

Le projet du LITIS 2012-2016

INSA de Rouen, Université de Rouen, Université du Havre

RÉSUMÉ : *Ce document présente le projet du LITIS, les priorités stratégiques du laboratoire, ses ambitions et les actions envisagées pour atteindre ses objectifs.*

Sommaire

1	Le LITIS : Projet de recherche 2012-2016	5
	Auto-analyse	5
	Projet et objectifs scientifiques de l'unité	8
	Mise en œuvre	9
2	Équipe C&A : Combinatoire et algorithmes	13
	Auto-analyse	13
	Projets et objectifs scientifiques	14
	Mise en œuvre	15
3	Équipe TIBS : Traitement de l'Information en Biologie Santé	17
	Auto analyse	17
	Projets et objectifs scientifiques	18
	Mise en œuvre	20
4	Équipe STI : Systèmes de Transport Intelligents	23
	Auto analyse	23
	Projets et objectifs	24
	Mise en œuvre	27
5	Équipe QuantIF : Quantification en Imagerie Fonctionnelle	29
	Auto-analyse	29
	Projets et objectifs scientifiques	30
	Mise en œuvre	32
6	Équipe DocApp : Documents et apprentissage	35
	Auto-analyse	35
	Projets et objectifs scientifiques	36
	Mise en œuvre	39
7	Équipe MIU : Modèles, interactions et usages	41
	Auto-analyse	41
	Projets et objectifs scientifiques	42
	Mise en oeuvre	44
8	Équipe RI2C : Réseaux d'Interactions et Intelligence Collective	45
	Auto-analyse	45
	Projets et objectifs scientifiques	47
	Mise en oeuvre	48

Le LITIS : Projet de recherche 2012-2016

Auto-analyse

Points forts

Un laboratoire affirmé

Après quatre ans d'existence dans sa structure actuelle, le LITIS affiche des résultats en nette progression sur la plupart des indicateurs (niveau et nombre de publications, citations, nombre de doctorants, nombre de projets financés, rayonnement).

Il est vrai que l'évaluation précédente est survenue (printemps 2007) alors que le LITIS achevait à peine son travail de structuration débuté peu avant sa création (janvier 2006). Il s'en est suivi une évaluation jugée sévère pas les membres du laboratoire, eu égard au travail réalisé et à l'ambition du projet impulsé par les établissements. Les résultats obtenus sur la période doivent être examinés dans cette perspective.

Des équipes bien identifiées au niveau national et international

Le LITIS a pu bénéficier des points forts des différentes composantes qui ont contribué à sa création. Ces éléments les plus dynamiques contribuent sans conteste au rayonnement international du laboratoire et jouent un rôle d'émulation important (voir le niveau de citation de nos meilleures publications).

Une activité de recherche partenariale riche

L'activité partenariale des équipes, collaborations industrielles, projets ANR, projets du GRR TL-TI, atteste du dynamisme des équipes. Le LITIS totalise 85 projets financés au cours de la période pour un montant total de plus de 4 M€ (auquel il faut ajouter 15 financements CIFRE).

Une couverture thématique large et complémentaire

Trois équipes du LITIS (TIBS, QuantIF et STI) regroupent plusieurs sections du CNU. Par ailleurs la couverture thématique de l'ensemble des équipes permet de positionner le LITIS sur quelques défis scientifiques importants en embrassant des problématiques complémentaires.

Soutien des établissements au projet du LITIS

En dépit de la mauvaise évaluation du projet en cours, le soutien des établissements est resté important au cours de la période. Les établissements ont permis au LITIS de renforcer son projet par l'affectation de moyens humains en enseignants chercheurs et ingénieurs de recherche. Dans le même temps, des dotations spécifiques ont été obtenues sous forme de BQR (Installation des nouveaux arrivants de l'équipe C&A, et mise à niveau du système de sauvegarde du LITIS à l'université de Rouen).

L'émergence régionale des Technologies de l'Information

Dans le cadre du Contrat de Projet Etat Région, le Grand Réseau de Recherche Régional (GRR) Transport Logistique – Technologies de l'Information (TL-TI) a facilité la structuration des projets relevant des Technologies de l'Information. Un soutien fort a été obtenu pour le développement de la Plateforme d'Indexation Régionale (PlAIR), financement FEDER sur 3 ans en collaboration avec le CHU de Rouen. De même, le projet RISC (Réseaux d'Interactions et Systèmes Complexes) s'est structuré au cours de l'année 2010 en élément fédérateur régional. Son rôle doit s'amplifier à brève échéance.

Points faibles

Un niveau de publication à améliorer

Rendre incontournable la publication en revue pour l'ensemble des équipes. Les pratiques de publication sont naturellement variables en fonction des équipes, mais cet objectif doit être mieux inscrit dans toutes les équipes du LITIS.

L'effectif en personnel administratif

Si le soutien des établissements en moyens humains (enseignants chercheurs) s'est maintenu durant la période, il faut toutefois souligner que le déficit en personnel d'encadrement administratif est particulièrement criant. De plus, l'éloignement géographique est un frein à la structuration du secrétariat. L'organisation est essentiellement structurée par établissements (une personne ressource par établissement) et chapeauté par un secrétariat de direction. Les échanges sont permanents mais les agents sont affectés à la gestion d'un

établissement. Ils ne peuvent donc pas se répartir les tâches et constituent les personnes ressources de chaque site, indispensables aux équipes.

L'animation scientifique

L'animation scientifique est principalement développée dans les équipes, elle reste cependant inégale, et devrait être développée de manière à favoriser les échanges entre équipes. Naturellement la répartition sur 5 sites est un obstacle à l'organisation fréquente de séminaires de laboratoires, mais un travail doit être mené en ce sens.

L'implication dans des projets Européens

Le LITIS est peu impliqué dans des projets européens. Ce point pourrait être amélioré en tirant partie de son savoir faire dans l'implication dans les nombreux projets (ANR, industriels,...). Cette faiblesse est sans aucun doute liée au manque de collaborations internationales en général au sein des équipes. Il faut donc favoriser les échanges internationaux en premier lieu.

Néanmoins, face au déficit en encadrement administratif, force est de constater que le LITIS n'est pas « armé » pour prendre à son compte l'administration de projets scientifiques importants (projets européens par exemple). Pour l'amélioration de cet aspect précis, il semble incontournable de se rapprocher des grands organismes de recherche (CNRS – INSERM – INRIA). Des demandes de financement ERC (European Research Council) semblent de ce point de vue plus simples à piloter. Une demande est envisagée dans la période à venir.

Opportunités

Le projet de fédération Norm@STIC

Ce projet encore en gestation, contribuera à renforcer les échanges avec les équipes du laboratoire GREYC qui est l'homologue du LITIS en basse normandie.

Le projet de PRES Normandie Université

En dépit des incertitudes, ou des craintes parfois légitimes, le projet de PRES normand est pour le LITIS une réelle opportunité car il permettrait à terme une meilleure identification de nos thématiques et vraisemblablement l'émergence d'un pôle Normand des STIC par des rapprochements avec les équipes du GREYC. C'est l'absence d'une telle stratégie qui peut en partie expliquer les obstacles au projet d'école doctorale Maths-STIC. Plus généralement, il existe actuellement un ensemble de contacts, et ou de collaborations entre la Haute et la Basse Normandie qui peinent à trouver un soutien inter-régional. A titre d'exemple, le Grand Réseau Régional TL-TI (Transport Logistique – Technologies de l'Information) reste inscrit dans le périmètre Haut Normand alors qu'il y aurait moyen d'impulser des thématiques plus ambitieuses en tirant profit des complémentarités des deux pôles Haut et Bas Normands. Enfin, ce projet de PRES Normandie Université est une opportunité pour affirmer notre identité face à la menace d'un pôle Parisien omniprésent.

Le pôle de compétitivité mondial Mov'eo

Ce pôle de compétitivité, fortement soutenu par les constructeurs automobiles très présents le long du bassin de la seine, représente une opportunité pour l'ensemble des laboratoires de Haute Normandie. L'équipe STI trouve naturellement son positionnement dans ce pôle de compétitivité et développe des collaborations avec l'INRETS, et l'INRIA-Rocquencourt. Elle est impliquée dans la création du laboratoire commun Mov'eo Lab qui correspond à un centre d'essai de véhicules.

L'équipe STI est également impliquée dans la demande d'Institut d'Excellence dans le domaine des énergies décarbonées (IEED porté par l'IFP) dans le cadre du Grand Emprunt sur le volet "véhicule intelligent".

Le pôle de compétitivité Nov@log

Ce pôle de compétitivité pilote les activités de la filière Logistique et Transport, très présente en Haute Normandie du fait de l'activité du port du Havre. Les équipes RI2C et MIU sont fortement impliquées dans la demande de création d'un IRT Logistique Maritime et Portuaire ainsi que dans la demande d'équipement d'excellence qui s'adosse à ce projet d'IRT.

Le projet d'Equipex PLEIAD

L'équipe DocApp développe depuis de nombreuses années des techniques de reconnaissance de formes pour la reconnaissance des documents écrits. Suite à plusieurs projets de collaboration nationales (ACI Madonne puis ANR Navidomass et ANR DigiDoc (actuellement en liste d'attente du programme ContInt) elle est actuellement impliquée dans le projet PLEIAD (Plateforme d'Excellente Industrielle-Académique de Dématérialisation). Ce

projet vise à rassembler les principales équipes Françaises, des industriels majeurs du secteur de l'archivage et du traitement des documents numérisés (ATHOS Origin, IBM, IRIS France, APROGED) pour constituer une plateforme de Recherche et Développement de systèmes de lecture automatique, d'Indexation et d'Archivage de fonds numérisés. Une demande d'équipement d'excellence portée par le L3I (Université de La Rochelle) a été déposée, les industriels et les laboratoires se positionneront sur le volet Contenus Numériques du Grand Emprunt.

Le projet d'Equipex PIRCE

Le projet PIRCE (Peptide Investigation Resource Center (www.pirce.org)) a pour objectif de proposer aux équipes de recherche régionales, nationales et internationales, une infrastructure multidisciplinaire (public/privé) qui dépasse la plupart des moyens actuels dédiés à la recherche sur les Peptides. Porté par l'Institut Fédératif Multidisciplinaire sur les Peptides (IFRMP), il regroupe une douzaine de laboratoires Français. L'équipe QuantIF est impliquée pour développer et tester de nouveaux traceurs de peptide, fluorescent ou radioactifs, pour le diagnostic moléculaire et le suivi des traitements, depuis le modèle animal à l'homme.

Vers une filière TIC en région

Suite à un diagnostic de l'activité socio-économique en Haute Normandie, l'état et la région ont décidé d'aider à la structuration d'une filière TIC régionale. Cette aide a pris la forme de trois projets emblématiques : le laboratoire des usages *Normandy living lab*, la pépinière seine innopolis et le cluster marétic.

Risques

La pluridisciplinarité

Cette richesse peut aussi être un frein à l'unité dans un contexte de plus en plus compétitif et parfois instable. L'attractivité d'autres projets disciplinaires (médecine par exemple) peut tenter certains membres du laboratoire. L'affirmation du projet du LITIS au cours du prochain quadriennal sera un élément déterminant. Il faut également veiller à ce que l'ouverture des projets pluridisciplinaires ne soit pas un obstacle à l'excellence thématique. Il y a parfois un risque à ce que les STIC soient considérées comme pourvoyeur d'outils par les autres disciplines. Toutefois ce risque est de mieux en mieux identifié par nos tutelles.

L'émiettement territorial

La répartition des membres du laboratoire sur 5 sites (2 au Havre – 3 à Rouen) est un obstacle naturel au croisement thématiques des équipes. Il faut veiller à ce que les recrutements prennent en compte la géographie des thématiques afin de ne pas amplifier encore davantage la dispersion territoriale.

Les politiques des établissements

Elles peuvent parfois devenir antagonistes à l'unité de gouvernance du LITIS. Le projet de création d'un département d'informatique à l'IUT de Rouen (entrant en concurrence avec le département d'informatique du Havre) est particulièrement emblématique des difficultés qui peuvent survenir entre les membres du laboratoire, la direction du LITIS, et l'université de Rouen, en l'absence de messages communs des établissements.

L'équipe de direction du laboratoire doit tout particulièrement veiller à l'affirmation de sa politique scientifique. A cela s'ajoute un contexte de plus en plus compétitif au sein des établissements dans lequel le projet du LITIS doit encore s'affirmer. A cet égard la présente évaluation sera déterminante.

L'attractivité des centres Parisiens

Dans le contexte du LITIS en recherche d'affirmation et dans un paysage régional lui-même en phase de structuration lente (projet de PRES Normandie-Université), la proximité du pôle Parisien demeure autant un atout (facilité des échanges) qu'une menace par l'attractivité qu'il représente, tant du point de vue de l'enseignement que de la recherche.

La régionalisation de la recherche

L'autonomie des établissements d'enseignement supérieur, associée à la concentration des moyens sur les pôles d'excellence, sont des éléments qui défavorisent le LITIS face aux grandes structures de recherche régionales. Le diagnostic STRATER (Service de la Coordination Stratégique et des Territoires) relatif à la Haute et la Basse Normandie, met en lumière le paysage de la recherche régionale et ses pôles d'excellence en Chimie, Matériaux et Santé. Le travail important mené dans le cadre des GRR par les membres du LITIS est un élément de visibilité des STIC. Mais il n'en reste pas moins vrai que la tradition industrielle de la région (Chimie, Médicament, Automobile,...) n'est pas principalement portée vers le développement des nouvelles technologies.

Projet et objectifs scientifiques de l'unité

Fédération des forces de recherche en Sciences et Technologies de l'Information et de la Communication

Le bilan très positif du contrat en cours met en évidence la pertinence de la stratégie qui a prévalu à la création du LITIS. C'est cette stratégie que nous nous attacherons à renforcer et amplifier au cours du prochain quadriennal, autour des quelques axes forts qui émergent du bilan. C'est autour des plateformes et des projets fédérateurs que nous projetons de structurer la politique scientifique auprès des équipes. Il s'agit pour nous d'aider à renforcer l'identité des équipes, de viser l'excellence thématique tout en contribuant à quelques uns des défis technologiques et sociétaux des STIC.

L'accès à l'information

Cette question, abondamment traitée, reste un des défis conceptuels importants. Depuis l'image d'un document d'archive numérisé, jusqu'au document structuré XML, en passant par une représentation numérique ad-hoc comme un document au format pdf, l'interrogation de ces multiples sources d'information pose des questions fondamentales à la reconnaissance des formes, à la recherche d'information, à l'informatique théorique, au traitement de la langue, à la représentation des connaissances et la modélisation de l'utilisateur. Plusieurs de ces questions sont abordées par les équipes Document et Apprentissage (DocApp), Traitement de l'Information en Biologie-Santé (TIBS), et Modélisation des Interactions et des Usages (MIU). Le projet de Plateforme d'Indexation Régionale (PlaIR) débuté en 2009 permet de mettre en synergie les contributions fondamentales des équipes en se donnant les moyens de les intégrer dans une plateforme technologique à fort potentiel d'innovation.

Plusieurs objectifs complémentaires sont visés. Tout d'abord la plateforme est un élément d'affirmation du potentiel d'innovation du LITIS dans le développement économique régional et même national (projet CISMEF, contribution au projet d'Equipex PLEIAD PLATEFORME d'Excellence Industrielle-Académique de Dématérialisation (porté par La Rochelle, partenaires : LORIA-LIRIS-LABRI-LI Tours-Paris5-LITIS). Par ailleurs, les premiers domaines d'utilisation ciblés (accès aux archives numérisées (archives de Seine Maritime), interrogation des sources juridiques) ouvrent des perspectives inter-disciplinaires riches susceptibles de rapprocher certains axes des Sciences Humaines et Sociales, et du Droit, dans le cadre du projet de PRES Normand notamment.

Enfin, cette plateforme ne sera pas sans retombée sur nos travaux fondamentaux. Par exemple, dans le domaine de la reconnaissance des formes et de la Recherche d'Information, si les techniques relevant de l'apprentissage occupent désormais une place importante, la question de l'adaptation des systèmes à de nouvelles données reste un point d'achoppement notamment par le manque de données annotées disponibles. Dans le cadre de PlaIR, l'accès à des données annotées par les utilisateurs permettra de mieux étudier l'apport des techniques d'apprentissage semi-supervisé, et toutes les questions relevant de la sélection de modèles avec peu de données étiquetées. Parallèlement, l'accès à des données d'interaction grâce à cette plateforme, devrait permettre de développer de nouvelles techniques de modélisation de l'utilisateur in situ, et de confronter les modèles cognitifs théoriques aux modèles expérimentaux. Enfin la plateforme PlaIR devrait également permettre d'intégrer les modèles de l'informatique théorique, automates, et techniques d'indexation compactes et rapides.

Dans le domaine de l'indexation, la plateforme PlaIR permet d'aborder la question de l'indexation multi-terminologique dans le domaine de la santé. Une généralisation du savoir faire de l'équipe CISMEF (Catalogue et Index des Sites Médicaux Francophones) à d'autres domaines (juridique, ingénierie...) est envisagée, ainsi que des extensions multi-langue.

Les projets INTERREG DocExplore 1 (2009-2010) puis 2 (2010-2013) portés par l'équipe DocApp sont l'occasion de développer une première collaboration avec les SHS de l'Université de Rouen, des collègues informaticiens de l'Université du Kent à Canterbury (avec lesquels nous avons maintenant une longue expérience de collaborations) ainsi que les archives de la Cathédrale de Canterbury.

Le traitement de l'information bio-médicale

L'une des spécificités du LITIS est son fort encrage pluri-disciplinaire alliant informatique et biologie d'une part au sein de l'équipe TIBS, et l'imagerie médicale d'autre part au sein de l'équipe QuantIF qui regroupe des médecins et des spécialistes du traitement d'images. Cette richesse est incontestablement un atout, sur lequel le laboratoire a investi avec l'appui de l'université de Rouen. L'équipe TIBS est l'équipe qui a le plus travaillé la pluridisciplinarité dès la création du LITIS, et avec un certain succès si l'on en juge par les résultats scientifiques. L'équipe TIBS est l'une des équipes constituant le futur Institut de Recherche et d'Innovations Biomédicales (IRIB, regroupement de l'IFRMP et de l'IHU) ce qui lui permettra de nouer plus naturellement de nouvelles collaborations.

Tout récemment (septembre 2010), l'équipe QuantIF a pu recruter un professeur spécialiste de traitement d'images (Su RUAN, section 61). Les collaborations inter-disciplinaires déjà amorcées entre QuantIF et l'équipe DocApp lors de la période écoulée vont donc pouvoir se développer à un autre rythme et au sein même de l'équipe QuantIF tout en maintenant des contacts étroits avec l'équipe DocApp. Soulignons par ailleurs les travaux de l'équipe DocApp en reconnaissance de formes pour des applications aux interfaces cerveau-machine. Sur cet axe thématique, le LITIS compte plus particulièrement contribuer aux techniques de séquençage à haut débit en biologie, aux techniques de fusion d'information en imagerie 3D/4D anatomique et fonctionnelle. Sur le volet médical, l'équipe QuantIF continuera ses développements en Imagerie fonctionnelle en TEP/TDM et TEMP/TDM, ainsi qu'en imagerie optique confocale. L'équipe QuantIF est impliquée dans le montage de la plateforme nationale pour l'étude des peptides (projet d'Equipex PIRCE). Elle est par ailleurs fortement impliquée dans l'animation du Cancéropôle Nord-Ouest.

L'intelligence Ambiante

L'intelligence ambiante constitue le troisième axe fort dans lequel le LITIS souhaite contribuer de manière renforcée au cours de la prochaine période. Dans ce vaste domaine déjà investi par de nombreuses équipes Françaises, nous pouvons situer les spécificités du LITIS. Depuis de nombreuses années, l'équipe STI a fait du véhicule intelligent son objet communiquant par excellence. La création du LITIS a été l'occasion de fédérer les compétences en vision embarqués (section 63 et 61) et les forces regroupées autour des bases de données (section 27) avec une spécificité sur les bases de données temps-réel. L'équipe STI porte donc aujourd'hui son attention sur l'ensemble du véhicule communicant avec une vision pluri-disciplinaire affirmée. Dans le cadre du pôle de compétitivité Mov'eo elle est impliquée dans le projet Mov'eoLab et la filière de développement du véhicule intelligent décarboné (demande d'un IEED)

Parallèlement, au cours de la période écoulée, l'équipe RIIC a mûri son projet scientifique en identifiant des applications porteuses de son expertise dans l'analyse des systèmes complexes. Elle porte aujourd'hui son attention sur la gestion des risques, la protection des populations, la modélisation des réseaux d'interactions, et l'étude des dynamiques urbaines. Fortement impliquée dans une collaboration avec des géographes elle positionne aujourd'hui ses contributions dans le domaine de l'intelligence territoriale dans le cadre du projet d'IRT Logistique, et en lien avec le pôle de compétitivité Nov@log.

Mise en œuvre

Comme évoqué précédemment, il s'agit pour nous de renforcer l'identité des équipes et de viser l'excellence thématique par un accompagnement dans le projet défini. Cette politique scientifique se décline assez logiquement en un ensemble d'actions qui seront développées en interne et avec le soutien attendu des établissements, et en externe dans le cadre d'actions régionales ou nationales.

Éléments de politique scientifique interne

Améliorer l'animation scientifique

– Séminaire du LITIS

Actuellement, l'animation scientifique est principalement réalisée au sein des équipes, par l'organisation de séminaires d'équipes. Cette situation, qui est en partie imputable à l'éloignement géographique, limite les échanges, les croisements interdisciplinaires, et donc l'émergence de collaborations internes. Pourtant, l'éloignement géographique n'est pas un frein à l'organisation des conseils de laboratoire mensuels entre Rouen et Le Havre. Il y a donc moyen de s'en inspirer pour mettre en place un rythme régulier de séminaires de laboratoire.

– Favoriser l'animation internationale

En l'état actuel les moyens propres du laboratoire sont concentrés sur le personnel de secrétariat, les frais fixes du laboratoire, les missions des équipes, le renouvellement des équipements. Les financements sur projets quant à eux permettent un apport supplémentaire d'équipements dits lourds (serveurs), le renouvellement partiel des équipements, et facilitent les déplacements en missions. 20% des moyens affectés aux équipes sont attribués sur projet. Ils sont consacrés en priorité à financer la participation aux écoles d'été et premières conférences des doctorants. Globalement cela laisse peu de levier pour impulser en interne des projets scientifiques de plus grande portée.

L'accroissement des moyens propres pourrait permettre d'impulser une politique scientifique plus volontariste, par exemple par l'affectation de moyens sur projets en visant plus particulièrement les échanges

internationaux que nous avons identifiés comme l'un des points faibles (financement de séjours courts à l'étranger de membres du LITIS, financement de l'accueil de chercheurs étrangers en complément du soutien des établissements).

Implication dans les formations

– *Masters*

Au cours de la prochaine période le LITIS sera porteur du master IGIS (Informatique, Génie de l'Informatique et des Systèmes) à l'Université de Rouen, et du master Informatique et Mathématiques à l'Université du Havre. Ces masters accueillent chaque année de nombreux étudiants de l'INSA (département ASI et GM) et de l'ESIGELEC (Ecole d'Ingénieurs en Electronique du campus consulaire). Les équipes du LITIS sont par ailleurs fortement impliquées dans la formation d'ingénieurs INSA (départements Architecture des Systèmes d'Information et Génie Mathématique).

Cette implication dans l'ensemble des formations de niveau Master permet au LITIS d'attirer de bons étudiants motivés par la recherche pour préparer un doctorat. Ceci n'interdit pas le recrutement de bons candidats externes.

– *Formation doctorale*

Le LITIS sera impliqué dans l'organisation de l'école doctorale SPMII par la participation de Abdelaziz Bensrhair. Le positionnement Haut-Normand de la formation doctorale ne devra pas être un obstacle à notre volonté de fédérer les travaux des équipes du LITIS avec celles du GREYC. Le projet de fédération Norm@STIC sera un outil important pour cela, auquel nous envisageons également d'associer certains moyens de l'école doctorale SPMII, par la proposition de sujets co-encadrés avec le GREYC. Une action réciproque pourrait être menée par les collègues du GREYC au sein de leur école doctorale SIMEN.

Les propositions des équipes de demandes d'allocations doctorales sont validées par le conseil de laboratoire. Les moyens étant restés modestes sur la période (2 allocations MESR par an en moyenne) nous mettons l'accent en priorité sur les sujets des équipes impliquant des jeunes collègues MCF dans l'encadrement, tout en veillant à une répartition équitable entre les équipes. Pour l'avenir, l'implication des sujets dans les trois axes de notre projet scientifique sera naturellement examinée.

Le LITIS est particulièrement vigilant au bon déroulement des thèses en cours. Ceci se traduit par l'organisation annuelle (juin ou septembre) d'une journée des doctorants qui mobilise l'ensemble des personnels du LITIS et dans laquelle les élus doctorants au conseil de laboratoire sont particulièrement impliqués. Les doctorants de première année présentent leur sujet sous forme de poster, tandis que les doctorants plus avancés présentent leurs travaux à l'oral. C'est une journée très riche qui permet de bien percevoir l'activité du laboratoire. Le conseil de laboratoire favorise le déplacement en conférences des doctorants, ainsi que leur participation à des écoles d'été. 20% du budget des équipes est consacré en priorité à ces actions. Les doctorants participent également à la formation dispensée au sein de l'Ecole Doctorale SPMII (Sciences Physiques Mathématiques et de l'Information pour l'Ingénieur). Les participations aux écoles d'été sont validées comme module de formation. Les doctorants, et les équipes du LITIS, participent à l'organisation des soutenances à mi-parcours de l'école Doctorale.

Soutien aux développements logiciel

Pour capitaliser sur ces développements logiciels le LITIS va renforcer son potentiel humain sur cette activité, mutualiser ses compétences et ses moyens et établir des partenariats notamment avec Sage et Scilab.

Politique de recrutement

Pour attirer de bons candidats le LITIS compte sur la pertinence de ses thématiques et sur le développement de sa politique de soutien aux nouveaux arrivants (moyens financiers et humains -bourse de thèse-). Il va aussi chercher à augmenter la proportion de ses recrutements extérieurs afin de bénéficier de la dynamique apportée par les nouveaux arrivants, notamment en soutien des équipes identifiées déficitaires, tout en veillant à pérenniser les spécificités des équipes et de capitaliser la compétence. Une gestion pluri-annuelle des demandes de postes a été mise en place afin d'anticiper les départs à la retraite. Nous veillerons à l'actualiser.

Gestion des moyens financiers

La gestion des moyens financiers comprend la gestion de la dotation du laboratoire et la gestion des contrats de recherches (industriels, ANR, etc...). Environ 60 % de la dotation est consacrée au financement des équipes dont 20% sont gérés sous forme de BQR par le conseil de laboratoire. La gestion des contrats occupe une part importante de l'activité du secrétariat. Un prélèvement de 10% des contrats industriels est mis en place par le laboratoire, ce qui permet notamment de recruter des moyens complémentaires en personnels de secrétariat.

Il serait souhaitable que l'activité financée par les projets ANR et génératrice de prélèvements des établissements (4% de frais de gestion + gestion du préceptif ANR) puisse profiter au laboratoire.

La démarche qualité en recherche

Le LITIS souhaite poursuivre dans l'expérimentation de la démarche qualité qu'il vient d'initier. Il veillera notamment à construire un tableau de bord d'indicateurs et l'actualiser périodiquement.

Éléments de politique scientifique externe

Politique scientifique

Le projet de fédération Norm@STIC est un projet que nous souhaitons mettre en place à l'horizon du parcours de la prochaine contractualisation. Ce projet pourra s'intégrer dans le PRES Normandie Université.

En région nous envisageons de continuer notre action structurante des STIC grâce aux Grands Réseaux de Recherche. L'articulation pluri-disciplinaire avec les différents réseaux régionaux (Sécurité Environnement Risque et Sciences Humaine et Sociale, Chimie Biologie Santé) sera renforcée sur chacun des trois axes de notre projet.

Notre implication dans quatre projets du Grand Emprunt devrait nous permettre d'amplifier notre positionnement et nos collaborations nationales en prenant appui sur les pôles de compétitivité. Elle devrait aussi renforcer les collaborations entre les équipes du laboratoire dans des secteurs clé. Le projet d'Equipex PLEIAD en ce qui concerne le document et l'accès à l'information, le projet d'Institut de Recherche et d'Innovations Biomédicales dans le domaine du biomédical et en ce qui concerne pour nous l'intelligence ambiante les projets d'IRT et d'Equipex en logistique et intelligence territoriale et l'IEED sur Satory sur le véhicule intelligent.

Valorisation

Le laboratoire continuera à s'impliquer au niveau régional dans de monde socio économique et notamment par le suivi des projets « Normandy living lab », « seine innopolis » et du cluster marétique opérations en cours. Il veillera à supporter les initiatives de transfert de technologie et d'innovation sur les projets dans lesquels il est impliqué. Par exemple par la participation aux « entrepreneuriales » (1er prix en 2009 sur un projet de numérisation porté par Damien Grosgean (étudiant du Master IGIS), actuellement doctorant au LITIS.

Diffusion de l'information

Le LITIS continuera aussi à participer à la fête de la science et aux manifestations liées à l'innovation dans le domaine des TIC comme « big tallent » et « le carrefour des possibles ».

SPECIAL LE HAVRE

La recherche met le turbo

Labos. Les chercheurs de l'université du Havre sont sur tous les fronts.



Éocomposites. Aircelle et le CNRT Matériaux créent de nouveaux matériaux pour les réacteurs d'avion.

Des réacteurs et des éoliennes bio
Il y a dix-sept ans, Joël Bréard fut de ceux qui firent du Havre une université pionnière dans la recherche sur les matériaux composites. Codirecteur du CNRT Matériaux, laboratoire commun aux universités du Havre et de Caen, à l'EnsiCaen et au CNRS, il pilote plusieurs projets dans le domaine des éocomposites. « Dans un contexte de développement durable, il faut repenser la mise en œuvre des matériaux composites et assurer la maîtrise de leur fin de vie », explique le chercheur. Le laboratoire travaille ainsi avec Aircelle sur des procédés moins coûteux et plus respectueux de l'environnement concernant les inverseurs de poussée des réacteurs d'avion. Il contribue également à la création de matériaux 100% naturels. Avec la société havraise La Tôlerie plastique (LTP), le CNRT a ainsi récemment été primé pour une éolienne à base de fibre de lin et de biopolymères. Il planche aussi sur la création d'éoliennes urbaines, sortes de cheminées à vent intégrées de façon innovante aux bâtiments. Cinq thèses sont par ailleurs en cours sur l'étude du lin. « Construire une éolienne ne suffit pas à faire de ces matériaux

des vecteurs de nouveaux marchés, explique Joël Bréard. Il faut prouver qu'ils ont des propriétés remarquables. »

Les cartes intelligentes du Litis
Cernée par seize établissements classés Seveso seuil haut, la population havraise serait elle bien protégée en cas d'accident ? Depuis trois ans, à l'université du Havre, l'équipe Réseaux d'Interactions et intelligence collective du Laboratoire d'informatique, du traitement de l'information et des systèmes (Litis) s'intéresse à la question. Les travaux sont menés en

partenariat avec le Conseil régional et la Communauté de l'agglomération havraise, qui dispose de cartes pour décider de l'évacuation ou du confinement de la population en cas d'incident. « Mais aucune de ces cartes ne prend en compte l'évolution des flux de population », souligne le chercheur Damien Olivier. Or la décision ne sera pas la même en cas d'embouteillage ou de perte de voies de communication. Grâce aux techniques d'intelligence collective et à de puissants calculateurs, le Litis imagine donc des cartes dynamiques prenant en compte l'évolution du réseau. Il a également mis au point un logiciel permettant de maintenir les communications en cas de destruction des infrastructures. Et il planche même sur la modélisation des phénomènes de panique.

Les innovations d'une shakespearieenne
Les sciences humaines aussi savent valoriser leurs recherches. Le projet Univerneen est la preuve. Lancé par Sarah Hatchuel, qui dirige depuis deux ans à l'université le Groupe de recherche Identités - Cultures, ce site reliera les chercheurs en arts aux acteurs de la vie culturelle. « Leur mise en relation ne fonctionne aujourd'hui que par réseau ou bouche-à-oreille », explique la spécialiste de Shakespeare. L'an dernier, Julien Bourguignon, responsable de l'action culturelle du festival Automne en Normandie, eût bien du mal à la contacter pour lui proposer de participer à un débat. C'est alors que leur est venue l'idée du projet. Le site, dont la mise en ligne est prévue début 2011, proposera un annuaire des universitaires et des professionnels de la région. Les premiers trouveront des contacts pour leurs colloques et leurs recherches d'archives ; les seconds, des chercheurs pouvant les aider dans l'écriture du programme d'un festival ou la préparation d'une mise en scène. Les deux porteurs du projet réfléchissent aussi à la mise en place d'un pass qui inciterait les étudiants à se rendre davantage dans les salles de théâtre. ■ **HUGUES DAVANT**

Annuaire. Être ou ne pas être sur le site Univerneen ?

Prévoyant. Damien Olivier modélise les flux d'évacuation en cas d'accident majeur.

**ATTENTION :
issu strictement interdit
(sauf en cas d'urgence)**

XVIII | 17 juin 2010 | Le Point 1970

Équipe C&A : Combinatoire et algorithmes

Auto-analyse

Essentiellement issue de l'ancien laboratoire LIFAR, l'équipe combinatoire et algorithmes est née de la fusion des différentes équipes du LIFAR (Laboratoire d'Informatique Fondamentale et Appliquée de Rouen). Elle est donc composée d'enseignants chercheurs en informatique fondamentale, mais qui travaillant sur des sujets divers. L'équipe possède néanmoins une culture scientifique commune et la découper en deux sous équipes nous semble arbitraire, d'autant plus que de nombreuses passerelles existent entre les différentes thématiques. Cependant, le projet de recherche qui suit présente un réel effort pour renforcer la cohérence scientifique.

Points forts de l'équipe

- L'équipe possède une reconnaissance internationale dans les domaines de la Combinatoire Algébrique et de l'étude des Automates finis ; De plus cette reconnaissance est en nette augmentation, comme en témoigne entre autre les invitations à des conférences. Elle est également bien reconnue sur le plan national, en particulier par son implication dans le GDR Informatique Mathématique.
- Entre 2006 et 2010, le nombre de publications est en très nette augmentation.
- La bonne implication des membres de l'équipe dans les filières d'enseignement permet de diffuser les connaissances et facilite la recherche de doctorants.
- Grâce au travail effectué sur le logiciel Sage, l'équipe est l'un des meilleurs laboratoires mondiaux dans le développement de logiciel d'expérimentation en combinatoire. C'est très certainement un important atout pour favoriser les collaborations internationales.

Points faibles

- Homogénéité des recherches de l'équipe : sur les quatre ans passés, il n'y a pas eu vraiment de projet de recherche commun ; quelques membres de l'équipe travaillent isolément ;
- La quantité de publications est très diverse d'un membre à l'autre de l'équipe. L'équipe tente de remédier à ces deux derniers problèmes en organisant des séminaires et des groupes de travail.
- Le nombre de doctorants de l'équipe est faible ; La principale raison est le manque de financements.

Opportunités

- Recrutement du professeur Jean-Gabriel Luque dont les thématiques de recherches recouvrent la combinatoire algébrique la théorie des automates et les applications à la physique. L'unité de l'équipe en est fortement renforcée.
- Des liens peuvent être consolidés avec le GREYC (équipe algo) en particulier dans le cadre de la fédération NORMASTIC et du GDR Informatique Mathématique.
- Suite à plusieurs départs en retraite, l'équipe devrait avoir des opportunités de recrutement en particulier de professeurs.
- Entre un et trois membres de l'équipe devaient passer leurs habilitations à diriger des recherches dans un avenir proche.

Risques

- Difficultés à trouver des financements pour les doctorants. Plus des deux tiers des diplômés du Master Informatique Théorique et Application trouvent une bourse de thèse, souvent dans des équipes concurrentes (IGM en particulier).
- Difficultés à trouver des contrats de type ANR. Les recherches fondamentales menées par l'équipe rendent encore plus difficile la recherche de contrats industriels.
- Le départ à la retraite de tous les professeurs sur un période courte fragilise fortement l'équipe.

Projets et objectifs scientifiques

L'équipe va poursuivre ses activités dans deux axes clairs de recherches : la théorie des langages et ses applications d'une part et la combinatoire énumérative et algébrique d'autre part. Elle prévoit également de favoriser les collaborations possibles entre les deux thématiques. Le thème automate prévoit de se concentrer sur les applications en vérification, en théorie des jeux et en apprentissage. La thématique combinatoire algébrique travaillera sur les fonctions symétriques et leurs généralisations ainsi que sur les applications en mathématique, en physique et en particulier sur l'information quantique.

Théorie des langages et Automates Le domaine de la théorie des langages est une thématique historique du LIFAR. Du point de vue de la recherche fondamentale, les problèmes étudiés portent sur le monoïde libre, les monoïdes partiellement commutatifs (les monoïdes de trace de Cartier-Foata) et la théorie des codes. Cette recherche a des applications à l'optimisation des programmes en compilation et en model checking. Cependant une importante part de l'effort de recherche est principalement axé sur les applications. Trois axes d'applications sont envisagés :

Le premier concerne les langages géométriques. Ce sujet est un sujet novateur, dont l'équipe est à l'origine. Leurs principale motivation était la généralisation d'une classe de langages utilisée en validation hors-ligne d'applications temps-réel. Ceci peut être fait en utilisant un modèle basé sur des langages rationnels. Mais, à cause d'une propriété particulière de ces langages, on peut aussi utiliser un modèle basé sur la géométrie discrète et diminuer ainsi drastiquement le temps de calcul. Le but est le suivant : si on peut interpréter en termes de langages les propriétés qui font que le modèle géométrique est si efficace, on peut alors espérer obtenir de nouveaux algorithmes sur les automates avec de meilleures complexités. Outre les questions de théorie des langages et de géométrie discrète, cette problématique pose également des questions de combinatoire énumérative (énumération de polycubes par exemple).

Dans une autre direction, l'équipe prévoit de continuer à explorer les applications en théorie des jeux et en apprentissage. Les méthodes à noyaux sont largement utilisées dans les techniques d'apprentissage statistique comme les SVMs (Support Vector Machines). Cortes, Haffner et Mohri ont introduit une famille générale de noyaux basée sur les transducteurs pondérés, les noyaux rationnels, qui étendent les méthodes à noyaux à l'analyse des séquences de longueur variable avec ou sans poids. Cette nouvelle technique de classification prend en compte la contrainte sur la longueur variable des séquences qui apparaît dans de nouvelles applications en traitement du texte et en reconnaissance de la parole. L'explosion du volume de données à traiter (extraction, classification, ...) dans des domaines comme le traitement automatique des langages naturels, l'ingénierie du web, ou encore l'ingénierie des documents digitaux, a suscité de nombreux travaux sur l'inférence grammaticale des séries rationnelles de mots et d'arbres. Il paraît donc d'un grand intérêt d'étudier les noyaux rationnels des langages rationnels stochastiques et ceux des langages rationnels d'arbres (pondérés ou non), en vue de l'apprentissage des familles de langages correspondants.

Les automates d'arbres à multiplicités sont une généralisation à la fois des automates de mots à multiplicités et des automates d'arbres. Ils sont parfaitement adaptés pour la modélisation des structures de données arborescentes (pages web, documents digitaux) et sont de plus en plus utilisés comme modèles de langages en traitement automatique des langages naturels. Les défis posés par l'analyse de grandes masses de données, en extraction automatique en particulier, ont suscité de nouveaux travaux sur l'inférence grammaticale de ces automates.

L'expérience montre que, pour un type d'objets donné, la mise au point d'une algorithmique efficace est indispensable pour construire des boîtes à outils de qualité. D'autre part la disponibilité de tels outils logiciels rend possible le développement de nombreux projets de recherche. L'objectif dans ce domaine est donc de développer une algorithmique efficace des automates d'arbres avec ou sans multiplicités et de construire une boîte à outils libre sur les automates d'arbres qui favorisera le développement de nombreux projets de recherche lié aux automates d'arbres.

D'autre part, grâce aux travaux liés aux algèbres de Hopf combinatoires, l'équipe a une grande expérience de la combinatoire algébrique des arbres et des séries d'arbres qui sont le pendant algébrique des automates à multiplicités. Des collaborations ont déjà commencées sur ce point et devraient permettre de focaliser les différentes forces de recherche de l'équipe.

Algèbre de Hopf combinatoire et calcul symbolique multivarié Cette recherche se situe à l'interface entre la combinatoire énumérative et algébrique et l'analyse d'algorithmes. La combinatoire énumérative

a pour but de compter le nombre de configurations d'une structure combinatoire donnée. Dans cette théorie, l'outil algébrique principal est la notion de série génératrice qui encode la suite des nombres de configurations en fonction de leur taille dans une expression algébrique de type polynôme, fraction rationnelle ou série. Le principal défaut de cette méthode est que l'information encodée dans l'exposant d'un monôme est très pauvre : c'est un nombre entier. En conséquence, avec la méthodologie des séries génératrices, on sait seulement répondre aux questions du type : quel est le nombre d'opérations si l'entrée de l'algorithme est de taille n où n est un entier ?

Pour s'affranchir de cette limitation, la combinatoire algébrique s'est intéressée depuis quelques années à l'étude d'objets algébriques appelés algèbres de Hopf combinatoires. Il s'agit essentiellement de polynômes ou séries dont les entiers qui sont en exposant des variables sont remplacés par des objets combinatoires tels que les partitions d'entiers ou d'ensembles, ordonnées ou non, les graphes ou encore les arbres binaires ou planaires. Hivert a montré en 2003 avec Novelli et Thibon que ces structures ont des liens très forts avec les algorithmes. En particulier, l'algorithme d'insertion d'un mot dans un arbre binaire de recherche intervient de manière cruciale dans la définition de plusieurs algèbres sur les arbres binaires.

L'effort de recherche porte sur deux points : pour le moment l'essentiel du travail dans ce domaine porte sur les calculs de produits. Pour les applications, il faut aussi réussir à définir sur ces objets algébriques nouveaux les opérations usuelles sur les polynômes comme par exemple les substitutions et les dérivations. D'autre part, si elles ont bien expliqué un certain nombre de résultats énumératifs (formules des équerres, énumération de tableaux à la Littlewood-Richardson), les algèbres de Hopf combinatoire ont pour le moment fournies relativement peu de résultats nouveaux. Des applications dans l'étude des algorithmes sont à rechercher. Un co-encadrement de thèse est en cours avec J.-C. Novelli de l'université de Marne-la-Vallée autour des questions d'équilibrage des arbres binaires.

Un autre volet du travail concerne les polynômes multivariés. En effet, ces derniers possèdent de nombreuses applications en cryptographie, en l'informatique quantique et en physique. En cryptographie, les résultats portent sur la complexité des attaques algébriques et en particulier l'algorithme F5 de J.-C. Faugère qui permet de résoudre efficacement les systèmes polynômes. D'autres travaux portent également sur les courbes elliptiques.

Le recrutement récent de Jean-Gabriel Luque a permis de développer le calcul des polynômes multivariés et applications à la physique. Il a décrit un algorithme, dont la complexité est indépendante du nombre de variables, permettant de calculer des intégrales de type Selberg. Cet algorithme, basé sur la combinatoire des polynômes de Jack, nous a permis de conjecturer des résultats sur l'asymptotique de ces intégrales lorsque le nombre de variables tend vers l'infini. Rappelons que les polynômes de Jack et de Macdonald sont le cœur de la plus grande part des efforts de recherche en théorie des polynômes symétriques, dont est issue la théorie des algèbres de Hopf combinatoires. Ces travaux ont des applications dans les domaines de la physique statistique (l'effet de Hall fractionnaire quantique) et des matrices aléatoires.

Les perspectives sont les suivantes : Outre la généralisation au cas général de la conjecture sur l'asymptotique des intégrales de type Selberg, il faut s'attendre à avoir des applications en théorie de l'effet de Hall quantique : Il existe des connections apparentes entre la théorie des polynômes de Jack symétriques invariants par translation et la théorie des polynômes de Jack non-symétriques singuliers. Nous envisageons d'étudier les liens entre les deux notions et d'étendre les résultats aux polynômes de Macdonald. Le but est de trouver et de décrire de nouvelles fonctions d'ondes "candidates" pour entrer dans la description de l'effet de Hall quantique fractionnaire.

Enfin, ces calculs s'appliquent très bien à l'étude de l'information quantique. L'expérience acquise permettrait de reprendre les calculs de classification des orbites des systèmes de bits quantiques, réalisés par Jean-Gabriel Luque et Jean-Yves Thibon il y a quelques années, en y intégrant des notions issues de la géométrie algébrique dans le but de simplifier les calculs et d'accélérer les algorithmes. Il faut également remarquer que les calculs d'invariants matriciels sont très proches de ceux mis en place dans la théorie de Valiant pour l'attaque du problème $P=NP$. C'est une perspective de recherche à long terme.

Mise en œuvre

Une bonne part de la recherche de l'équipe et en particulier en combinatoire repose sur le fait de pouvoir expérimenter sur des machines. Pour ceci il est crucial d'avoir une bonne plate-forme d'expérimentation. Forte de l'expérience de MuPAD-Combinat, l'équipe a décidé de porter le travail sur Sage qui est un logiciel libre.

Lancé par Nicolas Thiéry de l'université Paris XIII et Florent Hivert de l'équipe en 2008, le projet est en pleine croissance et de nombreux contributeurs le rejoignent régulièrement. Ils sont au nombre de 38 dont 19 contribuent très régulièrement à ce jour. L'activité du projet peut également se mesurer aux 4 ateliers de développement autour de la combinatoire qui ont été organisés cette année (CIRM-Marseille / Fields Institute-Toronto / Paris XIII / LACIM-Montreal). C'est certainement une très bonne opportunité pour créer de nouvelles collaborations internationales.

L'effort immédiat portera sur les problèmes suivants :

- énumération et génération automatique de structures combinatoires récurrentes : objets décomposables à la Flajolet, théorie des espèces, structures arborescentes ;
- monoides, automates finis, automates d'arbres ;
- algèbre de Hopf combinatoires, Fonctions symétrique usuelle et non-commutatives, fonctions quasi-symétriques.

Il est à noter que l'effort ne se limite pas à développer des algorithmes. Grâce à la souplesse du langage Python et au langage Cython, intermédiaire entre le C/C++ et Python, une bonne partie du travail consiste à intégrer dans Sage des bibliothèques déjà existantes. C'est déjà le cas par exemple avec l'intégration de Semigroupe de Jean-Éric Pin. C'est en ce sens que Sage n'est bien plus qu'un logiciel mais plutôt une distribution de logiciels pilotée par Python. Ce dernier point possède également un fort potentiel pour favoriser les collaborations.

Équipe TIBS : Traitement de l'Information en Biologie Santé

Auto analyse

Points forts

La nomination récente de J.-F. Gehanno sur un poste de PUPH permettra de mieux stabiliser l'équipe.

La bonne implication des membres de l'équipe dans les filières d'enseignement permet de diffuser les connaissances et facilite le recrutement de stagiaires et de doctorants.

La deuxième année de la spécialité Bioinformatique du master Biologie-Santé qui se déroule en deux ans en alternance nous permet de nouer des contacts avec les laboratoires, notamment locaux, qui recrutent des stagiaires ou des diplômés. C'est le cas par exemple sur un projet exome avec l'U614 de l'INSERM qui a recruté une diplômée (S. Coutant) de 2009 qui travaille dans nos locaux à Mt-St-Aignan ou même de J. Grosjean (même promotion) recruté dans l'équipe sur le projet PSIP.

Les nombreux projets financés au niveau national et européen de l'équipe lui assurent une bonne intégration dans le paysage français et européens de la recherche.

La pluridisciplinarité constituée par la provenance des membres de plusieurs sections CNU permet d'aborder les problèmes de différentes façons afin de mieux les résoudre.

La participation de l'équipe parmi les équipes constituant le futur Institut de Recherche et d'Innovations Biomédicales (IRIB, regroupement de l'IFRMP et de l'IHU) permettra à l'équipe de nouer plus naturellement de nouvelles collaborations.

La plateforme CISMef connue au delà du monde de la recherche constitue un des points forts de l'équipe.

L'équipe possède une reconnaissance internationale dans des domaines fondamentaux.

Une grande force de l'équipe réside dans le fait de posséder des membres associés actifs donc publiants.

Points faibles

Le départ à la retraite de Dominique Cellier risque de priver l'équipe de l'approche statistique.

L'éparpillement des sites d'enseignement pose un problème au quotidien.

Le caractère non pérenne de la majorité des membres de CISMef (documentalistes & ingénieurs) implique que la recherche de partenariats industriels qui a démarré depuis un an maintenant va s'amplifier dans le prochain quadriennal pour maintenir voire augmenter le nombre d'ingénieurs de développement au sein de CISMef (en mars 2010, pour les 24 prochains mois, nous avons obtenu un financement pour trois ingénieurs de développement et un post-doctorant).

Opportunités

L'émergence des technologies de Séquençage à Haut Débit (SHD) a influé un second souffle pour les applications de l'algorithmique du texte en bioinformatique et les technologies qui vont encore évoluer dans les années à venir accentueront les besoins des biologistes en solutions spécifiques.

L'achat de matériel de SHD envisagé à Rouen (ex IFRMP, futur IRIB) et Caen (Centre François Baclesse) et le fait que les biologistes sont conscient des énormes besoins en bioinformatique que vont générer les analyses de toutes les données issues des ces appareillages constitue une opportunité unique pour notre équipe.

Des liens peuvent être consolidés avec le GREYC (équipes TAL et algo) ce qui est rendu possible notamment grâce à l'intégration de Dominique Dutoit comme membre associé au sein de l'équipe.

La potentialité de recruter de nouveaux enseignants-chercheurs dans le prochain quadriennal pour renforcer l'équipe TIBS, permettrait d'effectuer une demande de labélisation INSERM au sein de l'Institut Multithématique Technologies pour la Santé, dirigé par Jacques Grassi¹.

Sur le plan stratégique, le versant « informatique médicale » de l'équipe se rapprochera en 2010 du Centre d'Investigation Technologique (CIC-IT) du CHU de Lille, spécialisé en « biocapteurs et e-santé : Innovation et Usages » (<http://www3.univ-lille2.fr/cic-it/>). Ce rapprochement s'effectuera dans le contexte du G4

1. www.inserm.fr/thematiques/technologies-pour-la-sante/enjeux et extranet.inserm.fr/instituts-thematiques/technologies-pour-la-sante/organisation-de-l-institut

(regroupement des quatre centres hospitalo-universitaires (CHU) de l'inter-région Nord : Amiens, Caen, Lille et Rouen).

L'équipe collabore déjà activement avec l'équipe INSERM UMRS872 équipe 20, dirigée par Mme Marie-Christine Jaulent (DR), notamment dans le projet Urgences financé par l'ASIP Santé. Cette équipe va dans l'avenir regrouper l'équipe du Pr Alain Venot de Paris 13 (avec laquelle nous collaborons dans le projet L3IM) et l'équipe du Pr. Rodrigues de l'Université de Saint-Étienne (avec laquelle nous collaborons dans le projet InterSTIS). Dans ce cadre, nous espérons renforcer nos collaborations avec cette équipe

Depuis mars 2009, le CHU de Rouen a embauché une personne en charge de la valorisation industrielle (Fabien Liéval), qui nous a permis en moins d'un an de finaliser le dossier de valorisation avec GSK et le projet PlaIR financé par des fonds FEDER. L'équipe a planifié une réunion mensuelle avec M. Liéval pour évaluer au cours du temps les opportunités avec les différents industriels régionaux, et (inter)nationaux, en particulier dans l'industrie pharmaceutique.

Risques

La forte concurrence internationale sur les sujets abordés par l'équipe constitue à la fois un danger et une forte motivation.

Projets et objectifs scientifiques

L'équipe TIBS va poursuivre ses activités dans deux axes clairs de recherche, à la fois sur des aspects appliqués et sur des aspects fondamentaux, pour les 4 ans à venir : l'analyse des données produites par SHD et la poursuite du développement de la plateforme d'indexation régionale (PlaIR) qui est en passe de devenir non seulement multiterminologique mais aussi multidisciplinaire. Une nouvelle direction sera le passage au multilinguisme.

Les interactions entre informaticiens médicaux et bioinformaticiens seront privilégiées.

Multiterminologies, multidisciplines, multilingues

Le principal axe de recherche de TIBS en informatique médicale n'est plus centré sur un catalogue de ressources, mais sur le développement d'outils et de méthodes innovantes, notamment en terme d'alignements, pour créer une plateforme d'indexation régionale, où sont intégrées les principales terminologies de santé disponibles en français qui peuvent interopérer sémantiquement (projet InterSTIS, thèse de T. Merabti). En mars 2010, déjà 20 terminologies y étaient intégrées et sont accessibles par un portail terminologique de santé (PTS) (<http://pts.chu-rouen.fr>), interfacé avec un serveur multiterminologique de santé (SMTS), développé en collaboration avec la société Mondeca et le LERTIM de l'Université d'Aix-Marseille (Michel Joubert et Marius Fieschi) depuis 2007. Intégré au PTS, un outil de connaissance contextuelle (CISMeF InfoRoute) a été développé (<http://inforoute.chu-rouen.fr>). Cet outil peut être intégré à de nombreuses applications des systèmes d'information hospitaliers et de santé. De nombreux alignements entre terminologies sont à réaliser et à intégrer dans CISMeF InfoRoute. Dans le cadre du projet PlaIR (Plateforme d'Indexation Régionale), financé par le Conseil Régional de Haute-Normandie et le fond européen FEDER, nous développons une plateforme multiterminologique multidisciplinaire, en intégrant les sciences de l'ingénieur et le droit du transport. Nous prévoyons d'ajouter une troisième dimension : le multilinguisme qui doit dépasser le bilinguisme français/anglais actuel. Plusieurs outils d'indexation multiterminologique (F-MTI et ECMT, projet ALADIN financé par l'ANR) ; l'extracteur de concept multiterminologique (ECMT ; <http://doccismef.chu-rouen.fr/Interpreteur.html>) permet de trouver les concepts terminologiques d'une phrase quelconque. Cette phrase peut provenir de n'importe quelle application distante, notamment de recherche clinique par un service XML et bientôt un Web service. Ces deux outils ont été largement évalués dans différents contextes et pour plusieurs terminologies. Plusieurs outils de recherche d'information multiterminologiques sont et seront évalués dans un double contexte (documentation – catalogue CISMeF – ou soin – dossier électronique du patient – ; thèse de S. Sakji & A.-D. Dirieh-Dibad). Ces outils ont été récemment intégrés à notre outil de recherche Doc'CISMeF, ce qui, à notre connaissance, l'a rendu le premier outil de recherche multiterminologique dans le monde de la santé. Dans les trois prochaines années, de nouvelles terminologies de santé seront intégrées au backoffice de CISMeF, ce dernier sera sans doute utilisé à l'échelle européenne (Norvège, Danemark) voire au Japon : le caractère multilingue deviendra un enjeu majeur des cinq prochaines années pour notre équipe. Les outils de recherche de notre équipe devraient être également utilisés par tous ces acteurs (étrangers en santé

multilingue et français multidisciplinaires). Des évaluations formelles seront entreprises à chaque étape (selon la discipline, selon la langue et selon le thesaurus). Dans ce projet, nous aborderons également l'intégration d'ontologies formelles, comme la FMA d'anatomie, qui vont nous poser des problèmes de représentation des connaissances différents de ceux posés par les terminologies et classifications. Il nous faudra adapter notre modèle générique pour pouvoir intégrer les ontologies dans le SMTS. CISMef continuera à être une application génératrice à la fois de nombreuses innovations, de recherche fondamentale dans les cinq prochaines années, dans une approche ascendante : les ressources terminologiques CISMef permettent de fournir des outils d'accès contextuel à la connaissance à partir d'application de gestion de dossiers patients ou de productions de soins, mais aussi peuvent être le support de nouvelles visualisations des éléments de dossier.

En outre nous faisons évoluer nos travaux d'une thématique sur les modalités de l'accès à l'information biomédicale à une approche, toujours bibliométrique, de l'analyse de la qualité et de l'utilisation de l'information, en particulier dans le domaine de la santé environnementale et professionnelle. Nous sommes en effet une des rares équipes à travailler sur ce sujet au plan international, et nous avons d'ailleurs instauré des collaborations avec l'équipe de Derek Smith, de l'université de Newcastle (Australie) et surtout avec l'équipe de Patrick Loisel, de l'université de Toronto (Canada), avec lequel un projet de recherche, financé par les CIHR Canadiens, vient de débiter sur le thème du transfert de connaissance dans le domaine de la prévention des troubles musculo-squelettiques. Des projets de partenariat avec le centre Cochrane de santé au travail d'Helsinki (Joss Verbeek) sont également actuellement en discussion. Une doctorante, médecin spécialiste de santé au travail, travaille actuellement sur ces questions qui nécessitent, outre une compétence en bibliométrie, une compétence médicale pour approcher à la fois les questions de rappel et de précision des recherches documentaires mais être également en mesure d'analyser la qualité et la pertinence médicale du résultat des recherches.

Séquençage à haut débit

La technologie de séquençage Sanger utilisée lors des 30 dernières années est peu à peu remplacée par les nouvelles technologies de séquençage à haut débit (SHD) (ou *Next Generation Sequencing* en anglais). Ces nouvelles technologies vont révolutionner la biologie et la médecine. En effet elles vont permettre de produire à des coûts abordables, même pour des laboratoires de petite taille, des quantités gigantesques de données dans des temps assez courts (10 à 15 jours en fonction des technologies en incluant les temps de préparation du matériel biologique à séquencer). Ces données se présentent actuellement sous la forme de millions de courts fragments (*reads*) de 50 à quelques centaines de bases pour les technologies les plus répandues. Il est donc absolument indispensable de pouvoir traiter ces données automatiquement. Même si les constructeurs de séquenceurs fournissent des pipelines pour traiter ces données, en fonction de la problématique biologique de départ, il existe un gap entre les sorties de ces pipelines et les résultats attendus par les biologistes.

Dans ce cadre nous avons déjà débuté une collaboration avec Dominique Campion (INSERM U614) sur un projet de séquençage d'exomes de huit individus précocement atteints de la maladie d'Alzheimer (avant 50 ans). En effet plutôt que de séquencer la totalité du génome il est possible de se concentrer sur l'exome constitué par les parties codant pour des protéines soit environ 3% du génome humain pour un peu près 180 000 exons connus. Des kits de biologie moléculaire existent qui permettent de capturer et séquencer environ 80% de ces régions d'intérêt.

Le but du présent projet (voir bioinfo.litislab.fr/EVA) est d'identifier des gènes candidats responsables de la maladie qui seraient altérés pour ces patients et exprimés dans le système nerveux central en se concentrant sur certaines variations ponctuelles ou SNP (Single Nucleotide Polymorphism).

On peut donc envisager de développer un outil générique indépendant de la pathologie étudiée. Pour cela on peut considérer les données issues du pipeline d'IntegraGen (www.integragen.com) qui effectue le séquençage et/ou les données brutes. D'un côté on peut s'attendre à environ 12 000 variations par individus et de l'autre il faut être capable de traiter des millions de courts fragments accompagnés de leur score de qualité (format *fastq*).

Les biologistes ont besoin d'outils spécifiques : il sera alors nécessaire de spécialiser les méthodes développées en fonction des technologies utilisées (Illumina/Solexa, Roche 454 ou Solid).

Dans le cadre d'un projet de transcriptome, nous collaborons avec Éric Rivals de l'équipe MAB du LIRMM. Les chercheurs de cette équipe ont déterminé qu'il est possible de faire la recherche des fragments sur le génome de référence à l'aide de segments de longueur 19 à 21 de manière exacte plutôt que d'utiliser la recherche approchée sur la totalité des fragments. En utilisant une structure d'indexation de type tableau des suffixes pour les fragments il est alors possible de les replacer sur le génome de référence et de discriminer plusieurs situations : erreurs de séquençage, SNP, phénomènes d'épissage...

Dans le cadre du projet 1000 génomes (www.1000genomes.org) il faudra être capable de stocker efficacement des données fortement similaires. Les structures de données compressées dynamiques peuvent servir d'appui pour résoudre cette question.

De plus dans le cadre d'une collaboration avec le Laboratoire de Microbiologie du Froid de l'université de Rouen (EA 4312) sur le thème « Perception de l'hôte par le pathogène opportuniste *Pseudomonas aeruginosa* : régulation de l'expression du gène *oprF* » nous mènerons une étude bioinformatique et biostatistique visant à analyser des régions promotrices de plusieurs gènes d'intérêt, à mettre en évidence des motifs communs sans *a priori* et à rechercher ceux-ci dans le génome complet bactérien. Ces travaux mettront en œuvre des méthodes d'alignements multiples, de construction de profils markoviens et leur utilisation pour la recherche de régulateurs. Les régulateurs sont des ensembles de gènes dont l'expression est dépendante du même facteur de transcription se fixant sur la région promotrice amont. La mise en évidence de régulateurs permet de mieux comprendre les circuits de régulations intimes qui gouvernent les réponses des bactéries à leur environnement. Cette étude conduira au développement d'un outil logiciel générique de recherche de régulateurs dans les génomes de bactéries.

Comme il n'est pas de bonne recherche appliquée sans bonne recherche fondamentale, nous poursuivons nos travaux d'informatique théorique et notamment :

- Nous continuerons notre collaboration avec Simone Faro de l'université de Catane en Italie pour étudier les problèmes de recherche de motifs dans des textes encodés avec des symboles encodés sur 2 ou 4 bits en étudiant la recherche multiple et la recherche approchée.
- Toujours en collaboration avec S. Faro, nous mettrons à jour le site « Exact String Matching Algorithms » et publierons une seconde édition du « Handbook Exact String Matching Algorithms » en incluant les algorithmes publiés dans les années 2000.
- Ce dernier site étant à caractère pédagogique, nous allons créer en collaboration avec S. Faro et Jan Holub (CTU Prague) un site regroupant les implantations les plus rapides connus des algorithmes de recherche exact de mots, afin d'offrir un large accès à ces techniques pour favoriser l'émergence de techniques encore plus rapides. Ce site sera hébergé à Prague.
- Nous continuerons l'étude statistiques de la répartition des répétitions, en particulier les répétitions chevauchantes en collaboration avec l'université de Tizi-Ouzou (Algérie).
- Nous poursuivons le travail commencé avec le MIS (Amiens) et le LIRMM (Montpellier) sur la caractérisation des palindromes.
- Nous avons initié un groupe de travail sur les périodes abéliennes.

Mise en œuvre

L'équipe TIBS poursuivra sa politique de R&D en répondant à certains appels d'offres spécifiques (ANR programme TecSan et PCRD).

En termes de politique de recrutement nous espérons recruter :

- en 2011, un maître de conférences sections CNU 26/64 pour remplacer Dominique Cellier ;
- en 2011, un maître de conférences section CNU 27 pour renforcer la partie « informatique médicale » ;
- en 2012, un ingénieur d'étude en bioinformatique pour mettre en œuvre les développements de l'équipe pour les équipes de biologie locales ;
- en 2013, un professeur d'université sections CNU 26/64 pour prendre en charge les aspects liés aux études statistiques de l'équipe ;
- en 2013, un professeur d'université section 27 pour prendre en charge une partie des activités d'informatique médicale.

Nous escomptons, dans le cadre du projet PlaIR, obtenir deux nouvelles allocations régionales pendant le prochain quadriennal et d'éventuels bourses CIFRE avec EADS et sa plateforme WebLab qui présente des synergies fortes avec notre plateforme PlaIR. Nous essaierons de continuer de privilégier les encadrements en co-tutelle avec des pays francophones. Deux nouvelles thèses ont démarré en 2010 :

- thèse de M. Zied Moalla sur les systèmes de question-réponses multilingues dans le domaine de la santé (co-tutelle avec l'université de Sfax, Tunisie). Zied Moalla est en cours d'évaluation des distances de Levenshtein et Stoilos dans notre algorithme de recherche d'information.
- thèse de Mlle Nazek Mohammad sur l'évaluation économique des systèmes d'information de santé.

Une collaboration avec le Pr. Christine Golbreich (LIRRM, Montpellier) a démarré en novembre 2009 sur la problématique d'OWLisation de la FMA avec OWL 2.0. Sur cette problématique des ontologies formelles, nous allons également travailler de nouveau avec Lina Soualmia, ancienne doctorante de TIBS, depuis MCF au LIM&BIO (Paris 13), dirigé par le Pr. Alain Venot. L'équipe TIBS est renforcée depuis l'arrivée en septembre

2008 d'Élise Prieur qui fait le pont entre bioinformatique et informatique médicale. Nous espérons pouvoir faire revenir des États-Unis Aurélie Névéol (en postdoctorat depuis février 2006 à la National Library of Medicine du NIH) au sein de notre équipe.

Les projets autour des nouvelles technologies de séquençage haut débit s'appuient sur plusieurs demandes de financement :

- une demande de financement conjointe avec le Professeur Costas Iliopoulos du King's College London à la Royal Society ;
- une demande de poste de Professeur invité pour un mois pour C. Iliopoulos à l'université de Rouen ;
- participation à une demande de financement de réseau au Leverhulme Trust ;
- participation à une demande de subvention d'équipement de séquençage à haut débit à la FRM (Fondation pour la Recherche Médicale) ;
- participation à une demande de subvention d'équipement dans le cadre des projets d'excellence « Equipex ».

De plus, avec O. Kulekci (Turquie), J. Tarhio (Finlande), J. Holub (République Tchèque), S. Faro (Italie) et É. Rivals (France) nous rédigeons un projet visant à développer des algorithmes de recherche de motifs efficaces sur des architectures multi-cœurs. Ce projet sera déposé dans le cadre du FP7 (probablement FET-OPEN).

Nous avons le projet d'accueillir pendant 10 mois un doctorant inscrit en deuxième année à l'université technologique de Hefei (Chine) pour travailler sur la recherche multiple approchée avec jokers.

Le comité de pilotage de l'*International Workshop On Combinatorial Algorithms* (www.iwoca.org) nous a confié l'organisation de cet évènement en 2013.

Équipe STI : Systèmes de Transport Intelligents

Auto analyse

En 2006, suite à la création du LITIS, les membres des anciens laboratoires, LIH (Université du Havre) et PSI (INSA de Rouen), se sont concertés afin d'unir leurs compétences. Nous nous sommes alors regroupés en 2007 dans l'équipe Systèmes de Transport Intelligents (STI), qui comprend des enseignants-chercheurs appartenant à plusieurs sections CNU (27, 61 et 63).

Points forts de l'équipe

- Mise en place de projets pluridisciplinaires : l'équipe utilise les compétences *Bases de données - Perception et Vision* de ses membres pour élaborer de projets pertinents
- Thèses en co-tutelles (Allemagne, Roumanie, Maroc, Tunisie) et d'autres collaborations internationales (Italie, Suisse, USA)
- Collaborations industrielles (Valeo, Egis-mobilité, Renault, Visionic, ...)
- Plusieurs brevets (nationaux et internationaux) déposés
- Membre fondateur de MOV'EO Lab. qui comprend, entre autres, les institutionnels (L'INRETS, L'INSA-Rouen, L'école des Mines de Paris, l'Université Saint Quentin en Yvelines), ainsi que des industriels comme Renault, Peugeot, Valeo, Véolia
- Complémentarité des compétences des membres de l'équipe
- Existence de liens transversaux (inter-équipes LITIS : avec MIU, RIIC, DocApp, Quantif)
- Collaboration avec le CETE-NC depuis plusieurs années : obtention en juin 2010, d'un projet ANR

Points à améliorer

- Trouver davantage de financements de thèses
- Augmenter le nombre de publications référencées
- Renforcer les co-directions de thèses entre les différents sous-thèmes de l'équipe (cf. suite du document)
- Proposer davantage de projets de type ANR
- Améliorer notre coopération avec des équipes de recherche à l'échelle internationale, en particulier européenne
- Encourager la soutenance de HDR de collègues maîtres de conférences (4 soutenances sont prévues pour le prochain quadriennal)

Opportunités

- Renforcement de nos activités au sein des pôles de compétitivité MOV'EO et NOV@LOG
- Participation au projet de création d'un IEED (Institut d'Excellence en matière d'Energie Décarbonnée) : Institut VeDeCoM (Institut du véhicule décarbonné communiquant et de sa mobilité), dans le cadre de nos activités au sein de MOV'EO Lab., en lien avec l'équipe Images du laboratoire GREYC (Caen) et l'équipe Documents et Apprentissage du LITIS
- Coopération avec le consortium LARA (INRIA, INRETS, ARMINE)
- Collaboration avec des équipes du laboratoire GREYC dans le cadre de la création du PRES Normand

Risques

- Cloisonnement des recherches lié à la nature multi-sites de l'équipe et à l'origine de ses membres.
- Orientation trop généraliste et plus assez spécialiste dans nos disciplines (27-61-63)

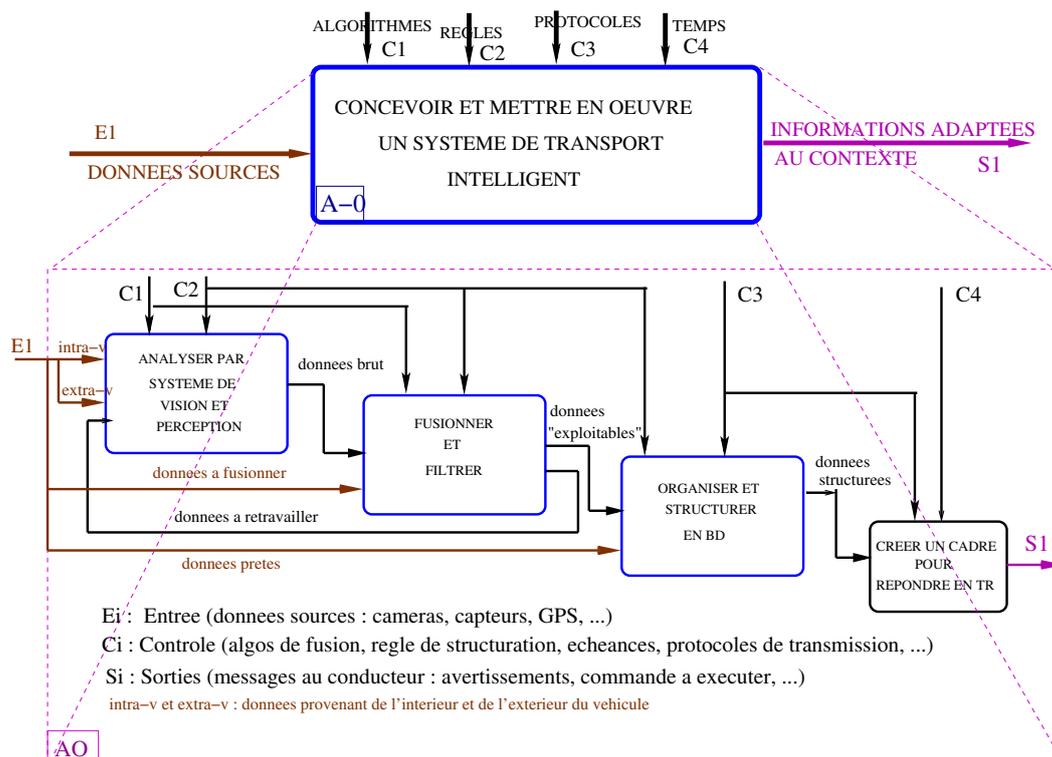


FIGURE 1 – Description des principales actions du projet de l'équipe STI

- Réorientation des recherches de l'équipe suite à la "recomposition" du paysage de la recherche en STIC en Normandie

Projets et objectifs

Introduction : Au sein de l'équipe STI, nous avons défini un projet scientifique et un projet stratégique.

D'une part, nous proposons d'adopter un projet commun associé à la chaîne de traitement des problématiques liées au transport, avec des applications orientées vers les systèmes embarqués en milieux dégradés, dotés de faibles infrastructures. Nous utilisons pour cela les informations issues de la perception de scènes routières qui seront structurées en bases de données temps réel. L'association de la perception de scènes et des bases de données, issue du rapprochement des deux équipes (anciennement Perception et SGBD temps réel), nous permet de développer une approche locale plus dense, et une modélisation globale de la scène routière plus robuste et ainsi de proposer une aide à la conduite et une analyse des situations accidentogènes dans les STI plus efficaces.

D'autre part, nous visons à positionner le projet stratégique de l'équipe parmi les projets de grande envergure du laboratoire LITIS et à faire reconnaître ses compétences en vision et communications dans les STI à l'échelle régionale, nationale et internationale.

Projet scientifique de l'équipe STI

Rappels : Sur le schéma de la figure 1, nous avons, au niveau supérieur, l'activité globale du système qui consiste à concevoir et mettre en œuvre un système de transport intelligent. Cette activité se décompose en quatre sous-systèmes interdépendants, dont les travaux de recherche correspondent aux trois sous-thèmes suivants qui décrivent le projet scientifique de l'équipe :

1. **Sous-thème 1** : Perception et l'analyse de scènes routières :

Concevoir un système de transport intelligent implique la construction d'un modèle de l'environnement qui soit précis, robuste et dynamique. Cette construction nécessite une perception la plus complète possible de l'environnement et requiert alors de nombreuses données sensorielles afin d'obtenir un résultat exploitable et fiable. Ces données peuvent émaner soit de capteurs différents (caméras, laser, GPS, capteurs infra-rouge, ...) embarqués dans un véhicule ou bien distribués sur plusieurs véhicules et/ou des éléments de l'infrastructure routière. L'objectif est de proposer des méthodes de fusion, de coopération et de traitement des données issues de ces capteurs homogènes ou hétérogènes pour la localisation, la détection d'obstacle, la détection et le suivi de véhicules, ...

Le projet s'articulera autour des problématiques suivantes :

- Nous souhaitons poursuivre les travaux que l'on a déjà initiés sur les capteurs stéréoscopiques et les étendre à une plus vaste gamme de capteurs visuels tels que les caméras omnidirectionnelles qui ont déjà démontré leur intérêt dans les ADAS. En effet, malgré la nature homogène de l'information traitée (l'image), une problématique importante consiste à adapter les différents traitements permettant la mise en commun des informations issues de ces différents capteurs.
- Un autre aspect concerne la fusion d'images classiques et infrarouges pour la détection d'obstacles routiers. L'utilisation conjointe des composantes couleur et infrarouge liées à chaque pixel, renforce d'une manière considérable la robustesse des algorithmes de détection de formes. L'une des problématiques importantes concerne alors la mise en correspondance des données entre les caméras.
- Dans le cadre de la coopération entre véhicules, nous souhaitons également développer les travaux initiés sur l'utilisation de la communication et des systèmes visuels pour des tâches telles que le suivi de véhicules ou encore la détection d'obstacles.
- Nous comptons poursuivre et approfondir les travaux sur la détection automatique du contexte routier (type de chaussée, détection des voies de circulation, conditions météorologiques telles que la pluie, le brouillard, ...), qui a déjà obtenu de nombreux résultats au sein de l'équipe.
- Modélisation de capteurs : la conception et la mise au point des applications mécatroniques multi-capteurs nécessitent de réaliser un univers de simulation particulier. Dans le domaine de la route intelligente, nous contribuons à la modélisation de capteurs de vision pour la conception de nouveaux *Simulateurs de systèmes d'aide à la conduite*. Ces simulateurs doivent, à partir d'un environnement connu, reconstruire les signaux capteurs synchrones avec une base de temps.

Actuellement, 5 thèses financées concernant ce sous-thème sont en cours.

2. **Sous-thème 2** : Bases de données temps réel pour la sécurité routière :

L'idée est de gérer toutes les données relatives à la sécurité routière à partir d'une base de données embarquée dans le véhicule. La première étape consiste à récolter les données depuis le véhicule et l'infrastructure environnante. Ces données peuvent être des images saisies depuis une (ou plusieurs) caméra(s) embarquées dans le véhicule. Des techniques de perception et d'analyse d'images nous permettent de sélectionner les données pertinentes qui seront ensuite stockées en base de données. D'autres données proviennent des capteurs dans le véhicule. Elles nous renseignent sur les valeurs de certains paramètres de conduite comme la vitesse, l'accélération, ... Ces données seront également stockées dans la base. L'une des problématiques est de définir un cadre pour le stockage des différentes données : certaines sont statiques comme l'identifiant du véhicule, certaines sont dynamiques ou temps réel, c'est à dire qu'elles varient avec le temps, à cause notamment de la mobilité du véhicule. Une fois que la base est constituée, il faudrait mettre à jour les données et exploiter la base. Là également, intervient le temps. En effet, certaines valeurs deviennent obsolètes avec le passage du temps, d'autres peuvent être tolérées même si dans l'absolu elles sont périmées. On utilise dans ce cas le critère de similarité qui consiste à tolérer une valeur pour une version de donnée à condition qu'elle soit similaire à la valeur de la version précédente. La définition du critère de similarité est dans ce cas primordiale. Par ailleurs, l'exploitation (manipulation) de la base nécessite également la prise en compte de contraintes temporelles pour certaines requêtes, qui doivent fournir un résultat avant une échéance donnée. Cette échéance est souvent dictée par la position du véhicule qui doit être dans un certain périmètre car au delà de ce périmètre

la réponse fournie par le système devient inutilisable. C'est en grande partie pour cette raison que des réponses approximatives ou incomplètes obtenues dans les temps peuvent parfois être préférables à des réponses exactes ou complètes obtenues en retard. En outre, on peut considérer qu'une partie de la base est une base de données de capteurs avec tous les problèmes qui y sont associés. On peut citer par exemple l'agrégation de valeurs de capteurs semblables qui doivent fournir une mesure homogène. 6 thèses de doctorat financées sont en cours sur ces aspects.

3. **Sous-thème 3** : BDTR pour une perception collaborative :

L'objectif de cette partie qui constitue une priorité de l'équipe consiste à associer plus étroitement les bases de données temps réel et la perception pour définir des méthodes collaboratives d'analyse et d'interprétation de scènes routières. En 2009, une thèse de Doctorat sur le sujet a été proposée. Le doctorant est Y. Regheb, financé par une entreprise allemande. L'idée de la thèse est d'exploiter les données provenant de plusieurs capteurs et caméras pour les structurer en base de données temps réel. L'objectif final est d'élaborer de nouveaux moyens pour la surveillance et la collecte d'informations du trafic routier (info-traffic). Deux principales technologies co-existent et semblent indispensables aujourd'hui pour la gestion de l'info-traffic : les réseaux de capteurs sans fils (WSN) et les réseaux de communication inter-véhicules (VANET pour Vehicular Ad-Hoc Network).

Dans les WSN l'utilisateur interroge le réseau et attend une réponse. Ceci nécessite la génération et le routage des données dans le réseau. Par analogie, ces aspects sont semblables à l'enregistrement des données et l'exécution de requêtes dans les bases de données, d'où, la notion de base de données de capteurs. Une base de données de capteurs (BDC) est une modélisation du réseau de capteurs en base de données virtuelle. On peut ainsi l'interroger grâce à un langage de type SQL ou un langage proche. Les données ne sont pas stockées dans cette base : elles sont instantanées, transitoires, et fréquemment modifiées. Généralement les capteurs sont alimentés par des batteries, qui pose le problème aux concepteurs de la gestion optimale de l'énergie. Il faut essayer d'effectuer le plus de traitements possible en local (au sein du capteur) car la transmission des données est l'opération la plus gourmande en énergie, notamment quand il s'agit de transmettre des quantités importantes. Ces caractéristiques peuvent être considérées comme un handicap pour la surveillance du trafic routier qui est très dynamique et demande la gestion d'une grande masse de données. Dans la technologie basée sur les VANET, il s'agit de développer un système de communication inter-véhiculaires afin de permettre la distribution rapide de données pour le bénéfice des utilisateurs en matière de sécurité et de confort (une thèse de doctorat est en cours sur cet aspect). Ces recherches se basent principalement sur le potentiel des systèmes de transport actuels qui ont des fonctions évoluées : des capacités sensorielles, de la mémoire, de la communication et du traitement de l'information. Loin des limites de l'énergie, les VANET peuvent être considérés comme des extra-WSN où le seul souci est la bonne gestion de l'aspect transitoire et variable des données.

Cette vision de ces deux domaines nous a permis d'aboutir à l'élaboration de notre projet qui consiste à concevoir une nouvelle approche pour la gestion de l'info-traffic en temps réel en exploitant nos recherches sur la conception de bases de données temps réel, appliquées à l'info-traffic (BDITTR). Le travail consiste à développer une nouvelle approche pour la collecte de l'info-traffic en vue de la détection des situations d'engorgement et autres incidents sur les routes. L'approche que nous proposons se base sur l'exploitation des potentiels des bases de données de capteurs et des systèmes de transport intelligents (VANET). Cette approche nous a permis de définir dans un premier temps une architecture à trois couches qui distingue trois niveaux de traitement des données dans les VANET. Cette représentation en couches permet de bien définir le type des données générées et le traitement à effectuer dans chaque couche, ce qui permettra d'identifier les applications potentielles. Cette architecture apporte une contribution dans l'exploitation de l'infrastructure existante (WSN, VANET), offrant ainsi une information pertinente en temps réel sur le trafic routier qui servira pour plusieurs applications telles que la détection des situations d'engorgement ou autres. Notons que les sources de données sont également fournies par le travail en amont sur la perception et la vision dans les véhicules depuis des caméras embarquées et/ou des systèmes de localisation tels que le GPS.

À terme, d'autres thèses seront proposées sur ces aspects qui intègrent les bases de données temps réel et la perception. Par ailleurs, des collaborations avec l'équipe RIIC du LITIS et avec des membres de MOV'EO Lab. sont prévues.

Projet stratégique de l'équipe

Les trois activités de l'équipe se regroupent à travers de nombreux projets menés ces dernières années. Nous trouvons ainsi les projets FUI, tels que EMOTIVE, PUMAS, Co-Drive, ou ANR comme Metramoto :

- Dans le projet EMOTIVE, il s'agit d'enrichir une base de modèles de capteurs pour y introduire de nouveaux modèles de capteurs optiques paramétriques et engager une réflexion sur les méthodes de calibration et de validation des modèles.
- Dans le projet PUMAS, l'idée est de créer une plateforme de renseignement de temps de parcours pour les villes et les agglomérations, à mettre, à terme, sur le marché international.
- Le projet METRAMOTO s'intéresse aux thèmes de la *sécurité routière*, appliquée aux *deux roues motorisés (2RM)*. Il vise à proposer des solutions technologiques pour une meilleure prise en compte de ces usagers vulnérables.
- Dans le projet CoDrive, il s'agit d'embarquer dans les véhicules des boîtiers *intelligents* où on y reporte des informations issues du véhicule (vitesse, temps de parcours, freinage d'urgence, déclenchement air-bag, ...) et d'autres issues de l'infrastructure (télépéage, signal d'obstacle, ...). Ces données sont ensuite compilées et interprétées pour transmettre au conducteur les actions à entreprendre.

De manière plus générale, nous souhaitons :

- Appliquer, évaluer et valider nos recherches en participant à des projets de recherche nationaux et internationaux où les mots clés principaux sont mobilité, sécurité, transport.
- Développer une plate-forme logicielle commune à tous les membres de l'équipe, exploitable de façon modulaire.
- Alimenter la plate-forme logicielle multi-sites par des données de différentes natures (acquises depuis des caméras, des capteurs, ou par GPS) : images, valeurs de capteurs, ...
- Gérer de manière efficace ces données (les organiser et les structurer en base de données) : organisation optimisée et manipulation en tenant compte des contraintes temps réel, de mobilité et géolocalisation.
- Créer des outils et des bibliothèques communs pour la présentation, le calcul, la simulation, ...

Les points suivants décrivent les orientations à long terme de l'équipe :

- Nous nous appuyons sur les Projets INSA Certifiés (PIC) qui permettent, dans le cadre de la démarche qualité ISO 9001, le développement de prototypes avec un partenaire industriel (VALEO, EGIS-mobilité, Techniques Nouvelles, ...)
- Nous comptons approfondir notre coopération avec des équipes de l'INRETS, de l'INRIA-Rocquencourt et avec les membres du pôle MOV'EO. Dans ce cadre, on a déjà commencé à mettre en place un laboratoire commun (MOV'EO Lab) comportant un centre d'essai de véhicules et des pistes, dont les membres de l'équipe peuvent disposer.
- Nous sommes en train de discuter pour établir des projets internationaux, principalement européens, par exemple, avec l'Italie et l'Allemagne.

Mise en œuvre

Des moyens importants, pour renforcer l'équipe, seront demandés dans le cadre de ses activités futures au sein du laboratoire MOV'EO Lab. et sa participation en réponse à l'appel d'offre relatif aux investissements d'avenir du projet IEED *Institut du Véhicule Décarboné Communicant et de sa Mobilité* (VeDeCoM). Cet institut est l'outil qui va permettre à l'industrie automobile française de relever le défi et de viser un leadership mondial pour une mobilité individuelle durable, décarbonée, sûre, fluide et connectée et à coût abordable. Les partenaires impliqués dans ce projet, dont le LITIS, partagent la vision européenne de la *Green car initiative* dont l'ambition en R&D sera d'accélérer la mise sur le marché de véhicules électriques et hybrides rechargeables dotés de l'intelligence embarquée et des moyens de communication avec les infrastructures les plus performants. Une des thématiques prioritaires de cet institut, intitulé *Usages et mobilités*, sera au centre des préoccupations de notre équipe. Elle concerne l'étude des paramètres nécessaires pour la mobilité, la communication et coopération infrastructure-véhicule ainsi que les systèmes de conduite assistée ou automatisée et de mise en convoi de véhicules. Notre collaboration scientifique suivie et de longue date avec Valeo

et le consortium LARA est un atout majeur pour la pérennité de nos activités de recherche au sein de ce groupe.

Pour la mise en œuvre de ce projet, en plus des moyens dont nous disposons (10 permanents, 2 associés, 13 doctorants et des stagiaires INSA et Master, ainsi que du matériel et logiciels, notamment des PC portables et quelques logiciels dédiés), nous envisageons de faire l'acquisition :

De moyens matériels :

- un nouvel environnement matériel : Ordinateurs PC, logiciels d'acquisition d'images, caméras, et améliorer les fonctionnalités d'une moto instrumentée (dont nous disposons en coopération avec le CETE-NC). La plupart de ces moyens seront financés dans le cadre des projets.

De moyens humains

- Il s'agit de renforcer notre potentiel en moyens humains : recrutement de 2 ou 3 professeurs et plusieurs maîtres de conférences. Nous envisageons également le recrutement d'un gestionnaire de projets pour renforcer la partie administrative, ainsi qu'un ingénieur de recherche et des post-doctorants. Certains doctorants et tous les post-doctorants demandés seront financés par les projets.

De locaux

- nous pensons dédier des locaux spécifiques sur le site du Madrillet (INSA et Université de Rouen), pour assurer la proximité avec les sites du Havre et de Satory.

Équipe QuantIF : Quantification en Imagerie Fonctionnelle

Auto-analyse

Points forts

- Jeune équipe dynamique, avec une production scientifique et un rayonnement international significatifs.
- Équipe pluridisciplinaire, très complémentaire, ayant les ressources humaines et matérielles pour aller de l'acquisition d'images médicales chez l'homme et l'animal jusqu'à leur traitement et analyse.
- Mise en place d'un M1 de « traitement et analyse des images médicales ».

Points faibles

- Petite équipe avec peu de chercheurs statutaires et sans ingénieur de recherche rendant plus difficile la conduite de travaux fondamentaux en traitement d'images et de collaborations avec d'autres équipes du LITIS.
- Thématique concurrentielle sur le plan international et national, en particulier avec les équipes d'imagerie de Caen.
- Exigüité des locaux.
- Pas de M2 de « traitement et analyse des images médicales ».

Opportunités

- Recrutement d'un PU61 en septembre 2010 (Pr Su Ruan) permettant de renforcer le développement de travaux fondamentaux en traitement et analyse d'images, ainsi que les collaborations avec les autres équipes du LITIS.
- Opportunité à terme de promotion en PUPH pour deux membres associés actuellement PH (Dr C. Savoye-Collet et Dr M. Salaün), et d'une promotion universitaire pour le Dr P. Bohn (pharmacien).
- Contact pris avec le président de l'université de Rouen pour la création d'un poste d'IATOS sur le quadriennal 2012-2016.
- Développement du PRES normand devant favoriser et renforcer les collaborations existantes avec Caen dans le cadre du Cancéropôle Nord-Ouest.
- Développement de la plateforme d'imagerie du petit animal par le renforcement de la collaboration avec l'Institut Fédératif de Recherches Multidisciplinaires sur les Peptides et l'Unité INSERM 644 (Pharmacologie des dysfonctionnements endothéliaux et myocardiques) et les demandes de financement de matériels dans le cadre du Grand Emprunt.
- Structuration de l'axe 3 du Cancéropôle nord-ouest et grand ouest sur la thématique poumon.
- Faisabilité de création d'un M2 pendant le plan quadriennal 2016-2020 en commun avec Caen et/ou avec Amiens à étudier.
- Continuer l'effort de convergence de la thématique scientifique engagé depuis 2005, avec la restructuration vers une thématique unique.
- Continuer de viser l'augmentation de la qualité scientifique des publications par l'augmentation des facteurs d'impact.

Risques

- Fragilité numérique des chercheurs statutaires pouvant mettre en cause la très forte complémentarité et pluridisciplinarité de l'équipe.

Projets et objectifs scientifiques

Contexte général

À ce jour, des convergences thématiques fortes se sont mises en place entre l'équipe des Pr P. Vera et L. Thiberville. Au contraire, les convergences sont de plus en plus grandes entre les thématiques du Pr J.-N. Dacher et de l'Unité INSERM U644. Ainsi, nous avons proposé que le Pr J.-N. Dacher rejoigne l'U644 et que les Pr L. Thiberville et P. Vera construisent un projet commun autour de **l'imagerie en 4D, en particulier au niveau du poumon**, mais également au niveau de l'imagerie abdomino-pelvienne.

Sous l'angle médical et biologique, nous avons le projet de développer des **marqueurs moléculaires fluorescents et radioactifs pour le diagnostic des pathologies du poumon profond**, utilisables en parallèle de l'imagerie endomicroscopique confocale et de l'imagerie TEMP ou TEP. La preuve du concept a été montrée récemment avec le développement d'un traceur de l'infection aspergillaire (projet en collaboration entre le Dr P. Bohn et le Pr L. Thiberville). Nous envisageons dans les quatre ans, d'utiliser ces traceurs chez l'homme dans le cadre d'étude de phase I (unité de phase précoce spécifiquement développée en 2010 au FNCLCC Henri Becquerel dans le cadre d'un appel d'offre INCa en 2010²).

Afin de développer ce projet nous allons répondre avec les équipes de l'IFR-MP23 à l'appel d'offre « Infrastructures Nationales de Biologie et Santé », dans le cadre du Grand Emprunt, pour compléter notre plateforme d'imagerie in vivo - microTEP et imageur optique (Molecular imaging platform for the development of radioactive and optical labelled peptides dedicated to diseases diagnosis and treatment monitoring, dans le cadre du projet plus large intitulé « Peptide Investigation Resource Center »).

L'Imagerie fonctionnelle en **TEP/TDM et TEMP/TDM** et **l'imagerie optique confocale** seront utilisées pour étudier des maladies du poumon chez l'animal et chez l'homme, et plus particulièrement pour caractériser les lésions cancéreuses à travers de nouveaux traceurs radioactifs et optiques. Les méthodologies développées serviront également pour certaines pathologies digestives.

Afin d'atteindre ces objectifs, il est impératif de développer de nouveaux algorithmes en traitement d'images. Les thématiques de notre équipe dans ce domaine seront donc en interaction très forte avec les médecins de l'équipe pour développer des méthodologies à la fois fondamentales et applicatives afin de répondre aux besoins de leur recherche en médecine.

Imagerie fonctionnelle en TEP/TDM et TEMP/TDM

Les études cliniques en cours (RTEP2 et RTEP3, Dr Edet-Sanson) vont se poursuivre. Dans les suites de nos travaux sur l'hypoxie et la prolifération tumorale (étude RTEP4), nous déposerons un projet national sur l'intérêt de l'imagerie TEP/TDM de l'hypoxie pendant la radiothérapie pulmonaire (PHRC2011 - Pr P. Vera, Pr B. Dubray, Pr L. Thiberville). Grâce à la présence du Dr C. Savoye-Collet et de son expertise en imagerie radiologique digestive et du petit bassin, nous pourrions étendre la thématique de l'optimisation de la balistique en radiothérapie aux cancers du rectum, de la prostate et du col de l'utérus.

Grâce à la microTEMP/TDM (et au projet d'anneau TEP petit animal), les nouveaux traceurs de l'hypoxie et de l'angiogénèse développés par l'IRCOF (UMR6014) devront être étudiés chez l'animal (y compris pendant une irradiation et après traitement par thérapie moléculaire ciblée - Projet du Dr S. Becker conjoint avec le Dr P. Bohn et le Pr L. Thiberville). Ce projet continue à l'heure actuelle sous la forme d'une collaboration entre l'IRCOF et l'IMRE (UCL, radiochimistes, Pr V. Gregoire). Parallèlement la possibilité offerte par le microTEMP/TDM d'étudier l'inflammation in vivo en association à une analyse morphologique TDM sera aussi envisagée à différents étages du tube digestif (projets développés par le Dr C. Savoye-Collet - collaboration ADEN EA4311). Les résultats obtenus en microTEMP seront confrontés à l'IRM (projet M2 co-encadrement ADEN).

Imagerie optique confocale

L'imagerie moléculaire *in situ* et *in vivo* des lésions cancéreuses broncho-pulmonaires proximales et distales se poursuivra. Ce travail de recherche, initié par le Dr M. Salaün au Fox Chase Center en 2009-2010 (Philadelphie, USA) sera prolongé au laboratoire dès le mois de novembre 2010. Il sera réalisé à partir d'un modèle murin d'adénocarcinome pulmonaire. Son objectif est de décrypter les phases précoces de la cancérisation du poumon distal, et de déterminer les marqueurs moléculaires de progression, qui pourraient alors être mis en évidence chez l'homme grâce à la technique d'imagerie par alvéoscopie développée dans l'équipe.

2. <http://www.e-cancer.fr/les-appels-a-projets/aac-labellisation-centres-l-phase-precoce-r>

Nous continuerons la validation d'une technique d'imagerie endomicroscopique multiphotonique. Ce travail financé par l'ANR, est une collaboration avec l'équipe du Dr G. Bourg Heckly, CNRS Paris 6, l'entreprise Mauna Kea, et l'Université de Limoges (Xlim). Il testera sur différents modèles animaux un système de microscopie confocale non linéaire, ou multiphotonique, dont un premier prototype sera disponible au début de l'année 2011 (Pr L. Thiberville).

En clinique humaine, notre recherche sera orientée vers l'évaluation et la mise au point de techniques endoscopiques innovantes d'exploration du poumon :

- Imagerie microscopique confocale *in vivo* des pathologies alvéolo-interstitielles du poumon profond. Il s'agit de la poursuite du travail en cours qui consistera à valider la sémiologie des images de microscopie confocale *in vivo* dans les pathologies pulmonaires périphériques (classification) et de tester leur valeur pronostique.
- Caractérisation des altérations microemphysémateuses du poumon profond par analyse de la microcirculation alvéolaire, *in vivo*, *in situ*. La trame capillaire étant accessible à l'examen confocal endoalvéolaire *in vivo*, nous souhaitons tester cette technique en comparaison des données fonctionnelles respiratoires, des tests de diffusion alvéolocapillaires et de la quantification de l'emphysème sur les images de TDM thoracique.
- Enfin, nous appliquerons à l'homme les marqueurs moléculaires et de traceurs fluorescents mis au point sur différents modèles animaux, pour le diagnostic des affections du poumon profond. Ce travail se fera en collaboration avec l'entreprise Quidd (implantée à Rouen), pour la production des traceurs et leur valorisation industrielle éventuelle.

Traitement et analyse d'images médicales

L'analyse des différentes images médicales de tomographie (TEP, TEMP, TDM) ou imagerie optique, chez l'homme, et maintenant chez l'animal, nous ont permis d'identifier trois problématiques en traitement d'images : la nécessité d'identifier de manière pertinente des zones d'intérêt (*segmentation*) sur des images multiples et complémentaires (*fusion d'informations*), et de classer des images pour aboutir à un meilleur diagnostic (*classification*).

Segmentation

Le but de notre segmentation est de séparer les tissus sains des tissus pathologiques sur des images de tomographie d'émission (TEP) et de transmission (TDM) en trois et quatre dimensions. Dans la suite de nos travaux sur la segmentation en TEP des hyperfixations au ^{18}F FDG, nous poursuivrons par la segmentation des lésions tumorales hypoxiques en imagerie TEP au F-Miso (thèse de P. Onoma, Directeurs Pr P. Vera et I. Gardin). Le challenge et la différence majeure entre l'observation des lésions tumorales au ^{18}F FDG et au F-Miso est le très faible contraste des images au F-Miso (rapport 1,3 entre les tissus sains et pathologiques vs 5 pour les images ^{18}F FDG). Ce travail sera abordé sous l'angle des fonctions de croyance et de la marche aléatoire. L'amélioration du contraste et de la quantification des données TEP est également en cours, par correction de l'effet de volume partiel, particulièrement important en imagerie d'émission, surtout si l'on considère le mouvement respiratoire (thèse de J. Anouan, Directeurs S. Hapdey et I. Gardin).

Par ailleurs, nous prolongerons nos travaux de segmentation des tissus sains par l'utilisation des fonctions de croyance sur les images TDM du thorax. Nous avons pour but de compléter la segmentation automatique déjà réalisée des deux poumons, de la trachée et du canal médullaire, par des organes plus difficiles à segmenter comme l'œsophage. Compte tenu de la forte expertise médicale de l'équipe, notre approche consistera à utiliser des connaissances a priori (connaissance médicale et atlas anatomique de référence) par modélisation de formes et de relations spatiales et fonctionnelles. Un grand nombre de problèmes en segmentation peuvent se reformuler en termes de minimisation d'énergies. Notre démarche sera de concevoir et étudier des critères adaptés dans des situations générales ou spécifiques à nos données en intégrant des modèles de connaissance a priori de manière intelligente et efficace. Nous étudierons également des méthodes d'optimisation récentes, telles que le graph-cut et la marche aléatoire afin de réaliser de manière automatique ou semi-automatique la segmentation des différents organes d'intérêt (c'est à dire les organes à risque en radiothérapie externe) et des tissus pathologiques (tumeurs et ganglions). Une thèse va débiter sur ce sujet (D. Grosgeorge, Directeurs Pr S. Ruan et C. Petitjean).

Ces méthodes devront permettre de segmenter les tissus sains, principalement sur les images anatomiques et les tissus pathologiques sur les images fonctionnelles. Les mouvements des organes devront être pris en compte pour ces segmentations. L'objectif final est que plusieurs méthodes de segmentation des organes sains

puissent être disponibles, en fonction de la complexité des organes à segmenter, ceci de manière simultanée sur une même console.

Nous chercherons également à développer le transfert industriel de nos travaux de segmentation. Une collaboration avec la société Aquilab (Lille, France) est en cours afin de valoriser les travaux de l'équipe dans le cadre d'un projet européen EuroStars en collaboration avec les équipes de l'université de Fribourg (Dr U. Nestle, Allemagne), la société UMANIA (Italie) et l'Unité INSERM U703 (M. Vermandel Lille - Institut de Technologie Médicale).

Fusion de données

Les images 3D/4D anatomiques et fonctionnelles sont autant de sources d'informations contributives à la prise de décision médicale. Chacune d'elles est porteuse d'une information présentant son propre degré de pertinence et de fiabilité. La prise en compte des informations hétérogènes nécessite une modélisation en adéquation avec la problématique posée, suivie d'une fusion d'informations pour en unifier la cohérence et l'interprétation. L'équipe poursuivra ses travaux sur la théorie de l'évidence et des fonctions de croyance dans le domaine de la fusion d'informations. Les travaux se situeront dans le cadre d'aide à la décision médicale lors de la délimitation du volume cible tumoral en radiothérapie externe. Ils porteront sur la fusion à partir d'informations anatomiques (TDM), d'informations du métabolisme du glucose (imagerie TEP au ¹⁸FDG), de la prolifération cellulaire (imagerie TEP au ¹⁸FLT), et de l'hypoxie (imagerie TEP au F-Miso). L'accent sera mis sur la modélisation et l'extraction des informations (thèse de B. Lelandais, Directeurs I.Gardin et L. Mouchard).

Au cours du prochain quadriennal, nos travaux de fusion d'informations continueront sur l'étude des fonctions de croyance, notamment la fusion de Dempster-Shafer dans les modèles de Markov cachés (HMM). Cela permet de combiner la généralité de la fusion de Dempster-Shafer et l'efficacité des HMM pour traiter nos données qui sont de nature multi-modalités et fortement corrélées entre les points. Ces travaux se feront en collaboration avec le Pr W. Pieczynski (École TELECOM SudParis).

Classification d'images

L'analyse des images en microscopie confocale fibrée est un sujet important dans l'étude du thorax. Cette technique d'imagerie étant nouvelle, il n'existe pas de réelle expertise sur ces images, qui présentent une forte complexité. Devant la difficulté de leur interprétation, nous concevons et développons, dans cette thématique, des outils efficaces d'aide au diagnostic, basé sur une classification automatique des images alvéoscopiques. Les algorithmes développés s'appuient, comme tout système de reconnaissance de formes, sur une étape de caractérisation des images et une étape de classification. L'objectif ici est de trouver les descripteurs les plus pertinents permettant de discriminer au mieux les images de sujets sains de sujets pathologiques. Les classificateurs utilisés sont développés en collaboration avec l'équipe Document et Apprentissage et font actuellement l'objet d'une thèse (C. Désir, co-encadrement C. Petitjean, Pr L. Thiberville, Directeur Pr L. Heutte).

Mise en œuvre

Moyens humains

- Obtention d'un poste de MCU 61 et un (idéalement deux) postes de PU (section 61 et/ou 27) dans le quadriennal
- Nomination de deux PU-PH (Dr C.Savoye-Collet et Dr M.Salaün), obtention d'un poste universitaire pour Dr P.Bohn.
- Ingénieur de recherche pour l'imagerie microTEMP/TEP/TDM et optique
- Doctorants et post-doc dans le cadre des projets ANR, appels d'offre divers. Un plan de financement des doctorants et post-doc est en cours de développement au CLCC Henri Becquerel dans le cadre de la politique scientifique du centre mené par le Pr P. Vera (Directeur des affaires scientifiques).

Moyens matériels

- MicroTEP (900 K€) ;
- Imageur optique type QUOS ou VISEN pour petit animal (350 K€) ;
- Système d'imagerie confocale fibré pour l'imagerie du petit animal dans le proche infrarouge (150 K€) ;

- Unité de Bonne Pratique de Fabrication (BPF) (450 K€);
- Agrandissement des locaux du laboratoire;
- Structuration du laboratoire et de la plateforme d'imagerie du petit-animal.

Équipe DocApp : Documents et apprentissage

Auto-analyse

Points forts

- Evolution positive, en référence à la dernière évaluation, de l'activité de publication (en nombre et niveau de publications) grâce à un travail important de l'équipe pour sélectionner les revues et conférences de qualité dans lesquelles l'équipe incite les doctorants à publier ;
- Bonne dynamique de l'équipe grâce au recrutement de 4 MCF sur les 4 dernières années, la constitution de groupes de travail et la collaboration intra-équipe sur des thématiques précises permettant de monter des projets ANR ;
- Fort potentiel d'*HDRisables* : 3 ou 4 collègues MCF devraient présenter leur HDR sur les deux ou trois prochaines années qui sont un signe de maturité des recherches conduites dans l'équipe ;
- Tous les collègues MCF participent à l'encadrement d'une thèse au moins. Ceci est le résultat d'une politique volontariste de l'équipe qui vise à impliquer très tôt les jeunes MCF recrutés dans le co-encadrement des doctorants ;
- Activité contractuelle et industrielle riche en partenariats à travers la participation de l'équipe à un réseau d'excellence (Pascal 2), 7 projets ANR sur le dernier quadriennal (dont 3 coordonnés par l'équipe), 2 programmes européens transfrontaliers Interreg IVa (dont un est porté par l'équipe) et un projet du GRR TL-TI avec EADS (projet PlaiR, structurant pour l'équipe et le laboratoire) ;
- Visibilité régionale, nationale et internationale des membres de l'équipe : 50% des publications de rang A sont co-signées avec des équipes internationales, 67% avec des chercheurs extérieurs au LITIS, 3 prix du meilleur papier ; activités éditoriales importantes, représentation dans les sociétés scientifiques internationales, organisation des conférences du domaine et de bons résultats dans des compétitions internationales ;
- Collaborations inter-équipes fortes : TIBS et MIU (plateforme d'indexation régionale), STI (apprentissage dans les systèmes de vision), QUANTIF (segmentation et classification d'images endomicroscopiques) ;
- Bonne insertion des docteurs formés par l'équipe : sur 16 docteurs formés dans l'équipe, 4 sont maintenant MCF, 2 sont actuellement ATER, 9 sont ingénieurs de recherche en entreprise (EADS, EMC,...) et 1 est ingénieur recherche sur contrat ANR.

Points à améliorer

- Bien que l'équipe DocApp ait déjà participé à un ReX, deux projets européens transfrontaliers Interreg sur le dernier quadriennal et soit le coordinateur d'un autre projet Interreg sur 2010-2013, il lui manque encore un grand projet européen fédératif et structurant impliquant tous ses membres qui lui permettrait d'avoir une meilleure visibilité internationale ;
- Améliorer encore la coopération avec les équipes de recherche internationales, en particulier européennes, par l'obtention auprès des établissements de tutelle de postes de professeurs invités pour accueillir nos collègues étrangers avec lesquels nous souhaitons collaborer. Favoriser l'embauche de post-doc étrangers sur les projets ANR ;
- Taux de publication encore hétérogène par rapport au potentiel de l'équipe et des thématiques abordées par l'équipe : il reste notamment un contributeur que l'équipe voudrait rendre publiant au cours du prochain quadriennal ;
- Animation scientifique : un travail a été mené dans ce sens mais il reste à organiser un vrai séminaire d'équipe (visible de l'extérieur) et à maintenir un rythme régulier de réunions.

Opportunités

- La création du PRES Normandie Université est un élément moteur pour renforcer nos liens avec le GREYC et développer des collaborations dans un cadre structurant notamment pour la proposition de formations co-habilitées de Master ;

- Le projet de Fédération de Recherche Norm@STIC avec le GREYC doit permettre de renforcer nos échanges scientifiques autour d'un axe Reconnaissance de Formes entre nos deux laboratoires donnant une meilleure visibilité régionale et nationale à nos activités de recherche. Deux projets ANR ont déjà été déposés dans ce sens ;
- Le volet Contenus Numériques du Grand Emprunt doit nous permettre d'être encore mieux identifié au niveau national sur les problématiques d'analyse automatique de documents en nous associant avec les principales équipes françaises du domaine et les industriels majeurs de l'archivage et du traitement de documents. Un projet (EquipEx PLEIAD) vient d'être déposé dans ce sens avec le L3i (Université de La Rochelle).

Risques

- Gestion des compétences : au moins 3 ou 4 habilitations à diriger les recherches vont être soutenues dans le prochain quadriennal (probablement 2 en 2011) et Y. Lecourtier (PR) part à la retraite en 2011. L'équipe doit réussir à obtenir l'ouverture de postes de rang A dans ses thématiques pour augmenter sa capacité d'encadrement tout en maintenant sa cohérence scientifique. L'équipe devra donc veiller à favoriser le recrutement extérieur pour développer de nouvelles compétences tout en maintenant celles qui ont été développées localement ;
- La proximité de laboratoires d'excellence à Lille, Rennes et Paris, qui peut être un élément moteur pour le développement de collaborations scientifiques, peut être aussi un facteur de marginalisation de nos activités de recherche et d'enseignement dans un contexte de régionalisation de la recherche (autonomie des établissements, concentration des moyens sur les pôles d'excellence).

Projets et objectifs scientifiques

Contexte général

Rappelons que la démarche de l'équipe DocApp vise de manière générale à équilibrer les aspects fondamentaux, le développement d'outils et leurs applications : les problèmes concrets spécifiques au traitement de grandes masses de données nous amènent à nous poser des questions fondamentales sur les méthodes d'apprentissage dont les réponses vont enrichir nos pratiques expérimentales. Notre démarche s'articule donc autour d'un aspect fondamental, un aspect appliqué et entre les deux, le souci constant de développer des algorithmes génériques comme solutions à certains aspects du problème d'apprentissage.

En particulier, l'équipe DocApp est reconnue nationalement et internationalement pour les compétences qu'elle a su développer dans deux domaines applicatifs particuliers que sont les interfaces cerveau-machine (BCI) et l'analyse automatique de documents :

- Dans le cadre des BCI, notre expertise est essentiellement reconnue grâce à notre savoir-faire pour la mise en oeuvre de méthodes d'apprentissage robustes à la variabilité des données. Cependant, malgré les efforts conjoints des différents groupes de recherche travaillant sur les BCI, il reste encore beaucoup à faire pour aboutir à un usage des BCI hors-laboratoire. En effet, une interface cerveau-machine totalement fonctionnelle devrait être suffisamment robuste et fiable pour être exploitable durant plusieurs heures sans intervention humaine. De telles interfaces se doivent également d'être efficaces quant à la reconnaissance des motifs EEG qu'un patient souhaite émettre pour commander la BCI et suffisamment adaptables pour ne nécessiter qu'un faible temps de calibration. Du point de vue de l'apprentissage statistique, cela pose des questions fondamentales notamment : comment représenter les signaux (quelles caractéristiques utiliser) et comment apprendre à s'adapter aux évolutions. Ces questions, à l'intersection entre la théorie statistique de l'apprentissage et le traitement du signal, sont intimement liées dans le cas des interfaces cerveaux machines.
- En ce qui concerne l'analyse automatique de documents, notre expertise est essentiellement reconnue grâce à notre savoir-faire dans la mise en oeuvre de systèmes de traitement : système d'extraction d'informations dans les documents manuscrits ou techniques, système de catégorisation par le contenu, système d'indexation de documents anciens dégradés.... Même si des solutions originales ont été apportées dans certains contextes, l'analyse automatique de documents reste l'un des champs d'application de la reconnaissance de formes les plus difficiles mais aussi les plus prometteurs. Un exemple illustrant bien les difficultés en termes de recherche mais aussi les enjeux considérables d'un point de vue économique est la préservation du patrimoine écrit : à l'heure où les bibliothèques, les musées, les services d'archives,

numérisent en masse leurs archives papier pour les mettre à disposition sur le web, les moyens de consultation offerts aux utilisateurs, qui devraient faciliter la navigation dans ces corpus, sont toutefois limités par la nécessité d'avoir indexé au préalable ces masses de documents, problème d'autant plus difficile à résoudre que les supports numérisés sont de qualité très inégale (dégradation des encres, des papiers...), que les contenus sont variables (impression ancienne, écriture manuscrite, mise en page, illustrations, schémas...) et que l'information textuelle elle-même peut être de différente nature (ancien français, littérature, revue, journaux, notes personnelles...). Les techniques d'apprentissage offrent alors un paradigme pertinent pour adapter les modèles aux données tout en tenant compte de l'incertitude sur ces données.

Ces deux applications sont donc des terrains d'expérimentation privilégiée particulièrement riches pour y développer des méthodes d'apprentissage à partir d'exemples puisque le passage à l'échelle qu'implique le traitement de ces grands masses de données, variées de par leur structure, leur dimensionnalité, leur stationnarité, et issues de contextes hétérogènes, pose le problème de la généralisation des modèles et de leur adaptabilité. Nous entendons par généralisation des modèles la capacité à concevoir des modèles à partir d'exemples en limitant le plus possible l'intervention humaine (*l'a priori*) dans la boucle de conception du système. Cette question fondamentale de l'apprentissage, au centre des préoccupations de l'équipe, sera abordée à travers l'*apprentissage conjoint représentation/décision* qui offre un paradigme pertinent pour essayer de lever ce verrou. De même, par *adaptabilité des modèles*, nous entendons la capacité des modèles à s'adapter à des données nouvelles (dans des cadres supervisé, semi-supervisé, incrémental ou évolutif) à partir de modèles déjà appris. Dans une problématique de traitement de grandes masses de données, cette question cruciale de savoir comment apprendre à moindre coût sera également un axe de recherche important de l'équipe. Ces deux verrous seront abordés sous différents angles d'attaque en continuant de développer les cadres théoriques, méthodologiques et algorithmiques d'apprentissage dans les deux contextes applicatifs pour lesquels l'équipe DocApp a des compétences reconnues et sur lesquels elle concentrera ses travaux.

Apprentissage conjoint représentation/décision

Nous aborderons le problème de la représentation des données en proposant des méthodes statistiques extrayant automatiquement des caractéristiques pertinentes et optimales pour une tâche de catégorisation donnée. Bien que l'importance de l'extraction automatique de caractéristiques pertinentes ait déjà été identifiée, dans la plupart des approches de reconnaissance de formes, ces caractéristiques sont construites manuellement en fonction de la tâche. Cette approche manque clairement d'adaptabilité puisque dès que la tâche varie, de nouvelles caractéristiques doivent être élaborées. Notre programme de recherche vise alors à développer des classifieurs plus efficaces et plus adaptables, en intégrant le processus d'extraction de caractéristiques pertinentes dans l'apprentissage même du classifieur, lui permettant ainsi de s'adapter à une tâche donnée sans intervention humaine aucune. Nous envisageons d'étudier ce point à travers les paradigmes de l'apprentissage de dictionnaires de représentation, de l'apprentissage de noyaux et de l'apprentissage par architecture profonde. Ces paradigmes seront appliqués dans le contexte des BCI principalement mais nous souhaitons également investiger le couplage architecture profonde et champs aléatoires conditionnels pour des tâches de *keyword spotting* dans les documents manuscrits.

Dans le domaine du traitement du signal, la question de la représentation de signaux est abordée via la décomposition de ceux-ci dans des dictionnaires adaptés à cette tâche. Un formalisme mathématique qui fournit de nombreux dictionnaires est celui des ondelettes, qui sont particulièrement opportunes lorsque des codages parcimonieux des signaux sont recherchés. Un thème de recherche très actuel est par ailleurs celui de l'apprentissage de dictionnaire, et donc de représentations, qui permettent de dépasser le cadre classique de décomposition de signaux. Notre programme de recherche dans ce cadre vise à développer de nouveaux algorithmes permettant l'apprentissage de dictionnaires sous différentes contraintes (dictionnaire parcimonieux, apprentissage en-ligne, apprentissage supervisé). L'extension de ce cadre à plusieurs niveaux amène naturellement à l'apprentissage par architecture profonde. Ici notre objectif serait d'étudier des problèmes d'optimisation inhérent à l'apprentissage à architecture profonde prenant en compte les aspects génératif et/ou discriminant des réseaux, les différents termes de régularisation aux différentes couches de l'architecture, et le partage de couches entre différents réseaux. En outre, le cadre inclura les paradigmes d'apprentissage supervisé, semi-supervisé, et non-supervisé.

La sélection de modèles peut être vue également comme un problème d'apprentissage conjoint représentation/connaissance. Les travaux que nous menons dans ce sens sont de nature théorique et algorithmique et sont abordés sous l'angle de la modélisation parcimonieuse. Les questions fondamentales à résoudre concernent le type de modèle que l'on recherche, le choix des critères à optimiser et l'algorithmique à mettre en oeuvre

dans un souci d'efficacité. Les critères classiques de sélection de modèle, fondés sur des hypothèses souvent irréalistes, sont des critères de pénalisation de contraste minimum à pénalités fixées. Un de nos objectifs principaux est de fournir des critères de sélection de modèle efficaces dont les termes de pénalité sont liés aux données. Dans ce contexte, nous cherchons à améliorer la panoplie des critères de la sélection statistique de modèle dans ses aspects à la fois théorique et pratique. Du point de vue théorique, nous visons à fournir des critères de sélection bien fondés et, du point de vue pratique, nous avons pour but de traiter des problèmes réels et complexes tel que la sélection automatique de variables ou le choix du modèle de mélange le plus adapté à un jeu de données.

En ce qui concerne l'apprentissage en présence de données structurées, de nombreux travaux traitent des techniques d'appariement de graphes (calcul de distance ou recherche d'isomorphisme de sous-graphe) en présence d'un étiquetage nominal, ce qui suppose d'avoir procédé à une classification des données représentées par les nœuds et les arcs du graphe. Les éventuelles erreurs issues de cette classification, qui pose elle-même le choix de l'espace de représentation des données, se révèlent préjudiciables pour la phase ultérieure de mise en correspondance. Nous souhaitons traiter l'appariement de graphes de façon globale, depuis l'espace de représentation initial, comme un problème d'optimisation d'une fonction de coût s'appuyant notamment sur des opérations d'édition de création et de suppression ou d'appariement multivoque afin de prendre en considération, outre les différences dans l'étiquetage des données, les différences topologiques des graphes comparés. Ainsi, le problème du choix de l'espace de représentation des données est intégré au processus de décision et abordé sous l'angle de l'apprentissage des fonctions de coût. Nous appliquerons ces travaux à l'extraction et la localisation d'informations dans les données structurelles notamment pour la recherche de symboles graphiques dans les documents techniques.

Adaptabilité des modèles

Nous allons aborder l'adaptabilité des modèles à travers la question plus générale de l'adaptation d'un classifieur existant vis-à-vis de données d'apprentissage et de test dont les distributions sont différentes de celles utilisées pour apprendre le classifieur. En particulier, nous comptons étudier dans quelle mesure les méthodes de type apprentissage multi-tâche peuvent prendre en charge de telles situations. En effet, l'apprentissage multi-tâche permet d'entraîner des classifieurs sur différents jeux de données de manière conjointe, de façon à transférer certaines connaissances d'un classifieur à un autre. Par exemple, dans le cas des BCI, l'hypothèse sous-jacente à l'utilisation d'un algorithme d'apprentissage multi-tâche avec un a priori de groupes, est que les utilisateurs de la BCI peuvent être regroupés selon un critère de ressemblance de leurs données d'apprentissage. De ce fait, un nouvel utilisateur pourrait être automatiquement considéré comme "similaire" à d'autres utilisateurs. Ainsi, un transfert de connaissance entre ces utilisateurs améliorerait les performances du classifieur associé au nouveau patient. Par définition, un problème multi-tâches met naturellement en oeuvre une optimisation à objectifs multiples. Dans ce cadre, nous proposons d'étudier deux problèmes spécifiques à l'apprentissage multi-tâches lorsque celui-ci est abordé par des méthodes à noyau : le choix de la norme mixte pour le terme de régularisation, et la pondération respective des tâches. Nous comptons également renforcer nos recherches sur cette problématique d'apprentissage multi-tâches dans le domaine de l'analyse automatique de documents pour l'extraction de structures physiques/logiques, la catégorisation par le contenu ou la reconnaissance de l'écriture manuscrite où des transferts de connaissances entre modèles de documents, lexiques ou scripteurs peuvent permettre d'améliorer les performances d'un nouveau système à partir d'un système déjà entraîné. Des résultats prometteurs ont déjà été obtenus dans ce sens pour le développement de systèmes OCR robustes : en étendant les algorithmes MLLR et MAP, nous avons mis au point des procédures originales d'adaptation de la structure des modèles markoviens. Nous comptons poursuivre dans cette voie pour l'adaptation des systèmes de reconnaissance de l'écriture manuscrite dans un contexte semi-supervisé c'est-à-dire en présence de données partiellement étiquetées ainsi que pour l'extraction d'information dans les images de documents en étudiant les modèles de Champs Aléatoires Conditionnels (CRF) 2D à composantes séparables.

Un autre angle d'attaque concernant l'adaptabilité des modèles repose sur l'idée que la plupart des systèmes de reconnaissance de formes sont aujourd'hui conçus pour optimiser un critère scalaire là où de multiples objectifs devraient être considérés : compromis entre performances en apprentissage et capacité de généralisation, entre sensibilité et spécificité, ou entre prise en compte d'un grand nombre d'exemples et rapidité de la prise de décision,... L'apprentissage automatique est donc par essence un problème d'optimisation multi-critère, notamment lorsque l'apprentissage doit être réalisé dans des environnements mal définis, pour lesquels les effectifs des classes sont déséquilibrés et les coûts de mauvaise classification sont inconnus. Dans ce contexte, il est bien connu qu'un unique critère d'apprentissage ne permet pas de construire un classifieur adapté à toutes les

situations. Nous proposons de poursuivre des travaux déjà initiés s'appuyant sur l'espace ROC pour construire non pas un unique classifieur mais une population de classifieurs optimisant chacun un compromis entre ces différents critères. L'objectif est de généraliser l'approche existante pour qu'elle s'adapte à des problèmes multi-classes d'une part, et à grande échelle d'autre part au travers d'une approche en-ligne.

Enfin, un dernier angle d'attaque s'intéresse à l'optimisation des hyper-paramètres dans les méthodes d'ensembles de classifieurs et en particulier pour les forêts aléatoires, une méthode d'ensemble d'arbres de décision très utilisée en classification, régression et data mining et toujours très bien classée dans les compétitions internationales. Comme toute autre méthode, la performance des forêts aléatoires dépend d'une façon complexe et souvent difficilement prévisible de la nature du problème attaqué, du nombre et de la nature des données utilisées pour l'apprentissage du modèle, et des valeurs choisies pour les hyper-paramètres de l'algorithme. Nous avons déjà proposé des solutions basées sur l'induction dynamique de forêts qui permettent de s'affranchir en partie de ces difficultés. Nous souhaitons poursuivre nos travaux sur cette problématique afin d'une part, de mieux comprendre et caractériser le fonctionnement des forêts aléatoires, en termes de convergence, de performance, et de sensibilité aux divers hyperparamètres, et d'autre part d'exploiter les connaissances acquises sur ce fonctionnement pour adapter, sous contraintes, l'inférence des forêts à la nature du problème traité.

Mise en œuvre

La mise en œuvre de notre projet s'appuie sur un ensemble d'actions développées en interne au sein de l'équipe et en externe dans le cadre de collaborations régionales, nationales et internationales.

Éléments de politique scientifique interne

Ces actions visent à améliorer l'animation scientifique de l'équipe et à renforcer les échanges scientifiques au sein de l'équipe.

Même si nous avons déjà mis en place des réunions d'équipe régulières alternant séminaires et réunions de politique scientifique, nous souhaitons adopter un rythme plus régulier et plus soutenu pour le séminaire d'équipe en invitant les chercheurs de la communauté extérieurs au LITIS à présenter leurs travaux mais aussi en alternant ces présentations avec des séminaires internes de présentation des projets en cours. Nous souhaitons mettre aussi un accent plus fort sur l'implication des doctorants dans ces séminaires en les incitant à faire part de l'avancement de leurs travaux devant l'équipe pour renforcer les échanges scientifiques.

Nous souhaitons également généraliser les groupes de lecture et groupes de travail de l'équipe. Si de tels groupes ont déjà été mis en place et fonctionnent de façon assez régulière au sein de l'équipe, les collaborations que nous avons avec les autres équipes du LITIS (TIBS, MIU, STI et Quantif) peuvent être soutenues par des groupes de travail inter-équipes réguliers. A ce titre, un des membres de l'équipe, contribuant, s'engage à mettre en place un groupe de lecture régulier sur la problématique de l'apprentissage des profils utilisateurs (collaborations avec l'équipe MIU). C'est une des actions qui doit lui permettre à terme d'améliorer sa production scientifique. Enfin, nous souhaitons renforcer la participation de nos doctorants aux écoles d'été. Cela passe par une gestion du budget de l'équipe un peu différente de celle adoptée jusqu'à maintenant.

Éléments de politique scientifique externe

Ces actions visent à renforcer l'excellence scientifique de l'équipe et à améliorer notre visibilité à l'international notamment.

Il s'agit d'abord de construire des partenariats forts au niveau régional en renforçant notre présence au sein du GRR TL-TI. Le projet PlaIR est à ce titre un élément moteur pour nous positionner clairement au niveau régional comme un des acteurs du développement des Technologies de l'Information. Nous envisageons de donner une suite au projet PlaIR dans le prochain quadriennal. Nous souhaitons également consolider nos liens avec le GREYC. La création du PRES Normandie Université et le projet de Fédération de Recherche Norm@STIC avec le GREYC, dans lequel un axe "Reconnaissance de Formes" a été identifié, seront des éléments structurants pour développer des collaborations plus poussées, notamment en privilégiant la proposition de sujets de thèse en co-tutelle.

Au niveau national, le Grand Emprunt doit nous permettre de nous positionner clairement et d'améliorer notre visibilité sur les problématiques liées à l'analyse automatique de documents. L'équipe est notamment impliquée dans le projet PLEIAD (Plateforme d'Excellence Industrielle-Académique de Dématérialisation), déposé dans le volet Contenus Numériques du Grand Emprunt, qui vise à rassembler les principales équipes françaises (L3i La

Rochelle, LORIA, LIRIS, LABRI, LI Tours, Paris5, LITIS) et les industriels majeurs du secteur de l'archivage et du traitement des documents numérisés (ATHOS Origin, IBM, IRIS France, APROGED) pour constituer une plateforme de recherche et développement de systèmes de lecture automatique, d'indexation et d'archivage de fonds numérisés. Nous souhaitons également nous rapprocher des grands organismes, notamment de l'INRIA. Deux actions sont en cours sur lesquelles l'équipe peut s'appuyer pour opérer un tel rapprochement. D'une part l'équipe est impliquée dans le projet ARC INRIA MABI sur les interfaces cerveau-machine associant les équipes Sequel (INRIA Lille) et TAO (INRIA Saclay); d'autre part, l'équipe est en discussion avancée avec l'INRIA pour porter la boîte à outils SVM-KM sur la plateforme libre et gratuite de calcul numérique SCILAB. Enfin, au niveau international, notre visibilité peut être améliorée par une participation plus active au montage de projets européens, l'un des points faibles identifiés de notre équipe. Toutefois, les projets européens transfrontaliers dans lesquels l'équipe est impliqué avec l'Université de Kent, UK, DocExplore1 (2009-2010) et surtout DocExplore2 (2010-2013) pour lequel nous sommes chef de file, devraient nous permettre d'acquérir l'expérience nécessaire pour le montage de programmes européens plus ambitieux. Notre participation aux Réseaux d'Excellence Pascal et Pascal2 nous donne également l'occasion de renforcer nos collaborations européennes et nous aider dans ce sens. Enfin, l'amélioration de notre visibilité internationale passe aussi par l'obtention auprès des établissements de tutelle de postes de professeurs invités pour accueillir nos collègues étrangers avec lesquels nous collaborons déjà et passe aussi par l'embauche de post-doc étrangers sur nos projets ANR.

Équipe MIU : Modèles, interactions et usages

Auto-analyse

Points forts de l'équipe

Notre équipe s'est constituée après la définition d'un axe de recherche (verrou scientifique et angle d'attaque). Elle a permis de regrouper pour collaborer des membres du Havre et de Rouen.

- Le thème abordé par l'équipe est transversal au sein du LITIS. Cette transversalité se traduit par de nombreuses collaborations inter-équipes, dont les retombées scientifiques sont bilatérales. On notera des interactions particulièrement fortes avec les équipes TIBS (aide à la recherche de documents, algorithmique du texte et extraction d'information), Doc&App (apprentissage supervisé et non-supervisé appliqués à l'analyse d'interactions), RI2C (un des modèles d'agent de l'équipe est utilisé dans le cadre de la recherche opérationnelle appliquée à la logistique des terminaux à conteneurs) et STI (interaction utilisateur avec un Système d'Information Géographique). Cette osmose se traduit en particulier par des thèses co-encadrées au sein du laboratoire facilitant les échanges de connaissances et points de vue, ainsi que des soumissions de projets.
- La thématique de l'équipe ainsi que sa répartition sur deux établissements (INSA de Rouen et Université du Havre) lui permettent de diversifier ses sources de financement de thèses. Sur la période on retrouve donc 3 bourses MESR, 4 financements CIFRE et 2 co-tutelles. Cette diversité se maintient avec l'ensemble des 4 thèses qui débiteront en Septembre 2010 : 1 bourse MESR, 1 bourse régionale, 1 financement CIFRE et 1 co-tutelle.
- L'ensemble des membres de MIU participe à des enseignements de niveau Master 2 afin de diffuser notre savoir-faire, d'encadrer des étudiants en stage recherche et de proposer des sujets de thèses (1 à 2 doctorants ainsi recrutés par an pour MIU).
- L'équipe bénéficie d'une bonne visibilité au niveau régional et particulièrement dans les GRR MRT (Projet Open Risk training dans l'axe Protection des populations) et TL-TI (Projets PlalR et passage portuaire). Cette implication régionale permet de faciliter l'obtention de contrats de recherche et d'allocations de bourses régionales.
- Enfin, proportionnellement à sa taille réduite, l'équipe bénéficie de financements de type contractuel relativement satisfaisant.

Points à améliorer

L'équipe fait face principalement à une productivité en deçà de sa valeur scientifique réelle. Cette faible productivité est directement liée à un ensemble de points à améliorer.

- Depuis le départ à la retraite de J. Labiche (PU émérite), non remplacé, nous devons faire face à une sous-dotation en PU.
- 2 membres de l'équipe sont impliqués dans des responsabilités administratives conséquentes : direction du département ASI (N. Delestre) et direction des Systèmes d'Information de l'INSA (J.-P. Pécuchet).
- Nous devons également déplorer une absence de politique de publication concertée et tournée vers les productions de rang A.
- Enfin, l'ensemble des contrats finançant l'équipe n'a pas eu l'impact scientifique escompté.

Cette faible productivité a d'ailleurs pour principale conséquence une visibilité internationale trop réduite, malgré plusieurs collaborations (co-tutelles passées et en cours) avec la Roumanie (Université de Babes-Bolyai, Cluj) et le Maroc (Université de Tanger).

Au niveau du fonctionnement interne de l'équipe, à la demande d'une majorité d'entre nous et afin d'augmenter nos interactions, la fréquence de nos séminaires sera augmentée. Nous projetons ainsi de la tripler dès la rentrée 2010 et d'y inviter des chercheurs reconnus dans notre domaine.

Opportunités

- L'équipe, en plaçant sa thématique dans un créneau novateur et émergent, permet d'apporter son savoir faire à d'autres thématiques. Elle commence, après trois ans d'activité, à avoir une visibilité auprès du

GDR I3 et de ses partenaires. Cette visibilité s'est traduite par plusieurs soumissions de projets ANR, qu'il faudra pérenniser.

- Fortes de collaborations pluridisciplinaires (publications, organisation de colloques, soumissions de projet) avec le PSY-NCA, le pôle Modesco de la Maison de la Recherche en Sciences Humaines, le SIRTAL UMR IDEE et le LiDiFra (anciennement Dyalang), MIU bénéficie de l'apport scientifique des laboratoires SHS de Haute et Basse Normandie.
- La création en cours de la fédération NormaSTIC intégrant le LITIS et le GREYC doit permettre de renforcer nos interactions existantes avec l'équipe D*LU, issu de la réunion de certains membres des équipes DoDoLa et ISLaND.
- L'équipe bénéficie d'une forte présence dans les départements ASI et GM de l'INSA qui permet de tisser des liens avec le milieu industriel du domaine. Ceux-ci ont déjà permis plusieurs thèses CIFRE et nous souhaiterions renforcer nos contacts industriels en devenant partenaire dans des projets collaboratifs européens. Pour cela nous nous appuyerons notamment sur l'expérience des projets ANR soumis (7 soumissions de projets ANR refusés au cours du quadriennal) en partenariat avec ces mêmes entreprises.
- La création de l'ISCN (Institute for Complex Systems in Normandy), auquel participe le LITIS et en particulier les équipes RI2C et MIU, devrait faciliter les collaborations sur les thématiques de l'équipe.
- L'implication de MIU dans les thèmes liés à la logistique (GRR TL-TI, pôle de compétitivité Nov@Log) devrait permettre des applications issues des concepts développés dans l'équipe.
- Enfin, un poste de PU est prévu à l'INSA pour venir renforcer l'équipe en septembre 2011.

Risques

Les risques identifiés par l'équipe comme vecteurs de dysfonctionnement, s'ils se vérifiaient, empêcheraient son développement scientifique.

- La pluridisciplinarité de nos travaux de recherche et nos fortes collaborations avec des laboratoires SHS, en raison de politiques de communication différentes, peuvent ne pas permettre une bonne valorisation de la production dans nos communautés respectives. Une surveillance accrue des soumissions en interne ainsi qu'une politique de publication clairement définie doit permettre de limiter ce risque.
- Enfin, malgré l'utilisation de la visioconférence, la distance entre Le Havre et Rouen peut gêner le fonctionnement de l'équipe.

Projets et objectifs scientifiques

Les mots clefs de l'équipe sont dans son acronyme mais peuvent s'enrichir de communication augmentée, description et analyse de la trace de l'utilisateur, modélisation des interactions, agentification des usagers avec stéréotypage de leurs interactions.

La figure 2 synthétise parfaitement la démarche adoptée dans l'équipe. La première étape concerne la collecte de traces des usagers issues de diverses sources (dialogues entre humains, environnements d'apprentissage, etc.). En tenant compte de la représentation de l'environnement, nous analysons une partie de ces données pour proposer un modèle de l'utilisateur. L'autre partie des données est utilisée pour valider ce modèle à l'aide, par exemple, de tests statistiques ou "à la Turing". Ensuite, ce modèle est implanté dans un système pour une mise en production afin d'améliorer la qualité des interactions avec l'utilisateur dans le cadre, par exemple, d'un système d'aide à la décision.

Nos travaux futurs visent à améliorer chacune des étapes de notre démarche.

- Aujourd'hui, l'analyse des traces est principalement faite manuellement ou à l'aide d'outils statistiques *ad hoc* et minimalistes. Nous souhaitons nous doter d'outils d'analyse performants et génériques comme ceux proposés dans le projet ACAMODIA soumis à l'ANR. Il s'agit de méthodes et algorithmes d'extraction automatique ou semi-automatique de connaissances et de fouille de données sur des dialogues annotés, afin d'extraire des motifs comportementaux récurrents. Par leur généralité, les méthodes proposées permettent, même si elles doivent encore être améliorées, la recherche de régularité indépendamment des types d'annotation (multimodalité, émotions, etc.).
- Nous souhaitons également enrichir et généraliser nos modèles de l'utilisateur et de l'environnement en complexifiant et en variant les structures de données que nous utilisons pour augmenter leur pouvoir d'expressivité (graphes ou collections d'automates). Les mesures de similarités utilisées doivent conjointement être précisées. Cela se traduit dans le domaine de l'évaluation des apprenants par la confrontation de nos méthodes au domaine de l'algorithmique en définissant une mesure de similarité

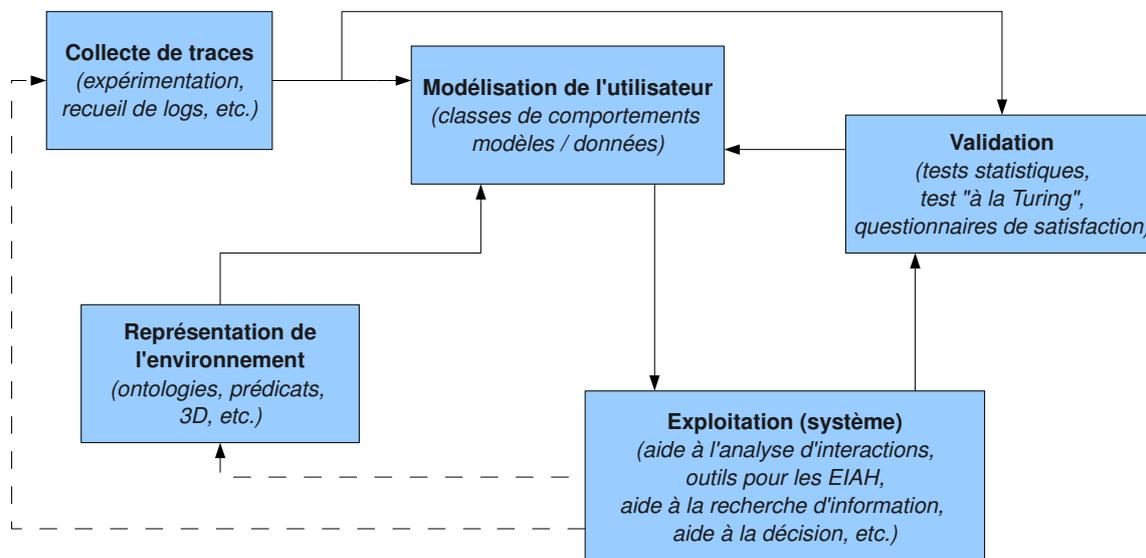


FIGURE 2 – Approche adoptée par l'équipe MIU

entre algorithmes, représentés à l'aide d'arbres syntaxiques abstraits ou d'automates. Dans le domaine de l'analyse de dialogue, la caractérisation des profils de comportements se fera à l'aide d'ensembles d'automates caractéristiques de motifs interactionnels et la classe d'un utilisateur sera déterminée en utilisant une distance entre collections d'automates. Nous souhaitons également augmenter la richesse des interactions en prenant en compte la déictique sur des interfaces tactiles partagées dans le cadre de la formation à la gestion de crise. Les traces seront alors des familles d'automates représentant la structure des dialogues contraints par le type d'interface et par les objets en interaction dans cette interface. Toujours avec l'idée d'enrichissement de nos modèles, nous souhaitons représenter les données complexes et structurées que l'on trouve sous forme de tableaux ou de graphes sous forme linéaire. L'objectif est de permettre la génération de phrases descriptives (structure linéaire par excellence) de structures non linéaires telles que sus mentionnées. L'intérêt applicatif est double : rendre accessible à un public aveugle des données autrement difficiles d'accès et présenter différemment des données à un public apprenant.

- Nous avons également pour but de pousser la phase d'exploitation afin d'obtenir une rétro-action sur nos modèles. Pour cela nous souhaitons accorder plus d'importance à l'environnement en considérant des environnements virtuels et augmentés comme décrit dans le projet équipex récemment soumis (Grand Equipement Numérique en Recherche Logistique Maritime et Portuaire). Dans ce cadre, il s'agit de concevoir des simulateurs de navigation maritime et fluviale pour la formation des officiers de la marine marchande. Une des clefs de voute de notre recherche est l'obtention de données expérimentales qui vont nous permettre de valider nos modèles. Il est alors très important que nous les mettions en commun afin de diversifier nos validations et de comparer nos méthodes. En effet, même si la thématique de l'équipe MIU est bien définie et cohérente, un des problèmes de l'équipe est la dispersion des applications. Le projet équipex, s'il est financé, nous permettra d'obtenir une application commune source de sujets d'étude pour nos différents intérêts scientifiques.
- Finalement, afin d'améliorer la prise en compte de l'environnement, nous allons également poursuivre le développement d'un système multiagent composé d'agents factuels. Ce système reçoit des informations d'un environnement qui n'est pas modélisé ; les traits sémantiques entrants participent d'une approche phénoménologique : ce sont les actions et observations issues de l'environnement qui permettent de le percevoir et de créer un modèle dynamique et émergent au sein du SMA par les interactions entre agents. Ce modèle a déjà été développé et appliqué à un cas de gestion de crise (thèse de Fahem Kébaïr soutenue en 2009). Le modèle est repris dans le cadre d'une thèse commencée en septembre 2010 sur la gestion du risque portuaire où l'environnement est perçu par voisinage avec des agents situés au sein d'un automate cellulaire spatial (3D). Le modèle comportemental des agents est basé sur des

automates de type ATN (Augmented Transition Network) et les communications sur un modèle CATn (coupled ATN). Ainsi, nous espérons pouvoir tester l'influence de la prise en compte des profils d'agents (d'apprenants) dans les changements comportementaux et l'acquisition de connaissances.

Mise en oeuvre

Pour atteindre tous ces objectifs, il nous faut surtout inciter les actuels producteurs, contributeurs et associés dont la production scientifique n'est pas reconnue, à mieux publier, en privilégiant les revues internationales (JAMAAS, JAIR, IJAIT, IJHCS, User Modeling and User-Adapted Interaction, *etc.*) et les conférences internationales de rang A (Interact, IJCAI, ECAI, AAMAS, HCI, *etc.*). En particulier pour les non producteurs cela nécessite la redistribution des tâches administratives, la réduction des services d'enseignement (ces objectifs dépassent le cadre de l'équipe) et l'aide à l'obtention de CRCT.

De plus, il est absolument nécessaire d'augmenter le nombre de PU dans notre équipe (l'équipe MIU est de ce point de vue l'équipe du LITIS la plus déficitaire). Dans ce cadre, l'INSA de Rouen recrutera en septembre 2011 un PU 27 qui rejoindra notre équipe. Le profil demandé porte sur la modélisation des interactions (multimodales ou non) avec l'utilisateur et/ou la conception de simulateurs ou d'environnements virtuels. Nous devons également inciter fortement les MdC de l'équipe à soutenir leur HDR dans les années qui viennent.

Comme nous l'avons déjà mentionné précédemment, nous devons multiplier les séminaires extérieurs et les séjours de professeurs invités. Cela permettra de renforcer nos contacts académiques et industriels pour essayer d'obtenir des financements de projets ANR et européens.

Avec la récente convention de partenariat signée entre l'INSA et l'Ecole Nationale de la Marine Marchande du Havre pour développer des recherches communes, nous pourrions investir de nouveaux champs d'applications avec les simulateurs de formation à la navigation et les environnements virtuels pour l'étude des facteurs humains. Ce dernier sujet nous donnera également une meilleure visibilité au sein du Normandy Living Lab.

Enfin, nos développements devront être pérennisés et rendus plus facilement partageables au sein de l'équipe, par exemple en les mettant au format WebLab (plate-forme de traitement de documents multimédias développée par notre partenaire EADS-Cassidian de Val de Reuil). Leur diffusion auprès de la communauté devra également être améliorée en faisant des dépôts dans le logiciel libre.

Nous souhaitons conclure sur le fait que nous avons les moyens de nos ambitions puisqu'en plus des actuelles thèses en cours, nous avons 4 nouvelles thèses financées pour la rentrée 2010 autour des différents projets de l'équipe. Guillaume Dubuisson Duplessis débute sa thèse sur l'utilisation d'interfaces partagées dans le cadre de la formation à la gestion de crise, avec un financement régional (GRR SER PdP). Rick Moritz travaillera en convention CIFRE chez Orange Labs Meylan à la définition de nouvelles méthodes d'interactions tangibles avec des contenus géolocalisés dans un contexte de mobilité. La thèse d'Ovidiu Serban, dirigée en co-tutelle avec l'université de Babes-Bolyai, vise à fusionner les deux approches de l'équipe, par les modèles et par les données, afin d'identifier des motifs émotionnels dans des dialogues. Enfin Mansoriya Hamidou vient de commencer sa thèse en bourse MESR, dans le cadre du projet passage portuaire (GRR TL-TI), sur le couplage entre automates cellulaires et système multiagent pour le contrôle de systèmes dynamiques.

Équipe RI2C : Réseaux d'Interactions et Intelligence Collective

Auto-analyse

L'équipe RI2C est à la croisée d'une évolution complexe que ses membres ont fait inscrire en cohérence à la fois avec le développement du laboratoire et de sa visibilité nationale et internationale mais également avec le souci de respecter et de soutenir des politiques d'établissements (dont leur positionnement dans les PRES en construction) qui sont aujourd'hui essentielles dans le paysage national de l'enseignement supérieur et de la recherche.

Points forts

- Une augmentation très nette de la visibilité de l'équipe RI2C est assez évidente sur la période 2006-2010, grâce aux montages ou participations à des projets identifiés au niveau régional (GRR - Grand Réseau de Recherche) ou nationaux (ANR, PREDIT), mais également avec une participation accrue à des conférences (Co-organisations de conférences en Angleterre et au Canada) dont plusieurs, à visibilité internationale, ont été organisées au Havre, faisant venir à chacune d'elle plus de 100 ou 150 participants dont plus du tiers en provenance de pays étrangers. Cette visibilité a eu pour conséquences scientifiques de renforcer les invitations de collègues étrangers ou encore nos propres invitations dans des conférences nationales ou internationales chez nos collaborateurs. La participation de RI2C dans des réseaux de recherche (Réseau national sur les systèmes complexes, Complex Systems Society, GDR-européen "S4 - spatial simulation for social systems", société scientifique internationale "SCS - The Society for Modeling & Simulation International") en a fait une entité visible et active auprès de ces communautés. Une forte activité éditoriale autour d'ouvrages à caractères internationaux et nationaux, dans des collections spécialisées et de référence chez Hermès et chez Springer, a permis également de montrer la reconnaissance de nos actions au niveau international ;
- L'équipe a augmenté de manière significative (plus que tripler) ses publications de rang A sur la période 2006-2010 par rapport au bilan de ses membres sur leurs équipes de rattachement sur le précédent contrat d'établissement. Une majorité de ces publications se fait sur la base de collaborations nationales ou internationales : le taux de publications dont l'un des auteurs est un collaborateur hors-LITIS, correspond à 55% ; celles où ce collaborateur est membre d'un laboratoire étranger correspondent à un taux de 30%.
- L'équipe a largement augmenté, en conséquence de l'accroissement de sa visibilité, ses activités contractuelles avec la participation à deux projets ANR et un projet ANR-Predit, mais également un projet FUI dont l'enveloppe financière globale approche les deux millions d'euros. L'augmentation des activités contractuelles au niveau des grands réseaux régionaux a sans doute été l'une des progressions les plus importantes de l'équipe avec la coordination d'un axe de recherche (Protection des populations) ou encore la coordination de nouveaux projets structurants qui vont être amenés à se développer lors du prochain contrat d'établissement (RISC du GRR TLTI -Transport, Logistique et Technologies de l'Information). Parallèlement et de manière complémentaire à notre politique de projet, l'investissement dans les pôles de compétitivité Nov@log et Mov'eo est notable ; Mhamed Itmi est membre du CA et de la commission projet de Nov@log.
- Sur les aspects de la formation doctorale, il est à noter la diversification des supports financiers de thèses (MESR, allocations régionales doctorales et industrielles, bourses du gouvernement français, CIFRE) caractérisant une implantation de notre politique scientifique sur plusieurs dimensions, fondamentales, applicatives ou encore à portée collaborative se basant sur nos échanges internationaux. Il est à noter le développement significatif de conventions CIFRE industrielles (RATP) ou sur des partenariats avec des collectivités locales (Communauté d'Agglomération Havraise - CODAH), sur des sujets au cœur de nos thématiques de recherche, approche systémique et complexe de la gestion du risque et des mouvements collectifs dans les espaces de transport.

Points faibles

- Il est nécessaire d'améliorer encore la politique de publications en revues. Même si on constate une forte hausse sur 2006-2010, il est nécessaire de structurer la politique de publication par une identification des revues et un meilleur ciblage ;
- La période 2006-2010 a malheureusement vu une mise à disposition de locaux très insuffisante au niveau de l'université du Havre où sont situés majoritairement les membres de RI2C. La forte volonté des chercheurs a permis d'aller au-delà de cette gêne qui empêche une bonne intégration de l'équipe, parfois dispersée sur des lieux différents et en conséquence rend la dynamique de recherche plus difficile. La situation devrait s'améliorer à partir de 2011, au prix de nombreux efforts et négociations qui ont pu monopoliser l'énergie des personnes en charge de suivre le dossier d'évolution des locaux ;
- Même si nos efforts en terme de politique contractuelle ont été importants et ont conduit à une progression par rapport au précédent contrat d'établissement, il nous est nécessaire de renforcer le nombre de projets ANR et européens. Nos tentatives (dont la majorité sur invitation à des montages coordonnés par des partenaires) sur ces deux volets n'ont pas toujours abouti, même si malgré tout, elles ont permis de consolider nos collaborations. Nous restons toutefois attentifs à préserver un équilibre permettant de conserver une politique scientifique qui ne soit pas uniquement gouvernée par une stratégie opportuniste de projets.

Opportunités

- Dans le souci de soutenir les démarches initiées par nos établissements de tutelle, nous avons noté le développement de la logistique à l'université du Havre et nous cherchons à accompagner celui-ci en proposant nos compétences. Plusieurs structures et projets se mettent en place dont le campus logistique et différents projets IRT, Equipex et Labex du programme national "Investissement d'Avenir". Le renouvellement et le déploiement dans ce contexte du pôle de compétitivité Nov@log sur les aspects valorisation nous offrent aussi des opportunités de valorisation. Nous devons nous positionner avec pertinence, en conservant la spécificité de nos approches ;
- Au niveau structurel, la création de l'Institut des Systèmes Complexes en Normandie (ISCN) en tant que troisième nœud du Réseau National des Systèmes Complexes (après Paris et Lyon) va nous servir à renforcer la dynamique que nous avons initiée en co-fondant cet institut et à améliorer la structuration de la recherche autour des systèmes complexes et de leurs nombreux domaines d'applications. Nous utiliserons notamment cette structure pour renforcer les volets intelligence territoriale et les approches intégratives de modélisation qui trouvent leur pertinence avec la collaboration de collègues d'autres disciplines, notamment en SHS. Les expériences déjà convaincantes sur ce type de collaboration, nous laissent présager de l'intérêt de cette opportunité ;
- Suite à nos efforts chroniques pour améliorer la situation en terme d'espace de recherche sur l'université du Havre, nous avons réussi à obtenir une mise à disposition de locaux supplémentaires en 2011 qui concerne à la fois le LITIS et le laboratoire de mathématiques appliquées du Havre. Nous avons défini une stratégie d'occupation et de regroupement d'un axe math-info sur les systèmes complexes qui pourrait ainsi devenir un point de référence et d'accueil local de l'ISCN ;
- Nous travaillons sur des perspectives à plus long terme concernant notamment le positionnement dans la fédération NormaSTIC entre le GREYC et le LITIS, dans un projet d'école doctorale autour des mathématiques et des STIC et dans le PRES Normandie Université qui est en phase de structuration opérationnelle.

Risques

- Dans le cadre d'une dynamique relativement récente tant sur les projets de l'équipe que sur les structures d'appui, institut des systèmes complexes en Normandie, IRT Logistique Maritime et Portuaire, la jeunesse relative du développement des projets et leur besoin de stabilisation à long terme peuvent présenter une part de risques mais c'est la combinaison de l'ensemble de ces opérations qui rend la stratégie plus viable ;
- Au sein des évolutions actuelles de nos établissements de tutelle, il nous faut réussir à bien faire identifier l'équipe dans les structures émergentes : fédération avec le GREYC, école doctorale Mathématiques et STIC, PRES Normandie-Université ;
- Nous devons être attentifs à la gestion de nos demandes d'allocations doctorales en accompagnement de

nos projets et structurations émergentes. L'expérience montre que l'association avec des collaborateurs ne produit pas toujours une augmentation des ressources disponibles. C'est la raison pour laquelle, la diversité de nos financements de thèse nous donne une meilleure garantie de pérennité.

Projets et objectifs scientifiques

Politique scientifique

Le projet scientifique de l'équipe, *"Interactions et réseaux complexes pour la dynamique territoriale et la logistique"*, se construit autour des quatre soutiens suivants :

- le soutien scientifique se développe sur les actions menées au niveau de la Région Haute-Normandie, présentées ci-dessous et qui trouvent des développements nationaux complémentaires (REMUS2, Robocoop) et que l'on va essayer d'étendre à l'échelle européenne et internationale, en se basant sur nos collaborateurs en Europe (Angleterre, Luxembourg, Espagne, Hongrie, Russie et Roumanie), aux USA (Arkansas et Californie) et au Moyen-Orient (Jordanie, Yémen, Kurdistan irakien) ;
- le soutien structurel se fait grâce à l'Institut des Systèmes Complexes en Normandie et au projet d'IRT Logistique Maritime et Portuaire ;
- le soutien logiciel est assuré par la bibliothèque GraphStream qui a vocation à devenir la plate-forme logicielle phare de l'équipe ;
- le soutien financier et humain se fait grâce à nos projets CPER, ANR, Predit et FEDER, nous permettant l'obtention d'allocations régionales, de post-doctorants ou encore d'ingénieurs recherche. Nos échanges internationaux nous permettent aussi de déployer des ressources humaines sous forme de collaborations sur des thèses ou encore par l'obtention de bourses soutenues depuis les pays de nos collaborateurs.

Politique de projets

Notre politique scientifique sur les projets de l'équipe passe par la confirmation de l'ancrage régional et la coordination de nouvelles opérations structurantes :

- "Protection des populations", action coordonnée par Damien Olivier, reste une action importante qui a été confirmée depuis 2007. Il s'agit de conserver cette identification et de renforcer notre visibilité au-delà de la Haute-Normandie (où le GRR de rattachement de cette action est situé) et de renforcer notamment nos collaborations avec le GREYC de Caen, dans le cadre de la fédération NormaSTIC. Il est envisagé de mettre en place un GIS sur la gestion du risque autour des collaborations entre géographes et informaticiens.
- RISC – Réseaux d'Interaction et Systèmes Complexes – est un nouveau projet structurant du GRR TL&TI coordonné par Cyrille Bertelle, qui s'est mis en place pour une première phase de trois ans à partir d'octobre 2010. Ce projet réunit plus de 60 partenaires (évalués à plus de 40 hommes.années sur les 3 ans) en provenance de plus de 11 laboratoires de recherche, écoles et structures de valorisation dans des disciplines variées, informatique, mathématique, géographie, STAPS, management, droit. Son objectif est de développer une base de modèles s'appuyant sur les réseaux d'interaction et les systèmes complexes pour étudier principalement divers aspects de l'intelligence territoriale, en dynamique urbaine, réseaux logistique ou encore environnement et usages. Le soutien financier est globalement de 2 millions d'euros, en provenance de la Région Haute-Normandie et de fonds européens FEDER.
- "Passage portuaire" est également un projet structurant du GRR TL&TI, coordonné par nos collègues du laboratoire de mathématiques appliquées du Havre et qui s'intéresse à la logistique portuaire. Il réunit également un grand nombre de laboratoires, écoles et structures de valorisation. Plusieurs membres de l'équipe, notamment Eric Sanlaville et Habib Abdulrab, y participent et représentent une partie importante du volet recherche opérationnelle et simulation.

Ces projets doivent renforcer de manière significative, à la fois nos collaborations, nos soutiens financiers et humains.

Par ailleurs, nous avons souhaité accentuer la participation du LITIS et de RI2C dans la politique scientifique de l'université du Havre. Ceci se traduit par une participation active au montage de l'Institut de Recherche Technologique Logistique, Maritime et Portuaire (IRT LMP), projet de l'université du Havre et Nov@log dans le cadre du programme national "Investissements d'Avenir" : le laboratoire LITIS et notamment des membres de l'équipe RI2C se sont investis dans la préparation de ce dossier, en intégrant l'équipe de pilotage du projet. Une dynamique de recherche et de valorisation se met en place dans le cadre de ce projet et sera pérennisée,

dans tous les cas, grâce à une structuration de la recherche en logistique qui n'est pas sans intersection avec la thématique de l'intelligence territoriale, thème à développer dans le cadre de l'institut des systèmes complexes en Normandie. L'articulation de nos participations à cette politique universitaire se fait donc dans le cadre de l'intégration de nos spécificités notamment autour de la recherche opérationnelle et des systèmes complexes. Elle permet de renforcer à nouveau des réseaux de collaborations et nous donne une meilleure visibilité.

Mise en oeuvre

La mise en oeuvre de notre politique passe par l'intégration des réseaux régionaux, nationaux et internationaux, notamment autour de la logistique, de la gestion des risques et de l'intelligence territoriale. L'ensemble de ces activités est amené à trouver un socle conceptuel et pluridisciplinaire autour des systèmes complexes ou encore des systèmes de systèmes. À cet effet, nous avons donc été moteur dans le lancement du troisième institut du réseau national sur les systèmes complexes, l'ISCN – Institut des Systèmes Complexes en Normandie – au sein duquel nous souhaitons inscrire et rendre visible la majorité de nos activités. L'enjeu est avant tout de faire connaître plus largement nos démarches et nos thématiques de recherche ainsi que de favoriser nos contacts et collaborations régionaux et internationaux puisque de nombreux collègues, d'une vingtaine de pays, se sont déjà déclaré intéressés pour participer aux travaux de cet institut qui est par ailleurs bien identifié et soutenu par le réseau national des systèmes complexes (RNSC). Cet institut nous permet également de trouver des collaborations pluridisciplinaires indispensables au développement de nos thématiques qui nécessitent un positionnement pertinent à la croisée des disciplines telles que la géographie, l'économie, la biologie et les sciences des modèles.

Une nouvelle approche conceptuelle est amenée à se développer et concerne les systèmes de systèmes. En effet, sous la pression de l'évolution de la technologie, les projets à réaliser sont de plus en plus complexes. Ils mettent en relation des groupes autonomes formés d'organisations, de systèmes technologiques, tous connectés et qui doivent répondre dans leur ensemble à des besoins de coordination ou de réorganisation en cas d'incident. Il s'agit là de systèmes de systèmes (SoS). Nous nous intéressons à de tels systèmes que nous appellerons SoS auto-adaptatifs dans lesquels tous les éléments ont une autonomie relative et variable. Nous les étudions selon des points de vues théoriques et appliqués.

Une orientation importante de notre projet concerne aussi le développement et déploiement de la bibliothèque Graphstream : mise en place d'une équipe de développement stratégique (positionnement dans les projets et réseaux de collaborations) et de développement technique. L'évolution attendue de cette bibliothèque consiste à pouvoir manipuler des graphes de plus grande ampleur en fournissant un système de distribution sur des grappes ou grilles de machines, en s'intéressant en particulier aux organisations existant au sein des réseaux d'interactions représentés par des graphes. On s'intéressera ainsi à la distribution des algorithmes et métriques dynamiques sur une telle architecture. Dans cette optique, on étendra le système de visualisation afin de fournir un système de méta-visualisation permettant d'explorer le graphe à plusieurs niveaux de représentation.

Par ailleurs, nous souhaitons accompagner notre démarche de développement de nos recherches par des partenariats et des collaborations avec des pays émergents, en profitant de compétences locales (Habib Abdurab est chargé de mission pour le développement de collaborations avec le moyen-orient, depuis de nombreuses années). Notre volonté est, au-delà d'appuis complémentaires sur nos thématiques et la mise en place de noyaux de recherche sur nos thématiques au niveau international, de partager une vision politique et humaniste de la recherche à une échelle internationale, respectant les démarches que notre pays a su développer et faire rayonner depuis de nombreuses années par l'intermédiaire de l'UNESCO. Nous commençons ainsi à mettre en place des écoles CIMPA³ ; l'une d'elle sera organisée par l'équipe RI2C, en 2012 à Sulaimani, au Kurdistan d'Irak. Nous allons également accueillir des étudiants via des accords bilatéraux de collaborations internationales (nous avons été retenu centre d'accueil d'étudiants kurdes en formation doctorale en informatique). Notre contribution est d'accompagner des chercheurs dans ces pays en développement, de proposer des thématiques de recherche permettant de répondre également à des besoins de développement à la fois scientifique et économique par des approches intégratives (formation-recherche-valorisation) sur des sujets tels que l'intelligence territoriale.

3. Le CIMPA (Centre International de Mathématiques Pures et Appliquées) est une association internationale (loi de 1901) créée à Nice (France) en 1978, avec le soutien de l'UNESCO. Son objectif est de promouvoir la coopération internationale au profit des pays en développement, dans le domaine de l'enseignement supérieur et la recherche en mathématiques et dans les disciplines connexes, informatique notamment.