

Résumés des présentations lors de la journée environnement et numérique

Mardi 19 octobre 2021

Traitement et analyse de données génomiques

- **Indexation de données génomiques, du moteur de recherche générique à l'identification de souches bactériennes**

Pierre Peterlongo ([GENSCALE](#))

Supposons que vous soupçonniez avec bonheur que “ACGGCAT...CGGATAG” est un morceau de gène majeur dans un domaine de recherche fondamental, par exemple la résistance des plantes à la sécheresse ou l'adaptation du plancton au changement climatique. L'étape suivante consisterait à vérifier si ce gène ou une version très similaire a déjà été vu quelque part ou non. Et si c'est le cas, les étapes suivantes consistent à vérifier les différences potentielles, le contexte dans lequel ce gène a été observé et, dans le meilleur des cas, s'il était déjà associé à une espèce ou une souche spécifique, et à un phénotype particulier. Aussi surprenant que cela puisse paraître, ces recherches ne sont pas réalisables à l'échelle mondiale aujourd'hui. Dans cet exposé, j'expliquerai ce contexte, et je montrerai les travaux en cours dans

l'équipe pour proposer des moteurs de recherche capables de répondre à de telles requêtes.

- **Détection de variants génomiques dans des populations naturelles pour étudier l'évolution et l'adaptation de diverses espèces d'insectes**

Claire Lemaitre ([GENSCALE](#))

L'analyse des différences de séquences entre les génomes d'individus provenant de populations naturelles permet de mieux comprendre comment les espèces évoluent et s'adaptent à leur environnement. Les technologies actuelles de séquençage permettent d'acquérir à faible coût des données de séquences pour de nombreux individus. Cependant, les données produites sont massives, fragmentées et bruitées et requièrent des traitements informatiques spécifiques pour reconstruire l'information génomique et détecter les variations de manière efficace. Dans l'équipe Genscale, nous sommes impliqués dans plusieurs projets d'écologie et de génomique d'espèces d'insectes, tels que le puceron du pois et des papillons mimétiques de l'Amazonie. Nous présenterons plusieurs méthodes que nous avons développées dans ce cadre pour mieux détecter et interpréter les variations génomiques dans ces données.

- **Intégration et analyse de données biologiques hétérogènes par exploitation de graphes multicouches pour comprendre et prédire des variations d'efficacité alimentaires chez le porc**

Camille Juigné ([DYLISS](#))

La filière porcine est aujourd'hui confrontée à diverses problématiques liées à la concurrence, la compétitivité ainsi qu'à l'environnement. Améliorer l'efficacité alimentaire (EA) est une bonne solution pour traiter ces problématiques. L'efficacité alimentaire (EA) est la valorisation des ressources pour produire de la viande de qualité en maintenant l'ensemble des fonctions physiologiques des animaux, en réduisant les effluents et en optimisant l'utilisation des ressources. Ainsi l'objectif de cette thèse est de développer des méthodes et analyses d'intégration de données multi-omiques pour identifier les mécanismes responsables des différences d'EA et les bases biologiques de cette variabilité inter-individuelle ou entre lignées.

Systèmes cyberphysiques respectueux de l'environnement

- **SeDuCe (Sustainable Data Centers), un noeud Grid'5000 alimenté pour partie par des énergies renouvelables.**

Jean-Marc Menaud ([STACK](#))

SeDuCe une infrastructure recherche dédiée à l'étude scientifique des problématiques croisées des centres de données et de leur alimentation électrique par des sources d'énergies renouvelables.

Cette infrastructure est physiquement hébergé à l'IMT Atlantique, campus de Nantes, et est l'un des sites Grid'5000.

Tout utilisateur de Grid'5000 peut ainsi réserver et utiliser les serveurs du cluster écotype en utilisant les outils habituels de Grid'5000.

Ce cluster est conçu pour permettre des recherches sur la gestion énergétique et thermique dans les centres de données.

Pendant une expérimentation, un utilisateur peut accéder en temps réel aux informations concernant la température des serveurs impliqués dans son expérimentation, et obtenir la consommation électrique de n'importe quelle partie du cluster, (serveurs, switch, systèmes de refroidissement, etc.), ou contrôler certains paramètres, tels que la définition des températures du système de refroidissement.

SeDuCe est en partie alimenté par des énergies renouvelables.

Pendant la journée, les panneaux photovoltaïques produisent suffisamment d'énergie pour notre cluster.

Différentes orientations et technologies pour les panneaux photovoltaïques sont installées, permettant ainsi de comparer les approches.

Cette présentation décrit cette infrastructure.

- **Des architectures de calcul efficaces en énergie grâce à l'accélération matérielle.**

Olivier Sentieys ([TARAN](#))

L'efficacité énergétique est désormais l'une des principales exigences pour la quasi-totalité des plateformes informatiques. Les architectes de machines sont toutefois confrontés à de nouveaux défis pour les prochaines décennies, le plus important étant la fin de la loi de Moore. Nous pensons que la clé de

l'amélioration durable des performances (vitesse et énergie) réside dans les calculateurs spécifiques à un domaine. Dans cette nouvelle ère, le processeur sera augmenté d'un ensemble d'accélérateurs matériels destinés à exécuter des tâches spécifiques de manière bien plus efficace. Notre équipe de recherche se concentre sur la conception d'accélérateurs qui peuvent s'avérer économes en énergie et tolérants aux pannes. Un autre ingrédient clé de ces architectures frugales est le calcul juste précis. Cela consiste à minimiser le nombre de bits utilisés pour un calcul donné. La conversion, par exemple, d'un code en virgule flottante au calcul avec quelques bits se traduit par des gains énormes. Dans cet exposé, nous abordons les défis des accélérateurs matériels et nous les illustrons par des exemples concrets de plateformes de calcul à faible consommation d'énergie ou à très faible puissance qui peuvent être utilisées dans des applications liées à l'environnement et à l'écologie.

- **Vers des systèmes cyber-physiques intelligents.**

Benoît Combemale ([DIVERSE](#))

Les systèmes cyber-physiques impliquent de nombreux composants interconnectés qui détectent ou agissent sur le monde physique pour offrir des services logiciels personnalisés aux utilisateurs. Du point de vue de l'ingénierie des logiciels et des systèmes, ces systèmes sont essentiellement des systèmes sociotechniques complexes qui conduisent au développement de systèmes dynamiquement adaptables. L'adaptabilité par rapport à l'environnement physique provient d'une boucle de rétroaction (e.g., la boucle MAPE-K) assimilant les données des capteurs, construisant un modèle de l'environnement, planifiant ou éventuellement prédisant de nouveaux scénarios, et sollicitant les actionneurs en conséquence, sous la forme d'une séquence d'actions.

Comme tout système sociotechnique, le processus de planification est généralement semi-automatique, avec une forte interaction avec les utilisateurs finaux afin de leur offrir la meilleure expérience possible. Divers services logiciels ont été développés au cours de la dernière décennie, par exemple en s'appuyant sur d'importants framework élaborés par la communauté de l'internet des objets (protocoles et passerelles, par exemple), et ont donné naissance à un large éventail de systèmes intelligents dans les domaines de l'énergie, des systèmes de production, de la robotique, des transports, de la santé et de l'agriculture, entre autres. L'intelligence du système provient de la capacité à introduire de

l'intelligence dans la boucle de rétroaction. Cette intelligence s'appuie principalement sur l'assimilation et le traitement des données acquises. Ce besoin est actuellement soutenu par le concept de jumeau numérique. Cependant, en tant que système sociotechnique, il est de la plus haute importance de prendre également en compte les préoccupations physiques, économiques, sociales et environnementales plus larges dans lesquelles les systèmes et les utilisateurs finaux sont impliqués. Comme il est difficile d'obtenir de telles informations à partir de capteurs ou de les coder en dur dans le logiciel lui-même, des informations supplémentaires doivent être combinées aux données disponibles pour fournir une vue holistique et systémique du système et de son environnement, permettant de prendre des décisions éclairées.

Lorsque vient le temps de concevoir une telle boucle de rétroaction, la modélisation apparaît comme clé. La modélisation est essentielle pour capturer toute sorte de connaissance sous la forme de modèles descriptifs construits à partir d'observations ou de données acquises, et la modélisation est également essentielle pour conduire le développement et l'évolution de systèmes complexes sous la forme de modèles prescriptifs réduisant la complexité de l'ingénierie accidentelle. L'écart entre les modèles descriptifs et les modèles prescriptifs peut être fait manuellement, ou automatiquement par le biais de modèles prédictifs.

Dans cet exposé, je passe en revue les différents types de modèles nécessaires à la conception intelligente de systèmes cyber-physiques, et je discute des différents rôles que ces modèles jouent dans le cycle de vie global. Je présente les opportunités pour la communauté de la modélisation, ainsi que les défis ouverts à relever pour réaliser une telle vision. En particulier, j'explore les bases de modélisation communes nécessaires pour combiner de manière transparente les différents types de modèles, et les environnements de modélisation associés pour soutenir un engagement plus large et prendre des décisions éclairées.

Mercredi 20 octobre 2021

- **L'optimisation de forme : une solution technique pour l'avion du futur ?**

Enora Denimal ([I4S](#))

Malgré la crise climatique, le nombre de voyageurs par voies aériennes ne cesse d'augmenter chaque année. Selon certains scénarios, ce nombre pourrait être multiplié par trois d'ici 2035. Pour autant, l'industrie aéronautique prévoit pour 2050 des émissions de CO₂ deux fois plus faibles que celles de 2005, laissant une grande place à l'innovation technique et scientifique. L'une de ces solutions consiste à alléger différents composants tout en augmentant leur efficacité, rendue maintenant possible grâce à l'arrivée de la fabrication additive (impression 3D). La présentation se focalise sur des travaux réalisés dans ce contexte afin d'optimiser la forme et/ou la topologie de composants de turbines afin d'améliorer leurs performances vibratoires non-linéaires en s'intéressant tout particulièrement aux amortisseurs par frottement.

- **Vers un pilotage intelligent des infrastructures et des bâtiments à des fins d'optimisation énergétique.**

Jean Dumoulin ([I4S](#))

Dans le cadre de la transition énergétique (i.e. économie d'énergie et diminution de la consommation en énergies fossiles), de nombreuses études et développements ont été conduits sur l'utilisation des énergies renouvelables (eau, vent, solaire, vagues). De plus, en dehors des actions de réhabilitation du patrimoine bâti (gros consommateur d'énergie), des pistes d'optimisation énergétique couplant modèle et mesure pour la gestion dynamique de tels systèmes sont aussi explorées. Ces pistes peuvent également inclure l'ajout de nouvelles fonctionnalités à des infrastructures existantes pour tenter de réduire l'impact du changement climatique (par exemple les îlots de chaleur urbains) voire tempérer les températures de surface en hiver. Dans le cadre de cette présentation, après avoir brièvement introduit le cas des bâtiments, nous nous focaliserons sur l'ajout de fonctionnalités aux infrastructures de transport soumises à un environnement naturel variable. Quelques exemples de

travaux sur de telles structures allant de la récupération d'énergie au contrôle de la température de surface seront abordés.

- **Sur l'utilisation du langage Modelica pour la modélisation des réseaux énergétiques : le défi du passage à l'échelle.**

Benoît Caillaud ([HYCOMES](#))

Le langage Modelica est parfaitement adapté à la modélisation des réseaux d'énergies. Le paradigme sous-jacent au langage Modelica, les équations différentielles algébriques multimodes (mDAE), permet d'adopter une méthodologie de modélisation à base de composants physiques. Ces modèles multimodes sont souvent traités de manière incorrecte par les outils Modelica du marché, en raison de l'analyse structurelle approchée effectuée lors de la compilation des modèles Modelica: le modèle est considéré comme n'ayant qu'un mode de fonctionnement, et le code généré par les compilateurs Modelica est le plus souvent incorrect. L'équipe Hycomes a conçu un algorithme d'analyse structurelle des mDAE qui réalise à la compilation une analyse structurelle exacte, pour tous les modes, sans avoir à recourir à une énumération des modes. Cet algorithme repose sur des techniques symboliques, à base de Diagrammes de Décision Binaires (BDD). Nous aborderons lors de cette présentation, une technique d'analyse structurelle compositionnelle, qui permet de traiter de grands systèmes de mDAE, notamment ceux utilisés pour modéliser des réseaux énergétiques.

- **Modélisation et optimisation des Smart Grids.**

Anne-Cécile Orgerie ([MYRIADS](#))

Les Smart Grids s'appuient sur les infrastructures ICT pour optimiser conjointement la production, la distribution et la consommation d'électricité. Comprendre ce couplage étroit entre le réseau électrique et son infrastructure de contrôle des ICT est crucial pour assurer la performance et avancer vers les objectifs ultimes de

réduction de la consommation globale et d'intégration d'une plus grande part d'électricité issue de sources d'énergie renouvelables. Cet exposé présentera nos travaux sur la co-simulation des Smart Grids et notre étude sur l'impact de la latence du réseau sur les stratégies de délestage en ligne.

Informatique frugale

- **Systèmes distribués frugaux : de la mesure de leur consommation énergétique à la quantification et à la réduction de leur impact environnemental**

Anne-Cécile Orgerie ([MYRIADS](#))

Le cloud computing et les systèmes distribués s'étendent de plus en plus dans le monde entier, avec des services numériques hébergés partout dans le monde et appartenant souvent à des systèmes complexes, utilisant eux-mêmes de nombreux autres services et ressources matérielles. Cette augmentation s'accompagne d'une croissance alarmante des appareils Cloud et de leur consommation en énergie associée. Malgré la complexité des systèmes distribués, comprendre comment ils consomment l'énergie est important pour chasser les Joules gaspillés. Cet exposé présentera nos travaux sur la mesure et la compréhension de la consommation énergétique des systèmes distribués, ainsi que sur la quantification de leur impact environnemental. Cette étape est essentielle pour proposer des stratégies efficaces pour les systèmes distribués frugaux.

Modélisation stochastique de l'océan

- **Modélisation stochastique d'écoulements océaniques.**

Etienne Mémin, ([FLUMINANCE](#))

Dans cet exposé, je décrirai un formalisme, appelé modélisation sous incertitude de localisation (LU), pour dériver de manière systématique des représentations

stochastiques à grande échelle de la dynamique des écoulements fluides. Cette modélisation permet de prendre en compte dans les lois d'évolution les effets à petite échelle négligés par l'introduction d'un champ aléatoire.

La dynamique résultante est construite à partir d'une représentation stochastique du théorème de transport de Reynolds. Ce formalisme permet, de la même manière que dans le cas déterministe, une dérivation physiquement pertinente (c'est-à-dire à partir des lois usuelles de conservation) des lois d'évolution recherchées. Nous montrerons en particulier comment dériver systématiquement une représentation stochastique de la dynamique des écoulements. Nous donnerons plusieurs exemples de simulations obtenues par un tel système et comment un ensemble de telles réalisations peut être utilisé dans l'assimilation de données ou pour la quantification des incertitudes.

- **Un satellite de nouvelle génération pour un problème d'envergure : SWOT et les ondes internes océaniques.**

Noé Lahaye ([FLUMINANCE](#))

SWOT, un satellite embarquant un altimètre de nouvelle génération, est attendu comme l'une des innovations majeures de la décennie à venir en océanographie spatiale. Les attentes et les défis associés seront abordés, notamment en lien avec les ondes internes — des perturbations en courants et densité omniprésentes dans l'océan. Ces ondes, qui jouent un rôle essentiel dans la circulation océanique globale et sur le climat, ont une signature à la surface. Dans le contexte d'estimations basées sur des observations, ceci constitue à la fois un verrou pour l'estimation des courants en proche surface, mais aussi une opportunité pour une meilleure estimation de leur dynamique. Ces défis appellent des méthodes de couplage modèle/observations innovantes.

Agriculture numérique

- **Contrôle haptique partagé d'équipes hétérogènes d'humains et robots pour le suivi de l'environnement.**

Claudio Pacchierotti ([RAINBOW](#))

Les équipes de robots sont utiles dans plusieurs applications, notamment la recherche et le sauvetage en milieu urbain, la réponse aux catastrophes et la surveillance de l'environnement. Alors que la plupart des approches considèrent la seule présence d'équipes robotiques, nous nous sommes récemment concentrés sur le contrôle d'équipes hétérogènes composées d'un nombre arbitraire de robots mobiles (drones et robots terrestres) ainsi que d'humains, pour réaliser en collaboration des tâches plus complexes et de manière distribuée. Les utilisateurs humains font physiquement partie de l'équipe, se déplaçant dans le même environnement que les robots et recevant des informations sur la connectivité de l'équipe par de signaux de feedback haptique ou audio. Pendant que l'humain explore l'environnement, les robots se déplacent pour maintenir l'équipe connectée et accomplir des tâches spécifiques, par exemple, visiter une partie de la carte pour enregistrer un paramètre environnemental d'intérêt.

- **AgriSense : une plate-forme pour l'agriculture de précision**

Frédéric Weis ([EASE](#))

De très petits objets connectés, de la taille d'un micro-contrôleur, sont désormais en mesure d'exécuter des codes logiciels, de piloter différents types de capteurs, d'échanger des données avec d'autres dispositifs proches et de prétraiter les données pour les nettoyer/agréger. L'objectif de la plate-forme AgriSense est d'étudier l'utilisation de ce type d'objets pour le suivi des cultures pour l'agriculture. Dans cet exposé, nous mettrons en avant les principaux verrous technologiques qui ont du être levés : disposer d'un nœud ouvert au niveau logiciel, gérer finement l'énergie du nœud tout en exploitant différents types de capteurs. Enfin, nous montrerons les premiers résultats expérimentaux obtenus avec cette plate-forme en collaboration avec l'équipe DEMECOLOGY de l'INRAE.

- **Analyse de données de capteurs issues de l'élevage de précision.**

Véronique Masson ([LACODAM](#))

De plus en plus d'élevages équipent leurs animaux et bâtiments de capteurs afin d'avoir un suivi plus fin des paramètres clés de chaque animal du troupeau.

Dans le cadre de l'Institut de Convergence #DigitAg, l'équipe Lacodam collabore avec l'UMR Pegase d'Inrae Rennes pour analyser les données issues de ce type de capteurs. Ces travaux portent sur l'analyse de séries temporelles uni ou multi-variées, dans le contexte d'élevages laitiers ou porcins. Les problématiques traitées vont de la reproduction des animaux (détection d'ovulation) à l'évaluation de leur bien-être. L'exposé montrera ce que la data science peut apporter dans chacun de ces cas.

Sciences et technologies du numérique et écologie

- **La plate-forme LivingFog, de la mesure de consommation d'eau à l'observation de l'environnement.**

Guillaume Pierre ([MYRIADS](#))

La plate-forme LivingFog a été développée par sept doctorants talentueux dans le cadre du projet européen FogGuru. Elle permet à ses utilisateurs de traiter des données de l'Internet des Objets transmises grâce à la technologie de réseau sans fil LoRa, à proximité immédiate des "gateways" LoRa qui reçoivent les données. Ceci permet d'obtenir de faibles temps de réponse, de réduire la consommation de ressources réseau longue-distance, et de mieux répondre à des dysfonctionnements réseau. La plate-forme a été initialement développée pour traiter des données provenant de compteurs d'eau intelligents, puis étendue pour permettre la tenue d'un hackathon IoT dans la Marina de València (Espagne). Elle constitue désormais un composant important d'un nouveau projet PIA3 sur le monitoring environnemental.

- **SmartSense : une plateforme expérimentale de réseau de capteurs pour la gestion de l'énergie et des usages dans les bâtiments et au-delà**

Olivier Sentieys ([TARAN](#)) présenté par Guillaume Pierre ([MYRIADS](#))

Avec 150 nœuds déployés dans les bâtiments des laboratoires de l'INRIA/IRISA (Lannion et Rennes), SmartSense permet de récolter de nombreuses de données sur l'énergie consommée et les usages dans les bâtiments. Ces données permettent un grand nombre d'application, notamment en fouille de données, désagrégation de charges électriques ou en traitement de données capteurs à base d'IA. Chaque nœud comprend environ vingt capteurs : image, infra-rouge, audio, spectre radio, centrale inertielle, humidité, pression, température, lumière (rouge, vert, bleu, blanc, UVA, UVB), radar de distance à précision centimétrique, CO2+COV. Une version développée pour un déploiement en environnement extérieur est en cours de test dans le cadre d'une collaboration avec l'OSUR/Géosciences Rennes (projets SmartObs et Terra Forma).

- **Interprétabilité / Explicabilité des IA pour la Méta-Modélisation. Application au domaine de l'agro-écologie.**

Tassadit Bouadi ([LACODAM](#))

Dans le contexte de la modélisation et de la simulation de systèmes dits complexes, le modèle est vu comme une abstraction du système en question, et la simulation comme un plan d'expérience permettant l'exploration d'une multitude de scénarios. Ces modèles souvent très complexes sont vus par l'utilisateur comme des boîtes noires.

Le modèle agro-hydrologique TNT2, développé par l'UMR SAS d'Inrae Rennes, est un exemple de ce type de modèles.

Dans le cadre de ce travail, l'équipe Lacodam collabore avec l'UMR SAS pour développer un méta-modèle sur la base des simulations de TNT2.

L'objectif est d'exploiter un corpus de simulations afin d'en déduire des ensembles de règles de fonctionnement qui constituent ainsi un modèle simplifié du modèle initial (i.e. un méta-modèle). En d'autres termes, nous souhaitons ouvrir la boîte noire afin d'en comprendre le fonctionnement, et de proposer une représentation simplifiée.

Plus précisément, nous proposons de transposer notre problématique de méta-modélisation au problème d'interprétabilité/explicabilité des IA.

- **Écologie microbienne: comprendre et exploiter la diversité des fonctions bactériennes.**

Anne Siegel ([DYLISS](#))

L'écologie microbienne et systémique vise à comprendre les interactions entre un consortium de microbes et un organisme hôte, où un "hôte" peut être une plante, une algue ou un humain. Cette question est de plus en plus cruciale à l'échelle environnementale puisqu'il est prouvé que les microbes interagissent avec leur environnement afin de fournir des nutriments spécifiques et de faciliter l'adaptation de l'hôte aux perturbations extérieures. Dans cet exposé, nous détaillerons les enjeux soulevés par l'interprétation de la diversité microbienne (métagénomique) dans des modèles fonctionnels à grande échelle de consortiums hôtes-microbes et les solutions nécessaires pour analyser ces systèmes.