



# Projet de recherche

Laboratoire informatique

AGROCAMPUS OUEST Centre de Rennes

## Table des matières

<b>1</b>	<b>Préambule</b>	<b>2</b>
<b>2</b>	<b>Partie administrative</b>	<b>2</b>
2.1	Ressources humaines . . . . .	2
2.2	Moyens financiers . . . . .	2
<b>3</b>	<b>Projet de recherche</b>	<b>3</b>
3.1	Problématique : <i>Gestion des écosystèmes par la connaissance</i>	3
3.2	Objectifs scientifiques . . . . .	3
3.3	Thématiques . . . . .	4
3.3.1	Theme 1- Extraction automatique d'informations spatio-temporelles . . . . .	5
3.3.2	Theme 2- Extraction automatique de connaissances à partir de l'expertise . . . . .	7
3.3.3	Theme 3- Construction et exploration de modèle qualitatif d'écosystème . . . . .	8
3.4	Résultats attendus, conclusion . . . . .	10
3.5	Principaux contrats . . . . .	11
3.6	Collaborations scientifiques . . . . .	12
3.7	Auto-évaluation . . . . .	12
3.8	Publications scientifiques (4 dernières années) . . . . .	14

# 1 Préambule

Ce document a pour objectif de présenter le projet de recherche récemment défini par le laboratoire informatique d'AGROCAMPUS OUEST Centre de Rennes dans le cadre d'une demande de rattachement auprès d'une UMR ou d'une structuration sous le label "jeune équipe".

## 2 Partie administrative

### 2.1 Ressources humaines

Le laboratoire est dirigé par Christine Largouët depuis 2006. Il est constitué de deux enseignants chercheurs, maître de conférences, d'un ingénieur de recherche et d'une secrétaire à temps partiel (20%). Tout le personnel est titulaire.

*Liste nominative des enseignants chercheurs*

Nom Prénom	Né en	Etab.	Corps	CNECA	HDR	Arrivée dans l'unité
GUYET Thomas	1980	ACO <sup>1</sup> Rennes	MC	Cneca 3	non	01/02/2009
LARGOUËT Christine	1969	ACO Rennes	MC	Cneca 3	non	01/12/1996

*Liste nominative des ingénieurs, techniciens, administratifs et personnels ouvriers et de service*

Nom Prénom	Né en	Corps	BAP	Tx particip.	Etab.	Arrivée dans l'unité
BONNEAU Louis	1977	IR	E	1	AO Rennes	2002
GAILLARD Marie-France	?	?	?	0,2	AO Rennes	2008

En 2009, le laboratoire a accueilli un chercheur post-doctorant (sur un poste d'IE), Yves-Marie Bozec pour une durée de 6 mois. En 2010, un stagiaire de M2 en informatique, Yulong Zhao, a réalisé son stage au sein du laboratoire.

### 2.2 Moyens financiers

L'équipe dispose d'une dotation annuelle de 1000 euros par enseignant-chercheur.

## 3 Projet de recherche

### 3.1 Problématique : *Gestion des écosystèmes par la connaissance*

Par son positionnement au sein d'un institut supérieur en sciences agronomiques, agro-alimentaire, horticoles et du paysage, le laboratoire informatique a pour objectif de développer des approches de modélisation dans le but de "mieux comprendre pour mieux gérer" des systèmes complexes tels que les écosystèmes.

Le sentiment d'urgence face aux problèmes liés à l'environnement a accru la demande de modèles capables de comprendre le lien complexe existant entre les activités humaines, l'environnement et la capacité d'adaptation de l'écosystème. Ces modèles doivent nous aider à comprendre et à expliquer le fonctionnement global de l'écosystème dans le but de pouvoir mieux le gérer. Si les problèmes spécifiques émanent des chercheurs thématiques, les questions posées, qui peuvent souvent se généraliser à plusieurs classes de problèmes, constituent de vrais défis en informatique. Lorsqu'il s'agit de modéliser dans son ensemble des systèmes complexes tels que des écosystèmes, les difficultés consistent alors à intégrer des composants et leurs interactions qui sont parfois mal connus ou difficiles à mesurer. À ces incertitudes, se rajoutent la variabilité spatio-temporelle de tels systèmes et l'intégration de connaissances expertes. Par manque d'exhaustivité, le recours à la modélisation classique, dite mathématique, est généralement impossible. Dans ce contexte, le laboratoire informatique propose des approches alternatives permettant aux thématiques, de différentes disciplines, de créer et d'exploiter des modèles d'écosystèmes. Les questions de recherche abordées consistent donc à trouver la représentation adaptée au système à modéliser, à acquérir les connaissances permettant de construire le modèle puis à exploiter ce modèle pour comprendre et aider à gérer le système.

Les approches abordées par le laboratoire appartiennent au domaine de l'Intelligence Artificielle par le souci permanent de représenter des informations incomplètes et imprécises, d'intégrer des représentations expertes, d'approcher la modélisation qualitative, d'acquérir des connaissances symboliques, même si les formalismes de représentation soient parfois empruntés à d'autres communautés de l'informatique.

### 3.2 Objectifs scientifiques

L'objectif de recherche du laboratoire est de proposer des méthodes et des outils permettant de gérer les écosystèmes à l'aide de la connaissance. Dans cette démarche, les principaux centres d'intérêt sont les suivants :

- l'acquisition de connaissances à partir de données,

- l’intégration d’informations hétérogènes dans la construction de modèles globaux de systèmes complexes : données numériques issues de capteurs ou d’images, connaissances expertes ou extraites par apprentissage symbolique,
- la représentation des informations temporelles (dynamique, contraintes, granularité) par un formalisme approprié,
- la supervision des systèmes par l’exploration des modèles à l’aide de techniques efficaces afin de proposer une assistance intelligente, fiable et flexible, aux utilisateurs.

Pour les quatre années à venir, nous souhaitons traiter ces centres d’intérêt dans des approches théoriques, permettant de traiter plusieurs classes de systèmes à modéliser, mais également dans des applications spécifiques du domaine de l’environnement, en partenariat avec des équipes thématiques. Il s’agit de proposer des approches nouvelles du point de vue de la recherche en informatique puis de proposer des outils efficaces permettant leur validation sur des applications environnementales concrètes.

Les recherches menées par ce projet se situent dans la communauté de l’Intelligence Artificielle pour les aspects théoriques et les travaux viseront à être publiés dans les supports de ce domaine : conférences internationales (ECAI, IJCAI, KDD, ICML, ICDM, etc.), nationales (RFIA, EGC, IC, SAGEO, etc.), journaux (Artificial Intelligence, Journal of Artificial Intelligence Research, Applied Artificial Intelligence, Revue d’intelligence artificielle, etc.). Etant donné la nature pluridisciplinaire du projet de recherche et par souci de diffusion auprès des communautés concernées, une valorisation des outils développés sera également effectuée dans les supports appropriés tels que les journaux Ecological Modelling, Environmental and Modelling Software, etc.

### 3.3 Thématiques

Nous focalisons notre recherche autour de trois thèmes :

1. Extraction automatique d’informations spatio-temporelles,
2. Extraction automatique de connaissances à partir de l’expertise,
3. Construction et exploration de modèle qualitatif d’écosystème.

Pour construire des modèles de fonctionnement des écosystèmes, on acquiert des connaissances à partir de données de deux types : les données d’observation ou l’expertise. L’acquisition des connaissances à partir d’un important volume de données d’observations est réalisée par le thème 1. Ce point s’intéresse plus spécifiquement à l’extraction de données spatio-temporelles très caractéristiques des systèmes agro-écologiques. Le domaine scientifique auquel se rattache ce travail est celui de la “fouille de données” et de

l’“extraction de connaissances”. L’acquisition de l’expertise à partir des dires d’acteurs est réalisée par le thème 2. Dans le domaine de l’environnement, une grande connaissance issue d’enquêtes sur le terrain existe, mais l’absence de formalisation et l’hétérogénéité des données la rend généralement sous-exploitée par les modèles. Ce travail s’intéresse ici à l’extraction de la connaissance utile à partir d’un ensemble de graphes de concepts. Il se situe dans le thème scientifique de l’“ingénierie des connaissances et ontologies”. Les thèmes 1 et 2 peuvent être des préambules nécessaires à la construction de modèles.

Dans le thème 3, la démarche de modélisation qualitative intègre toute information symbolique pouvant être issue, soit des processus d’extraction de connaissances, soit de la discrétisation des résultats fournis par des modèles numériques connus. Cet axe de recherche se rattache aux thématiques scientifiques : “représentation des connaissances” et “modélisation”. Par ses aspects de prédictions et d’aide à la décision, on peut également le raccrocher à la “supervision des systèmes complexes”.

Outre leur appartenance à une démarche générale permettant de proposer des méthodes et des outils de gestion des écosystèmes, la cohérence scientifique est assurée par la prise en compte de la dimension temporelle dans la formalisation des connaissances dans chacun des trois thèmes ainsi que par les applications traitées. Les applications privilégiées sont aujourd’hui l’halieutique (thèmes 1 et 3) et l’agronomie, pour les aspects de paysages et de gestion des bassins versants (thèmes 2 et 3).

### **3.3.1 Thème 1- Extraction automatique d’informations spatio-temporelles**

#### **Problématique**

La recherche dans le domaine de l’extraction de connaissances (Knowledge Discovery and Delivery ou KDD) développe des outils interactifs de fouille de données pour extraire des structures sous-jacentes à des bases de données<sup>2</sup>. Ce domaine s’intéresse, d’une part, aux problèmes algorithmiques de la fouille de données et, d’autre part, à la conception de schémas d’interaction pour aider les utilisateurs à extraire des connaissances utiles.

Pour nos applications agro-écologiques (organisation spatiale de l’espace agricole, flux de capteurs géolocalisés), la nature spatio-temporelle des données doit être explicitement prise en considération. Cependant, l’information spatio-temporelle (quantitative ou qualitative) est encore sous-utilisée dans les processus classiques de KDD et sa prise en compte ouvre de nouvelles perspectives de recherche<sup>3</sup>.

---

2. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, “From Data Mining to Knowledge Discovery in Databases”, *AI Magazine*, 17(3) :1590-1592, 1996.

3. H. Miller and J. Han, “Geographic Data Mining and Knowledge Discovery - Second

La problématique de ce thème est de développer des processus d'extraction de connaissances spatio-temporelles. Ceci inclut, d'une part, le développement de méthodes de fouilles de données spatio-temporelles, c'est-à-dire où la dimension spatio-temporelle des structures extraites automatiquement est un résultat du processus d'extraction et, d'autre part, leur intégration dans des outils interactifs.

### **Programme des travaux**

Dans des travaux antérieurs, nous nous sommes intéressés à la fouille de séries chronologiques et de séquences d'événements temporisés. Nous avons travaillé en particulier à l'adaptation des méthodes d'extraction de motifs (algorithmes APriori, PrefixSpan ou Eclat<sup>4</sup>) au cas de données séquentielles et temporisées. Notre objectif est maintenant de proposer des algorithmes pour la fouille de séquences d'intervalles temporels, c.-à-d. une suite d'événements  $(A, B, \dots)$  associés à des intervalles de temps  $(\{(A, [1.3, 2]), (B, [2.1, 4.5]), \dots\})$ . Ce travail est réalisé en collaboration avec des chercheurs de l'INRIA (EPI DREAM). De tels algorithmes permettraient, par exemple, d'affiner la caractérisation des successions culturelles avec des dates réellement observées dans les itinéraires techniques.

Récemment, nous avons introduit l'utilisation de ces mêmes méthodes d'extraction de motifs pour la caractérisation d'un paysage agricole (organisation spatio-temporelle de l'occupation du sol d'une région agricole). Ce travail nous a conduit à proposer une représentation du paysage sous la forme d'un graphe de parcelles décrivant des informations spatiales riches (forme des parcelles, distances, etc.). L'originalité applicative de ce travail se trouve dans la caractérisation structurelle d'un paysage par l'utilisation de motifs locaux (c.-à-d. un sous-graphe du graphe de paysage ayant des propriétés informationnelles intéressantes sur le paysage). Le défi algorithmique que pose ce problème réside dans la proposition d'algorithmes efficaces pour extraire des motifs locaux (dont la structure reste encore à définir) à partir de graphes étiquetés par des informations complexes. Ce travail est réalisé dans le cadre d'un projet en collaboration avec des équipes de recherche en informatique et des équipes thématiques (cf. projet PayoTe).

Face à des données aussi complexes que les données temporelles ou spatiales, les experts comme les outils automatiques ne sont pas en mesure d'extraire simplement les informations pertinentes. Et pour tirer au mieux parti des compétences spécifiques des utilisateurs et des outils automatiques, il est nécessaire de concevoir des systèmes interactifs. Nous travaillons à l'intégration de nos outils de fouille de données dans des systèmes permettant une collaboration entre la machine et l'utilisateur inspirée de la théorie

---

edition", Taylor and Francis, 2009.

4. Pour une revue récente des méthodes d'extraction de motifs, voir : J. Han, H. Cheng, D. Xin, and X. Yan, "Frequent Pattern Mining : Current Status and Future Directions", *Data Mining and Knowledge Discovery*, 15(1) :55-86, 2007.

de l'Enaction<sup>5</sup>. Nous étudierons l'interaction collaborative au travers du développement d'un outil interactif pour la fouille de paysages.

### 3.3.2 Theme 2- Extraction automatique de connaissances à partir de l'expertise

#### Problématique

Pour la modélisation des systèmes naturels complexes, le savoir des acteurs locaux constitue une source essentielle de connaissance. Nous cherchons à intégrer cette expertise, exprimée par les acteurs lors d'entretiens de terrain, au sein d'une démarche d'aide à la concertation.

Pour un écosystème particulier, différentes visions sont ainsi recueillies sous la forme de graphes de causalités entre concepts, appelés cartes cognitives. Les concepts sont formés de mots ou de groupes de mots identifiant des événements, des activités, des organismes... Les relations de causalité portent une valuation traduisant l'intensité et la durée du phénomène transcrit.

Notre objectif est d'extraire la connaissance utile par l'identification de thématiques portant l'adhésion d'un groupe ou au contraire suscitant la division, la définition de classes d'acteurs, ou la production automatique de résumés. À cette fin, il est nécessaire de pouvoir rapprocher les concepts employés par les acteurs lors de leur description du système.

#### Programme des travaux

Produire une mesure de similarité entre concepts est une tâche délicate puisque les experts ne s'expriment pas avec la même précision ou le même champs lexical. De manière à pallier l'hétérogénéité linguistique rencontrée, une ontologie a été construite, intégrant tout le vocabulaire rencontré dans les cartes cognitives. Le calcul de similarité au sein d'un réseau sémantique a été largement étudié et différentes approches confrontées<sup>6</sup>. Cependant, ces mesures sont parfois biaisées par une approche taxonomique. Par ailleurs, les modèles linguistiques construits à partir des relations lexicales classiques (hyponymie, hyponymie, synonymie, etc.) ne traduisent généralement pas correctement les structures conceptuelles sous-jacentes<sup>7</sup>. Ainsi les calculs de similarité classiques ne prennent pas en compte les relations fonctionnelles du type "se produit dans", "habite", "utilise". Or ce type de relations est souvent utilisé implicitement par les personnes décrivant le fonctionnement d'un écosystème, le recours à la métonymie pour substituer un mot à un autre en est un bon exemple.

---

5. P. De Loor, K. Manac'h, J. Tisseau, "Enaction-Based Artificial Intelligence : Toward Co-evolution with Humans in the Loop", *Minds and Machines*, 19 :319-343, 2009.

6. A. Budanitsky, G. Hirst, "Evaluating WordNet-based Measures of Lexical Semantic Relatedness", *Computational Linguistics*, 32 :13-47, 2006.

7. C. Roche, "Le terme et le concept : fondements d'une ontoterminologie", *Terminologie & Ontologie : Théories et Applications*, 2007.

Aussi l'approche proposée consiste à construire une mesure de similarité reposant sur la connaissance modélisée par notre réseau sémantique, incluant les relations lexicales et les liens fonctionnels, et de confronter les résultats de cette mesure à des affirmations d'experts. Cette mesure étant ensuite utilisée pour comparer les cartes cognitives entre elles : soit de manière directe, en recherchant des appariements partiels de graphes, soit en comparant des ensembles d'attributs décrivant ces graphes. Chaque carte cognitive pouvant, par exemple, être vue comme un sous-ensemble de l'ensemble des thématiques abordées.

### **3.3.3 Theme 3- Construction et exploration de modèle qualitatif d'écosystème**

#### **Problématique**

Si les modèles de simulation peuvent être utilisés dans un contexte d'aide à la décision, par exemple pour prédire ce qui va résulter d'actions décidées sur le terrain, ou pour suggérer les actions qui pourraient améliorer la situation, il est souvent difficile de modéliser dans son ensemble l'écosystème et son environnement (naturel ou anthropique). Les modèles qualitatifs permettent d'éviter certains problèmes de représentation lorsque les données sont difficiles à acquérir, imprécises ou incomplètes. Ces connaissances sont souvent plus faciles à acquérir auprès des experts, ce qui les rend, de plus, adaptées à la construction d'outils d'aide à la décision.

Ce travail propose une nouvelle approche pour modéliser puis explorer un écosystème afin d'améliorer la compréhension du système, de prédire les effets à terme d'actions sur le terrain, de déterminer les valeurs critiques des variables impliquées et ainsi de mieux évaluer les actions possibles en terme de gestion et de gouvernance. Ces problèmes ne sont pas simples surtout pour les systèmes dynamiques dont l'évolution dans le temps est complexe comme dans le cas du contrôle de population. Nous proposons de modéliser le système de manière qualitative et de définir un langage de haut-niveau afin d'interroger le modèle. Le comportement de l'écosystème est représenté par un formalisme de type système à événements discrets (SED) tels que les automates temporisés, les réseaux de Petri, les modèles hybrides ou l'algèbre de processus. Dans ce contexte, l'originalité consiste à proposer une formalisation qui, en fonction du type d'écosystème, caractérise au mieux la dynamique du système et de son environnement. De manière corrélée, un langage de haut-niveau sera proposé afin de permettre aux acteurs d'interroger le système en fonction des différents scénarios (politiques appliquées, évolution de l'environnement) qu'il souhaite explorer.

#### **Programme des travaux**

Une première approche, encore en cours, est développée sur cette thématique. Elle présente une approche générique pour représenter l'écosystème comme

un ensemble d'automates temporisés synchronisés<sup>8</sup> sur lequel viennent se greffer des automates contextuels représentant les forces contrôlables ou non, agissant sur le système (pressions anthropiques, catastrophes naturelles, habitat). Ce travail définit un langage de haut niveau s'appuyant sur une logique temporelle<sup>9</sup> permettant à un utilisateur d'exprimer, sous une forme similaire à une requête de base de données, le scénario qu'il veut explorer. Les techniques de model-checking<sup>10</sup> permettent d'obtenir des réponses à des scénarios prédictifs qui ont pour objectif d'interroger le système sur ses états futurs en fonction d'une politique de gestion proposée et d'un environnement climatique envisagé. Ce travail est réalisé en collaboration avec Marie-Odile Cordier de l'équipe DREAM de l'IRISA. Après avoir exposé l'approche sur un écosystème jeu, une modélisation sur un écosystème corallien de la Nouvelle-Calédonie a été effectuée.

Une première étape va consister à valoriser ces nouveaux travaux par des publications, à diffuser une première plateforme de modélisation, et à recueillir les retours des thématiciens sur l'utilisation, à la fois, de la démarche et du produit proposé. À ce sujet des collaborations sont engagées avec des chercheurs australiens (Rich Little du CSIRO et Craig Johnson d'UTAS) pour que cette approche de modélisation qualitative soit utilisée sur de nouvelles applications de gestion des pêches.

Une perspective consiste à augmenter les possibilités d'interaction dans la modélisation. On pourrait, par exemple, se focaliser sur une sous-partie du système ou intégrer plus précisément le changement d'échelle. Une autre évolution est d'intégrer dans le modèle des éléments évoluant de manière dynamique dans l'espace (l'habitat en particulier) ce qui nécessite de repenser le formalisme de représentation.

En ce qui concerne les aspects théoriques, un nouveau champ d'étude va porter sur le traitement des scénarios proactifs. Si pour une politique donnée, on est capable de savoir ce qui peut se passer, il manque cependant un scénario essentiel pour les décideurs, qui répondrait à la question "Quelle politique de pêche appliquer pour que l'écosystème atteigne un état souhaité?". Ce type de requête peut être résolue en appliquant la théorie de la synthèse de contrôleur sur les automates<sup>11</sup>. Un premier travail va consister à adapter la modélisation de l'écosystème afin de pouvoir définir les états "souhaités" et "interdits" requis par la synthèse de contrôleur. Un autre axe s'intéresse à un nouveau type d'interaction avec l'utilisateur (extension du

---

8. R. Alur and D.L. Dill. "A theory of timed automata". *Theoretical computer science*, 126 :183-235, 1994.

9. P. Bouyer. "Model-checking timed temporal logics", *Proc. of 4th Workshop on Methods for Modalities*, 231 :323-341, 2009.

10. P. Schnoebelen, B. Bérard, M. Bidoit, F. Laroussinie, et A. Petit. "Vérification de logiciels : techniques et outils du model-checking". *Vuibert*, 1999.

11. P. J. Ramadge and W. M. Wonham. "Modular supervisory control of discrete event systems". *Proc. of 7th Int. Conf. Analysis and Optimization*, 83 :183-235, 1986.

langage de haut-niveau) pour recueillir ses requêtes quant à l'évolution souhaitée du système selon les échelles de temps. Enfin, un point important du travail est le traitement des résultats fournis par la synthèse de contrôleur afin d'en extraire une connaissance explicite sur les politiques à appliquer. Ce travail va faire l'objet d'une thèse entre l'UMR SAS, l'équipe DREAM de l'IRISA et le laboratoire informatique d'AGROCAMPUS OUEST. La démarche sera ici appliquée à un écosystème différent du réseau trophique, puisqu'il s'agit, en agronomie, des flux de nitrates dans un bassin versant à l'échelle de quelques décennies.

### 3.4 Résultats attendus, conclusion

Certains thèmes sont plus engagés que d'autres ce qui fait que du point de vue de la production scientifique, les exigences seront adaptées. Le thème 3, plus avancé, visera pour ces quatre prochaines années des publications dans des journaux internationaux tels que Applied Artificial Intelligence, Ecological Modelling ou Ecological and Modelling Software. Les thèmes 1 et 2 auront une stratégie de publications dans des conférences et journaux nationaux, au moins pour les deux prochaines années.

Un autre axe de valorisation est la production de logiciel, qui est déjà non négligeable étant donné la taille de l'équipe. Deux plateformes logicielles sont disponibles :

- OntoMap est un outil en ligne de comparaison de points de vue sur un système, exprimés par différents acteurs, sous forme de cartes cognitives. Son originalité consiste à employer une ontologie (définition de plus d'un millier de concepts) permettant la classification et la synthèse de cartes cognitives. Son objectif est de faciliter la concertation entre acteurs et la prise de décision.
- EcoMata est une plateforme permettant la modélisation d'un réseau trophique (marin pour le moment) et son exploration à l'aide de requêtes prédéfinies. La plateforme permet de définir des scénarios de pêches et des perturbations environnementales puis d'interroger le système sur son évolution dans le temps.

En ce qui concerne les champs d'application, OntoMap a été appliqué sur les cartes cognitives décrivant l'écosystème marin de la Manche. Etant donné la difficulté de recueillir les cartes cognitives, il serait intéressant d'établir des collaborations avec des sociologues et des thématiciens pour mener d'autres campagnes. Toujours dans le domaine halieutique, EcoMata a permis la modélisation d'un écosystème récifo-corallien. Une étude sur un écosystème non récifal, sur lequel des données d'observation sur plusieurs années sont disponibles est envisagé, en particulier avec nos partenaires d'AGROCAMPUS OUEST et du CSIRO (une thèse en écologie doit uti-

liser le logiciel). L'apprentissage de données spatio-temporelles est pour sa part appliqué à l'agronomie pour l'étude des paysages. Etant donné l'aspect pluridisciplinaire de notre travail, nous souhaitons poursuivre nos collaborations mais également initier de nouveaux projets afin de pouvoir appliquer nos méthodes sur des applications réelles. Dans ce but, il est nécessaire de se faire connaître dans les réseaux concernés (participation à des ateliers ou à journées très thématiques) puis de convaincre des partenaires de l'intérêt de méthodes alternatives, ce qui demande du temps, difficile à valoriser.

Concernant les ressources humaines, le laboratoire souhaite recruter un post-doctorant afin de renforcer la thématique d'extraction automatique de l'expertise. Autour de la modélisation d'écosystèmes, une thèse en co-tutelle est envisagée avec l'université de Hobart, en particulier sur l'intégration de modèles pour la représentation d'un écosystème marin. Concernant les aspects de synthèse de contrôleur, une autre thèse doit démarrer en octobre 2010 (les deux co-directrices sont Marie-Odile Cordier de l'équipe DREAM et Chantal Gascuel de l'UMR SAS et l'encadrement sera également assuré par Christine Largouët ). Enfin, le recrutement d'un AERC est espéré pour les années à venir.

### 3.5 Principaux contrats

Nous participons aux projets nationaux suivants :

- PayoTe2 (projet de collaboration INRA-INRIA) : ce projet se situe dans la continuité du projet PayoTe mais se focalise sur deux points :
  - la fouille et la caractérisation de paysages réels (formes et distribution spatiale des parcelles, distribution spatiale des occupations du sol ...);
  - la simulation de paysages virtuels dotés de caractéristiques identifiées dans 1.

Les équipes participantes sont, d'une part, des équipes de mathématiciens et informaticiens (INRA et INRIA), d'autre part, des équipes d'agronomes écologues (INRA). La durée du projet est de deux ans (2010-2011). Le soutien couvre les besoins pour les indemnités d'un stage de M2 et les déplacements (trois réunions et un colloque national).

- Ra2Brest (programme LITEAU du Ministère de l'Écologie, de l'Énergie, du développement durable et de la Mer) : ce projet a pour ambition d'apporter des connaissances aux acteurs de la zone côtière en proposant une modélisation et une scénarisation des activités humaines en rade de Brest. Le projet se déroule sur les années 2009 et 2010, et nous sommes impliqués, avec Guy Fontenelle du laboratoire halieutique d'AGROCAMPUS OUEST sur la construction des scénarios

d'évolution plausibles en fonction de différentes hypothèses de gestion. Le soutien couvre le financement d'un stagiaire de Master 2.

### 3.6 Collaborations scientifiques

Au niveau international, nous avons de bonnes relations avec les chercheurs australiens du CSIRO de Hobart. Nous envisageons une thèse en cotutelle avec Rich Little du CSIRO et un projet avec Jeff Dambacher autour de la modélisation qualitative. Les travaux se poursuivent avec Yves-Mare Bozec, post-doctorant, dans notre laboratoire et actuellement à l'Université de Queensland. Enfin des travaux devraient être initiés sous peu avec Craig Johnson de l'Université de Tasmanie.

Au niveau national, des liens sont établis avec le LORIA autour du thème d'extraction des connaissances (Florence Le Ber et Jean-François Mari).

En local, Christine Largouët et Thomas Guyet sont chercheurs extérieurs de l'équipe DREAM de l'IRISA. Des travaux sont menés régulièrement avec les chercheurs de cette équipe (Marie-Odile Cordier, René Quiniou, Véronique Masson) sur les thèmes 1 et 3.

Au sein d'AGROCAMPUS OUEST et de l'INRA de Rennes, des relations privilégiées sont établies avec le laboratoire halieutique (en particulier Guy Fontenelle et Etienne Rivot), l'UMR SAS pour les travaux portant sur la modélisation d'écoulement de nitrates.

### 3.7 Auto-évaluation

La structuration des activités de recherche du laboratoire informatique est un nouveau défi puisque, jusqu'il y a peu de temps, ses missions étaient essentiellement tournées vers l'enseignement. Cette évolution s'est concrétisée par le recrutement d'un nouvel enseignant-chercheur et par le souci de mener une recherche unifiée autour des "gestion des écosystèmes par la connaissance". L'équipe du laboratoire informatique est jeune et dynamique et ses membres sont connus dans la communauté française de l'Intelligence Artificielle. Cette année, des collaborations encourageantes avec des chercheurs étrangers ont été initiées offrant une nouvelle visibilité au laboratoire. Christine Largouët puis Thomas Guyet sont intégrés et participent activement aux activités de recherche d'une équipe de l'IRISA. Par ailleurs, le laboratoire est soutenu localement par la direction d'AGROCAMPUS OUEST.

Les faiblesses de notre équipe sont incontestablement sa taille (deux enseignants-chercheurs et un ingénieur de recherche) ainsi que l'absence d'HDR au sein du laboratoire. L'absence de cursus à forte composante informatique au sein d'AGROCAMPUS OUEST ne permet pas aux membres de l'équipe de proposer des stages ou des thèses à des étudiants de l'école (sauf pour les cas particuliers d'application thématique). Alors que la re-

cherche devient moins attractive pour les étudiants en informatique, il est difficile pour une petite structure (face à l'université de Rennes 1 principalement) de recruter les étudiants qui ne connaissent pas les enseignants. Enfin, la particularité même de notre recherche, implique une pluridisciplinarité revendiquée mais gourmande en temps qu'il est d'autant plus difficile de compenser que l'équipe est petite. Une dernière contrainte est la part importante d'enseignement réalisée par les membres de l'équipe qui, sans vigilance, pourrait empiéter sur les missions de recherche.

### 3.8 Publications scientifiques (4 dernières années)

#### Chapitres d'ouvrage

- [1] Christine Largouët et Marie-Odile Cordier, “Reconnaissance de l’occupation du sol à l’aide d’un automate temporisé”, Ouvrage *Raisonnements sur l’espace et le temps : des modèles aux applications*, Traité IGAT, série Géomatique, Lavoisier, Paris, 2007, Pages 317-332.
- [2] Jean-François Mari et Christine Largouët, “Quelques modèles graphiques pour le raisonnement temporel et spatial”, Ouvrage *Raisonnements sur l’espace et le temps : des modèles aux applications*, Traité IGAT, série Géomatique, Lavoisier, Paris, 2007, Pages 249-271.

#### Revue avec comité de sélection

- [3] Thomas Guyet, Catherine Garbay et Michel Dojat. “Knowledge construction from time series data using a collaborative exploration approach” *Journal of Biomedical Informatics*, Vol. 40, August 2007, Pages 672-687
- [4] Julie Scopélitis, Serge Andrefouët et Christine Largouët, “Modelling coral reef habitat trajectories : evaluation of an integrated timed automata and remote-sensing approach”, *Journal Ecological Modelling*, Vol. 205, Issues 1-2, July 2007, Pages 59-80

#### Congrès d’audience internationale avec comité de sélection

- [5] Christine Largouët, Marie-Odile Cordier et Guy Fontenelle, “Scenario templates to analyse qualitative ecosystem models”, *18th World IMACS Congress and MODSIM09 (International Congress on Modelling and Simulation. Modelling and Simulation)*, July 2009, Pages 2129-2135.
- [6] Thomas Guyet et René Quiniou. “Mining temporal patterns with quantitative intervals”. *The 4th International Workshop on Mining Complex Data, (IEEE ICDM Workshop)* 2008, Page 10.
- [7] Wei Wang, Thomas Guyet and Svein Knapskog. “Autonomic Intrusion Detection System”. *Proceedings of the International Symposium On Recent Advances In Intrusion Detection (RAID)* : 2009 (Poster)
- [8] Wei Wang, Florent Massegli, Thomas Guyet, René Quiniou et Marie-Odile Cordier. “A General Framework for Adaptive and Online Detection of Web attacks”. *Proceedings of the 18th international World Wide Web conference (WWW 2009)* : Pages 1141-1142. 2009 (Poster)

### **Congrès d'audience nationale avec comité de sélection**

[9] Christine Largouët and Marie-Odile Cordier, “Patrons de scénarios pour l’exploration qualitative d’un écosystème” , *RFIA 2010 (Reconnaissance des Formes et Intelligence Artificielle)*, Caen, Janvier 2010, Pages 58-66.

[10] Thomas Guyet, Catherine Garbay et Michel Dojat. “Interprétation collaborative de séries temporelles”. *Colloque de l’Association pour la Recherche Cognitive (ARCo’09)*, 2009, Page 10.

### **Ateliers**

[11] Marie-Odile Cordier, Thomas Guyet, Christine Largouët, Véronique Masson et Henri-Maxime Suchier, “Apprentissage incrémental de règles de décision à partir de données d’un simulateur”, Atelier INFORSID : Systèmes d’Information et de Décision pour l’Environnement, Toulouse, 2009, Pages 84-93.

### **Démonstrations logicielles**

[12] Thomas Guyet, René Quiniou and Marie-Odile Cordier. “LogAnalyzer : monitoring adaptatif d’un flux de données”. *RFIA 2010 (Reconnaissance des Formes et Intelligence Artificielle)*, Caen, Janvier 2010, Pages 881-882.

[13] Louis Bonneau de Beaufort, Guy Fontenelle et Christine Largouët. “OntoMap : ontologie et cartes cognitives”. *RFIA 2010 (Reconnaissance des Formes et Intelligence Artificielle)*, Caen, Janvier 2010, Pages 877-878.