

Réseau Mixte Technologique

Science des Données et Modélisation pour l'Agriculture et l'Agroalimentaire

Contexte

Le monde de la production agricole et de la transformation est en pleine mutation due à l'arrivée massive des technologies du numérique (applications, capteurs, ...) depuis quelques années. Ces transformations sont à l'origine de nombreuses données collectées, à l'instar des autres secteurs d'activité (UsineNouvelle 2019). Les activités de recherche, développement et d'enseignement sont également impactées par ces changements (Brun et al. 2016).

Grâce à l'apparition de nouveaux capteurs, à la réduction du coût des images satellitaires et au développement des systèmes d'information, des masses importantes de données sont désormais collectées au sein des exploitations agricoles (ex. projet EUXDAT, European Commission. 2017), dans les usines de transformation et auprès des consommateurs. Les données disponibles ne sont cependant que très rarement directement utilisables pour l'aide à la décision car trop volumineuses et hétérogènes. Elles doivent être au préalable filtrées et transformées avant toute utilisation pratique : ces prétraitements soulèvent souvent de nombreuses questions méthodologiques (ex. dans pour la sécurité sanitaire, Marvin et al 2017). Enfin, il reste ensuite à montrer les apports de ces sources de données pour répondre aux grandes problématiques de nos secteurs (réduction des intrants, changement climatique, qualité, sécurité sanitaire, analyse de risque, ...).

En parallèle de cette tendance, l'expérimentation agronomique, bien que très ancienne, tend également à évoluer depuis quelques années, notamment avec l'essor des "On-Farm Research Networks" (Laurent et al. 2019), essentiellement pour l'instant aux USA (en particulier dans l'Iowa). Ce type de réseau expérimental est constitué d'expérimentations réalisées chez et par les agriculteurs. Chaque expérimentation comporte plusieurs répétitions d'un traitement de référence (une pratique agricole standard) et d'un traitement innovant (une pratique agricole nouvelle). Les mesures sont réalisées directement par les agriculteurs à l'aide de capteurs connectés à une base de données centrale. L'ensemble des données du réseau sont ensuite analysées et les résultats sont communiqués auprès des agriculteurs et conseillers régionaux. Cette approche est très prometteuse mais soulève des questions à la fois méthodologiques et logistiques.

Enfin, lorsque des références basées sur des observations directes font défaut, les experts sont sollicités pour quantifier des variables d'intérêt dans des études d'analyse de risque. L'élicitation probabiliste d'expert permet de quantifier le jugement d'un expert en les formalisant ces connaissances par une distribution de probabilités (O'Hagan et al, 2006).

En parallèle de ces évolutions, de nombreuses méthodes issues des mathématiques appliquées deviennent plus accessibles grâce à l'augmentation de la puissance de calcul des ordinateurs, mais aussi à la diffusion de certains algorithmes facilitant le développement de modèles prédictifs et la visualisation des données massives. Les secteurs de l'agriculture et de l'agroalimentaire ont pour défi de s'approprier ces évolutions, d'évaluer la pertinence des méthodes aujourd'hui disponibles en fonction des questions posées, et comment mobiliser ces méthodes sur nos données et modèles.

Le transfert de ces approches innovantes vers le terrain et la montée en compétence de l'ensemble des acteurs doivent également être appréhendés. Dans ce contexte, on observe également une évolution des modes d'animations, avec l'émergence de défis qui peuvent constituer à la fois des moyens d'animation et de formation de la communauté et des actions de communication vers l'extérieur. Cette opportunité contribuera

à faire évoluer les métiers du conseil qui s'orientent vers un accompagnement des agriculteurs avec des outils d'aides à la décision nouvelles générations qui doit garantir un conseil adapté au territoire et aux objectifs des agriculteurs. Ainsi, les futurs agriculteurs et conseillers doivent aussi bénéficier dès leur formation d'apport pédagogique leur permettant de mieux appréhender la question des données agricoles, de leur valorisation et de l'utilisation des outils d'aide à la décision.

Historique des dynamiques partenariales précédentes.

Une convention concernant la modélisation a été signée en février 2003, pour une période de 3 ans, entre l'INRA, l'ACTA et 9 instituts techniques. Le dispositif mis en place était double. Il comprenait un club modélisation et une plate-forme opérationnelle avec 2 ingénieurs en informatique.

Par la suite, ce partenariat s'est poursuivi dans le cadre du RMT Modélisation et logiciels d'intérêt commun appliqués à l'agronomie (septembre 2007-décembre 2013) animé par l'ACTA, l'INRA et ARVALIS - Institut du végétal avec les instituts techniques agricoles (Terres inovia, Institut de l'élevage, ITB, IFIP, CTIFL, IFVV, Terres d'Innovation (association fédérant l'ANITTA, ASTREDHOR, IFPC, ITEIPMAI et ITL) et l'enseignement agricole (EPLEFPA de Toulouse Auzeville et ENFA) comme partenaires.

En janvier 2014, le réseau s'est élargi en RMT Modélisation et Analyse de données pour l'agriculture. Les retours des partenaires et les évaluations font apparaître un bilan très positif des productions de ce précédent réseau. Aujourd'hui, de par l'arrivée des données massives dans les secteurs de l'agriculture et de l'agroalimentaire, et, pour répondre à cette mutation révolutionnaire, nous proposons de revoir en profondeur nos orientations du point de vue des méthodes, mais aussi des champs d'application, du partenariat et enfin de l'organisation de l'animation. Au niveau du partenariat, on peut noter notamment l'intégration de l'ANSES et du CNIEL pour mieux couvrir la partie analyse de risque et transformation agroalimentaire, l'INRIA pour son expertise sur les méthodes en science des données, mais aussi les instituts techniques comme la FNAMS, la FN3PT, l'ASTREDHOR, l'ITEIPMAI et la CRA de Normandie pour compléter le dispositif de R&D agricole.

Ces propositions, en adéquation avec les enjeux méthodologiques d'actualités, vont nous permettre de répondre au mieux aux besoins des ingénieurs, chercheurs et enseignants des partenaires.

Objectifs du réseau

Le Réseau Mixte Technologique proposé porte sur l'application des méthodes de la science de données et de la modélisation dans le domaine agricole et de la transformation agroalimentaire.

Les objectifs du réseau sont d'animer la communauté autour de ces méthodes et leurs applications sur la diversité des thématiques en organisant les échanges, en conduisant des travaux d'approfondissement ou de mutualisation, en apportant de nouvelles compétences via des formations et en apportant l'expertise nécessaire aux projets des partenaires présentant des actions importantes sur de science de données ou de modélisation.

Pour cela, le programme prévisionnel à 5 ans pour 2020-2024 se décompose en 3 volets.

- Volet 1. Nouveaux modes d'acquisition d'informations
- Volet 2. Méthodes pour la science de données et la modélisation
- Volet 3. Vers une data science participative favorisant la diffusion des méthodes et des innovations

Nous proposons également un calendrier des deux premières années permettant de détailler les actions envisagées sur la phase de début du réseau.

Programme proposé

● **Volet 1. Nouveaux modes d'acquisition d'informations**

Responsable du volet : François Brun

Notre RMT prendra en considération les nouveaux modes d'acquisition d'information dans les domaines agricoles et agroalimentaires : 1) données acquises directement dans le contexte des exploitations agricoles ou des usines de transformation, 2) expérimentations menées directement en exploitations agricole, 3) collecte d'informations quantitatives auprès des experts. Nous étudierons en particulier comment les nouveaux modes d'acquisition de données changent les pratiques de recherche et comment filtrer, transformer et analyser les masses de données disponibles pour répondre à des questions opérationnelles et construire de nouveaux services utiles aux agriculteurs et industriels.

Pour les années à venir, nous proposons de travailler sur les 3 points suivants:

- **V1.1 Caractérisation des informations acquises directement dans le contexte des exploitations agricoles ou des usines de transformation.**

Comme indiqué ci-dessus, des données massives issues de capteurs et de systèmes d'information sont dorénavant collectées au sein des exploitations agricoles et des ateliers de transformation. Nous proposons d'évaluer l'intérêt de ces sources d'informations en complément ou en remplacement des expérimentations classiques en station expérimentales. Nous considérerons les sources d'informations suivantes :

- Données collectées dans le cadre d'**observatoires pilotés** (ex: réseau d'épidémiologie-surveillance, Siné et al. 2010).
- Données issues des **différentes unités de production** ("Big data", Acta, 2016), acquise avec des capteurs ou saisis dans les systèmes d'information comme en production laitière (robot laitier, données génomique, ...) ou les usines de transformation connectées avec le concept d'usine 4.0 (RFL, 2018; Vitagora 2017).
- Données acquises à **plus larges échelles** comme les données de télédétection, les corpus de texte issus de réseaux sociaux et *text mining*, ou les données des consommateurs (ex: consommation de produits alimentaires).

Il s'agira notamment de caractériser ces données pour évaluer les conséquences du point de vue du choix des méthodes à mobiliser (volet 2).

Conscient des difficultés de définition des données dans ce cadre, la question de la modélisation des connaissances sera forcément évoquée dans le réseau, mais elle ne fera pas a priori l'objet d'investissement du RMT. De même, les questions d'infrastructures informatiques ne seront pas traitées spécifiquement dans le cadre de ce réseau. Par contre, l'ensemble des problèmes méthodologiques posés par l'analyse et l'utilisation de ces données sera étudié.

- **V1.2 Définir les conditions pour mener des expérimentations en exploitations agricole.**

La démarche d'**expérimentation en exploitations agricole** (concept de *On-Farm Research Network*) repose sur une collaboration entre agriculteurs et organismes scientifiques et techniques pour mettre en place des expérimentations au champ pilotées par les agriculteurs. L'agriculteur est l'acteur principal du management de son essai, de sa mise en place à la collecte des observations. Ce type de dispositif nécessite l'utilisation de nouvelles technologies, notamment des capteurs de rendement embarqués. Les organismes scientifiques et techniques jouent également un rôle important dans les *On-Farm Research Network* car ils s'assurent de l'organisation et l'animation du réseau ainsi que de l'analyse des données et la communication des résultats.

Un avantage important de cette approche est qu'elle permet d'acquérir des données dans des conditions réelles de production, en prenant en compte une diversité des situations.

- **V1.3 Acquisition d'information auprès des experts**

L'**élicitation probabiliste d'expert** permet de quantifier le jugement d'un expert, et de formaliser ses connaissances sur les valeurs probables d'une quantité inconnue d'intérêt en représentant ces connaissances par une distribution de probabilités. Cette méthode, généralement utilisée en cas d'absence ou de manque de données ou lorsque l'expérimentation est impossible, permet de présenter le jugement d'un expert de manière quantitative à travers la distribution de probabilités. La distribution de probabilités décrit l'incertitude de l'expert sur la quantité estimée. Des travaux dans le cadre du projet L-i-cite (Carnot Plant2Pro) et SMART-PIC (CASDAR AAP RT) et thèse de Mathilde Chen, Chen et al. 2018) ont permis de tester cette approche sur des cas concrets (prévision de rendement et protection des cultures) et de proposer un outil facilitant ce processus. Les résultats ont suscité l'intérêt de plusieurs partenaires, notamment d'Arvalis, de l'Anses et du CIRAD.

Dans le cadre du RMT, nous souhaitons organiser des formations afin de faciliter le transfert de cette méthode ainsi que l'utilisation du logiciel développé dans le cadre du projet L-i-cite. Nous proposerons également de nouvelles applications de cette méthode en collaboration avec nos partenaires.

- **Volet 2. Méthodes pour la science de données et la modélisation**

Responsable du volet : David Makowski

Notre RMT prendra en considération les nombreuses méthodes issues des mathématiques appliquées permettant de répondre aux questions suivantes : (1) Comment construire un modèle prédictif à partir d'un jeu de données en utilisant des méthodes de type "apprentissage automatique" ? (2) Comment extraire des connaissances ou des références consolidées des données ? (3) Peut-on remplacer un modèle de processus complexe par un émulateur beaucoup plus simple et moins coûteux à utiliser ? (4) Comment corriger les simulations des modèles par des données acquises en temps réelles ?

Nous développerons des actions sur chacun de ces enjeux en collaboration avec les partenaires.

Pour les années à venir, nous proposons de travailler sur les 4 points suivants

- **V2.1. Apprentissage sur les données pour proposer des outils prédictifs**

L'objectif est de rendre accessible les principales méthodes de **machine learning (apprentissage automatique)** disponibles et de les appliquer à des cas concrets co-construits avec les partenaires. Dans ce but, nous avons entamé la rédaction d'un ouvrage collectif mobilisant plusieurs partenaires du RMT intitulé "Science des données pour l'agriculture" en lien avec une première formation proposée (Acta, 2018). Cet ouvrage nous servira d'outil pédagogique pour organiser des sessions de formation sur la data science en collaboration avec nos partenaires. Nous contribuerons ainsi à la diffusion de méthodes de régression avancées (pénalisée, PLS, GAM, sélection de variable), des réseaux neurones et leur extension multicouches, des arbres de régression et des méthodes d'ensemble (random forest, boosting, bagging) ainsi que de diverses méthodes de classifications supervisées et non-supervisées.

- **V2.2. Etablir et synthétiser les connaissances**

Les méthodes de **fouille de données** visent à comprendre ou découvrir un phénomène grâce aux données (Tan et al 2018). De nombreuses méthodes existent et sont de plus en plus accessibles, mais il reste difficile d'évaluer le résultat. Par ailleurs, lorsque les méthodes de machine learning sont mobilisées pour prédire, il est souvent difficile de comprendre l'origine des prédictions et elles sont ainsi souvent perçues comme des boîtes noires. De nombreuses recherches sont conduites actuellement pour faciliter l'exploration de ces outils afin de permettre aux utilisateurs de mieux comprendre leur fonctionnement. Il s'agit ainsi de rendre la boîte noire un

peu plus grise de telle sorte qu'elle nous permette de mieux comprendre les processus à l'œuvre dans le phénomène observé. Pour alimenter ces réflexions, nous profiterons d'une nouvelle collaboration Acta-INRIA-INRA (début prévu fin 2019) autour de thèse "RegEpi" (DigitAg et Ecophyto) avec une application sur les données d'épidémiologie avec méthode de règles hybride (Galárraga et al. 2015) facilitant les interactions avec les experts.

La **synthèse des données de réseaux expérimentaux** doit permettre de mieux valoriser les données d'essais expérimentaux. Nous profiterons de la dynamique initiée dans le RMT précédent depuis 2016 (Makowski et al 2018) en continuant à proposer des formations sur la base de l'ouvrage existant. Nous pensons qu'il y a notamment des applications sur la thématique de la protection des cultures.

La **Méta-analyse et la méta-régression** peuvent être utilisées pour développer des "*policy briefs*" auprès des décideurs. Comme dans le RMT précédent, nous continuerons à organiser des ateliers sur la méta-analyse et la méta-régression pour synthétiser de manière quantitative les données disponibles dans la littérature scientifique. La place de cette approche dans le nouveau RMT sera cependant un peu moins grande que dans le précédent car cette approche a maintenant atteint une certaine maturité, en témoigne l'augmentation exponentielle du nombre de publications scientifiques utilisant ces approches et les nombreux outils disponibles pour les mettre en œuvre, mais nous pensons poursuivre les efforts de formation en lien avec l'ouvrage publié (Makowski et al 2018).

○ **V2.3. Méta-modélisation pour simplifier l'utilisation des modèles**

Les modèles de processus complexes décrivent finement le fonctionnement des systèmes. Du fait de leur complexité, ces modèles sont difficiles à évaluer et à calibrer. Il est également difficile de comprendre leur fonctionnement et de hiérarchiser l'importance de leurs variables d'entrée. Par ailleurs, du fait de leur temps de calcul et du nombre souvent élevé de leurs variables d'entrée, il est difficile de les utiliser de manière opérationnelle, notamment à grande échelle (nationale ou plus) ou pour optimiser des décisions. **La méta-modélisation** (Favre et al. 2013) peut potentiellement répondre à ces défis car elle permet de développer des modèles simplifiés imitant le modèle complexe original. Cette approche mobilise les méthodes d'apprentissage du volet V2.1 mais dans un but très différent. Ces travaux seront conduits en collaboration avec le réseau MEXICO (Méthodes pour l'EXploration Informatique des modèles COmplexes, reseau-mexico.fr).

○ **V2.4. Prise en compte des incertitudes dans les prédictions et correction pour les réduire**

Les modèles décrivant le fonctionnement d'un système sont des représentations simplifiées des systèmes réels considérés sur nos différentes thématiques. Ainsi, de nombreuses incertitudes persistent sur les formalismes et les entrées. Aussi, il est indispensable de réaliser des **analyses d'incertitude** et des **évaluations** de ces modèles (Wallach et al. 2018). Nous proposons notamment de profiter de la thèse initiée entre l'Acta, Météo France et l'INRA pour prendre en compte les incertitudes des prévisions météorologiques (Aleksavska et al. 2018).

Ces incertitudes bien définies et prises en compte, on peut aussi se poser la question de la **prise de décision en contexte incertain**. Il s'agit d'apporter des méthodes pertinentes pour aider à faire un choix optimal entre plusieurs décisions quand les décisions sont basées sur des informations incertaines. On imagine bien ces questions dans le choix des variétés en grande culture ou de pilotage des traitements phytosanitaires.

Enfin, avec la présence de données acquises en cours de simulation comme celles issues de capteurs au champ, sur troupeau ou encore dans les chaînes de transformation ou celles issues de l'imagerie satellitaire à grande échelle (cf volet V1.1), il est légitime d'étudier les possibilités **d'assimilation de ces données pour corriger les simulations en cours de campagne** (Wallach et al. 2018). Des travaux sur ce sujet sont notamment prévus en lien avec l'UMT MAGNUM, intéressée par ces questions sur la partie modèle de culture.

- **Volet 3. Vers une data science participative favorisant la diffusion des méthodes et des innovations**

Responsable du volet : François Brun

Le transfert de méthodes et d'outils d'aide à la décision issus de la R&D vers les utilisateurs finaux reste une étape délicate, trop rarement couronnée de succès. Dans ce RMT, nous souhaitons mettre en œuvre des animations participatives donnant la possibilité aux participants de tester eux-mêmes des méthodes et outils sur des jeux de données directement liés à des questions pratiques. Un effort de diffusion vers les futurs acteurs agricoles via l'enseignement agricole est également prévu.

Trois types d'animation seront proposés :

- **V3.1 Data challenges : des compétitions pour développer des outils prédictifs**

Un **data challenge** est un concours visant à comparer les performances de méthodes statistiques et d'apprentissage automatique pour prédire une variable d'intérêt (voir V2.1). Lors du lancement du data challenge, une base de données est mise à disposition des participants (jeu de données d'entraînement), avec un objectif de prédiction clairement défini (ex: prédire l'occurrence de pertes de rendement) lié à un problème opérationnel (ex: développer un système d'assurance couvrant les risques de perte de rendement). Les participants mobilisent alors leurs compétences pour développer des algorithmes prédictifs et transmettent à l'organisateur les prédictions générées par leurs algorithmes. Les prédictions sont alors évaluées par l'organisateur à l'aide d'un jeu de données indépendant (jeu de données test) en calculant un critère quantitatif mesurant la performance des algorithmes développés par les participants. Les meilleurs algorithmes sont ensuite identifiés.

Différents data challenges pourront être proposés sur la vie du RMT, pour couvrir différents domaines d'application.

- **V3.2 Hackathons : de nouveaux services construits sur des données**

Un **hackathon** est également d'un concours de science de données, mais dans ce cas, diverses données sont mises à disposition et les participants font preuve d'inventivité pour proposer un nouveau service qui valorise ces données acquises à l'origine pour un autre objectif (voir V2.2). L'objectif n'est pas défini au départ et ne correspond pas forcément à un objectif de prédiction. Il s'agit plutôt de définir un nouveau service informatique basé sur la valorisation de données déjà acquises.

Différents Hackathons pourront être proposés sur la vie du RMT, pour couvrir différents domaines d'application (ex : reconnaissance de symptômes de maladies ou de ravageurs à partir d'images). Ils pourront se faire en partenariat avec d'autres acteurs, afin de rajouter la composante informatique dans les prototypes proposés.

- **V3.3 Renforcer la mise à disposition gratuite des outils opérationnels pour la formation**

Il s'agit de mieux apprendre aux **futurs agriculteurs ou conseiller à mobiliser les outils d'aide à la décision** vecteur de diffusion des bonnes pratiques et la **maîtrise des données** sur leur exploitation.

Il s'agit de mieux référencer les outils disponibles sur le marché notamment pour le domaine agricole. Sur cette base, les acteurs de l'enseignement agricole contribueront à l'identification d'outils pouvant s'insérer dans les programmes des différents cursus (BTS notamment). Un travail avec les fournisseurs des services considérés sera alors entrepris afin de les mettre à disposition gratuitement, dans le cadre des enseignements. Un critère sera aussi la disponibilité de la documentation de ces outils. A titre d'exemple, on peut citer Taméo pour la protection des cultures (Arvalis), la Réglette azote colza® (Terres inovia), le guide Ecoherbi (Acta, action prévue en lien avec le RMT FLORAD).

Il s'agit aussi d'accompagner la montée en compétences des (futurs) agriculteurs et conseillers en facilitant la prise en compte des enjeux autour de « la data agricole » avec des apports sur ces sujets.

Nous travaillerons à la construction de séquence pédagogique et la construction de module de formation pour notamment le BTS APV et la licence pro COSYA autour de l'utilisation des outils d'aides à la décision et la gestion des données agricoles.

Moyens d'animation et d'action

- **Capitaliser les informations et les faire circuler**

Le **site internet** permettra de présenter les orientations du réseau et enrichi par l'ensemble des travaux au fur et à mesure de la vie du réseau. Le site actuel www.modelia.org pourra éventuellement être conservé, avec un changement prévu de plateforme, en gardant seulement une partie des contenus capitalisés afin de continuer à profiter de son audience et des contenus (séminaires et formations) mis à disposition depuis 2007. Le **calendrier** du site relayera les événements organisés par le réseau, mais aussi les événements externes intéressant notre communauté.

Une **liste de diffusion** permettra aux partenaires de s'inscrire pour recevoir les actualités par email sous la forme de **lettres d'information**.

Les **réseaux sociaux professionnels** (twitter, linkedin, ...) seront mobilisés également pour diffuser les événements et des informations de veille.

- **Construire la communauté et défricher de nouveaux sujets**

Premier moyen d'animation, l'organisation de **séminaires** ouverts sera indispensable pour la vie de la communauté, en brassant les idées et en amenant un premier niveau d'information.

L'organisation de la **veille collective** devrait aussi permettre notamment de faire le point sur les nouvelles sources de données (volet 1) et les méthodes, ainsi que les solutions logicielles afférentes, (volet 2).

Des **ateliers** participatifs (ou workshop) permettent quant à eux de pousser plus loin la réflexion afin de préciser ensuite les contours d'une action commune.

Enfin, des événements de type concours comme les **Data Challenge** ou **Hackathon** (décrits dans le volet 3) viendront compléter le dispositif avec des formats d'animation nouveaux permettant à la fois de promouvoir l'innovation participative grâce aux méthodes et de communiquer largement sur les actions du réseau.

- **Apporter de nouvelles compétences au réseau**

L'investissement dans la création de **formations** sera conséquent, avec la création de supports adaptés permettant de faciliter l'appropriation des nouvelles méthodes par les participants du réseau.

Il s'agira aussi de travailler à la création d'**ouvrages** de référence, vecteur complémentaire pour diffuser les méthodes dans notre communauté de manière plus formelle et complète.

- **Approfondir des sujets d'intérêt**

La mise en réseau des compétences des participants permettra de formuler de **nouvelles questions d'approfondissement** qui pourront faire, par la suite, l'objet de **projets** plus ambitieux, notamment en réponse à des appels d'offre (AAP ANR, CASDAR, europe, ...), permettant de mobiliser des moyens supplémentaires.

Programme 2020-2021.

Les actions du RMT seront complétées et affinées au fur et à mesure de la vie du réseau, néanmoins voici ce que l'on planifie sur les années 2020-2021, à ce stade.

Période	Moyen	Titre	Commentaire
2020-T1	Formation	Working with Dynamic Crop Models.	2 jours, en anglais, dans le cadre du colloque iCROP2020
2020-T1	Data challenge	Concours de prédiction, "construire le meilleur méta-modèle de culture" ?	remise des prix lors du colloque iCROP2020
2020-T2	Séminaire	Approche de modélisation et data science pour traiter les questions "de la fourche à la fourchette"	En collaboration avec le RMT QUALIMA
2020-T3	Formation	Data Science pour l'Agriculture	
2020-T4	Formation	Définition avec l'enseignement agricole de séquences pédagogiques	
2021-T1	Ouvrage	Data Science pour l'Agriculture	
2021-T2	Workshop	Synthèse de connaissances pour la production animale et l'alimentation	Enquêtes, réseaux d'expérimentations, méta-analyse. Idele, Anses....
2021-T4	Workshop	Text mining: méthodes et applications pour analyser les comportements des agriculteurs et des consommateurs	

Ponts avec les autres RMT et réseaux

Nous établirons des collaborations et des échanges avec divers réseaux mobilisant les méthodes couvertes par le périmètre de notre réseau. Ces collaborations pourraient conduire à des co-organisations d'événements ou de montage de projets sur des sujets d'intérêt commun.

Réseau	Point d'accroche	Volet
RMT QUALIMA (maîtrise de la qualité microbiologique des aliments)	Méthode pour la modélisation multi-critères et appui statistique.	V2
RMT FLORAD (Gestion de la flore adventice)	Expérimentation à la ferme. Enseignement agricole.	V1.2 V3.3
RMT IDEAL (Innover en Diagnostic Et contribuer aux Analyses des risques en santé des végétaux)	Modélisation de prévision des dégâts et l'analyse des risques. Valorisation des données d'épidémiosurveillance.	V2.1 V2.2
RMT Biodiversité & Agriculture	Accès à des données sur les niveaux de potentiel de régulation biologique via des analyses simplifiées en ligne sur un site web interactif.	V1 V2
RMT Al-chimie (Contaminations chimiques de la chaîne alimentaire)	Travaux sur la modélisation des transferts de contaminants dans les matrices animales et sol-plante.	V2
UMT CAPTE (CAPteurs et TELédétection)	Données drone et satellite pour le suivi des grandes cultures. Machine learning sur image.	V1.1 V2.1
UMT MAGNUM (Modélisation et agriculture numérique)	Assimilation de données des modèles de cultures.	V2.3 V2.4
Institut de convergence #DigitAg	Financement de thèses en Science des données.	V2.1,V3.2
Institut de convergence CLAND	Modélisation	V2.3
le réseau RNSC-EXTRAM (Extraction de connaissance pour la méta-analyse)	Synthèse de connaissance.	V2.2
les réseaux thématique PIC, ATMOSPHERE, MODELISAD de l'INRA	Méthodes de modélisation.	V2.3 V2.4

la plateforme projet RECORD (plate-forme de modélisation des agro-écosystèmes, INRA)	Méthodes de modélisation.	V2.3 V2.4
le réseau MEXICO (Méthodes pour l'EXploration Informatique des modèles COMplexes, INRA-IRSTEA-IFREMER)	Méthodes de modélisation.	V2.3 V2.4
GDR MADICS AGEE (De la gestion à l'analyse de données pour l'agriculture, l'écologie et l'environnement)	Données à l'échelle de l'exploitation agricole. (à préciser - action en cours de montage)	V1.1 V2.1

Références

- Acta, 2016. Le Big Data Agricole : quelle réalité aujourd'hui? 10 mars 2016. <http://www.modelia.org/moodle/course/view.php?id=72>
- Acta, 2019. Formation Data Science pour l'agriculture" (3 jours, septembre 2019). <http://www.modelia.org/moodle/course/view.php?id=77>
- Aleksovska I, Brun F, Raynaud L, Robert F, Raynal M, Deudon O. 2018. Prendre en compte l'incertitude des prévisions météorologiques dans les OAD utilisées pour gérer les maladies et ravageurs des cultures. Végéphyll – 12e conférence internationale sur les maladies des plantes - Tours (France) - 10, 11 et 12 décembre 2018, Tours, France; 12/2018
- Brun F, Siné M, Gallot S, Colinet J, Cimino M, Haezebrouck TP, Lauga B, Besnard J, ACTA - Les Instituts Techniques Agricoles. 2016. *L'accès aux données pour la Recherche et l'Innovation en Agriculture. Position des Instituts Techniques Agricoles*. Edited by ACTA - Les Instituts Techniques Agricoles, 10/2016; ACTA - Les Instituts Techniques Agricoles., ISBN: 978-2-85794-298-6
- Chen M, Brun F, Raynal M, Debord C, Makowski D. 2018. Estimer la date d'apparition d'apparition du mildiou de la vigne grâce à l'élicitation probabiliste d'expert. Végéphyll 12e CONFÉRENCE INTERNATIONALE SUR LES MALADIES DES PLANTES, TOURS; 12/2018
- European Commission. 2017. EUXDAT project. European e-Infrastructure for Extreme Data Analytics in Sustainable Development. www.euxdat.eu
- Faivre R, Iooss B, Mahévas S, Makowski D, Monod H (eds). 2013. Analyse de sensibilité et exploration de modèles. Quae éditions.
- Galárraga L, Teflioudi C, Hose K, Suchanek F. 2015. Fast Rule Mining in Ontological Knowledge Bases with AMIE+. VLDB Journal. December 2015, Volume 24, Issue 6, pp 707-730.
- Laurent A, Makowski M, Brun F. 2019. Qu'est ce qu'un On-Farm Research Network ? Workshop On-Farm Research Network, 14 Mars 2019, Paris <http://www.modelia.org/moodle/course/view.php?id=79>
- Makowski D, Piraux F, Brun F. 2018. De l'analyse des réseaux à la méta-analyse. Quae.
- Marvin HJP, Janssen EM, Bouzembrak Y, Hendriksen PJM, Staats M. 2017. Big data in food safety: An overview, Critical Reviews in Food Science and Nutrition, 57:11, 2286-2295, DOI: 10.1080/10408398.2016.1257481
- O'Hagan A, Buck CE, Daneshkhah A, Eiser JR, Garthwaite PH, Jenkinson DJ, Oakley JE, Rakow T, 2006. Uncertain Judgements: Eliciting Experts' Probabilities. John Wiley & Sons.
- RFL, 2018. Oser l'usine interconnectée. Revue Laitière Française – décembre 2018
- Sine, M, Morin, E, Simonneau, D, Brochard, M, De Cosnac, G, and Escriou, H. 2010. VIGICULTURES—An early warning system for crop pest management. In Scientific and Technical Information and Rural Development IAALD XIIIth World Congress, Montpellier (pp. 26–29).
- Tan PN, Steinbach M, Karpatne A, Kumar V. 2018 Introduction to Data Mining (Second Edition). <https://www-users.cs.umn.edu/~kumar001/dmbook/index.php>
- UsineNouvelle. 2019. La révolution de la Data. L'usine nouvelle Avril 2019.
- Vitagora 2017. L'usine agroalimentaire du futur : à quoi vous attendre ? <https://www.vitagora.com/blog/2017/usine-futur-agroalimentaire>
- Wallach D, Makowski D, Jones J, Brun F. 2018. Working with dynamic crop models. Methods, Tools and Examples for Agriculture and Environment. Elsevier (third edition).

Annexe 2. Proposition de composition des instances de gouvernance

Animateurs et membre de la cellule d'animation

Membre	Institut	Email	Expertise	Volets
François Brun Animateur	Acta	francois.brun@acta.asso.fr	sources de données, modélisation agriculture numérique	V1,V2,V3
David Makowski Animateur	INRA	david.makowski@inra.fr	statistiques agronomie	V1,V2,V3
François Piraux	Arvalis	f.piraux@arvalis.fr	statistiques production végétal	V1, V2
Chris Roth	ANSES	chris.roth@anses.fr	statistiques analyse de risque	V2
Fanny Tenenhaus-Aziza	CNIEL	ftenenhaus@cniel.com	Statistiques, modélisation AQR Valorisation des données de la filière laitière	V1,V2
Elodie Doutart	IDELE	elodie.doutart@idele.fr	science des données élevage	V1, V2
Luis Galarraga	INRIA	luis.galarraga@inria.fr	science des données	V2
Samuel Buis	INRA	samuel.buis@inra.fr	méthodes modélisation agronomie	V2
Ronan Trepos	INRA	ronan.trepos@inra.fr	méthodes modélisation agronomie, plateforme Record	V2

Comité de pilotage restreint

Partenaires impliqués comme animateurs (bénéficiaire de financement)		
ACTA	Francois Brun	francois.brun@acta.asso.fr
INRA	David Makowski	David.makowski@inra.fr
Partenaires impliqués dans le groupe d'animation		
ARVALIS – Institut du végétal	François Piraux	f.piraux@arvalis.fr
Institut de l'Élevage	Elodie Doutart	Elodie.Doutart@idele.fr
ANSES	Chris Roth	chris.roth@anses.fr
CNIEL	Fanny Tenenhaus-Aziza	ftenenhaus@cniel.com
INRIA	Luis Galárraga	luis.galarraga@inria.fr
INRA	Samuel Buis	samuel.buis@inra.fr
INRA	Ronan Trepos	Ronan.Trepos@inra.fr
Représentants des partenaires bénéficiaire de financement		
ACTA	Philippe Vissac	philippe.vissac@acta.asso.fr
INRA-IRSTEA	Philippe Hinsinger	philippe.hinsinger@inra.fr
Représentant du Ministère en charge de l'Agriculture		
MAP, DGER	à définir	à définir

Liste des correspondants des autres partenaires

Autres partenaires : personne contact principal		
ITB	Fabienne Maupas	fabienne.maupas@itbfr.org
IFV	Xavier Delpuech	xavier.delpuech@vignevin.com
Terres Inovia	Sébastien Gervois	s.gervois@terresinovia.fr
CTIFL	Bertrand Alison	alison@ctifl.fr
IFIP – Institut du porc	Bernard Fostier	bernard.fostier@ifip.asso.fr
ITAVI	Gaëtan Laval	gaetan.laval@itavi.asso.fr
ASTREDHOR	Olivier Yzebe	olivier.yzebe@astredhor.fr
FNPPPT	Yves Le Hingrat Christophe Dargier	yves.lehingrat@fnpppt.fr christophe.dargier@fnpppt.fr
FNAMS	Camille Guérin	camille.guerin@fnams.fr
ITEIPMAI	Guillaume Frémondrière	guillaume.fremondriere@iteipmai.fr
CRA Normandie	Nicolas Gravelle	nicolas.gravelle@normandie.chambagri.fr
CIRAD	Philippe Letourmy	philippe.letourmy@cirad.fr
ENSFEA	Laurent Bedoussac	laurent.bedoussac@inra.fr
EPLEFPA Toulouse Auzeville	Frédéric Robert	frederic.robert@educagri.fr
société ITK	Philippe Stoop	philippe.stoop@itkweb.com

Annexe 3. Engagements des partenaires

Partenaires impliqués comme animateurs (bénéficiaire de financement)	
ACTA	François Brun - animateur, 5 mois par an Ivana Aleksovska (V2) Mathilde Chen (V2)
INRA	David Makowski - animateur, 2 mois par an Ronan Trepos - groupe d'animation Samuel Buis - groupe d'animation nombreux autres participants non listés des département Environnement et Agronomie, le département Physiologie Animale et Systèmes d'Élevage, le département Mathématiques et Informatique Appliquée et le département Santé des Plantes et Environnement et le département Sciences pour l'Action et le Développement.
Partenaires impliqués dans le groupe d'animation	
ARVALIS – Institut du végétal	François Piraux (V1, V2) - groupe d'animation Florent Duyme (V1, V2) Emmanuelle Héritier (V1, V2) Emmanuelle Gourdain (V1,V2,V3)
Institut de l'Élevage	Elodie Doutart - groupe d'animation Aurore Philibert (V2) Mohammed El Jabri (V2) Maxime Legris (V2)
ANSES	Chris Roth - groupe d'animation autres participants à définir
CNIEL	Fanny Tenenhaus-Aziza - groupe d'animation Bruno Roney (V1, V2) Florence Grimaud (V1,V2)
INRIA	Luis Galárraga - groupe d'animation autres participants de l'équipe-projet LACODAM à définir
Autres partenaires : personnes participantes	
ITB	Fabienne Maupas - correspondant Juliette Adrian (V2) François Joudelat (V1, V2)
IFV	Xavier Delpuech - correspondant Marc Raynal (V1, V2) Xavier Burgun (V1, V2) Christian Debord (V2) Valérie Lempereur (V3) Eric Chantelot (V3) Aurélien Hornoy (V1)

Terres Inovia	Sébastien Gervois (V1, V2, V3) - correspondant Frédéric Salvi (V1, V3) Célia Pontet (V2)
CTIFL	Bertrand Alison (V1, V3) - correspondant Jérôme Vibert
IFIP – Institut du porc	Bernard Fostier - correspondant Autres participants à préciser
ITAVI	Gaëtan Laval (V2) - correspondant Sylvain Gallot (V1) Pauline Creach (V1) Yann Guyot (V2)
ASTREDHOR	Olivier Yzebe - correspondant
FNPPT	Christophe Dargier (V1, V2) - correspondant Yves Le Hingrat (V1,V2) - correspondant
FNAMS	Camille Guérin - correspondant Benjamin Coussy (V1, V2, V3) Julie Gombert (V1,V2,V3)
ITEIPMAI	Guillaume Frémondrière - correspondant Benjamin Lemaire (V1, V2, V3)
CRA Normandie	Nicolas Gravelle - correspondant Laura Saxe (V1) Jean-Jacques Beauchamp (V1) Marine Louargan (V1) Pierre Cordel (V1, V3) Ugo Denis (V3)
CIRAD	Philippe Letourmy (V1, V2, V3) - correspondant Eric Gozé (V1,V2,V3) autres participants non listés des différentes unités
ENSFEA	Laurent Bedoussac - correspondant
EPLEFPA Toulouse Auzeville	Frédéric Robert - correspondant Ingrid BERJAUD (V1,V2,V3) Mathilde RAMOS (V1, V2, V3) Frédéric LACAZE (V1, V2, V3) Pauline MOLINIER (V1,V2,V3)
société ITK	Philippe Stoop - correspondant autres participants à préciser (V1, V2)

Annexe 5. Budget prévisionnel du RMT

Conformément au cahier des charges, les financements demandés sont fléchés sur les activités d'animation du réseau. Aussi, seuls les organismes de rattachement des deux animateurs (ACTA et INRA), recevront un financement de la part de la DGER qui fera l'objet d'une convention de financement ultérieur.

Les autres partenaires s'engagent uniquement à mettre à disposition un certain nombre de leurs agents afin de participer aux actions du réseau (séminaires, groupe de travail, formation, projets...).

En fonction des actions et en accord avec les animateurs du réseau, ils pourront bénéficier indirectement du financement (frais d'inscription gratuits ou réduits, remboursement des frais de déplacement des intervenants, co-encadrement de stagiaires, ...).

Nous proposons un budget sur 5 ans de **490750** euros, dont 380250 euros hors salaire public, et une demande de subvention de **303255** euros, soit un financement à **79.8%**.

Par action

Titre des actions	Volet 1,2,3	Valorisation	Total général
Coût total en €	392375	98375	490 750 €
Total hors salaire public	309500	70750	380 250 €
Total salaire public	82875	27625	110 500 €
Aide sollicitée CAS DAR	247655	55600	303 255 €
Autres concours financiers			0 €
Autofinancement	144720	42775	187 495 €

Par partenaire

Nom des partenaires	ACTA	INRA	Total général
Coût total en €	330250	160500	490 750 €
Total hors salaire public	330250	50000	380 250 €
Total salaire public	0	110500	110 500 €
Aide sollicitée CAS DAR	253255	50000	303 255 €
Autres concours financiers			0 €
Autofinancement	76995	110500	187 495 €

Détails

Action	Désignation des partenaires par catégorie	Nature de la dépense	Coût total (en euros)	Temps de travail techniciens, ingénieurs et cadres scientifiques (en jours de travail)	Aide sollicitée du ministère en charge de l'agriculture (en euros)	Auto-financement (en euros)
Volet 1,2,3	ACTA	François Brun, ingénieur (4 mois/an)	187 000	340	129 030	57 970
Volet 1,2,3	ACTA	Secretariat (10 jours par an)	12 500	50	8 625	3 875
Volet 1,2,3	ACTA	3 stage Master 2	15 000	315	15 000	0
Volet 1,2,3	ACTA	location salle	20 000		20 000	0
Volet 1,2,3	ACTA	Fonctionnement (déplacement dont autres partenaires,...)	25 000		25 000	0
Volet 1,2,3	INRA	David Makowski, chercheur (salaire publique) (1.5 mois/an)	82 875	128	0	82 875
Volet 1,2,3	INRA	5 stages Master 2 ou équivalent CDD	25 000	525	25 000	0
Volet 1,2,3	INRA	Fonctionnement (déplacement dont autres partenaires,...)	25 000		25 000	0
Valorisation	ACTA	François Brun, ingénieur (1 mois/an)	46 750	85	31 600	15 150
Valorisation	ACTA	site internet	7 000		7 000	0
Valorisation	ACTA	serveur d'application	7 000		7 000	0
Valorisation	ACTA	frais d'édition, traduction,...	10 000		10 000	0
Valorisation	INRA	David Makowski, chercheur (salaire publique) (0.5 mois/an)	27 625	43	0	27 625
		Total hors salaires publics	380 250 €		303 255 €	73 120 €
		Total Salaire public	110 500 €		- €	114 375 €
		Total Général	490 750 €		303 255 €	187 495 €

Annexe 6. Lettres d'intention et de signature de la convention

Partenaires et catégories	Participation
Institut Technique Agricole	
Acta, les instituts techniques agricoles	Animateur officiel
Arvalis – Institut du végétal.	Membre de la cellule d'animation
CNIEL	Membre de la cellule d'animation
ITB	
IFV	
Terres Inovia	
CTIFL	
Institut de l'Élevage	Membre de la cellule d'animation
IFIP	
Astredhor	
FNPPPT	
FNAMS	
ITEIPMAI	
ITAVI	
Chambre d'agriculture	
CRA normandie	
Organismes de recherche	
INRA	Animateur officiel
INRIA	Membre de la cellule d'animation
CIRAD	
Etablissement public	
ANSES	
Enseignement Agricole	
EPLEFPA Toulouse Auzeville	
ENSFEA	
Autres partenaires privés	
société ITK	

note : il a été décidé avec l'IRSTEA de ne pas les mettre dans le partenariat, sachant que la fusion INRA-IRSTEA est prévue au 1^{er} janvier 2020.