



agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Rapport du comité d'experts

Unité de recherche :

Laboratoire d'Informatique de Paris-Nord
(LIPN) – UMR 7030

de l'Université Paris 13



février 2008



agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Rapport du comité d'experts

Unité de recherche :

Laboratoire d'Informatique de Paris-Nord
(LIPN) – UMR 7030

de l'Université Paris 13



Section des unités
de recherche

Le Directeur

Jean-Jacques Aubert

février 2008



Rapport du comité d'experts)

L'Unité de recherche :

Nom de l'unité : Laboratoire d'Informatique de Paris-Nord (LIPN)

Label demandé : UMR

N° si renouvellement : 7030

Nom du directeur : M. Christophe FOUQUERE

Université ou école principale :

Université Paris 13

Autres établissements et organismes de rattachement :

CNRS

Date(s) de la visite :

30 janvier 2008



Membres du comité d'évaluation

Président :

M. Philippe FLAJOLET, INRIA-Rocquencourt/Paris

Experts :

M. Pierre-Louis CURIEN, PPS, Paris 7

M. Marc EL-BEZE, LIA - CERI, Avignon

M. Henri MAITRE, Telecom Paris

M. Rémi MUNOS, INRIA FUTURS

M. Christian RETORE, LaBRI, Bordeaux 1

M. Jacques TEGHEM, Université Mons, Belgique

Expert(s) représentant des comités d'évaluation des personnels (CNU, CoNRS, CSS INSERM, représentant INRA, INRIA, IRD...) :

M. Dominique ROSSIN, Comité National, LIAFA, Paris 7

M. Denis TRYSTRAM, CNU, Université Grenoble

Observateurs

Délégué scientifique de l'AERES :

M. Luis FARIÑAS DEL CERRO

Représentant de l'université ou école, établissement principal :

Représentant(s) des organismes tutelles de l'unité :

Mme Véronique DONZEAU-GOUGE

M. Jean-Michel MULLER, CNRS



Rapport du comité d'experts

1 • Présentation succincte de l'unité

Le Laboratoire d'Informatique de l'Université Paris-Nord (LIPN) a le statut d'UMR depuis 2001. Il comprend 55 chercheurs et enseignants chercheurs permanents : 15 sont professeurs, 36 sont maîtres de conférences et 4 sont chercheurs CNRS (de niveau Chargé de recherche). Sur l'effectif enseignant, 26 sont membres de l'Institut Galilée, tandis que 23 sont rattachés à un IUT. Ces chiffres correspondent à une bonne croissance maîtrisée, depuis la précédente évaluation : cette croissance de plus de 15% est certainement justifiée par l'importance socio-économique de l'informatique. Le personnel administratif comprend 5,5 ITA et IATOS.

Le nombre de doctorants est de 25, financés sauf 2, soit par bourse du ministère pour une petite moitié, soit par des contrats, industriels ou publics. Cet effectif relativement modeste représente déjà une amélioration de la situation précédente ; il s'explique au moins partiellement par la charge d'enseignement pour les personnels sur poste IUT, dont la proportion avoisine 40%, et la jeunesse de l'effectif MDC (voir plus bas). La durée typique d'une thèse, proche de la moyenne nationale, se situe entre trois et quatre ans.

Pour ce qui est de la Prime d'Encadrement Doctoral et de Recherche (PEDR), 10 sur les 15 professeurs en bénéficient, ce qui est un excellent score. Par contre, le nombre de maîtres de conférences habilités, 4 sur un effectif de 36, paraît faible : cela peut s'expliquer par le fait que la médiane de l'ancienneté dans la fonction MDC est de 5 ans et la moyenne de 7 ans, ce qui traduit la jeunesse de l'effectif MDC. Étant donnée la relève de génération qui devra se faire d'ici quelques années, l'habilitation des MDC en poste méritera une attention particulière.

Nombre de publiants : 43 publiants sur 55 EC, C. Si l'on considère comme publiant un EC/C qui publie au moins quatre articles de revues internationales ou communications à des conférences internationales avec comité de lecture et publication d'actes, on constate que 4 chercheurs CNRS (sur effectif de 4) publient ; 13 professeurs (sur 15 au total) vérifient le critère ; 1 est non publiant ; 1 représente une situation limite. Pour les MdC, nous estimons que 26 (sur 36) sont publiants au sens ci-dessus ; sur les 10 non publiants, 2 ont zero ou une publication internationale sur la période. Les 8 restant ont une activité de publication soit faible, soit trop fortement nationale, soit sans publication claire d'actes (voir cependant ci-après pour quelques réserves sur la significativité de ces estimations). A ce sujet, nous émettons la recommandation générale que les références bibliographiques soient systématiquement accompagnées d'indication claire de l'éditeur (« publisher »), des numéros de pages, voire également d'une référence web ou ISBN. Faute de ce faire, la réalité de la publication d'actes (plutôt que la réunion de simples communications orales) reste, dans nombre de cas, difficile à établir.

2 • Déroulement de l'évaluation

Nous tenons à souligner l'excellente préparation de ces journées par l'équipe de direction du laboratoire ainsi que le caractère parfaitement informatif des exposés oraux des différents thèmes. Le rapport d'activité 2003-2007 représente un effort tout particulier de présentation, de structuration, et de pédagogie que nous applaudissons. Nous tenons également à féliciter l'université et l'IUT pour leur soutien à la recherche, qui se manifeste notamment par les décharges horaires accordées aux nouveaux maîtres de conférences ainsi qu'aux porteurs de projet ANR. La gouvernance du laboratoire nous a paru tout à fait excellente. Le directeur est dynamique, son charisme et son pragmatisme sont à l'origine d'une bonne ambiance, ce qui permet de créer les conditions nécessaires à une bonne production scientifique. Le comité a trouvé tout à fait judicieux qu'un second mandat de 4 ans ait été confié récemment à son directeur.

L'évaluation a consisté en : (1) une présentation générale du laboratoire, de ses évolutions et de ses axes par le directeur ; (2) les exposés des quatre équipes de recherche, chacun correspondant à une présentation générale des activités par le chef d'équipe suivie d'un exposé scientifique du contenu de l'un des axes ; (3) une rencontre avec l'équipe de direction du laboratoire ; (4) des rencontres avec les personnels, les doctorants, et la présidence de l'université ; (5) un bref tour de table de conclusion du comité.



3 • Analyse globale de l'unité, de son évolution et de son positionnement local, régional et européen

Le LIPN représente un pôle de référence en informatique dans l'ensemble des universités situées dans le grand Nord-Ouest de Paris. Dans plusieurs de ses axes, se trouvent des personnalités scientifiques dotées d'une visibilité nette, comme l'atteste le rôle qu'elles jouent dans les conférences internationales de leurs communautés d'appartenance. Les domaines couverts par les recherches du laboratoire concernent à part sensiblement égale l'informatique "cœur" [= *core computer science*] et les domaines d'application. Le premier thème recouvre les domaines importants que sont l'algorithmique et les cadres logico-sémantiques de l'informatique. Le second thème, historiquement issu des approches liées à de l'intelligence artificielle, s'est largement restructuré dans une perspective plus scientifique et technique : il concerne actuellement d'une part l'apprentissage, d'autre part le traitement des langues naturelles. L'insertion dans le tissu industriel régional est assurée par un volume sensible de contrats industriels.

Le laboratoire est structuré en quatre thèmes (équipes), constitués en partie sur la base de l'héritage historique ; à savoir : A3 (apprentissage), LCR (logique), OCAD (algorithmique et combinatoire), et RCLN (langue).

L'évolution sur la période du dernier quadriennal a permis de renforcer les axes existants de l'équipe LCR. L'axe "apprentissage" de l'équipe A3 a été consolidé. Il y a eu un lancement d'un nouvel axe dans OCAD sur les grilles de calcul ainsi que la transformation de l'axe "Bases de données" en "entrepôts de données". Enfin, trois axes ont été abandonnés : réseaux, multi-agents et diagnostics. Globalement, ces restructurations conjuguées à une croissance bien maîtrisée sont appréciées par le comité.

4 • Analyse équipe par équipe et par projet

4.1. EQUIPE A3 (Apprentissage Artificiel & Applications)

Nous apprécions une réorientation de l'équipe et une focalisation en direction d'un thème, l'apprentissage, lequel possède, au moins en principe, la caractéristique de pouvoir s'appuyer sur des cadres théoriques solides (computational learning theory ; statistical learning theory, par exemple). Au LIPN, les recherches se déclinent selon deux axes : apprentissage numérique et apprentissage symbolique. Un exemple d'approche transversale prometteur est le domaine de l'apprentissage par renforcement (domaine émergent dans l'équipe, axe apprentissage numérique) appliqué à des représentations relationnelles (axe apprentissage symbolique).

Il s'agit d'une équipe dynamique, en particulier au regard des aspects suivants : 1) le nombre élevé de contrats (ANR, BQR, ARC, RNTL, RNRT, CIFRE) et de projets de coopération au niveau national et international dans lesquels l'équipe est impliquée ; 2) le nombre important de doctorants encadrés pendant la période 2004-2007. Dans le contexte de pénurie de bourses ministérielles, il est à noter que dans l'équipe A3, tous les doctorants sont financés par des contrats. Ces faits témoignent de la forte vitalité de l'équipe A3 et de son responsable.

La présentation de l'axe apprentissage numérique a mis l'accent sur le comportement d'insectes sociaux. Il n'est pas apparu clairement qu'en l'état actuel, son rôle aille bien au delà d'un rôle d'instrumentation (mesurer le mouvement de puces collées sur le dos de fourmis), ni que les données collectées aient une richesse particulière (le déplacement d'un grand nombre de fourmis dans trois chambres). Globalement, même si les manipulations sont susceptibles d'intéresser un laboratoire d'éthologie par l'outil technique, la pertinence de la problématique en informatique et la profondeur des recherches ne sont pas apparues nettement. On ne peut qu'encourager pour l'avenir le développement de recherches visant à analyser, expliciter la généralité des méthodes et renforcer les cadres algorithmiques et/ou statistiques sous-jacents, tout en limitant les approches ad hoc trop liées à des types particuliers de données.

L'axe d'apprentissage symbolique a le mérite de traiter des problèmes "bien posés" dans un cadre formel net, avec des repères, liés tant à la théorie de la complexité qu'aux théories logiques, qui sont bien établies. Il y a ainsi là une construction, dont les progrès sont mesurables et tangibles et à laquelle le comité est particulièrement sensible. On pourrait de plus espérer un effet d'entraînement intéressant de cet axe en direction des problématiques de l'équipe RCLN.



Aspect à améliorer au niveau global de l'équipe : le nombre de publications en revues internationales paraît faible et il y a assez peu de conférences internationales sélectives.

Point positif : il y a eu 2 dépôts de brevet.

Recommandation : Afin d'étoffer le formalisme théorique des travaux dans le domaine de l'apprentissage numérique, il pourrait être opportun de collaborer avec des groupes de travail en région parisienne dans le domaine de l'apprentissage statistique (ENS, Paristech, ...).

4.2. EQUIPE LCR (Logique, Calcul et Raisonnement)

Ce thème est un des points forts du laboratoire et il a été étayé récemment par des recrutements judicieux. Si la nouvelle direction de l'équipe doit encore trouver ses marques en termes de dynamisme et de vision, l'équipe comporte des scientifiques de qualité et animés d'une forte dynamique.

La composante "Calcul combinatoire" a donné lieu à une présentation surprenante centrée sur certains aspects du calcul formel sous-jacent à la théorie quantique et à ses liens avec la combinatoire. D'une certaine façon, l'exposé brillant qui a été donné est "en rupture" avec les approches traditionnelles et nous tenons à saluer une attitude hardie, en décalage avec les cadres traditionnels de la discipline. Il nous semble qu'il y a là, à la fois un sujet porteur, pour lequel le LIPN a une possibilité de faire émerger une problématique très originale, mais aussi des risques. Le risque à terme est que ces recherches se confinent à un cercle de qualité mais étroit de spécialistes en combinatoire algébrique. On serait alors loin des sujets importants de l'informatique contemporaine. L'espoir est que ces travaux, d'une valeur intrinsèque visible, effectuent une jonction avec l'étude des modèles de calcul, notamment distribués, répartis, ou concurrents. Les discussions avec des scientifiques du laboratoire nous ont convaincu qu'ils sont conscients de ces possibilités très porteuses ; il nous semble que ce point devra nécessiter une volonté soutenue en direction de la convergence des composantes en jeu.

Composante logique linéaire et complexité. Un groupe a été constitué autour de la logique linéaire depuis une quinzaine d'années, avec l'arrivée de jeunes recrues. Un thème particulièrement porteur est celui de la complexité implicite, où il s'agit de fournir des langages dans lesquels tout ce que l'on écrit (correctement) présente la garantie de s'exécuter en une complexité déterminée. L'approche de la logique linéaire à ces questions est théoriquement séduisante, mais nécessite d'être confrontée ou intégrée à d'autres approches. Avec le recrutement de deux personnes issues de la culture de la réécriture et de la complexité, ce groupe a acquis le bon spectre de compétences, et apparaît déjà comme l'une des équipes phare internationalement dans cette thématique.

Les autres directions liées à la logique linéaire qui sont développées dans l'équipe (planification, modules) semblent également prometteuses. Cependant, le comité de visite souhaite souligner que la pérennité de cet axe de recherche, et plus globalement de la partie de l'équipe LCR travaillant sur les fondements logiques de la programmation, doit être garantie rapidement par un recrutement d'un jeune senior de niveau A.

Composante réseau de Petri. Ce sujet fait une très bonne impression. La spécification formelle des réseaux de Petri en Coq représente une direction prometteuse, et doit permettre davantage de croisements entre les diverses cultures présentes dans l'équipe LCR. Dans cet esprit, les liens entre réseaux de Petri et logique linéaire pourraient être davantage exploités.

Composante temps. L'approche des S-mots pour le temps fait sens, mais son originalité est orthogonale à la communauté de la linguistique computationnelle. Il pourrait y avoir des rapprochements intéressants avec l'équipe RCLN (ou logique linéaire et grammaires catégorielles) ou encore avec des spécialistes de linguistique formelle hors du laboratoire.

Composante programmation et web. Il s'agit là d'un champ fécond pour des applications élégantes et utiles des idées de la théorie de la programmation. On pense ici notamment à l'usage des continuations pour programmer et contrôler des retournements dans le temps (cf l'action du bouton « back » des navigateurs). L'activité sur ce thème est portée par deux contrats européens et un contrat ANR jeunes chercheurs, ce qui témoigne de son dynamisme.

4.3. EQUIPE OCAD (Optimisation Combinatoire et Algorithmique Distribuée)

Cette équipe regroupe essentiellement des activités centrées autour de la combinatoire, c'est-à-dire de l'étude structurelle et de la l'algorithme des objets mathématiques discrets. Elle est organisée autour de deux axes principaux : optimisation combinatoire et algorithmique distribuée ; elle inclut de plus dans la géométrie actuelle des activités liées aux grilles de calcul. Il y a une bonne logique d'ensemble des deux axes principaux, logique qui est bien ressortie de la présentation générale. En effet, la composante algorithmique distribuée, si elle représente un thème largement lié au distribué/réparti, embrasse aussi des aspects plus vastes liés à l'étude quantitative des structures discrètes aléatoires et à leurs propriétés probabilistes. D'autre part, l'optimisation combinatoire repose dans ses développements récents sur des méthodes d'approximation et



d'exploitation fréquente de l'aléa discret. Il semble que cette équipe aurait vocation à se développer numériquement sur des thèmes peut-être élargis mais certainement voisins de ses centres d'excellence.

Axe optimisation combinatoire. L'équipe est dynamique et productive. Elle a un bon niveau sur un créneau original, notamment en ce qui concerne l'algorithmique des méthodes hybrides (mélange de méthodes exactes et heuristiques), et les méthodes de refactorisation. Il y a des mouvements en cours, et le départ à la retraite d'un leader scientifique, mais le laboratoire a engagé une réflexion pour anticiper cette situation : un poste de professeur est prévu dès cette année. Nous saluons cette initiative. Le renouvellement qui s'opère justifie un ajustement des thèmes de recherche, et nous appuyons les perspectives de s'intéresser davantage aux domaines porteurs que sont la théorie polyédrale, les métaheuristiques, ainsi que l'optimisation non linéaire en nombres entiers.

Axe algorithmique distribuée. L'équipe a incontestablement une expertise reconnue en algorithmique distribuée et les différents thèmes de recherche couvrent fort bien le domaine. L'équipe est dynamique et bien insérée dans le tissu national, où elle a une excellente visibilité. Ses compétences sont originales et distinctives, notamment dans le domaine de l'algorithmique probabiliste et celui de la quantification de l'aléa discret (lié tant au distribué qu'à l'optimisation combinatoire, voire occasionnellement aux réseaux). Un recrutement dans cette voie générale est encouragé.

Axe grille. Une présentation a enfin été faite d'une activité émergente dans le domaine du calcul parallèle et des calculs sur grille. Il nous semble que laboratoire envisage la possibilité d'investir plus dans cette direction. Nous exprimons une certaine réserve à ce sujet. Le calcul sur grille de quelques milliers de processeurs (au niveau national) a une finalité mal définie, notamment si l'on prend en compte les coûts rapportés aux bénéfices scientifiques et technologiques. En effet, si les approches distribuées peuvent être développées à surcoût logiciel, humain, et matériel faible, il n'en est pas de même pour le calcul parallèle : un équipement parallèle est largement lié à une génération de processeurs, une grosse infrastructure matérielle, et à un investissement système lourd et peu recyclable. Notons que le laboratoire et l'Institut Galilée ont investi sur ce thème ; un équipement de 89 machines est à cheval entre LAGA et LIPN. Le positionnement au LIPN est ici dans le développement d'outils logiciels pour la gestion efficace des grilles (sécurité, certification, placement des tâches, optimisation des communications telles que la diffusion), ce qui est en continuité avec les thèmes d'OCAD, et il devrait y avoir place pour des développements algorithmiques voisins des thèmes de l'axe « algorithmique distribuée ». Nous apprécions certainement le dynamisme du leader de cet axe, mais nous mettons en garde contre une possible évolution trop « technique » et dispersée de cet axe.

Notons ponctuellement le rattachement à cette équipe de quelques autres recherches isolées pour lesquelles nous encourageons la direction à mettre en place des mécanismes de fédération et d'intégration adaptés.

4.4. EQUIPE RCLN (Représentation des Connaissances et Langage Naturel)

Cette équipe est historiquement issue des domaines de l'intelligence artificielle, représentation des connaissances, et systèmes experts. Elle s'ouvre aujourd'hui vers la recherche d'information, l'extraction de connaissances, la terminologie. Elle bénéficie d'une direction dynamique et de qualité, récompensée par une médaille de bronze du CNRS en 2004, ce qui est particulièrement significatif dans le cas d'un enseignant chercheur. Pour un domaine d'interface, il importe en effet que la liaison avec l'"autre" discipline (ici, la linguistique) soit bien assuré. Globalement le niveau des recherches est assez bon, bien que son caractère international soit difficile à établir, (l'accent mis sur le traitement de données en langue française en est-il la cause ?). Le niveau des publications paraît assez faible et il n'y a pratiquement pas de publication dans les revues internationales.

L'exposé oral d'une des directions de recherche a laissé transparaître le risque qu'il peut y avoir à se focaliser sur l'assemblage de composants logiciels ; ce semble-t-il au détriment d'un examen de fond des principes logiques et/ou des méthodes algorithmiques véritablement nécessaires à un traitement de pointe des données linguistiques. Sur un plan général, nous avons ainsi du mal à identifier nettement les contributions scientifiques de l'équipe : il s'agit peut-être plus de développement, d'assemblage de composants logiciels que de recherche au sens strict. Le recrutement d'un chercheur, la nouvelle direction d'un professeur actif et dynamique annonce un mieux. Il nous semble qu'un point fort de cette équipe est en liaison avec l'apprentissage comme présenté dans l'exposé spécialisé de l'équipe AAA ("apprentissage symbolique").

Contrairement à ce qui a été annoncé, on peut regretter qu'il n'y ait pas davantage de contacts avec des équipes de linguistique, en particulier de sémantique formelle, ou de linguistique computationnelle plus orientées vers l'informatique fondamentale, et que cette équipe ne participe pas davantage aux actions nationales ou internationales. Il s'agit d'un point important pour lequel un changement est souhaitable. Notons également que l'équipe LCR connaît des liens entre logique linéaire et langues naturelles, qui ne sont plus exploités mais qui pourraient valablement l'être de nouveau.

Nos recommandations sont les suivantes : améliorer le niveau de l'activité de publication, notamment



international ; veiller à renforcer l'activité d'encadrement doctoral ; pratiquer une ouverture vers des approches (algorithmiques, formelles, statistiques) complémentaires ; enfin, puisqu'il a été consacré tant de temps et de moyens pour développer une plate-forme d'annotation, il serait utile de la soumettre à des tests lors d'au moins une des grandes campagnes internationales d'évaluation.

5 • Analyse de la vie de l'unité

Nous apprécions tout spécialement l'excellente gouvernance du laboratoire, le haut niveau d'implication, notamment des membres les plus actifs, les soutiens de l'Université et de l'Institut Galilée. Malgré quelques réserves ponctuelles, la communication et l'ouverture scientifiques des équipes sont à un très bon niveau.

6 • Conclusions

– Points forts :

Les points forts sont une très bonne dynamique du laboratoire avec des domaines d'excellence en informatique cœur : logique (LCR) et combinatoire (OCAD). Il y a une bonne dynamique en apprentissage (AAA) et traitement de la langue (RCLN), mais avec des inégalités entre les forces les plus vives et certaines des composantes. Sur ces deux dernières équipes, on peut conseiller de réserver une place importante aux cadres méthodologiques, formels, logiques, algorithmiques et/ou statistiques ; ce, peut-être, en s'appuyant sur les autres équipes du laboratoire. Pour RCLN, veiller à assurer un niveau suffisant de publications véritablement internationales.

Nous explicitons ci-dessous des recommandations plus ponctuelles. Elles concernent la communication dans le laboratoire (création de séminaires communs), la politique de recrutement (éviter un recrutement à la proportionnelle), le financement du laboratoire (la possibilité d'accroître dans des proportions raisonnables les ressources contractuelles ?), le rapprochement avec le L2TI (au moins « géographiquement »), une politique de réhabilitation du campus (laquelle dépasse le LIPN et est du ressort de l'Université).

Nous avons tous recueilli une très bonne impression de la gouvernance du laboratoire.

– Points à améliorer :

Les rencontres avec les personnels ont révélé l'existence de problèmes de sécurité sporadiques mais assez graves (vols avec violence) : c'est un point qui devrait être pris en compte sérieusement, sans doute avec le soutien actif des tutelles (plan de réhabilitation du campus, délimitation du périmètre universitaire et renforcement de la présence de personnels spécialisés, par exemple). L'état actuel limite l'attractivité du site, pourtant aux portes de Paris. Une amélioration très sensible conditionne la possibilité pour l'université d'attirer des scientifiques et des étudiants du meilleur niveau international.

– Recommandations :

Recrutements : Le plan de développement futur du laboratoire paraît quelque peu « frileux ». Nous recommandons de chercher à étayer prioritairement les forces du laboratoire dans le domaine de l'informatique cœur, de façon à en accroître la part. C'est à dire : éviter un recrutement à la proportionnelle de ce qui existe ; notamment limiter la croissance de l'intelligence artificielle "généraliste", ce qui a été déjà partiellement fait. Sur ces axes, privilégier des recrutements de personnalités scientifiques qui se fondent sur des cadres formels, logiques, statistiques, et algorithmiques précis. Envisager si possible un élargissement (par continuité) des thématiques de l'informatique cœur. Quelques pistes : algorithmique des données textuelles (dans le style de l'École de Marne la Vallée) ; algorithmique probabiliste voire algorithmique quantique (cf les groupes de travail du GDR-IM) ; logique et modèles de calcul (temps réel, circuits, ...).



Financements : Environ un tiers des financements provient du ministère. Ceci semble plus élevé que la moyenne nationale, mais nous paraît justifié par l'importance de la composante de recherche fondamentale au LIPN : ce soutien a bien porté ses fruits. Le financement direct du laboratoire par le CNRS est très faible. Le volume des contrats industriels est considéré, selon les membres du comité, soit comme étant, grosso modo, en adéquation avec le profil du laboratoire, soit comme pouvant être accru dans des limites raisonnables. Si ce second scénario est retenu, il faudra, en tout cas, veiller soigneusement à maintenir une politique ferme visant l'obtention de contrats susceptibles de contribuer à renforcer et diversifier les directions de recherche du laboratoire, tout en formant les doctorants de manière large et en profondeur, au tout meilleur niveau scientifique.

Rapprochement avec le L2TI : La question des rapports entre le LIPN et le L2TI se pose. Les communautés étant largement disjointes (sections différentes du CNU) et les activités très complémentaires, la question d'une éventuelle fusion ne paraît pas à l'ordre du jour. Cependant, des efforts de rapprochement et de convergence au moins partielle pourraient être bénéfiques. Par exemple, un séminaire ou colloquium commun LIPN--L2TI pourrait permettre d'élargir valablement les perspectives des doctorants sur les disciplines concernées. Au niveau des équipes de recherche, il y a possibilité d'interaction entre les activités en combinatoire du LIPN et l'équipe Réseaux du L2TI ; de même entre apprentissage et raisonnement au LIPN et Images au L2Ti. De plus, un certain rapprochement des formations liées fortement à chacun des deux laboratoires pourrait permettre un échange de cultures et une diversification intéressante pour les étudiants. Au niveau de l'Université, nous recommandons ainsi que soit envisagée la perspective d'une évolution des locaux permettant d'assurer à terme le rapprochement des trois laboratoires LAGA (Mathématiques), LIPN, et L2TI. (Le rapprochement LAGA--LIPN est déjà prévu ; la composante « Images » du L2TI collabore déjà très sensiblement avec le LAGA.) Un tel scénario pourrait permettre de mettre en place un continuum scientifique de forte visibilité au sein de l'Institut Galilée. (Note : une recommandation similaire figure dans le rapport d'évaluation du L2TI.)



CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

Sciences et technologies de l'information et de l'ingénierie (ST2I)

D/2008/269

Paris, le 08/04/2008

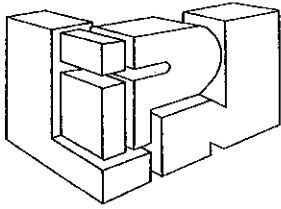
Monsieur le Directeur,

J'ai bien reçu le projet de rapport du Comité d'évaluation de l'UMR 7030 « LIPN » transmis par vos soins et je vous joins la réponse du directeur de l'unité.

Dans l'attente de l'avis définitif de l'AERES sur cette unité, je vous prie de bien vouloir agréer, Monsieur le Directeur, l'expression de mes salutations les meilleures.

Véronique DONZEAU-GOUGE
Directrice scientifique adjointe

Monsieur Jean-Jacques Aubert
Directeur de la section des unités
AERES
20 rue Vivienne
75002 Paris



Réponse du LIPN – UMR 7030 au rapport d'évaluation de l'AERES

L'équipe de direction du LIPN a pris connaissance du projet de rapport faisant suite à la visite du comité d'experts de l'AERES qui s'est tenue le 30 janvier 2008.

Nous nous félicitons que le LIPN ait été à cette occasion bien perçu par le comité d'experts, en particulier sur sa gestion globale et ses orientations scientifiques principales. Le comité d'experts a par ailleurs considéré qu'une partie substantielle des recherches menées était d'un excellent niveau et que celles-ci devaient avoir un soutien amélioré des autorités de tutelle.

Le rapport contient aussi des remarques dont nous tiendrons compte dans l'évolution du laboratoire. Il nous semble que, comme toute critique, celles-ci se justifient par le souci qu'a le comité d'experts de permettre au LIPN d'exprimer ses qualités intrinsèques. Il nous semble toutefois que certaines de ces critiques méritent une réponse de notre part. C'est le cas en particulier pour trois d'entre elles qui semblent remettre en cause une démarche scientifique que nous souhaitons totalement assumer.

- Apprentissage numérique : le rapport nous semble à tout le moins mal refléter la réalité des recherches effectuées au LIPN. S'il est vrai que l'exemple applicatif des insectes sociaux a été évoqué lors de l'exposé oral, il n'en demeure pas moins, comme cela a été apporté en réponse à une question orale par le responsable de l'équipe concerné, que la problématique des recherches menées se situe en amont. Le rapport écrit du LIPN (p.28 en particulier) résume clairement la démarche adoptée. Il s'agit en effet de concevoir de nouveaux algorithmes d'apprentissage génériques, de développer des méthodes hybrides en particulier dans le cadre de l'apprentissage non- ou semi-supervisé. Plusieurs travaux ont ainsi porté sur l'ajustement et le contrôle de la capacité de généralisation dans les systèmes d'apprentissage supervisé (pp. 29 à 32), des études sur l'extension de cette notion à la sélection de modèle et la stabilité dans le cas non supervisé sont en cours (pp. 32 et 33). Les applications effectuées ne servent qu'à valider, valoriser la démarche scientifique amont, en aucun cas s'y substituer. C'était le cas pour cet exemple d'application sur les insectes sociaux dont l'originalité vient par ailleurs de son ancrage dans une recherche pluridisciplinaire (en l'occurrence éthologie-informatique-statistique-électronique). Notons du reste qu'un brevet a été déposé pour la partie relative à l'analyse et la visualisation des flux de données générés par les dispositifs utilisant la RFID et décrivant des comportements dynamiques d'individus.
- Traitement du Langage Naturel : le domaine du Langage Naturel a subi de profondes mutations depuis une quinzaine d'années parce qu'il a fallu apprendre à traiter des masses de données textuelles conséquentes, une problématique que la linguistique informatique traditionnelle prenait mal en compte. Traitement et analyse du Langage Naturel utilisent des cadres formels et des méthodes différentes qu'il faudra *in fine* unifier et que nous ne souhaitons pas opposer. Le traitement de grandes masses de données textuelles requiert une évaluation expérimentale propre et ne peut en l'état se suffire des cadres formels linguistiques. La démarche adoptée par le LIPN consiste, dès lors, à mieux connaître les caractéristiques des données, à concevoir les outils nécessaires d'évaluation des logiciels sans pour autant négliger l'apport de la linguistique formelle. Cette démarche n'est du reste pas isolée dans la communauté, tant en France qu'à l'étranger.

Il est clair cependant que la complémentarité doit être de mise, aussi bien entre Traitement et Analyse, qu'avec ce qui relève de l'apprentissage symbolique ou numérique. C'est dans cette voie que le LIPN s'est engagé au travers des recrutements passés et dans le développement d'un axe transversal entre les équipes RCLN et A3. Cela ne doit pas nous empêcher, comme le suggère le rapport, d'avoir des liens plus étroits avec des laboratoires de linguistique (comme avec le LDI sur l'université Paris 13 par exemple).

- Développement concerté de l'informatique sans dissocier ce qui serait du « coeur » de ce qui relèverait de l'« appliqué » : il s'agit là d'une volonté affichée du laboratoire que de vouloir articuler ces deux optiques, plus que de les dissocier. Nous regrettons de n'avoir pas réussi à mieux faire passer ce message auprès du comité d'experts. Si l'on voulait avoir une lecture des équipes à l'aune de cette distinction, nous trouverions ces deux orientations dans chacune d'elles. Ainsi, l'équipe A3 valorise ses travaux au travers de projets divers, ceux-ci étant fondés sur des bases théoriques qui lui sont propres (analyse statistique, algèbre, treillis de Galois, etc.) ; l'équipe LCR étudie bien les fondements du calcul (et son versant logique) jusqu'à les décliner en termes de langages de programmation, de méthodologie de conception logicielle ; outre les travaux appliqués de l'équipe RCLN qui s'appuient sur la linguistique, la conceptualisation faite en analyse sémantique du langage est originale ; enfin, l'équipe OCAD porte ses efforts sur les méthodes de la combinatoire (optimisation, analyse d'algorithmes), tout en sachant étudier des problèmes concrets relevant tant de l'optimisation des transports ferroviaires que de la communication dans les capteurs mobiles.

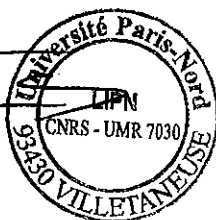
Il nous apparaît que pour les trois points évoqués précédemment, le LIPN entend bien conserver l'originalité de sa démarche scientifique. Le rapport est toutefois pour nous l'occasion de mieux saisir ces lignes de force. A nous de savoir les exploiter et les expliquer. D'autres éléments plus mineurs peuvent aussi être soulignés, car ne correspondant pas tout à fait à la réalité :

- Certains axes du laboratoire sont mentionnés alors qu'ils n'apparaissent pas en tant que tels dans le rapport du LIPN. C'est le cas, par exemple, pour la Composante appelée « réseaux de Petri » dans le rapport, alors que cet axe concerne tout à la fois la spécification et la vérification de logiciels (rapport du LIPN, partie 2.4, pp. 98 et sub.). *A contrario* l'analyse sémantique et l'ingénierie des connaissances ne sont pas mentionnées dans le rapport des experts mais ceci tient à une erreur de notre part : la section décrivant les orientations scientifiques de l'équipe RCLN a malencontreusement été omise dans la version du rapport d'activité qui a été adressée au comité de visite (voir document joint).
- Le rapport du comité d'experts ne mentionne pas les masters s'appuyant sur le laboratoire (mathématique-informatique et informatique), ni l'implication dans des masters externes.

Ces diverses remarques ne nous font pas oublier les éléments portés à notre attention et que nous nous attacherons à améliorer (signalétique et qualité des publications, vigilance quant à certaines recherches trop isolées, ou pour certaines orientations scientifiques, recrutements, collaboration avec d'autres équipes à et hors Paris 13).

le 1^{er} avril 2008

Christophe Fouqueré
Directeur du LIPN – UMR 7030
Université Paris 13 et CNRS



Le Président
de l'Université Paris 13

Jean-Loup SALZMANN

