japan, July 2003.
- [IC03b] Lamirel J.-C., Al Shehabi S., Toussaint Y.,
A hybrid classification method for database contents analysis,
Proceedings of FLAIRS 2003, Orlando, FL, May 2003.
- [IC03a] Lamirel J.-C.,
Optimized data analysis based on neural networks: experiments on the MultiSOM model in the framework of patent and web data analysis,
Proceedings of the 3rd Sino-Franco Workshop on Multimedia, Taipei, Taiwan, March 2003.

- [IC02a] Kammoun H., Lamirel J.-C., Ben Ahmed M.,
SARCI: système agent pour la classification et la recherche d'information,
Proceedings of CIFED 2002, Hammamet, Tunisie, October 2002.
- [IC02g] Lamirel J.-C., Al Shehabi S., Polanco X., Francois C., Toussaint Y.,
Global and partitioned classification methods for mapping science and technology: a hybrid methodology for comparison,
Proceedings of ICT 2002, Shiraz, Iran, October 2002.
- [IC02f] Lamirel J.-C.,
Using unsupervised learning for extracting and manipulating knowledge in a data mining context,
Proceedings of ICTEI Workshop, Ibadan, Nigeria, August 2002.
- [IC02e] Lamirel J.-C. , Hsiang J. & al.,
A neural-based approach for building an operational digital library model,
Proceedings of SCI 2002, Orlando, FL, July 2002.
- [IC02d] Lamirel J.-C.,
MultiSOM: a Multimap Extension of the SOM Model,
Proceedings of IJCNN 2002, Honolulu, Hawaii, May 2002.
- [IC02c] Lamirel J.-C.,
Intelligent browsing and discovery through images,
Proceedings of FLAIRS 2002, Pensacola, FL, May 2002.
- [IC02b] Lamirel J.-C., Hsiang J. & al.,
Using a background neural model in a digital library for assistance to user navigation and for enhancing the library evolutivity,
Proceedings of the 2nd Sino-Franco Workshop on Multimedia, Sophia Antipolis, March 2002.
- [IC02a] Lamirel J.-C., Toussaint Y.,
Association de méthodes symboliques et numériques pour l'analyse du contenu de bases de données,
Proceedings of RFIA 2002, Angers, France, January 2002.
- [IC01f] Lamirel J.-C., Toussaint Y., Ducloy J., Czysz C., Francois C.,
Réseaux neuronaux avancés pour la cartographie de la science et de la technologie : application à l'analyse des brevets”,
Proceedings of the Technological and Strategic Survey Conference – VSST'2001, Barcelone, Spain, September 2001.
- [IC01e] Lamirel J.-C., Toussaint Y., Francois C., Polanco X.,
Using a MultiSOM approach for mapping of science and technology,
Proceedings of ISSI 2001, Sydney, Australia, July 2001.
- [IC01d] Toussaint Y., Lamirel J.-C., D'Aquin M.,
Combining symbolic and numeric techniques for database content analysis,
Proceedings of IEA01, Budapest, Hungary, June 2001.
- [IC01c] Kammoun H., Lamirel J.-C.,
A classificatory system for information retrieval: SARCI,
Proceedings of SCI 2001, Orlando, Florida, USA, July 2001.

- [IC01b] Lamirel J.-C.,
MicroNOMAD model application to digital museum of butterflies,
Proceedings of the 1st Sino-Franco Workshop on Multimedia, Puli, Taiwan, March 2001.
- [IC01a] Lamirel J.-C.,
Using images for enhancing discovery tasks in a digital library context,
Proceedings of SPIE2001, San José, CA, USA, January 2001.
- [IC00g] Lamirel J.-C., Toussaint Y.,
Combining symbolic and numeric techniques for digital libraries contents classification and analysis,
Proceedings of the 1st DELOS Network of Excellence Workshop, Zurich, December 2000.
- [IC00f] Martinez M. M., Cuesta C.E., de la Fuentes P., Lamirel J.-C.,
Integracion de colecciones heterogeneas en bibliotecas digitales,
Proceedings of "Primida Jornadas de Bibliotecas Digitales" (JBDIDT'2000), Valladolid, Spain.
- [IC00e] Lamirel J.-C., Ducloy J., Kammoun H.,
Topographic combination of images and texts for information discovery in a digital library context,
Proceedings of ACM/SBBD 2000, Joao Pessoa, Brazil, October 2000.
- [IC00d] Lamirel J.-C., Ducloy J., Kammoun H.,
A self-organizing map (SOM) extended model for information discovery,
Proceedings of RSCTC2000, Banff Ridge, Canada, October 2000.
- [IC00c] François C., Polanco X., Lamirel J.-C.,
Information visualization and analysis for knowledge discovery using a multi self-organizing mapping,
Proceedings of PKDD2000, Lyon, France, September 2000.
- [IC00b] Lamirel J.-C., Polanco X., François C.,
Adaptive mapping model for data clustering and visualization,
Proceedings of S&T 2000, Leiden, The Netherlands, May 2000.
- [IC00a] Lamirel J.-C., Ducloy J., Oster G.,
Adaptive browsing for information discovery in an iconographic context,
Proceedings of RIAO2000, Vol. 2, Paris, France, April 2000.
- [IC97b] Lamirel J.-C., Nauer E., Ducloy J.,
Using of multiple data sources for information filtering: first approaches in the MedExplore project,
Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering, Budapest, Hungary, November 1997.
- [IC97a] Nauer E., Lamirel J.-C.,
Environnement d'investigation sur WWW : assistance à l'utilisateur par des connaissances fédérées,
Proceedings of the International Conference on Hypertext and Hypermedia, Saint-Denis, France, September 97.

[IC94d] Lamirel J.-C., Créhange M.,
Application of a symbolico-connectionist approach for the design of a highly interactive documentary database interrogation system with on-line learning capabilities,
Proceedings ACM-CIKM 94, Gaithersburg, Maryland, USA, November 94.

[IC94c] Ducloy J., Lamirel J.-C., Nauer E.,
A workbench for bibliographic or factual data handling,
Proceedings of CODATA 94, Chambéry, France, September 94.

[IC94b] Lamirel J.-C., Créhange M.,
NOMAD : un système d'interrogation de bases documentaires basé sur un modèle connexionniste intégrant des topographies neuronales multiples et la détection de nouveauté : aspect neuronaux,
Proceedings of Neural Modeling, Lyon, France, June 94.

[IC94a] Lamirel J.-C., Créhange M., Ducloy J.,
NOMAD: a documentary database interrogation system using multiple neural topographies and novelty detection,
Proceedings of the 3rd International ISKO Conference, Copenhagen, Denmark, June 94.

[IC93a] Lamirel J.-C., Grivel L.,
SDOC, an analysis tool for scientometric studies integrated in an hypermedia environment,
Proceedings of ICO, Montreal, Canada, May 93.

[IC91c] Grivel L., Lamirel J.-C.,
SDOC, a generator of hypertext structures,
Proceedings of the Multimedia Information Conference, Cambridge, UK, July 91.

[IC91b] Lamirel J.-C., Gandhilon T., Hamza N., Delacroix M.,
ZEN: an hypertext system on Unix,
Proceedings of RIAO91, Vol. 3, Barcelona, Spain, April 91.

[IC91a] Ducloy J., Grivel L., Lamirel J.-C., Polanco X., Schmitt L.,
INIST experience in hyperdocument building from bibliographical data,
Proceedings of RIAO91, Vol. 1, Barcelona, Spain, April 91.

7.10 Rapports techniques de projets européens

[TR05c] Castagnos S., Kassab R., Boyer A., Charpillat F., Lamirel J.-C.,
Selection of profiling, filtering and content analysis techniques,
SAT-At-ONCE Technical Report: A.1-D3 Analysis, August 2005.

[TR05b] Boyer A., Castagnos S., Kassab R., Lamirel J.-C.,
Proposal of valuation methods for the filtering algorithm,
SAT-At-ONCE Technical Report (2005).

[TR05a] Kassab R., Castagnos S., Lamirel J.-C., Boyer A.,
State of the art on information filtering methods,
SAT-At-ONCE Technical Report: A.1-D1 Bibliography, December 2004.

[TR03b] François C., Hoffmann M., Lamirel J.-C., Polanco X.,
Artificial neural network mapping experiments,
EICSTES (IST-1999-20350) Final Report (WP 9.4), 86 p., September 2003.

[TR03a] François C., Hoffmann M., Lamirel J.-C., Polanco X.:
Multi-maps SOM platform,
EICSTES (IST-1999-20350) Intermediary Report (WP 9.1), 92 p., September 2003.

[TR01a] Lamirel J.-C.,
User's requirements analysis,
SCHOLNET (IST-1999-20664) Intermediary Report (WP1), 50 p., February 2001.

7.11 Communications dans des conférences nationales avec comité de lecture

[NC05a] Kassab R., Lamirel J.-C., Nauer E.,
Une nouvelle approche pour la modélisation du profil de l'utilisateur dans les systèmes de filtrage d'information basés sur le contenu : le modèle de filtre détecteur de nouveauté,
Proceedings of CORIA 2005, Grenoble, France, March 2005.

[NC00b] Martinez M. M., Cuesta C.E., de la Fuentes P., Lamirel J.-C. ,
Consultas heterogeneas en bibliotecas digitales distributas,
Proceedings of SEID 2000, Ourense, Spain, September 2000.

[NC00a] Lamirel J.-C., Kammoun H., Ducloy J.,
Navigation adaptative et découverte d'information assistées par un fonds iconographique,
Proceedings of Ergo-IA'2000, Biarritz, France, October 2000.

[NC99a] Lamirel J.-C.,
L'approche multi-topographique MicroNOMAD : application à la navigation dans une base iconographique,
Proceedings of "Colloque de Bibliométrie Appliquée de l'Ile Rousse", France, May 99.

[NC95a] Nauer E., Lamirel J.-C.,
Recherche précise d'information sur WWW et Veille Technologique : utilisation de données structurées pour l'interrogation via les moteurs de recherche,
Proceedings of "Colloque de Bibliométrie Appliquée de l'Ile Rousse", France, September 1995.

[NC94a] Lamirel J.-C., Créhange M.,
NOMAD : un système d'interrogation de bases documentaires basé sur un modèle connexionniste intégrant des topographies neuronales multiples et la détection de nouveauté : aspect ergonomiques,
Proceedings of Ergo-IA'94, Biarritz, France, October 94.

7.12 Communications dans des conférences nationales sans comité de lecture

[NS07a] Lamirel J.-C.,
Vision scientifique et auscultatoire du monstre,
Proceedings of "Le corps monstrueux" Colloquium, Nancy, France, Octobre 2007.

H. AUTRES REFERENCES

- [Agra96] Agrawal, R., Mannila, H., Srikant, R., Toivonen, H. and Verkamo, A.I. (1996), ***Fast discovery of association rules***, In: Advances in knowledge discovery and data mining: 307-328, American Association for Artificial Intelligence.
- [Anke99] Ankerst M., Breunig M. M., Kriegel H.P. and Sander J. (1999), ***OPTICS: Ordering Points To Identify the Classification Structure***, Proceedings of the ACM SIGMOD'99, Int. Conf. on Management of Data, Philadelphia PA.
- [Alla98] Allan, J., Carbonell, J., Doddington, G., Yamron, J. and Yang, Y. (1998), ***Topic detection and tracking pilot study, final report***, Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop, Lansdowne, Virginia.
- [Arro03] Arroyo, N., Pareja, V. M. and Aguillo, I. (2003), ***Description of Web data in D3.1***, Eiscetes European Project IST-1999-20350: <http://www.eicstes.org>.
- [Azza04] Azzag, H., Picarougne, F., Guinot, C. and Venturini, G. (2004), ***Un survol des algorithmes biomimétiques pour la classification***, Revue des Nouvelles Technologies de l'Information (RNTI-C-1), Classification et fouille de données, pp 13-24, Cépaduès édition.
- [Azza05] Azzag, H., Picarougne, F., Guinot, C. and Venturini, G. (2005), ***Classification de données par automate cellulaire***, 12èmes Rencontres de la Société Francophone de Classification (SFC2005), Montréal, Canada, pp 47-50, Juin 2005.
- [Bata02] Batagelj, V. and Zaversnik, M. (2002), ***An $O(m)$ algorithm for cores decomposition of networks***, University of Ljubljana, Preprint: IMFM 797.
- [Bast06] Bastide, Y., Taouil, R., Pasquier, N., Stumme, G. and Lakhal, L. (2006), ***Mining minimal nonredundant association rules using frequent closed itemsets***, In: J. et al.. Lloyd (Ed.), In Proc. of the Computational Logic (CL'00), Vol. 1861 of Lecture Notes in Artificial Intelligence – LNAI: 972-986, Springer.
- [Bins02] Binsztok, H. and Gallinari, P. (2002), ***Un algorithme en ligne pour la détection de nouveauté dans un flux de documents***, JADT 2002: 6èmes Journées Internationales d'Analyse Statistique des Données Textuelles, Saint-Malo, Mars 2002.
- [Brei01] Breiman, L. (2001), ***Random Forests***, Machine Learning 45 (1): 5–32.
- [Cali74] Calinski, R.B. and Harabasz, J. (1974), ***A dendrite method for cluster analysis***, "Communications in Statistics", vol. 3: 1-27.
- [Chen03] Chen, C.C., Chen, Y.T., Sun, Y.S. and Chen, M.C. (2003), ***Life Cycle Modeling of News Events Using Aging Theory***, 14th European Conference on Machine Learning (ECML), Dubrovnik, Croatie, pp 47-59.
- [Chia04] Chiang, J.H., Hsiao-Sheng, L., Shih-Yi, C. and Cheng-Yu, C. (2004), ***Discovering Gene-Gene Relations from Sequential Sentence Patterns in Biomedical Literature***, Expert Systems With Applications, Vol. 33: 1036-1041 SCI/EI.
- [Chia06] Chiang, J.H., Jyh-Wei, S., Heng-Hui, L. and Chong-Liang, C. (2006), ***GeneLibrarian: an effective gene-information summarization and visualization system***, BMC Bioinformatics, 7:392 SCI.

- [Chua04] Chuang, S. and Chien, L. (2004), *A practical web-based approach to generating topic hierarchy for text segments*, In Proceedings of the 20th International Conference on Information and Knowledge Management.
- [Chu05] Chu, H. (2005), *Taxonomy of inlinked Web entities: What does it imply for webometrics research?*, Library & Information Science Research, Vol. 27: 8-27.
- [Cutt93] Cutting, D.R., Karger, D.R. and Pederson, J.O. (1993), *Constant interaction-time Scatter/Gather browsing of very large document collection*, In Proceeding of International ACM Conference on Research and Development in Information Retrieval.
- [Cuxa09] Cuxac, P., Lelu, A. and Cadot, M. (2009), *Suivi incrémental des évolutions dans une base d'information indexée : une boucle évaluation /correction pour le choix des algorithmes et des paramètres*, 2ème conférence Internationale sur les systèmes d'informations et Intelligence Economique SIIE 2009, Hammamet, Tunisie, Février 2009.
- [Davi79] Davies, D.L. and Bouldin, D.W. (1979), *A cluster separation measure*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1(2): 224–227.
- [Dema97] Demartines, P. and Héroult, J. (1997), *Curvilinear component analysis: A self-organizing neural network for linear mapping of data sets*, 8(1): 184-154.
- [Demi98] Demiriz, A. and Bennett, K.P. (1998), *Optimization approaches to semi-supervised learning*, Proceedings of Neural Information Processing Systems, Denver.
- [Demp77] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977), *Maximum likelihood for incomplete data via the em algorithm*, Journal of the Royal Statistical Society, vol. B-39: 1-38.
- [Derw90] Deerwester, S., Dumais, S., Furnas, G. W., Landauer, T. K. and Harshman, R. (1990), *Indexing by Latent Semantic Analysis*, Journal of the American Society for Information Science, 41(6): 391–407.
- [Erte03] Erten, C. and al. (2003), *Exploring the computing literature using temporal graph visualization*, Report, Department of Computer Science, University of Arizona.
- [Este96] Ester, M., Kriegel, H.-P., Sander, J. and Xu, X. (1996), *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*, In Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining (KDD'96): 226-231, AAAI Press, Menlo Park, CA.
- [Friz95] Frizke, B. (1995), *A growing neural gas network learns topologies*, Tesauro G., Touretzky D. S., Leen T. K., Eds., Advances in neural Information processing Systems, vol. 7: 625-632, MIT Press, Cambridge MA.
- [Falk09] Falk, I., Gardent, C., Jacquy, E. and Venant, F. (2009), *Sens, synonymes et définitions*, TALN'2009, Senlis, France.
- [Fish87] Fisher, D. (1987), *Knowledge Acquisition via Incremental Conceptual Classification*, Machine learning, Vol. 2, pp. 139-172.
- [Gab05] Gaber, M., Zaslavsky, A. and Krishnaswamy, S. (2005), *Mining Data Streams: A Review*, SIGMOD Record, 34(2).
- [Garf86] Garfield, E. (1986), *Introducing the ISI Atlas of Science: Biotechnology and molecular genetics, 1981/82 and bibliographic update for 1983*, Essays of an Information Scientist, Vol. 9, pp 42-47.

- [Gaud05] Gaudin, R. and Nicoloyannis, N. (2005), *Apprentissage non supervisé de séries temporelles à l'aide des k-Means et d'une nouvelle méthode d'agrégation de séries*, 5èmes journées d'Extraction et Gestion des Connaissances (EGC 2005), Paris, pp 201-212, Janvier 2005.
- [Glov89] Glover, F. (1989), *Tabu Search-Part I*, ORSA Journal on Computing, Vol. 1:190-206.
- [Gross87] Grossberg, S. (1987), *Competitive learning: From interactive activation to adaptive resonance*, Cognitive Science (Publication), vol. 11: 23-63.
- [Guen04] Guenoche, A. (2004), *Classification by vertex density in a Graph*, Meeting of the International Federation of the Classification Societies. Chicago, Classification, Classification and Data Mining, D. Banks et al. (Eds.), Springer, pp 15-23.
- [Hade03] Hader, S. and Hamprecht, F. A. (2003), *Efficient density classification using basin spanning trees*, Between Data Science and Applied Data Analysis, pp 39-48, Springer.
- [Hali90] Halin, G., Créhange, M. and Kerekes, P. (1990), *Machine Learning and Vectorial Matching for an Image Retrieval Model: EXPRIM and the System RIVAGE*, SIGIR: 99-114.
- [Hamm02] Hammer, B., Rechten, A., Strickert, M. and Villmann, T. (2002), *Rule extraction from self-organizing networks*, In Proc. of ICANN: 877-882, Springer.
- [Hamz08] Hamza H., Belaïd Y., Belaïd. A. and Chaudhuri B. B. (2008), *Incremental classification of invoice documents*, 19th International Conference on Pattern Recognition – ICPR.
- [Havr02] Havre, S. and al. (2002), *ThemeRiver: visualizing thematic changes in large document collections*, IEEE Transactions on Visualization and Computer Graphics, Vol. 8, N°1.
- [Hebb49] Hebb, D.O. (1949), *The Organization of Behavior: A Neuropsychological Theory*, Wiley, New York.
- [Heim06] Heimeriks, G., Hörlesberger, M. and Van den Besselaar, P. (2003), *Mapping communication and collaboration in heterogeneous research networks*, Eiscetes European Project IST-1999-20350.
- [Hinn98] Hinnburg, A. and Keim, D. A. (1998), *An Efficient Approach to Classification in Large Multimedia Databases with Noise*, Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining, (KDD98), Las Vegas, USA, August 24-27 1998.
- [Humb06] Humbert, P., François, C., Cuxac, P. and David, A. (2006), *La visualisation dynamique de connaissances scientifiques : intégration des besoins des utilisateurs*, Canadian Journal of Information and Library Science, 30, n°1/2, pp 38-55.
- [Ingw03] Ingwersen, P. (2003), « *Webometrics: basic concepts* », ISSI 2003, Beijing, China.
- [Jack79] Jackson, J. E. and Mudholkar, G. S. (1979), *Control Procedures for Residuals Associated with Principal Component Analysis*, Technometrics, 21(3): 341–349.
- [Joac99] [13] Joachims, T. (1999), *Transductive inference for text classification using support vector machines*, In International Conference on Machine Learning.
- [Kary99] Karypis, G., Han, E.-H. and Kumar, V. (1999), *CHAMELEON: A Hierarchical Clustering Algorithm Using Dynamic Modeling*, IEEE Computer, 32(8): 68-75.

- [Kask98] Kaski, S. (1998), *Dimensionality reduction by random mapping*, In Proc. Int. Joint Conf. on Neural Networks, vol 1: 413–418.
- [Kim02] Kim, J. H. and Beale, G. O. (2002), *Fault Detection and Classification in Underwater Vehicles Using the T2 Statistic*, AUTOMATIKA - Journal for Control, Measurement, Electronics, Computing and Communications. 43(1-2): 29–37.
- [Koho01] Kohonen, T. (2001), *Self-Organizing Maps*, 3rd ed. Springer Verlag, Berlin.
- [Kops98] Kopcsa, A. and Schiebel, E. (1998), Science and Technology Mapping: *A New Iteration Model for Representing Multidimensional Relationships*, Journal of the American Society for Information Science (JASIS), 49(1): 7-17.
- [Kret85] Kretschmer, H. (1985), *Cooperation structure, group size and productivity in research groups*, 7(1985): 39-53.
- [Lave07] Lavergne, J., Azzag, H., Guinot, C. and Venturini, G. (2007), *Construction incrémentale et visualisation de graphes de voisinage par des fourmis artificielles*, 7èmes journées d’Extraction et Gestion des Connaissances (EGC 2007), Namur. RNTI-E-9 Cépaduès-Éditions, Vol. 1, pp 135-146.
- [Leba82] Lebart L., Morineau, A. and Tabard, N. (1982), *Techniques de la description statistique*, Dunod, Paris.
- [Lee00] Lee, C.-Y. and Antonsson, E. (2000), *Dynamic partitional classification using evolution strategies*, Proceedings of the Third Asia-Pacific Conference on Simulated Evolution and Learning, Nagoya, Japan.
- [Lelu06] Lelu, A., Cuxac, P. and Johansson, J. (2006), *Classification dynamique d’un flux documentaire: une évaluation statique préalable de l’algorithme GERMEN*, Journées internationales d’Analyse statistique des Données Textuelles (JADT), Besançon, pp 617-630, Avril 2006.
- [MacQ67] McQueen, J.B. (1967), *Some methods of classification and analysis of multivariate observations*, L. Le Cam and J. Neyman (Eds.), Proc. 5th Berkeley Symposium in Mathematics, Statistics and Probability, vol. 1: 281-297, Univ. of California, Berkeley, USA.
- [Mart91] Martinetz, T. and Schulten, K. (1991), *A “neural-gas” network learns topologies*, In T. Kohonen, K. Mäkisara, O. Simula, and J. Kangas, editors, Artificial neural networks: 397-402, North-Holland, Amsterdam.
- [Merk03] Merkl, D., Shao, H.H., Dittenbach, M. and Rauber, A. (2003), *Adaptive hierarchical incremental grid growing: an architecture for high-dimensional data visualization*, Proceedings of the 4th Workshop on Self-Organizing Maps, Advances in Self-Organizing Maps, pp 293-298, Kitakyushu, Japan, September 11-14 2003.
- [Mfou98] Mfoumoune, E. (1998), *Les aspects algorithmiques de la classification ascendante pyramidale et incrémentale*, Thèse Université Paris IX Dauphine (E. DIDAY), 220 p.
- [Mich88] Michelet, B. (1988), *L’analyse des associations*, Thèse de doctorat, Université Paris VII.
- [Mitt04] Mitton, N., Busson, A. and Fleury, E. (2004), *Self-organization in large scale ad hoc networks*, The Third Annual Mediterranean Ad Hoc Networking Workshop, (MED-HOC-NET 04). Bodrum, Turkey.

- [Monm01] Monmarché, N., Laügt, D., Mestre, M., Labroche, N., Oliver, A. and Venturini, G. (2001), *Classification et visualisation dynamique de données par nuage d'insectes volants*, Neuvièmes rencontres de la Société Francophone de Classification, pp 254-261, Pointe à Pitre (Guadeloupe), Décembre 2001.
- [Mood01] Moody, J. (2001), *Identifying dense clusters in large networks*, Social Networks, Vol. 23, pp 261-283.
- [Mosc03] Moschitti, A. (2003), *A study on optimal parameter tuning for rocchio text classifier*, Proceedings of the 25th European Conference on Information Retrieval Research (ECIR) : 420-435.
- [Ould97] Ould, M.A. (1997), *Comparaison de méthodes neuronales avec des méthodes d'analyse des données dans le cadre d'ingénierie de l'information*, Mémoire de stage de D.E.S.S. en «Ingénierie mathématique et outils informatiques», CEC, UHP, Nancy, France.
- [Orwi97] Orwig, R.E., Chen, H. and Nunamaker Jr., J. F. (1997), *A graphical, Self Organizing Approach to Classifying Electronic Meeting Output*, JASIS, 48 (1): 157-170.
- [Park03] Park, H.W. and Thelwall, M. (2003), *Hyperlink analyses of the World Wide Web: a review*, Journal of Computer-Mediated Communication, 8(4).
- [Pasq99] Pasquier, N., Bastide, Y., Taouil, R. and Lakhil, L. (1999), *Efficient mining of association rides using closed itemset lattices*, Inf. Syst. 24(1): 25-46.
- [Pasq00] Pasquier, N. (2000), *Mining association rides using formal concept analysis*, In Proc. of the 8th International Conf. on Conceptual Structures (ICCS '00): 259-264, Shaker-Verlag.
- [Pope00] Popescul, A. and Ungar, L.H. (2000), *Automatic Labeling of Document Clusters*, Unpublished manuscript, available: <http://citeseer.nj.nec.com/popescul100automatic.html>
- [Pons06] Pons, P. and Latapy, M. (2006), *Computing communities in large networks using random walks*, Journal of Graph Algorithms and Application.
- [Pons04] Pons-Porrata, A. Berlanga-Llavori, R., Ruiz-Shulcloper, J. and Perez-Martinez, J.M. (2004), *JERARTOP: A new topic detection system*, Lecture notes in computer sciences – 9th Iberoamerican congress on pattern recognition, Vol. 3287, pp 446-453.
- [Prud04] Prudent, Y. and Ennaji, A. (2004), *Clustering incrémental pour un apprentissage distribué : vers un système évolutif et robuste*, CAP04, Montpellier.
- [Rask04] Raskutti, B. and Kowalczyk, A. (2004), *Extreme re-balancing for SVMs: a case study*, SIGKDD Explor. Newsl. 6(1): 60–69.
- [Rizz01] Rizzo, R. (2001), *LBG-m: a modified lbg architecture to extract high-order neural structures*, In: Proc. Of International Joint Conference on Neural Networks (IJCNN): 779–783.
- [Rodr06] Rodriguez, J. J., Kuncheva, L.I. and Alonso Carlos, J. (2006), *Rotation Forest: A New Classifier Ensemble Method*, IEEE Transactions on Pattern Analysis and Machine Intelligence (28)10: 1619-1630.
- [Rocc71] Rocchio, J.J. (1971), *Relevance feedback in information retrieval, In The SMART Retrieval System: Experiments in Automatic Document Processing*, Prentice Hall Inc., Englewood New Jersey.

- [Rume86] Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986), *Learning internal representation by error propagation*, Rumelhart, D.E. and McClelland, J. L., editors, *Parallel Distributed Processing: Explorations in the microstructures of cognition*, MIT Press: 318–362.
- [Samm69] Sammon, J.W.Jr. (1969), *A Nonlinear Mapping for Data Structure Analysis*, IEEE Transactions on Computers, C-18(5): 401-409.
- [Schol99] Scholköpfung, B., Platt, J., Shawe-Taylor, J., Smola, A. J. and Williamson R. C. (1999), *Estimating the support of a high-dimensional distribution*, Neural Computation, 13(7): 1443-1471.
- [Schi10] Schiebel, E., Hörlesberger, M., Roche, I., François, C. and Besagni, D. (2010), *An advanced diffusion model to identify emergent research issues: the case of optoelectronic devices*, Scientometrics, Vol. 83, N° 3, pp 765-781.
- [Sing96] Singhal, A., Buckley, C. and Mitra, M. (1996), *Pivoted document length normalization*, Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval: 21–29.
- [Szat06] Szathmary, L., Napoli A. and Kuznetsov, S. O. (2005), *ZART: A Multifunctional Itemset Miner Algorithm*, LORIA Research Report A05-R-013.
- [Thel02] Thelwall, M. (2002), *Evidence for the existence of geographic trends in university Web site interlinking*, Journal of Documentation, 58 (5): 563-574.
- [Thel03] Thelwall, M., Harries, G. and Wilkinson, D. (2003), *Why do web sites from different academic subjects interlink?*, Journal of Information Science, 29 (6): 453-471.
- [Trem79] Trémolières, R.C. (1979), *The percolation method for an efficient grouping of data*, Pattern Recognition, 11(4).
- [Thij10] Thijs, B. and Glänzel, W. (2010), *A new hybrid approach for bibliometrics-aided retrieval*, Proceedings of 6th International Conference on Webometrics, Informetrics and Scientometrics and 11th COLLNET Meeting, Mysore, India, October 2010.
- [Vapn98] Vapnik, V.N. (1998), *Statistical Learning Theory*, Wiley Inter-Science.
- [Vars92] Varsis, A. and Versino, C. (1992), *Clustering of Socio-Economic Data with Kohonen Maps*, Proceedings of third International Workshop on Parallel Applications in Statistics and Economics, Pragues, Czechoslovakia.
- [Verl04] Verleysen, M. (2004), *Learning High-Dimensional Data with Artificial Neural Networks*, LEARNING'04, Elche, Spain, 20-22 October 2004.
- [Watz64] Watzlawick, P. (1964), *An Anthology of Human Communication*, Text and Tape, Science and Behavior Book.
- [Wayn98] Wayne, C.L. (1998), *Topic detection & tracking (TDT): Overview & perspective*, Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop, Lansdowne, Virginia.
- [Wilk03] Wilkinson, D., Harries, G., Thelwall, M. and Price, L. (2003), *Motivations for academic website interlinking: evidence for the Web as a novel source of information on informal scholarly communication*, Journal of Information Science, 29(1): 49-56.