

# L'accès aux données pour la Recherche et l'Innovation en Agriculture

Position des Instituts Techniques Agricoles



**acta**

LES INSTITUTS  
TECHNIQUES  
AGRICILES #

## Editorial

Le livre blanc construit par l'ACTA et le réseau des Instituts Techniques Agricoles vient affirmer la position de ces acteurs majeurs du développement agricole et de l'innovation en France, sur un domaine en forte transition.

Les données constituent une composante de la transition numérique, qui est en réalité, pour l'agriculture, le fruit d'une innovation exogène et qui favorise l'implication sur l'agriculture d'un grand nombre de nouveaux acteurs.

La production et l'utilisation de données mobilisent des capteurs nouveaux pour produire de la donnée, souvent basée sur des processus physiques, des algorithmes pour traiter ces signaux, et des modèles pour transformer l'information en un support de décisions, ces différents éléments étant au cœur du métier des ITA. Le développement des données massives nécessite aussi des technologies nouvelles pour le transport à haut débit de l'information, interrogeant l'équité de toutes les régions de France et d'Europe face à cet accès.

Les réflexions partagées ici montrent clairement l'intérêt de favoriser les données ouvertes, car ceci permet de créer de la valeur, et ceci permet aussi aux acteurs qui ouvrent leurs données d'imposer des standards, et au-delà des standards, ce sont des cadres de réflexion qui s'imposent.

Les e-infrastructures et les démarches permettant l'interopérabilité doivent être adaptées ensuite pour permettre cette ouverture et l'accès aux données.

Les données et les e-infrastructures constituent une partie (seulement) de la transition qui se dessine. La première étape est celle de l'agriculture de précision, qui permet d'une part de capturer les hétérogénéités des ensembles sur lesquels on travaille (parcelles agricoles ou troupeaux) et d'autre part de tracer et d'éclairer toutes les performances, productives et environnementales, de la production agricole et agro-alimentaire. Mais il faut d'ores et déjà préparer l'étape suivante, celle de l'agriculture après généralisation du numérique. Et ce monde-là, avec ses acteurs nouveaux et ses relations nouvelles entre acteurs, ses réseaux sociaux, ses régulations et réglementations, reste totalement à construire.



**Christian Huyghe**  
*Directeur Scientifique  
Agriculture de l'INRA*



**Hervé Pillaud**  
*Agriculteur  
connecté*



**Philippe Lecouvey**  
*Directeur général de l'ACTA –  
les Instituts Techniques Agricoles*

En 2014, l'ACTA et les Instituts Techniques ont décidé de doter leur dispositif de Recherche Innovation d'un réseau numérique. Ce réseau a trouvé spontanément sa dynamique en s'appuyant sur les expériences acquises par chacun et en recherchant les synergies.

Il s'est préoccupé très rapidement de l'accès et de la valorisation des données avec deux réalisations fondatrices :

- Le projet API-AGRO, plateforme ouverte de données et de services agricoles qui se transforme en une société ouverte à d'autres acteurs,
- Le présent livre blanc, fruit d'un an de réflexions, qui propose 10 recommandations autour de 3 axes : « Innover », « Fluidifier », « Rassurer ».

Notre réseau se place volontairement dans une démarche d'Open innovation pour co-construire les services de demain en mobilisant de nouveaux partenariats et compétences, avec l'ambition de mobiliser les technologies adéquates pour fluidifier les échanges de données entre les acteurs.

En parallèle, nous proposons de prendre position, tous ensemble et dès à présent, sur les questions de propriétés des données et de transparence sur leurs usages pour rassurer les acteurs.

Dans ce nouvel univers du numérique et de l'économie numérique, l'ACTA et les Instituts techniques ont pris le parti d'être proactifs et innovants et ainsi de ne pas s'enfermer dans les peurs et l'immobilisme.

## Sommaire

<b>Les Instituts Techniques Agricoles se saisissent de la question des données .....</b>	<b>5</b>
Le réseau Numérique et Agriculture .....	5
Une prise de conscience du monde agricole.....	5
Pourquoi et comment favoriser l'accès aux données pour la recherche et l'innovation ?.....	6
<b>Pourquoi et comment valoriser les données des exploitations agricoles ? .....</b>	<b>9</b>
Capturer directement les données des exploitations.....	9
Les enjeux liés à la valorisation des données des exploitations.....	14
Méthodes pour valoriser les données.....	16
<b>Quelles garanties et moyens pour l'accès aux données des exploitations ?.....</b>	<b>23</b>
Garantir une transparence des usages et une utilisation vertueuse des données .....	23
Les moyens et services pour accéder aux données et en garantir les usages .....	28
<b>Innover – Fluidifier – Rassurer :</b>	
<b>10 recommandations pour favoriser l'accès et la valorisation des données.....</b>	<b>32</b>
Innover .....	33
Fluidifier .....	35
Rassurer.....	36
<b>Scénarios : quelle agriculture numérique dans le futur proche ?.....</b>	<b>37</b>
Contexte.....	37
Scénario 1 : La privatisation exponentielle des données produites sur les exploitations agricoles par les fournisseurs de l'agriculteur .....	38
Scénario 2 : L'agriculteur au cœur du système d'information .....	40
<b>Liste des personnes consultées .....</b>	<b>42</b>
<b>Informations sur la publication.....</b>	<b>43</b>

## Les Instituts Techniques Agricoles

Créés et pilotés par la profession agricole, les Instituts Techniques Agricoles (ITA) sont des organismes de recherche appliquée, d'appui technique, d'expérimentation, d'expertise, de formation et d'information. Ils ont une mission opérationnelle d'adaptation permanente des pratiques de terrain et ils jouent, à ce titre, un rôle essentiel dans la création et la diffusion du progrès technique en agriculture et vers l'amont et l'aval des filières.

Spécialisés par filières de productions (grandes cultures, élevage, fruits et légumes, viticulture, cultures spécialisées (cultures ornementales, médicinales...), agriculture biologique, cultures et élevages d'Outremer...), ils exercent des missions d'intérêt général et des missions de filières orientées par les professionnels agricoles et agro-industriels.

L'ACTA - Les Instituts Techniques Agricoles - est en charge de la représentation et de l'animation du réseau des ITA sur des thématiques partagées.



## Les Instituts Techniques Agricoles se saisissent de la question des données

### Le réseau Numérique et Agriculture

L'ACTA – le réseau des Instituts Techniques Agricoles (ITA) des filières animales et végétales – a créé fin 2014 un réseau « Numérique et Agriculture » afin d'étudier tous les enjeux et impacts du numérique dans le secteur agricole.

Ce réseau s'est saisi de plusieurs thématiques à fort potentiel pour le monde agricole et les activités de Recherche-Développement conduites par les Instituts :

- les usages et la propriété des données,
- les nouvelles modalités d'analyse et de valorisation des données (notamment du Big Data et Open Data),
- le développement du transfert numérique des savoirs (e-learning, réseaux sociaux...),
- l'émergence de nouveaux modèles socio-économiques alliant numérique et agriculture,
- les outils connectés et les infrastructures-réseaux,
- l'interopérabilité des systèmes d'informations agricoles.

Les objectifs de ce réseau sont de faire prendre conscience de l'évolution numérique aux Instituts Techniques Agricoles, de faire évoluer leurs pratiques de R&D, de mettre le numérique au service de la performance des filières agricoles et des activités de R&D des ITA, d'avoir plus de visibilité sur les projets/orientations/réalisations numériques et de monter des nouveaux projets dans le champ du numérique.

### Une prise de conscience du monde agricole

Les objectifs de R&D en agriculture se diversifient. Les objectifs de soutien de la compétitivité des filières se sont enrichis ces vingt dernières années d'une prise en compte des demandes sociétales autour des questions de préservation de l'environnement, de respect du bien-être animal et de qualité des produits. Cet élargissement des champs de compétences et l'évolution des technologies de l'information conduisent à repenser profondément les modèles et les dispositifs d'acquisition de références historiques : l'expérimentation reste un outil central de la R&D agricole, mais ne peut désormais se suffire à elle-même.

Pour poursuivre leurs missions, les ITA doivent intégrer cette émergence de nouvelles technologies, nouveaux outils, nouveaux acteurs, nouveaux modèles de développement et de collaboration. Des acteurs importants du monde agricole, de l'agroéquipement ou des nouvelles technologies se positionnent sur le marché de la gestion des données en France ou à l'international.

## Agriculture 2025 et le portail de données

Au croisement du « projet agro-écologique pour la France » et de la Stratégie Nationale de Recherche, la mission « Agriculture – Innovation 2025 », menée en 2015, avait pour objectif de poser les bases d'un plan ambitieux pour la recherche et l'innovation agricoles.

Les recommandations du rapport, formulées à travers 30 projets déclinés en près d'une centaine d'actions, visent à mobiliser les acteurs de la recherche et de l'innovation autour de grands enjeux fédérateurs : stimuler les transitions agroécologique et bioéconomique, favoriser l'innovation ouverte et soutenir les actions de formation, prendre le virage du numérique et favoriser l'essor de la robotique, mobiliser les biotechnologies et soutenir le biocontrôle. Parmi les dix recommandations, deux concernent directement le numérique et les questions d'accès et de valorisation des données.

1- Mettre en place un portail de données à vocation agricole pour l'innovation ouverte.

2- Structurer la recherche sur le Numérique en agriculture.

Pour en savoir plus :

<http://agriculture.gouv.fr/sites/minagri/files/rapport-agriculture-innovation2025.pdf>



Leurs stratégies sont basées sur des intérêts commerciaux centrés sur leur domaine d'activité avec une tendance prononcée à la privatisation de ces données issues des exploitations agricoles. Dans ce nouveau paysage, **les ITA souhaitent participer à cette dynamique porteuse d'innovation tout en rappelant l'importance d'une R&D collective au service des agriculteurs.**

Dans le même temps, depuis quelques années, on constate **une réelle prise de conscience de la valeur des données par les acteurs**

**institutionnels, économiques et les agriculteurs eux-mêmes** (Cf. charte des acteurs agricoles aux États-Unis, Mission Agriculture 2025, Mission OPECST sur le Big Data dans l'agriculture...).

## Pourquoi et comment favoriser l'accès aux données pour la recherche et l'innovation ?

Le développement du numérique dans les équipements, matériels et bâtiments des exploitations agricoles conduit à une situation paradoxale en matière de collecte des données sur les exploitations. La création de données se déconcentre très rapidement, les agriculteurs deviennent des fournisseurs de volumes toujours croissants de données. Mais dans le même temps, ces flux de données sont très fréquemment captés par des opérateurs privés, propriétaires des matériels générant ces informations, qui les valorisent auprès des exploitants agricoles sous des formes très synthétiques. Cette forme de valorisation répond bien aux attentes des producteurs, mais en parallèle ce processus conduit à une privatisation totale ou partielle des flux croissants de données brutes. Or, le potentiel de valorisation de ces données brutes par la R&D agricole, dans un objectif d'amélioration constante de la connaissance des systèmes agricoles et de conception de nouveaux services innovants, dépasse sans doute très largement les retours synthétiques proposés par les opérateurs qui

les collectent. Les producteurs n'ont pas la possibilité à ce jour de mettre ces données brutes générées sur leur exploitation à disposition des organismes de R&D pour conduire des travaux à vocation collective. Cette privatisation de la collecte de données fragilise ainsi le potentiel de recherche appliquée et d'innovation du secteur de la R&D collective, à l'encontre même de l'intérêt général des exploitants agricoles.

Dans ce document, les Instituts Techniques Agricoles (ITA) présentent leur analyse de la problématique de l'accès aux données agricoles pour les activités de recherche et d'innovation. A partir de ce diagnostic, plusieurs propositions sont formulées pour faciliter l'accès et la valorisation des données. Cela concerne des données déjà disponibles au sein des organismes de R&D collective, mais surtout les nouveaux gisements que constituent les exploitations agricoles connectées et l'ensemble des acteurs économiques du secteur.

Au-delà de la disponibilité des données, différentes propositions portent sur l'amélioration de la valorisation de ces données ouvertes, à travers un renforcement des compétences des ITA en matière de gestion et d'analyse de données et de coordination collective.

Enfin, nous pensons que notre analyse et nos propositions pourront éclairer les débats sur les données agricoles qui ont commencé à s'engager et qui seront au cœur de l'agenda agricole des prochaines années. La liberté des agriculteurs à décider de manière volontaire de réorienter, totalement ou partiellement, les flux des données collectées sur leur exploitation vers les acteurs de leur R&D collective nous semble un enjeu majeur pour les producteurs. Cet enjeu devra être clairement formalisé, par exemple dans une future charte clarifiant les conditions de captation des données, de partage et de valorisation. Cette réflexion a été initiée lors d'un workshop sur l'accès aux données organisé le 8 octobre 2015 par notre réseau afin de proposer un éclairage sur la question de l'accès aux données. L'objectif était d'entamer une réflexion collective des ITA et de leurs partenaires sur cette problématique, de partager un diagnostic et de proposer des moyens de valorisation des données pour la Recherche et l'Innovation.



### Zoom sur le Workshop du 8 octobre 2015

Le workshop comportait une session de présentations concrètes d'expériences de valorisation collective de données avec les problèmes rencontrés, les solutions trouvées, la valeur ajoutée des services proposés... Puis, nous avons organisé des ateliers par groupe pour traiter plus en profondeur les thèmes suivants : valorisation des données pour la recherche, valorisation des données pour l'innovation, transparence par rapport aux usages des données. Ces ateliers ont été animés par des échanges et structurés sous la forme d'une carte heuristique.

A la suite du workshop un groupe plus réduit a rédigé le présent document en réorganisant les idées échangées lors du workshop, mais aussi en complétant les illustrations et en développant certains points. Ce document a été proposé à la relecture des participants et aux directeurs des ITA pour leurs avis sur les propositions.

# Réseaux Technico-Économiques des élevages porcins



## Historique et contexte

L'élevage de porcs est un grand consommateur d'intrants, avec notamment l'aliment qui représente environ 70% du coût de revient. Cette production est dans une situation économique sensible. Le marché est très concurrentiel et les agriculteurs ne maîtrisent pas le prix de vente de leurs porcs et subissent les aléas des productions végétales dont sont issus les aliments consommés. L'optimisation des revenus passe par une meilleure efficacité de l'élevage via l'augmentation du nombre de kilos produits et/ou la diminution des charges alimentaires. Depuis 40 ans, deux dispositifs de collecte de données ont été mis en place : la Gestion Technique des Troupeaux de Truies (GTTT) et la

Gestion Technico-Économique (GTE). Les références constituées à partir de ces données permettent d'analyser les performances des élevages suivis.

## Les outils/ les services proposés

Plusieurs milliers d'élevages sont suivis dans ces dispositifs. La collecte des données est en général proposée par le technicien du groupement de producteurs ou de la chambre d'agriculture, sur la base du volontariat. La GTTT, qui permet d'évaluer la productivité des truies au sevrage, est suivie par 40% des élevages de plus de 50 truies en 2014, ce qui représente 53% des truies présentes au niveau national. La productivité de l'ensemble de l'élevage jusqu'à la vente de la viande par les charcutiers, est évaluée au niveau de la GTE qui concentre au niveau national les données de 44% des élevages de plus de 50 truies. Le dispositif GTE-TB (GTE et Tableau de Bord), qui permet de calculer le coût de revient de l'atelier porcin, est moins suivi (8% des élevages de plus de 50 truies).

Les références calculées au niveau national permettent aux éleveurs de situer leurs résultats, et d'identifier des pistes d'amélioration. Les bases de données constituées permettent des études complémentaires indispensables à la profession, et peuvent également fournir un éclairage aux pouvoirs publics sur certaines pratiques (ex. dépenses de santé).

## La question de l'accès aux données

Les données de gestion sont d'abord saisies dans des logiciels ad hoc à l'élevage ou au groupement. Le groupement réalise un premier contrôle des données et des analyses spécifiques, puis transmet à l'IFIP – Institut du porc les données des éleveurs suivis. Aucun transfert de données individuelles n'est réalisé à des tiers sans l'autorisation des éleveurs.

L'IFIP réalise les calculs de référence deux fois par an (printemps et automne). Les moyennes ne sont calculées que pour les groupes d'au moins 10 éleveurs. L'IFIP met en forme les résultats et diffuse les références publiques sous forme agrégée dans des revues annuelles, des plaquettes régionales et sur son site web ([www.ifip.asso.fr/fr/resultats-economiques-gttt-graphique.html](http://www.ifip.asso.fr/fr/resultats-economiques-gttt-graphique.html)).

L'IFIP enregistre ces dernières années une réduction de la remontée des données des éleveurs dans la base nationale, et le développement de systèmes internalisés au niveau des groupements de production. Cette évolution pourrait conduire à une privatisation des données, qui mettrait en péril le calcul de références nationales représentatives, et limiterait l'accès aux données pour les travaux de recherche. Afin de maintenir le dispositif, l'IFIP développe depuis peu une valorisation directe des données de gestion par les éleveurs via le portail web GT-Direct. Il propose aujourd'hui aux éleveurs suivis un accès gratuit à 5 modules simples d'utilisation, de la consultation des résultats de l'éleveur au calcul du coût de revient du kilo de carcasse. D'autres outils complètent ces informations, comme l'outil de Gestion Environnementale des Elevages Porcins (GEEP), ou un outil d'auto-évaluation de la consommation d'antibiotiques (à venir).

## Pourquoi et comment valoriser les données des exploitations agricoles ?

### Capter directement les données des exploitations

#### L'acquisition des données techniques par la R&D collective agricole

Les données mobilisées ou mobilisables dans les différents travaux de recherche et d'innovation des acteurs de la R&D collective agricole sont très diverses ; elles proviennent de différentes origines et sont issues de différents fournisseurs.

**Ces acteurs mobilisent beaucoup de moyens en expérimentation planifiée** pour répondre à leur mission d'acquisition de référence et d'évaluation des innovations. Ces expérimentations sont soit menées de manière répétée d'année en année selon un protocole standardisé (évaluation des nouvelles variétés, des nouveaux produits phytosanitaires ou de fertilisation...), soit liées à des actions de recherche plus ponctuelles. Elles sont notamment menées par les Instituts Techniques Agricoles et leurs partenaires de R&D (INRA, Chambres d'Agriculture...), dans des stations expérimentales (avec souvent des dispositifs de micro-parcelles), ou bien directement chez des agriculteurs sur de plus grandes zones d'observation.

Ces données sont souvent bien structurées et collectées dans des bases de données par projet et/ou organisme. Elles font systématiquement l'objet d'analyses statistiques pour évaluer la significativité des effets.

Cette démarche expérimentale permet de garantir une grande qualité des analyses mais pose la question de la transposition à d'autres situations que celles expérimentées, de la représentativité des conditions agricoles et du coût de mise en œuvre. Par ailleurs, l'expérimentation semble limitée pour traiter de certaines problématiques complexes auxquelles l'agriculture d'aujourd'hui doit faire face.

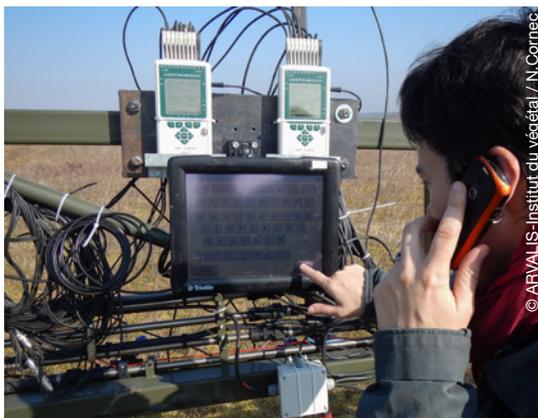


Afin d'enrichir cette approche expérimentale historique, des **données représentatives de la production agricole** sont collectées par différents organismes agricoles pour leurs besoins et leurs missions propres. Ces organismes (coopératives, négoce, centres de gestion...) collectent des données en nombre, et cette volumétrie s'accroît rapidement. Des données sont également collectées pour les travaux propres des organismes de recherche ou des ITA, via des enquêtes, réseaux d'observation... La majorité est informatisée et enregistrée dans des bases de données, avec des degrés de centralisation variables selon les thématiques et les organismes. Elles sont généralement analysées pour évaluer les tendances représentatives des exploitations françaises au niveau régional ou national.

Avec les expérimentations factorielles planifiées, on cherche à contrôler les facteurs pour comparer les différentes modalités et à gommer les hétérogénéités de sol ou de climat avec des répétitions expérimentales. Pour les observatoires basés sur des campagnes de collectes ou d'enquêtes, la représentativité des situations agricoles est recherchée en multipliant le nombre de situations et leur diversité, mais les conditions de culture ou d'élevage ne sont pas contrôlées et l'on n'a pas de réelle répétition au sens statistique du terme. Ces observatoires permettent un suivi quantitatif mais également réglementaire ou même de contrôle (exemple dans le cadre de la santé des plantes et des animaux). Les laboratoires sont aussi de grands pourvoyeurs de données sur la base d'échantillons collectés dans des expérimentations ou observatoires, mais aussi directement chez les agriculteurs ou les organismes stockeurs : analyse du sol, des grains, de la qualité du lait, des hormones, pesticides, antibiotiques, etc. Avec l'émergence des micro-stations d'analyse, par exemple embarquées dans les robots de traite, la donnée à valeur ajoutée est collectée en temps réel directement sur l'exploitation, accélérant les délais de prise de décision (adaptation de l'alimentation, action vétérinaire à déclencher,...).



© Terres Inovia / Jung



© ARVAL IS - Institut du végétal / N. Cornec

# France Génétique Elevage : Système d'information génétique

## Historique et contexte

Le dispositif de suivi de la génétique des élevages a été mis en place avec la loi sur l'élevage de 1966 avec une très forte implication des pouvoirs publics et un contexte fortement réglementé avec des monopoles de zones d'activité.

Ce système d'information est organisé selon un niveau régional et national. Dans les années 80, on assiste à la constitution de bases de données régionales rassemblant toutes les données collectées sur le terrain avec des accords mutuels d'accès aux données. Puis, dans les années 90, un Système d'Information Génétique (SIG) est mis en place pour aboutir à un système unique et collectif. Dans les années 2000, des acteurs se restructurent et s'équipent en système d'information. La donnée devient un enjeu commercial, à la base du conseil en élevage pour tous les acteurs. Les sources de données (capteurs, robots...) se multiplient et l'indexation génomique fait son apparition. Mais il est nécessaire de maintenir une caractérisation des phénotypes.

En 2006, la loi d'orientation agricole met fin au monopole des zones pour la mise en place de l'insémination artificielle. Récemment, en 2014 le contrôle de performance est ouvert à la concurrence. Une évolution réglementaire via l'arrêté SNIG du 24 mars 2015 oblige ces acteurs à recueillir, gérer et respecter le consentement éclairé des éleveurs. La restructuration des acteurs se poursuit. Un nouveau règlement zootechnique européen est en cours de rédaction : il confierait toutes les responsabilités aux Breed Societies avec la fin du monopole de l'INRA pour l'évaluation des reproducteurs et le passage à un dispositif contractuel.

## Objectif(s)

Le dispositif visait à maximiser les valorisations de données par tous les acteurs pour inciter à la collecte et fiabiliser la détection des erreurs, faire circuler les données entre les acteurs de la génétique des bovins et mettre à disposition ces données. Ce dispositif concerne 38 000 éleveurs au contrôle laitier, avec 180 millions d'identifications de bovins, 39 millions de pesées en vif et 140 millions d'inséminations...

## Les outils/ les services proposés

Beaucoup d'outils de valorisation de données ont été réalisés au niveau régional. Il s'agit d'outils pour les techniciens et éleveurs (portails, technico-économique...). Ces valorisations dépassent largement les objectifs initiaux.

## La question de l'accès aux données

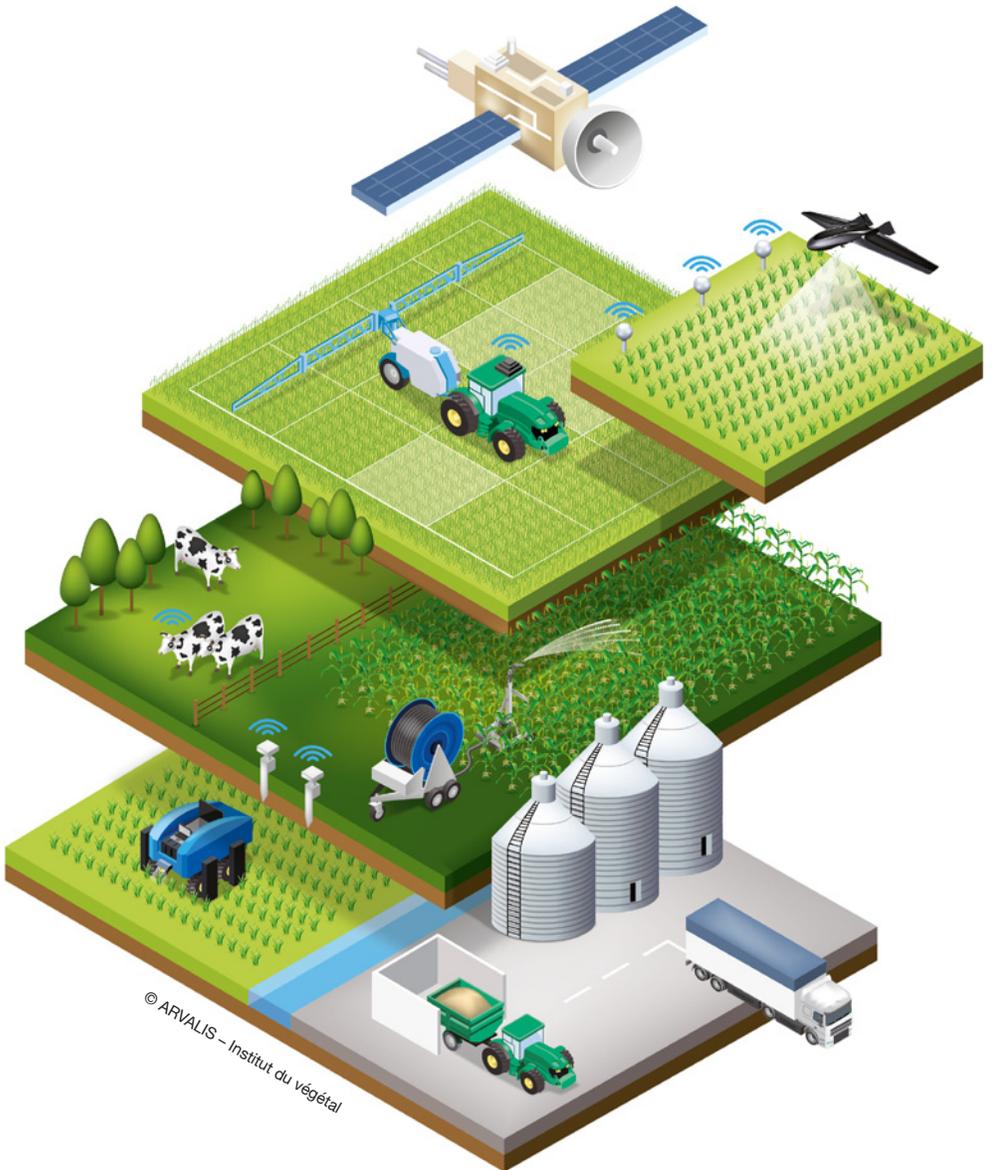
Une charte « Droits et devoirs » gère les grands principes, avec notamment une gestion fermée des apports de données et une gestion ouverte des lectures de données. Seuls les organismes agréés peuvent mettre à jour la base de données.

L'accès aux données devient un enjeu crucial avec la pression des organismes extérieurs pour l'accès aux données et une demande très forte des acteurs de renforcer la sécurité des données et de fermer l'accès. L'intérêt à la mutualisation du système d'information au niveau national est remis en cause du fait de la privatisation des données hors du système collectif.

Les évolutions réglementaires imposent aux acteurs de s'adapter pour continuer à fournir ce service national tout en répondant aux nouvelles contraintes (gestion du consentement...) et en prenant en compte le caractère privatif de certaines données. Cela nécessite notamment le développement de standards, de service web...

En savoir plus : <http://fr.france-genetique-elevage.org/Cahiers-des-charges-specifiques.html>





La plupart des acteurs de la R&D agricole utilisent une combinaison de ces moyens de collecte de données, avec pour le moment, un poids important accordé aux données issues des expérimentations planifiées pour comparer les pratiques, mettre au point des référentiels et développer des connaissances qui de plus en plus sont mises sous la forme de modèles mathématiques. Les données de simulation, issues de ces modèles intégratifs, sont des données de seconde génération dont l'importance grandit dans les travaux de R&D.

La contextualisation des données dans l'espace et dans le temps **est nécessaire à toute valorisation. Les données collectées sont représentatives de différentes échelles** avec une collecte ou agrégation aux échelles intra-parcellaire, parcellaire ou territoriale en fonction des objectifs et problématiques visés (agriculture de précision, production à l'échelle de la parcelle, économie à l'échelle de l'exploitation, environnement à l'échelle territoriale...). Une donnée peut par ailleurs être acquise et exploitée à différentes échelles de temps (du temps réel, à des échelles pluriannuelles).

Ces données peuvent être brutes (ex. le signal de fluorescence issu du capteur), **transformées** (ex. la teneur en chlorophylle) et plus ou moins **agrégées** (ex. carte intra-parcellaire de la teneur en chlorophylle).

### Une explosion de la quantité de données disponibles dans les exploitations

Il existe un autre gisement de données issues directement des exploitations agricoles et de leur activité de production. Elles sont collectées soit pour assurer le fonctionnement de l'exploitation, soit pour répondre à des demandes réglementaires ou comptables.

Dans ce cas, ces données sont collectées par l'exploitant lui-même et sont en grande partie numérisées. Elles sont notamment liées aux logiciels de gestion des exploitations (gestion comptable, gestion parcellaire, gestion des troupeaux...).

Force est de constater que, grâce aux capteurs autonomes, tracteurs connectés, drones, logiciels de gestion parcellaire, systèmes automatiques de gestion de l'ambiance dans les bâtiments d'élevage, échanges par courrier électronique... le nombre de données générées par les agriculteurs croît aussi vite que les évolutions technologiques et le déploiement des réseaux de communication. Cette tendance de la collecte des données « à la source » devrait d'ailleurs s'accroître avec l'avènement de l'IoT (« Internet Of Things » c'est-à-dire l'internet des objets connectés).

Cela se traduit par une augmentation des volumes de données collectées, mais aussi par une augmentation de leur hétérogénéité en ce qui concerne leur nature, leurs formats et leur qualité.

Grâce à ces outils puissants, l'agriculteur pourra mieux intégrer l'historique de ses interventions et les spécificités de son système de production. Son processus de décision sera facilité par des simulations en temps réel, intégrant par exemple objectifs de rendement, risques de maladies et coûts de production. Son matériel « connecté » pourra appliquer de manière plus précise et plus personnalisée des décisions.

Mais au-delà de cet usage personnel, pour certaines thématiques, la population des agriculteurs bien équipés pourrait constituer un observatoire de grande valeur sans précédent pour les travaux de la R&D collective agricole et une source d'innovation pour de nouveaux services.

**Cette révolution numérique replace l'agriculteur au centre du processus d'innovation. Il est à la fois fournisseur de données mais également utilisateur de services.** Cette révolution impose aux fournisseurs de services une marche forcée dans laquelle **l'enjeu n'est plus la donnée elle-même, mais plutôt la capacité à y accéder, à la traiter et surtout à la valoriser sous forme d'information utile.**

Fournisseurs et utilisateurs de data s'inscrivent ainsi de plus en plus dans une démarche de co-conception pour optimiser leurs chances de réussite.

### Mobiliser des données « externes » d'intérêt pour ajouter de la valeur

Les données météorologiques, les caractéristiques des sols, les données géographiques, cadastrales... sont les données essentielles pour le domaine agricole, notamment pour les productions végétales qui sont directement sous l'influence des conditions pédoclimatiques. Elles sont à la fois utiles pour piloter l'exploitation agricole, mais également - en les croisant aux données agricoles collectées - permettent de mieux comprendre les facteurs influençant les productions et servent à élaborer des modèles de prévision. Souvent, différents fournisseurs peuvent partager l'accès aux mêmes bases de données.

Les données économiques (énergie, mais aussi prix de vente...) sont également utilisés par les acteurs agricoles ont aussi besoin de références économiques pour élaborer leurs stratégies et plans d'action. Aussi, ces données économiques sont manipulées par les acteurs de la R&D depuis longtemps.

Avec les nouveaux défis de l'agriculture (dimension sociétale, environnementale, santé humaine, sécurité alimentaire, développement harmonieux des territoires...), d'autres dimensions rentrent désormais en ligne de compte. Cela nécessite de manipuler des données traditionnellement moins utilisées par les acteurs de la R&D.

Les principaux gestionnaires de données sont, historiquement, le Ministère de l'Agriculture, le Ministère de l'Ecologie et du Développement Durable, les industriels, les Instituts de recherche, les Instituts techniques agricoles, les coopératives, les marchés et autres opérateurs privés. Ces organisations ont développé au fil du temps des types de collectes différents en fonction de leurs missions respectives : collecte institutionnelle pour une valorisation collective, crowdsourcing, remontée automatique depuis des « bases privées »...

**Avec le captage des données à la source, l'exploitation agricole est devenue une source de données brutes incontournable et essentielle.** Les groupements de producteurs pourraient s'organiser en conséquence afin de prendre en compte et ajouter cette nouvelle contribution dans leurs missions plus historiques.

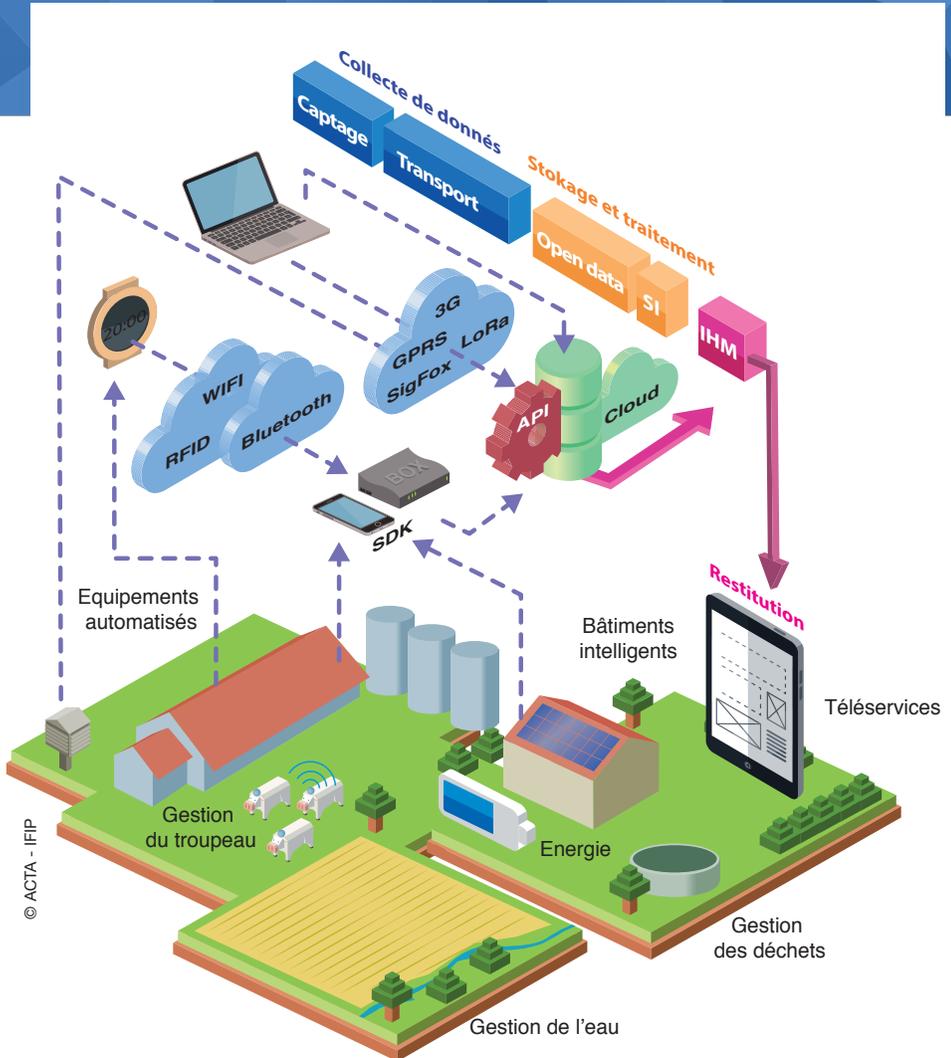
### Les enjeux liés à la valorisation des données des exploitations

**La capitalisation des données dans le temps et l'espace permet de construire des référentiels représentatifs de la diversité des situations.** Cela permet ainsi d'avoir une meilleure connaissance de la situation agricole et de son évolution. Ces informations sont précieuses pour mieux connaître les évolutions des pratiques et de leur diversité.

Mais ces référentiels sont surtout indispensables pour les exploitations qui peuvent ainsi mieux se situer par

# L'internet des objets – Internet of Things (IoT)

Cela représente les échanges d'informations et de données provenant de dispositifs présents dans le monde réel vers le réseau Internet. L'internet des objets est en partie responsable d'un accroissement exponentiel du volume de données générées sur le réseau, à l'origine du Big Data. Il pose également de nouvelles questions du point de vue de l'hébergement des données sur des serveurs (Cloud) et des capacités des réseaux de télécommunication à véhiculer ces données (3G, 4G, 5G, LoRa, Sigfox...).



rapport aux autres exploitations similaires. Se comparer est en effet essentiel pour progresser techniquement et économiquement.

**Si une donnée brute seule n'a que peu de valeur intrinsèque, croiser différentes sources d'information peut en revanche générer beaucoup de valeur ajoutée et augmenter considérablement la qualité des services rendus.** En croisant ces informations brutes avec d'autres données, comme les données météorologiques, sur les variétés ou la génétique et bien d'autres encore, les ingénieurs et chercheurs développent des modèles prédictifs.

En retour, l'agriculteur aura accès à des prédictions plus fiables à travers ses outils lui permettant de prendre les meilleures décisions stratégiques et tactiques en fonction de son contexte, du contexte économique et des prédictions météorologiques. À l'avenir, l'enregistrement de plus en plus automatique et géo-localisé des interventions techniques permettra un zonage précis pour exploiter pleinement le potentiel d'une parcelle. Mais pour cela, il faut être capable de croiser les différentes couches de données et de valoriser ces informations grâce aux référentiels validés.

Si cette capitalisation des données et leur valorisation font partie des missions des ITA et d'un certain nombre d'acteurs de la R&D collective avec des réseaux de collecte de données dédiés, les évolutions technologiques facilitent désormais la remontée directe des informations des exploitations et offrent des opportunités à d'autres acteurs de capitaliser et valoriser les données. Cette ouverture à de nouveaux acteurs est une importante source de renouvellement et doit pousser l'innovation. Cependant les acteurs de la recherche et de l'innovation au service des agriculteurs doivent prendre conscience de l'adaptation nécessaire de leur schéma actuel d'élaboration des préconisations et la diffusion du conseil. **Les acteurs de la R&D collective agricole doivent s'adapter à ces évolutions afin de continuer de fournir une expertise objective et indépendante sur les évolutions des systèmes agricoles et sur les services offerts par ces nouveaux acteurs.**

## Méthodes pour valoriser les données

### Du signal à la donnée

Avec la multiplication des capteurs et des dispositifs d'enregistrement, de plus en plus de données sont collectées automatiquement (ex. température ou pluviométrie au champ). La plupart des capteurs ne nécessitent pas ou peu d'intervention humaine. Ils peuvent être embarqués, par exemple dans des drones ou des agroéquipements plus conventionnels (tracteurs, moissonneuses batteuses...), ou installés dans des bâtiments d'élevage. Mais dans la majorité des cas, les données issues de capteurs nécessitent un traitement pour transformer le signal en une valeur interprétable. Ainsi, une confrontation à des référentiels est indispensable pour construire des modèles d'utilisation de ces données. Une plateforme de données, voire une plateforme IoT, peut alors être d'une grande utilité.

# Réseau d'agriculteurs : exemple du Farmers Business Network



## FARMERS

BUSINESS NETWORK

### Historique et contexte

Pour plus de 30% des agriculteurs, une source importante d'informations est l'échange entre agriculteurs (Enquête Agrinautes Agrisurfeurs 2014 NTIC AGRI CONSEIL – BVA). Pourtant, ces échanges restent peu formalisés. Aux Etats Unis, l'Open Ag Data Alliance (OADA) vise à mieux mutualiser les données et le Farmers Business Network s'inscrit dans cette initiative.

### Objectif(s)

Fondé par deux anciens salariés de Google avec une levée de fonds de 15 millions de dollars (principalement auprès de Google Ventures), Farmers Business Network mise sur un réseau d'agriculteurs indépendants. En collectant un maximum de données sur les exploitations américaines, cette société a pour ambition d'offrir aux agriculteurs les outils adaptés aux échanges d'expériences.

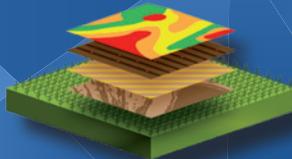
### Les outils/ les services proposés

Pour 500 dollars par an, l'agriculteur accède au stockage sécurisé des données numériques de son exploitation, en particulier, toutes les cartes issues des agroéquipements de précision (cartes de semis, de fertilisation, de rendement, ...). Puis, après un nettoyage des données par des algorithmes, FBN propose une analyse des données anonymisées et agrégées, comme la comparaison aux autres agriculteurs du réseau dans des situations pédoclimatiques similaires. En misant sur la loi des grands nombres, FBN propose également un outil d'analyse de la performance des variétés (à ce jour 1800 variétés de 110 semenciers) évaluées en conditions réelles dans le réseau.

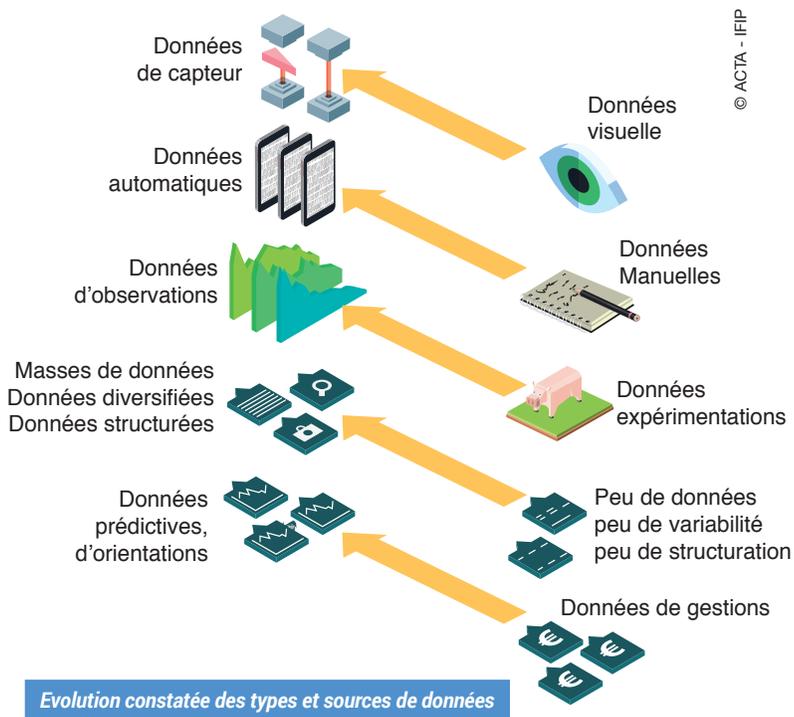
### La question de l'accès aux données

Dans les CGU, il est mentionné que les données sont anonymisées et agrégées et ne peuvent être revendues. Pour autant FBN n'a pas signé la charte américaine.

En savoir plus : [www.farmersbusinessnetwork.com](http://www.farmersbusinessnetwork.com)



*Le service propose une segmentation des exploitations pour permettre les comparaisons pertinentes*



Si on prend l'exemple des drones, ils représentent aujourd'hui un moyen d'observation en plein développement, notamment dans le secteur agricole. Et de fait, le nombre de sociétés qui fabriquent et/ou utilisent des drones pour des activités de conseil en agriculture de précision est en croissance rapide. Pour les cultures, il s'agit d'un outil souple d'utilisation, permettant d'acquérir des observations précises des parcelles agricoles au cours de la saison culturale avec une moindre dépendance aux conditions météorologiques que le satellite. Le drone constitue de ce point de vue un complément intéressant au satellite. Cependant, sa légèreté engendre des contraintes fortes sur la charge utile notamment.

La création d'Outils d'Aide à la Décision (OAD) combinés à ce type de vecteur nécessite un bon niveau d'expertise agronomique, qui en pratique est souvent limité et peut engendrer des conseils erronés. Les instituts, comme Arvalis - Institut du végétal associé à l'INRA, développent des collaborations avec des opérateurs de drones sur la partie chaîne de traitement de l'information pour combler ce déficit. Ces collaborations visent à développer des capteurs (caméras multi-spectrales...),

# Les données d'épidémiologie-vigilance végétale



## Historique et contexte

Dans le cadre du plan Ecophyto le Bulletin de Santé du Végétal a été mis en place en 2009, en remplacement des avertissements agricoles, pour assurer la maîtrise sanitaire des productions végétales et optimiser l'utilisation des pesticides. Ce document d'épidémiologie-vigilance public et gratuit est édité pour chacune des principales cultures dans chaque région à partir d'une analyse de différentes sources d'informations dont notamment un réseau régional d'observations. Cette édition est mise en œuvre par les chambres d'agriculture régionale avec les différents acteurs régionaux (ITA, Fredon, coopératives...).

## Description de la source de données

Les données d'observations sont des données collectées par des techniciens selon des protocoles harmonisés. Ainsi, toutes filières confondues, des milliers de parcelles font l'objet d'un suivi régulier chaque année. Il y a d'un côté les caractéristiques des parcelles (localisation, semis, variété, précédent...) et les données d'observations à une date donnée (stade de la culture, niveaux des différentes maladies et ravageurs, commentaires...).

En fonction des filières et régions, les systèmes d'information diffèrent (Vigicultures® pour les grandes cultures (Arvalis-Institut du végétal, Terres Inovia, ITB), Epicure pour Viticulture (IFV), Latitude pour viticulture et arboriculture en région PACA (CIRAME)...), mais globalement ces données d'observations sont saisies à travers des interfaces web ergonomiques et enregistrées dans un nombre limité de bases de données bien structurées (caractéristiques des parcelles et données d'observations...).

## Quelques chiffres-clés

- des milliers de parcelles sont suivies chaque année
- des millions d'observations sont collectées chaque année

## Les outils/ les services proposés

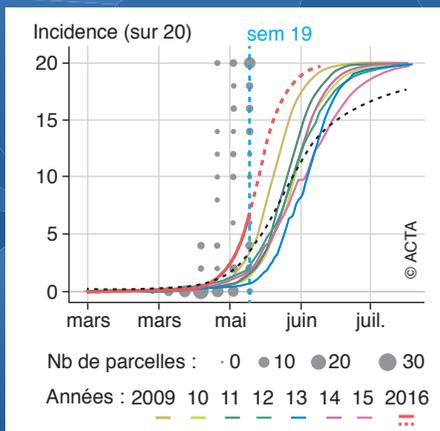
Au-delà de l'utilisation en routine de ces données pour l'édition du bulletin, quelques exemples de valorisation de ces données existent. Un premier exemple est la construction d'un outil d'analyse des dynamiques épidémiologiques pour la septoriose à destination des rédacteurs des BSV afin de les aider à réaliser leur analyse de risque (travaux ACTA-INRA-Arvalis-Institut du végétal). Ce service à destination des observateurs peut encourager la participation à ces dispositifs collectifs.

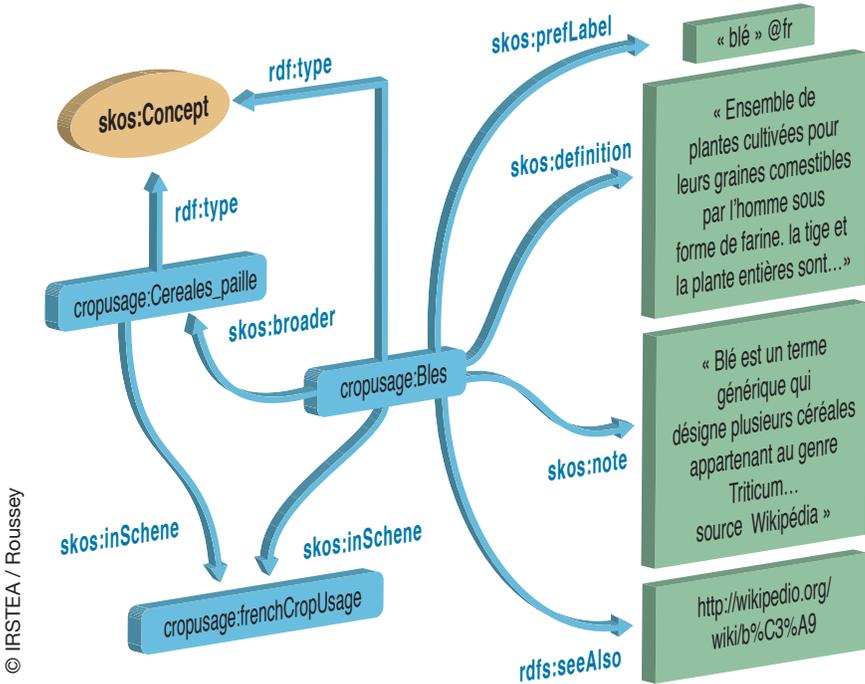
Un second exemple est la construction d'un modèle d'apparition des symptômes de la rouille jaune en fonction des conditions météorologiques, qui devrait être diffusé aux agriculteurs (Arvalis-Institut du végétal).

## Problématique d'accès aux données

Ce sont des données publiques-privées mais d'intérêt général avec une propriété dispersée entre les organismes des observateurs. Si l'agriculteur accepte d'avoir une zone de suivi dans ses parcelles, il n'est pas en charge de l'observation ni propriétaire des données. Ces données sont mises en commun pour l'édition du BSV et peuvent être exploitées pour des fins de recherche. Les possibilités d'accès à ces données restent à préciser.

En savoir plus : [www.vigicultures.fr](http://www.vigicultures.fr)





## Les ontologies pour la valorisation des données

Jean-Pierre Chanet, IRSTEA

Des capteurs apparaissent partout (champs, animaux...) et la définition d'un modèle de données commun est nécessaire pour favoriser les échanges et l'exploitation des données collectées par ces capteurs. Une telle ontologie est en cours de normalisation (Semantic Sensor Network) avec un volet agricole.

Ce travail nécessite de bien définir les spécifications et trouver les autres ontologies existantes pour favoriser la compatibilité avec l'existant.

Ces travaux permettent par exemple de publier en temps réel des jeux de données (Exemple : <https://datahub.io/fr>).

Une autre application est l'annotation de documents avec les ontologies (Exemple de la base de données documentaire des BSV).

Trois projets européens se sont emparés des ontologies pour le machinisme agricole ([www.agriopenlink.com](http://www.agriopenlink.com)), pour open data pour l'agriculture ([www.foodie-project.eu](http://www.foodie-project.eu)), ou pour l'agroalimentaire ([www.foodvoc.org](http://www.foodvoc.org)).

Pour le moment, dans le secteur agricole, le web sémantique n'est pas encore mis en œuvre à grande échelle.

une chaîne de traitement des données physiques (géométriques, radiométriques...) et enfin une chaîne de traitement agronomique permettant, grâce au couplage entre les données capteurs et les modèles agro-climatiques disponibles, de proposer des conseils sur le pilotage des cultures. Cet enchaînement de traitements fait intervenir différentes étapes et types de compétence. Les capteurs produisent des mesures brutes qu'il faut contrôler et stocker, puis corriger pour les transformer en variables biophysiques (facteur de réflectance converti en indice de végétation), et enfin y appliquer des modèles décisionnels combinés à l'interprétation des données agronomiques des parcelles (phénologie, conduite culturale...) pour aboutir à des préconisations.

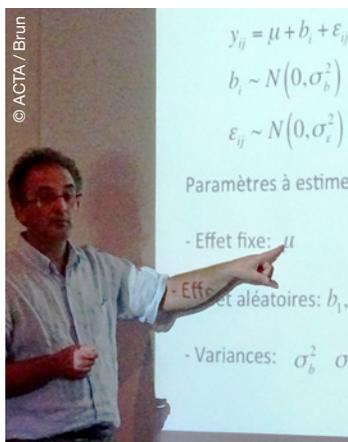
## Ontologie et web-sémantique

Les données issues de capteurs sont souvent évoquées, mais derrière le terme capteur se cache la problématique de la multiplicité des mesures. Un modèle de données commun d'échanges de ces données est nécessaire pour les valoriser (ontologie).

Par ailleurs, les bases documentaires de la communauté de R&D agricole rassemblent les documents scientifiques et techniques qui servent aussi de sources d'informations aux travaux en constituant un socle fondamental de connaissances des instituts de recherche et de la R&D collective agricole. Une meilleure exploitation de cette documentation scientifique et technique fait appel à des méthodes comme la méta-analyse ou d'analyse sémantique (ontologie, web sémantique...).

## Statistiques

**Pour valoriser le Big Data Agricole, les compétences en analyse de données deviennent centrales et les méthodologies doivent évoluer.**



Avec les données captées directement dans les exploitations agricoles, le niveau de contrôle se réduit. Si cela est proche des observatoires existants, de nouvelles difficultés apparaissent. Ces nouvelles données sont saisies directement par les agriculteurs ou leurs conseillers et souvent directement enregistrées par un



dispositif automatique. Avec cette collecte à grande échelle, l'expert qui collecte ou vérifie la donnée n'est plus présent dans ce début de chaîne. Les contrôles ne sont plus possibles sur les données individuelles brutes. Il faut donc mettre en place des algorithmes pour vérifier les données et leur cohérence et réussir à gérer les hétérogénéités sur les données et les données manquantes.

Par ailleurs, les volumes des données peuvent rendre non pertinentes des méthodes statistiques couramment employées (intervalles de confiance ou tests de comparaison) mais ouvrir des opportunités quant à la mobilisation de méthodes nouvelles. Il y a beaucoup d'attente par rapport à ces nouvelles méthodes qui permettraient de proposer des prédictions spécifiques en fonction d'un niveau de détails important et donc de valoriser la variabilité des situations. Mais avec ces approches, d'autres problèmes doivent être considérés (reproductibilité des résultats, perte de lisibilité, fort risque de confusion entre corrélation et causalité, bruitage, etc.).

Mais pour assurer une prédiction de qualité il faut chercher à profiter conjointement de sources de données complémentaires, avec d'un côté des données agricoles très contextualisées et de l'autre des références validées et expertisées issues de l'expérimentation ou des réseaux d'observations. Pour cela, les méthodes bayésiennes pourraient permettre d'intégrer ces différentes sources de connaissances.

## Quelles garanties et moyens pour l'accès aux données des exploitations ?

### Garantir une transparence des usages et une utilisation vertueuse des données

**Les acteurs agricoles, parmi lesquels se trouvent des agriculteurs, commencent à exprimer leurs interrogations et craintes** au travers de différents médias (conférences, réseaux sociaux, journaux spécialisés, think tanks, ...). Si, à travers la collecte massive des données, l'agriculteur se trouvait dépossédé des informations qu'il génère ? Et s'il y avait transfert de propriété des données vers l'entreprise qui les collecte (via les Conditions Générales d'Utilisation (CGU) des fournisseurs de services) ? Elle pourrait alors les utiliser à loisir pour générer de nouveaux revenus (par la vente de ces données, par de nouveaux services basés sur ces données). La question du contrôle réglementaire est également une forte préoccupation pour les agriculteurs. A quel degré de maîtrise des données, son producteur initial peut-il prétendre ? Le monde agricole prend conscience de ces problématiques de la propriété des données et du gisement de valeur ajoutée via leur réutilisation et la génération de nouveaux services.

Pour les acteurs de la R&D collective, c'est la question de l'accès à ces nouvelles sources de données pour mener à bien leurs missions de R&D qui est posée, notamment dans un contexte de rationalisation des efforts de collecte de données sur les dispositifs historiques (expérimentations et réseaux d'observation).

L'accès à ces données assure également une représentativité de ces exploitations agricoles connectées dans les références produites par ces acteurs et donc une prise en compte de leurs problématiques. A contrario, se focaliser uniquement sur ces données issues d'exploitations connectées risque de mettre de côté tout un pan des exploitations à court et moyen terme pour cette période de transition.

### Les parlementaires s'inquiètent également pour la souveraineté française

Récemment (2 juillet 2015), l'OPECAST (Office Parlementaire d'Evaluation des Choix Scientifiques et Technologiques, délégation interparlementaire de l'Assemblée Nationale et du Sénat) a organisé une audition publique sur « la place du traitement massif des données (Big Data) dans l'agriculture : situation et perspectives ». Cette séance a mis l'accent sur les questions de propriété, de souveraineté et sur les risques de privatisation de la donnée. La proposition d'une démarche de plateforme collective de données agricoles a notamment été faite afin d'intégrer la création d'un cadre clair pour l'utilisation des données issues des exploitations.

Cet événement parlementaire montre bien la prise de conscience de nos politiques sur les potentiels de valorisation de ces données, mais surtout sur les craintes de perdre la maîtrise des données agricoles.

# Charte américaine « Principes pour la confidentialité et la sécurité des données des exploitations agricoles »



Pour que l'agriculteur américain profite pleinement des nouveaux services offerts pour valoriser ses données en toute transparence, l'American Farm Bureau (syndicalisme américain) s'est emparé du sujet en 2014. Ce sont des raisons financières et de transparence des utilisations qui en sont à l'origine. En effet, aux Etats-Unis, certaines données des exploitations étaient transférées (vendues) aux banques / négociants, mettant ces derniers en position de force dans leur relation avec l'agriculteur. De plus, une loi

environnementale oblige toute entreprise qui a connaissance d'une infraction à dénoncer son auteur pour ne pas être accusée de complicité : cela peut s'appliquer à certaines réglementations agricoles (fertilisation ou phytosanitaire par exemple).

Le Farm Bureau se met à négocier avec les collecteurs de données autour d'une charte en brandissant la menace d'en faire un sujet de lutte frontale allant jusqu'à la réglementation auprès du Congrès si aucun accord ne peut être conclu. Cette menace est pertinente puisque ces problématiques sont très fortes aux Etats-Unis (avec les GAFA – Google, Apple, Facebook et Amazon). Les collecteurs ne voulaient pas se voir accusés de déloyauté ni d'intrusion au cœur de l'exploitation. Ils ont préféré être prudents et s'asseoir à la table des négociations. Le Farm Bureau a alors missionné une économiste en chef et un représentant syndical pour animer des groupes de travail.

Les discussions ont abouti à une charte (Privacy and Security Principles for Farm Data) qui, en plus de former le public agricole sur ces sujets, dicte les principes de rédaction des conditions générales d'utilisation (CGU) des logiciels des collecteurs de données. Les CGU doivent reconnaître la propriété des données aux agriculteurs, préciser les utilisations des données. 12 organisations (37 en mars 2016) ont signé cette charte en novembre 2014 (dont Monsanto, John Deere, Syngenta). Parmi les signataires il n'y a pas celles des éleveurs, des producteurs de lait, ni des producteurs de fruits et légumes. Cette démarche centrée grandes cultures en grandes plaines représente plutôt le « Middle West ». Il y a par exemple des semenciers et des gestionnaires de l'eau.

Les collecteurs de données peuvent utiliser l'adhésion à cette charte comme outil de communication et ainsi afficher une relative transparence grâce à ce label. Mais cela reste non contraignant juridiquement. Le Farm Bureau quant à lui affirme qu'il n'y a pas eu de modification des clauses qui liaient les collecteurs de données aux agriculteurs (CGU). Ils ont donc créé un site pour expertiser les contrats des signataires et constater le respect ou non de la charte. Cet « évaluateur de transparence » est lancé le 31 octobre 2015 pour guider les choix de l'agriculteur, à partir d'une grille de lecture de 10 questions simples remplies par le fournisseur de technologies. Les conséquences du non-respect de la charte par un signataire seront connues quand l'évaluateur sera publié et mis en œuvre

Il n'y a pas l'équivalence des ITA aux USA mais un conseil agricole. Un réseau d'universités agricoles (une par Etat (Land Grant Universities)) tient un rôle de conseil en s'appuyant sur des fermes expérimentales qui appartiennent à l'Etat. Les ingénieurs de ces universités sont les interfaces entre les producteurs et la Recherche. Les universités ne sont pas signataires de la charte d'origine mais sont acteurs du système. Par exemple, l'université de l'Ohio collabore avec les industriels sur l'analyse des données et Le Farm Bureau pense à cette université pour gérer une plateforme d'agrégation des données.

En savoir plus : [www.fb.org/tmp/uploads/PrivacyAndSecurityPrinciplesForFarmData.pdf](http://www.fb.org/tmp/uploads/PrivacyAndSecurityPrinciplesForFarmData.pdf)

Sources : Version du 13 novembre 2014 et les contributions de Christophe Malvezin (Ambassade de France aux Etats-Unis) et de Marie-Cécile Damave (Saf agr'iDées)

## Les modalités juridiques d'accès aux données

**La confiance des producteurs de données repose en partie sur la transparence vis-à-vis des usages de ces données : pas d'usage, direct ou indirect, qui pourrait s'avérer contraire aux intérêts de l'agriculteur producteur de données, pas de diffusion, à titre commercial ou gracieux, à des tiers sans l'avis du producteur de données.** La confiance s'appuie aussi sur la capacité à garantir la fiabilité et la qualité des données.

Les agriculteurs semblent demander des garanties sur le fait que leurs données ne seront pas utilisées contre eux (ex. données sanitaires, spéculations). La revente des données est aussi exclue. Les droits d'accès doivent être correctement gérés, ainsi que le consentement de l'agriculteur en fonction des usages de ses données.

Si un cadre juridique concernant les données personnelles existe, la plupart des données agricoles dépassent en général ce cadre et entrent dans le champ de la relation contractuelle, dans laquelle le rapport de force peut être très déséquilibré entre un agriculteur client et un fournisseur de l'agriculture.

Un contrat entre l'agriculteur et l'exploitant des données doit nécessairement être établi, avec un consentement préalable conscient et mesuré. Ce contrat devrait porter sur toute la chaîne d'utilisation de données: l'agriculteur > le technicien > le fournisseur de service. Les conditions de ce consentement doivent pouvoir être aisément consultables et le cas échéant modifiables par l'agriculteur.

Les CGU doivent être lisibles, opposables. Elles doivent permettre la réversibilité (la possibilité de récupérer ses données), l'anonymisation au cas par cas selon le choix de l'agriculteur, le passage par des pseudonymes pour un éventuel retour en arrière (ex. pour un problème sanitaire).

Un tiers de confiance, organisme habilité à mettre en œuvre des signatures électroniques, pourrait être utilisé pour assurer les fonctions liées à l'exploitation d'un entrepôt de données.

**L'usage de données et les services connexes qui en découlent doivent répondre au principe de réciprocité : ils doivent être gagnant-gagnant.** Ils doivent créer de la valeur aussi pour l'agriculteur. Les différents niveaux de service et de coût doivent envisager la rémunération de l'agriculteur.

Concernant **les questions de « propriété »**, le discours est souvent unanime pour dire que les données liées à l'exploitation agricole appartiennent à l'agriculteur. C'est le cas des données saisies par les agriculteurs dans les différents logiciels. Et certains éditeurs de logiciels agricoles s'engagent à n'y accéder qu'à des fins uniquement de correction de bugs. Mais cela dépend des conditions générales d'utilisation et dans certains cas les usages par les éditeurs sont plus larges. Il faut noter que pour les données acquises par des tiers sur l'exploitation, les règles de propriété peuvent être différentes. Ainsi, les données d'épidémiologie-surveillance ou des réseaux de conseil peuvent être la propriété des organismes en charge de la collecte de ces

## Contexte juridique de la protection des données personnelles

La Convention européenne des droits de l'homme (CEDH) dans son article 8, protège la vie privée et familiale, le domicile et la correspondance.

Le Pacte international de l'ONU relatif aux droits civils et politiques (abrégé : Pacte ONU II) ne protège que la vie privée, sans mentionner la protection des données dans son article 17. Toutefois, l'Assemblée générale de l'ONU a adopté une résolution le 14 décembre 1990 qui fixe certains principes généraux en matière de fichiers personnels informatisés.

En Europe, la Convention de 1981 (appelée : Convention 108) protège l'individu en obligeant les États « Partie » de la Convention à adopter certains principes minimaux de protection dans leur législation concernant la transmission de données à d'autres États « Partie » de la Convention.

Toujours en Europe, la Charte des droits fondamentaux de l'Union européenne consacre expressément la protection des données à caractère personnel dans son article 8.

La Directive 95/46/CE, adoptée avant la Charte, traite de la protection des données et fixe les grands principes applicables dans l'Union européenne. Cette Directive est en cours de révision au Parlement Européen.

### L'intervention de la CNIL

La Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, version consolidée au 24 janvier 2006 qui définit la donnée à caractère personnel : « Constitue une donnée à caractère personnel toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. ». Dès lors que ces données (noms, adresses...) concernent des personnes physiques ou qu'elles permettent d'en identifier une, la loi Informatique et Libertés s'applique. Elle définit également les notions de « traitement des données à caractère personnel » qui doivent répondre à plusieurs règles (collecte transparente, finalité explicite et légitime, conservation). Suivant le périmètre, il faut suivre les procédures de déclaration ou d'autorisation à la CNIL.

### Un cadre juridique en pleine évolution

Ainsi, récemment, en octobre 2015, la Cour de justice de l'Union Européenne (CJUE) a invalidé le régime du Safe Harbor qui permettait aux entreprises américaines d'importer aux USA des données personnelles de citoyens européens depuis 2000, en présupant que celles-ci bénéficiaient d'une « protection adéquate » dès lors que des engagements standards étaient respectés. C'est un cas de jurisprudence appelé « Schrems » du nom de ce citoyen autrichien opposé à l'autorité irlandaise de protection des données qui avait autorisé l'exportation de ses données à Facebook. Le Safe Harbor est jugé comme non valide en raison des révélations d'Edward Snowden sur le programme PRISM, par lequel la NSA accèderait aux données stockées aux USA. Cette invalidation va avoir des répercussions très importantes sur les transferts et stockages des données européennes par des entreprises américaines.

données. Enfin, **la chaîne de propriété est dynamique, tout au long de la chaîne de traitement de la donnée** (source de données > traitement de la donnée > producteur de services) et si la donnée brute (avec peu de valeur) appartient à l'agriculteur, il pourra en être autrement de la donnée agrégée ou recalculée.

Il faut aussi noter la différence entre les discours et la réalité juridique. En effet, la donnée brute en tant que telle n'est pas appropriable et la propriété intellectuelle est liée à un effort de production, sur des objets tangibles. Elle est liée à l'usage et donc à une finalité de traitement. Aussi, si la donnée individuelle est difficile à protéger, la loi reconnaît un droit de propriété intellectuelle sur le contenu d'une base de données au concepteur de cette base plutôt qu'aux utilisateurs des services (voir encadré).



Les licences ouvertes pour les services et les données sont également à considérer afin de favoriser la diffusion de services et de certaines données. Il s'agit notamment des licences Creative Commons (CC) et du mouvement sur l'Open Data.

### Autres textes encadrant l'usage des données

Le secret professionnel est précisé à l'article 26 alinéa 1er de la loi du 13 juillet 1983 portant droits et obligations des fonctionnaires. Le Code pénal, dans son article 226-13, dispose que « la révélation d'une information à caractère secret par une personne qui en est dépositaire soit par état ou par profession, soit en raison d'une fonction ou d'une mission temporaire est punie d'un an d'emprisonnement et de 15000 euros d'amende »

Le Code de la Propriété intellectuelle L. 112-3 (protection de la base de données par le droit d'auteur) : « Les auteurs de traductions, d'adaptations, transformations ou arrangements des œuvres de l'esprit jouissent de la protection instituée par le présent code sans préjudice des droits de l'auteur de l'œuvre originale. Il en est de même des auteurs d'anthologies ou de recueils d'œuvres ou de données diverses, tels que les bases de données, qui, par le choix ou la disposition des matières, constituent des créations intellectuelles. On entend par base de données un recueil d'œuvres, de données ou d'autres éléments indépendants, disposés de manière systématique ou méthodique, et individuellement accessibles par des moyens électroniques ou par tout autre moyen. » Depuis la Directive du 11 mars 1996 qui a défini un régime juridique spécifique, les bases de données peuvent bénéficier d'une double protection au titre du droit d'auteur et des droits des producteurs de base de données, ces différents régimes de protection pouvant se cumuler.

Le Code de la Propriété intellectuelle L. 341-1 (protection spécifique des producteurs de base de données) : « Le producteur d'une base de données, entendu comme la personne qui prend l'initiative et le risque des investissements correspondants, bénéficie d'une protection du contenu de la base lorsque la constitution, la vérification ou la présentation de celui-ci atteste d'un investissement financier, matériel ou humain substantiel. Cette protection est indépendante et s'exerce sans préjudice de celles résultant du droit d'auteur ou d'un autre droit sur la base de données ou un de ses éléments constitutifs. »

## Les conditions de partenariats

Les modalités juridiques d'accès aux données peuvent être abordées sous différents angles : la question de la gouvernance, celle de la transparence et l'équilibre dans la chaîne de valeur.

**Une gouvernance** adaptée est une condition sine qua non pour instaurer un climat de confiance entre les acteurs qui partagent et traitent des données. Il faut envisager deux grandes modalités associées à ces différents outils juridiques.

**L'accord de partenariat.** Il s'agit de construire un partenariat au cas par cas entre ceux qui possèdent les données (coopérative, groupement,...) et les utilisateurs qui les exploitent (organisme de recherche). Il peut s'agir d'un partenariat dans le cadre d'un projet de recherche ou autre, avec un plusieurs partenaires et souvent entre acteurs publics et privés. Dans ce cas, un protocole d'accord ou une convention est conclu entre partenaires afin de régir toutes les questions de propriété des données, de leur exploitation et de leur devenir en fin de collaboration. Cela peut s'appuyer sur une convention cadre unique, des conventions bilatérales, une charte...

**Une relation fournisseur/utilisateur.** Dans ce cas, les modalités d'accès aux données sont fixées par avance aux utilisateurs via l'établissement de contrats de licences, de Conditions Générales d'Utilisation (CGU) qui régissent les droits et devoirs des utilisateurs. Cette modalité a pour avantage de favoriser l'accès aux données en apportant de la lisibilité et des latitudes pour exploiter et réutiliser les données (exemple de Vigicultures® ou des licences d'utilisation des données Météo France). Il peut y avoir éventuellement différentes licences en fonction des usages (utilisation pour la recherche ou pour un service commercial). L'Open Data s'appuie sur des licences particulières (exemple de [data.gouv.fr](http://data.gouv.fr)).

## Les moyens et services pour accéder aux données et en garantir les usages

### La standardisation des données

Afin de favoriser l'interopérabilité entre systèmes d'informations, différents acteurs organisent et normalisent les échanges de données informatisés dans les secteurs agricoles et agro-industriels. Ce travail s'appuie notamment sur la mise au point d'ontologies. La normalisation des flux de données, notamment ceux des objets connectés qui ne suivent pas de standard aujourd'hui, est un prérequis à la valorisation mobilisant des données de différentes sources. Il est aujourd'hui indispensable de promouvoir l'utilisation des standards déjà existants, dont certains restent insuffisamment utilisés et d'identifier les besoins prioritaires en nouvelles normes.

## Les interfaces entre application (API)

Pour que ces données « ouvertes » soient exploitables notamment par des machines, une solution consiste à mettre en place une ou plusieurs API, ou Interface de programmation applicative (Application Programming Interface). Ces interfaces, à destination de développeurs, permettent de faire communiquer une application avec le système d'information d'une entreprise (utilisation de service ou accès à des bases de données).

On distingue deux types d'API. Les API publiques sont ouvertes à tout le monde. Différents niveaux de services peuvent être proposés et des restrictions d'utilisation sont souvent imposées. Elles diffusent des services généraux, utiles pour chacun. Les grandes entreprises du Web misent une partie de leurs revenus sur ce type d'API (Facebook, Twitter, Amazon). Les API privées sont développées pour des besoins en interne ou pour des partenaires. Les services ou les données proposées sont plus sensibles, et leur public plus restreint. Les API privées et publiques sont différenciées par la notion de contrat. Les API privées font l'objet d'un accord entre les deux parties (fournisseur et développeur) matérialisé par un contrat. Les API publiques prévoient uniquement des conditions d'utilisation garantissant le respect des droits du fournisseur.

Les plateformes de gestion d'API mettent à disposition d'un fournisseur un ensemble de services lui permettant de gérer toutes les contraintes liées à la publication d'une API (par exemple : gestion des accès et sécurité, monétisation). Elles agissent un peu comme un portail filtrant les communications entre l'API et les applications conçues par les développeurs externes.

## Passeport de la donnée

Aujourd'hui, l'innovation liée à la valorisation des données des agriculteurs risque de rester concentrée sur les seuls collecteurs de données. En effet, chaque fournisseur d'objet connecté propose sa plateforme d'hébergement des données générées associée. Souvent, l'application fournie pour valoriser ces données ne répond qu'à un besoin / une problématique précise. L'utilisateur doit alors manipuler plusieurs applications (et les comptes d'accès spécifiques à chaque fois !).

Si l'on considère que ce n'est pas forcément celui qui collecte les données qui en proposera la meilleure valorisation, et que la valeur

### L'expérience Terms of Service (TOS) : Didn't Read

Ce projet communautaire vise à analyser et classer les termes de politiques de service et de confidentialité des principaux sites et services Internet. Chaque aspect d'une TOS ou d'une politique de confidentialité est évalué comme positif, négatif ou neutre. Les services sont classés de A (meilleure) à E (pire) une fois qu'un consensus est trouvé entre les membres du projet. Le nom du projet « Didn't Read » fait référence au fait que généralement personne ne lit les conditions générales, trop longues et complexes.

# API-AGRO : plateforme ouverte et innovante au service de l'agriculture numérique



## Historique et contexte

API-AGRO vise à améliorer l'interopérabilité et les échanges de données entre les différents acteurs de l'écosystème agricole. L'idée est de mutualiser des références agronomiques au sens large pour faciliter l'évaluation des systèmes agricoles. Cette plateforme offre une approche basée sur les API (Application Programming Interface), canal de distribution permettant le partage de données avec des règles de diffusion et d'usage claires.

Cette plateforme a été conçue dans le cadre d'un projet de R&D (2013-2016) regroupant l'essentiel des ITA et piloté par l'ACTA. Elle a été conçue à partir de l'inventaire des références agronomiques présentes au sein des partenaires du projet.

## Les outils/ les services proposés

API-AGRO facilite la réutilisation des données de référence des ITA, notamment par des acteurs extérieurs. Elle permet d'autoriser l'accès d'une application à un ensemble de données, de créer de la valeur par la structuration, la connexion et la diffusion des données et d'offrir des possibilités de partenariats de co-développer des applications innovantes.

Elle propose un accès unifié à un jeu de données (public ou à accès restreint) ainsi qu'aux API publiques ou privées (réservées à des partenaires spécifiques ou certains clients).

A ce jour, la plateforme contient des dizaines de ressources pouvant contenir de quelques centaines d'enregistrements pour des référentiels (sol par exemple), à plusieurs millions d'enregistrements pour les données Météo France (modèles Arôme et Arpege).

La question de l'accès aux données

Après la mise en ligne de la plateforme de gestion d'API et de diffusion de données, des licences seront établies pour rendre transparente la diffusion de ces référentiels au plus grand nombre d'utilisateurs. Un consortium d'Instituts Techniques Agricoles portera la gouvernance de cette plateforme.

En savoir plus : [www.api-agro.fr](http://www.api-agro.fr)



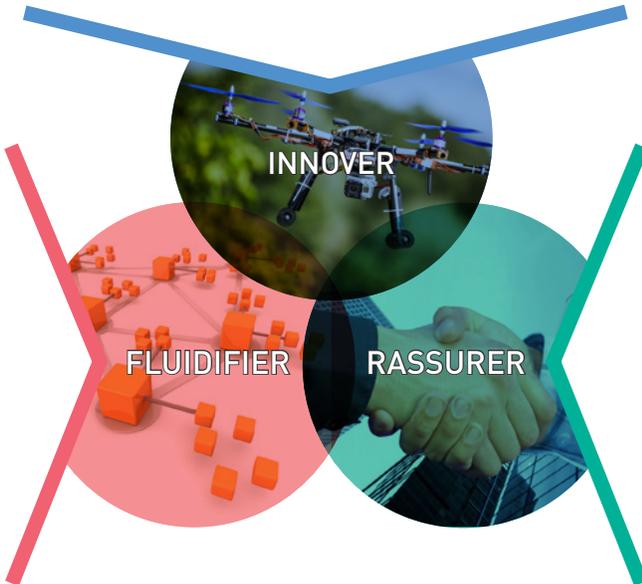
de ces données ne se réalisera qu'en considérant toutes les données à la fois, on imagine l'intérêt pour un utilisateur de donner l'accès à ses données à d'autres acteurs pour accéder à leurs services. En plus de bénéficier d'une plus large offre de services valorisant la créativité de la multitude, l'utilisateur reste maître de ses décisions en s'affranchissant de la dépendance d'un seul fournisseur.

Toutefois, cette créativité ne pourra s'exprimer que si différents gisements de données sont accessibles, interopérables tout en garantissant la maîtrise de ces données à leur propriétaire (dans notre cas, l'agriculteur). Ces propriétés pourraient être offertes par une passerelle de données agricoles à caractère collectif.

- Accessible signifie que les données sont rendues disponibles pour qu'un développeur puisse les consommer (sous réserve du consentement de son propriétaire) dans son application. Cela nécessite de parler un langage commun avec les applications clientes. La passerelle promouvrait les interfaces de programmation d'application (API) pour les échanges de données.
- Interopérable signifie que les données sont normalisées à l'entrée de la passerelle et que les applications qui les consomment peuvent en interpréter la structure et fédérer les données de sources diverses.
- La garantie de la maîtrise des données nécessite une totale confiance des utilisateurs dans les utilisations qui seront faites de ces données. Ceci sera rendu possible par l'implication d'un tiers de confiance, pour fédérer les authentifications d'accès aux données, gérer finement les consentements d'utilisations des données par tel ou tel fournisseur de services en affichant en toute transparence les conditions d'utilisation de ces services. La technologie émergente de la Blockchain (stockage et de transmission d'informations, transparence, sécurité, sans organe central de contrôle) est également à envisager et permet aux acteurs de se passer d'un tiers de confiance, la confiance étant assurée par un fonctionnement en réseau.

**L'agriculteur évoluerait au sein d'un espace personnalisé dans lequel les informations collectées sur son exploitation seraient accessibles, car récupérées avec son accord (authentification mutualisée) auprès des collecteurs de données par API. Il déléguerait à un tiers de confiance la gestion du consentement et l'authentification pour l'accès à ses données par des services tiers.**

## Innover – Fluidifier – Rassurer : 10 recommandations pour favoriser l'accès et la valorisation des données



Sur la base de l'ensemble de ces constats, les Instituts Techniques Agricoles établissent 10 recommandations pour proposer de **nouveaux services innovants** basés sur la valorisation des données, **fluidifier leur échange** tout en rassurant les acteurs agricoles. Ces recommandations s'adressent aux acteurs de R&D mais également aux autres acteurs de la filière dans un mode de partenariat ouvert et basé sur la confiance entre les tiers. Les agriculteurs et leurs représentants doivent également se saisir de la question et se mobiliser pour que ces recommandations se concrétisent. Les Instituts Techniques Agricoles resteront vigilants pour assurer les intérêts de la recherche et de l'innovation au service des agriculteurs en favorisant l'accès aux données agricoles et surtout leur valorisation.

## INNOVER

### Open Innovation : co-construire les services en mobilisant de nouveaux partenariats et compétences

Pour développer l'innovation ouverte, et diffuser rapidement des services innovants sur le terrain, il est nécessaire d'impliquer les utilisateurs finaux. L'ouverture à de nouveaux partenariats et compétences doit permettre aux acteurs de la R&D d'assurer les innovations de demain.

#### 1. Co-construire des services numériques utiles aux acteurs agricoles

La valorisation des données doit aboutir à des services innovants, utiles et ergonomiques pour les agriculteurs et les autres acteurs agricoles. Pour cela, nous devons mieux intégrer les utilisateurs finaux dans leur développement, dès les premières idées. Cela garantira des services plus personnalisés et adaptés aux usages opérationnels des acteurs de terrain. Cette démarche de co-construction peut s'appuyer sur des développements adaptatifs sur la base de prototypes et de tests successifs.

- un travail avec les agriculteurs sur les services du futur
- des services plus adaptés

#### 2. Evaluer les applications dans les exploitations

Il s'agit de renforcer l'implication des utilisateurs finaux des services en leur faisant tester ces nouveautés très en amont, éventuellement durant leur construction. Les réseaux d'agriculteurs (notamment des futurs *living labs*) et les fermes expérimentales connectées (Digifermes®) sont des dispositifs d'excellence méthodologique pour assurer le suivi des tests et ainsi évaluer les services de manière objective.

- des tests en conditions réelles
- des évaluations objectives des services



### 3. S'ouvrir à de nouveaux partenariats

Pour favoriser les travaux innovants et ouvrir de nouvelles perspectives, il faut s'appuyer sur des collaborations avec des entreprises et des organismes de recherche, au niveau national et international. Des partenariats privilégiés doivent être noués avec les organisations agricoles collectives (fédérations professionnelles, interprofessions, ...) afin de valoriser leurs données. Des échanges avec les entreprises des filières doivent permettre d'évaluer leurs besoins en valorisation des données. Ce partenariat doit pouvoir impliquer également des start-up afin de profiter de leur dynamisme, et des acteurs extérieurs au monde agricole pour compléter mutuellement les domaines de compétence et co-développer de nouveaux outils au sein des Digifermes®. Une veille collective doit être organisée afin d'enrichir ces partenariats. Les procédures de conventionnement doivent se simplifier afin de gagner en efficacité tout en assurant un partenariat équilibré. Des ateliers de créativité sur des thématiques précises pourraient compléter ces dispositifs.

- une veille stratégique
- un partenariat plus riche et ouvert
- un partenariat efficace et équilibré
- des ateliers de créativité

### 4. Renforcer les compétences pour mieux valoriser les données

Les compétences en gestion de données et analyse de données sont précieuses, en particulier pour exploiter le Big Data Agricole. Ces compétences, à l'interface entre statistique et informatique, mais aussi en modélisation et système d'information, sont à rechercher dans les futurs recrutements et dans des formations ambitieuses des personnels. Favoriser le transfert de compétence ou mutualiser des compétences entre les instituts est également un levier important. Un plan de formation doit permettre d'insuffler une culture commune et partagée entre les différents acteurs de la R&D sur les questions d'échange de données et de leur valorisation, mais aussi de former concrètement aux outils et méthodes nécessaires.

- un plan de formation ambitieux
- des collaborations avec les écoles
- des recrutements de nouveaux talents
- une mutualisation des compétences

## FLUIDIFIER

### Mobiliser les technologies pour fluidifier les échanges de données

Les nombreux progrès en informatique et sur la normalisation des données offrent aujourd'hui la possibilité d'améliorer les échanges de données.

#### 5. Inventorier les sources de données d'intérêt pour l'agriculture

Il s'agit de proposer un inventaire des plateformes de données ouvertes en relation avec les secteurs de la production agricole, de l'agroalimentaire, de l'agrofour-niture, de la dynamique du monde rural. Différents groupes (portail prévu dans la mission agriculture 2025, travaux du GIS Relance Agronomique, GIS Elevage Demain, Inventaire API-AGRO) ont déjà commencé cet inventaire mais une mise en commun et une promotion restent nécessaires.

- un inventaire des sources de données

#### 6. Favoriser l'interopérabilité entre les systèmes d'information

De nombreux services et données existent, mais leurs utilisations restent parfois limitées car imaginées dans le cadre d'outils très souvent fermés. A destination de développeurs, les API (Interface de programmation applicative) doivent permettre de rendre une application plus accessible pour d'autres usages que ceux prévus initialement et d'aider l'intégration de services en facilitant la communication de l'application avec les autres éléments du système d'information. La construction et la diffusion de standards de données ouverts et sans restriction d'accès ni d'utilisation doivent contribuer à cette interopérabilité.

- des données et services accessibles par API
- des services plus intégrés

#### 7. Mutualiser les systèmes d'information

Certains organismes ou communautés ont investi depuis des années sur des systèmes d'informations performants permettant d'enregistrer et d'extraire les données d'expérimentation ou des réseaux observations. Une analyse et une comparaison des solutions mises en œuvre doivent permettre d'évaluer les opportunités de mutualisation de certains services ou des transferts de compétences.

- des systèmes d'information performants et ergonomiques
- des données capitalisées

## RASSURER

### Clarifier les questions de propriété des données et de transparence sur leurs usages pour rassurer les acteurs

#### 8. Etablir des principes et des bonnes pratiques

Pour favoriser l'accès aux données et services, nous souhaitons qu'une charte soit établie en s'appuyant sur une réflexion collective des différents acteurs du monde agricole. Elle devra laisser une large place aux agriculteurs, principaux fournisseurs de données et utilisateurs de services. Pour éclairer le débat, des échanges sur les Conditions Générales d'Utilisation (CGU) ou licences existantes permettront d'identifier les marges de progrès pour gagner en transparence envers les utilisateurs, favoriser le partage de données et de services. Il s'agit notamment de favoriser l'accès à ces nouvelles données pour les travaux menés au service des agriculteurs et des filières.

- une charte sur l'accès et la valorisation des données agricoles
- des services et des données connectés en toute transparence
- un engagement des ITA à mettre en application ces bonnes pratiques

#### 9. Garantir la maîtrise des données et établir une chaîne de confiance

Au-delà du service rendu, l'accès aux données des agriculteurs est conditionné à une bonne compréhension de l'utilisation de leurs données qui doit se faire en toute transparence. Nous proposons de mettre en place des outils permettant de solliciter le consentement des agriculteurs pour assurer la maîtrise de leurs données. Il s'agit également de gérer la sécurité des accès aux données. Tester et promouvoir les outils permettant de constituer un véritable passeport de la donnée, c'est-à-dire qui assure la traçabilité de la propriété des données, mais également du consentement donné quant à leur utilisation par des tiers pour des usages définis. Ces outils doivent permettre aux agriculteurs de contrôler le plus simplement possible leurs droits et consentements afin de pouvoir enrichir leurs accès à des services ou au contraire révoquer certains accords.

- un passeport pour les données agricoles
- des utilisations transparentes et révocables

#### 10. Encourager l'ouverture des données des entreprises

L'ouverture des données peut occasionner des coûts de développement informatique sans forcément avoir de retours sur investissement immédiats. Des dispositifs incitatifs d'aide financière ou technique pourraient être mis en place pour accompagner ces acteurs dans l'ouverture de leurs données.

- un appui financier ou technique pour ouvrir les données

## Scénarios : quelle agriculture numérique dans le futur proche ?

Afin de rendre plus explicites les enjeux liés à la mise en œuvre de ces recommandations, nous proposons une mise en situation avec un exemple fictif des conséquences au niveau d'une exploitation agricole, de son environnement et des acteurs de la R&D collective. A court terme (3 ans), les contextes technologique et agricole sont relativement connus, mais, de leur côté, les fournisseurs de matériels agricoles et de services sont en pleine mutation.

### Contexte

Mars 2019 - Hervé est agriculteur sur une exploitation de 270 ha produisant des grandes cultures, du lait (troupeau de 150 vaches) et des poulets (3 bâtiments, 4500 m<sup>2</sup>). Ce matin, il rejoint ses 2 associés afin d'examiner le plan d'investissement de l'année à venir. L'exploitation s'est engagée depuis quelques années dans le numérique, en lien avec une automatisation croissante des travaux : machines agricoles connectées, estimation de l'état des cultures et des animaux, outils de prévision et d'aide à la décision...

Selon les scénarios, l'avenir de l'exploitation d'Hervé et de ses associés ne se dessine pas de la même manière...



## Scénario 1 : La privatisation exponentielle des données produites sur les exploitations agricoles par les fournisseurs de l'agriculteur

Les fournisseurs de l'agriculteur (fournisseurs d'intrants, de matériels, de services) développent et proposent aux agriculteurs des outils propriétaires qui visent à simplifier le travail et la relation client : les applications « tableaux de bord » se multiplient, plus ou moins couplées avec des outils d'aide à la décision. Tous rivalisent d'imagination pour proposer leur outil maison, qui sait faire ce que les autres ne savent pas encore réaliser. Cette frénésie est stimulante, les idées foisonnent, mais ces outils restent très sectorisés, ne portent que sur un équipement ou un produit d'un seul fournisseur. Par ailleurs, très peu d'informations sur leur fonctionnement ou leur qualité sont disponibles.

Loin de simplifier et d'améliorer l'analyse croisée des données de l'exploitation, cette situation tend à cloisonner encore un peu plus les données produites sur les fermes. Et pourtant, tout le monde sent bien que c'est le croisement de ces sources d'informations qui apportera une amélioration de la gestion des exploitations.

Pour tenter d'apporter une solution à cette tendance, des entreprises d'amont et/ou d'aval ont conclu entre elles des accords commerciaux pour offrir des gammes d'outils complémentaires et interopérables, créant des passerelles entre leurs systèmes propriétaires ou tableaux de bord communs. Mais ces accords, pour la plupart bilatéraux, n'ont pas réellement résolu le problème. Et dans les faits, une part croissante des agriculteurs a le sentiment de perdre une partie de sa liberté de décision : en effet, le choix d'un partenaire ou fournisseur entraîne l'obligation de travailler avec tous les autres partenaires du consortium si l'on souhaite pouvoir bénéficier de l'interopérabilité et de la convergence des outils numériques. Au final, le volume de données produites s'accroît toujours plus, mais leur interconnexion à l'échelle de l'exploitation et leur mise en valeur au service de l'agriculteur restent très imparfaites, au détriment des agriculteurs qui ne peuvent en exploiter tout le potentiel.



*Depuis plusieurs mois, Hervé et ses associés envisagent d'investir dans un robot de traite. L'idée est dans la tête des 3 associés depuis bien plus longtemps, mais le projet a été différé plusieurs fois, compte tenu de la difficulté des choix à réaliser. Les fournisseurs de matériel proposent en effet des solutions de suivi de troupeaux très intéressantes, qui présentent de surcroît la possibilité de s'interfacer avec les outils numériques fournis par les coopératives de collecte. Mais il reste un obstacle majeur, qui est l'objet une nouvelle fois de la discussion du jour entre les associés : l'équipement proposé par l'opérateur, lié par des accords commerciaux exclusifs à leur coopérative, présente des caractéristiques techniques mal adaptées à leur configuration de bâtiments. Un autre opérateur propose un matériel bien mieux adapté mais l'intégration au sein du système d'information de l'élevage suppose de changer de partenaire de collecte afin d'adopter des outils compatibles. Sur l'exploitation, personne ne souhaite rompre la relation établie avec cette coopérative depuis maintenant plus de 10 ans et qui donne entière satisfaction par ailleurs.*

*Hervé et ses associés ont bien essayé d'explorer les possibilités de mise en relation de bases de données, mais les formats propriétaires compliquent considérablement la tâche, et entraînent des coûts supplémentaires. Et le changement de coopérative, outre qu'il ne se justifie par aucun autre motif d'insatisfaction des 3 associés, se traduirait par la perte des données de collecte de 4 campagnes successives, base de données constituée par la coopérative qui revendique la propriété des données sans droit pour les éleveurs de les récupérer. Hervé avait dû lire un peu trop rapidement les conditions générales d'utilisation lors de la souscription du service, 4 ans plus tôt. Le débat risque d'être animé entre les associés, et la solution finale insatisfaisante dans tous les cas de figure. Hervé soupire, un brin fataliste : les solutions techniques existent pour donner à l'agriculteur la capacité de s'approprier pleinement le système d'informations de l'exploitation, mais les stratégies commerciales viennent encore une fois contrarier une partie des marges de progrès.*

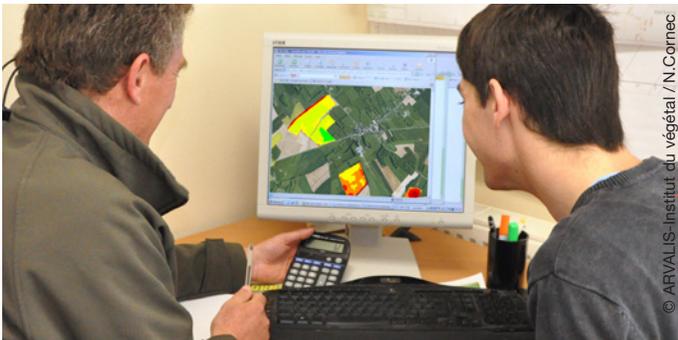
Le secteur de la R&D agricole collective subit également cette atomisation des sources de données. La construction de références et d'observatoires au service des producteurs et des filières se trouve compliquée par la multiplicité des acteurs et l'hétérogénéité des sources de données, de leur nature et de leurs formats. Sans compter la gestion des consentements. Le temps passé en négociation, collecte, standardisation, harmonisation, consolidation des bases de données est jugé très lourd par les acteurs de la R&D agricole, mais aussi par leurs partenaires, qui attendent des délais de valorisation toujours plus courts. Cette situation complique également l'intégration de ces données dans divers travaux de recherche (développement d'OAD, outils de diagnostic, etc.). Sans pouvoir mobiliser ces nouvelles données d'exploitations, la R&D agricole s'éloigne des préoccupations d'une partie de ces mandants, les agriculteurs engagés « pieds et poings liés » dans cette révolution numérique.

## Scénario 2 : L'agriculteur au cœur du système d'information

Les équipements et matériels connectés se sont multipliés ces dernières années sur les exploitations agricoles, et la masse de données collectées aussi. Ces données présentent bien entendu un intérêt en elles-mêmes, mais le potentiel d'analyse est décuplé dès qu'on les met en regard les unes des autres. Quelques opérateurs, partenaires des agriculteurs, ont bien compris cet enjeu, et ont également compris qu'ils ne parviendraient pas à contrôler toutes les chaînes de production de données, même en multipliant les partenariats commerciaux. Une stratégie de segmentation des métiers s'est alors mise en place. On a vu l'émergence de deux nouveaux acteurs :

- Des sociétés de « connexion de données », qui pour certaines ont pris le statut de coopératives, proposent aux agriculteurs de rassembler l'ensemble de leurs données (physiquement ou virtuellement) au sein de portails communs et ouverts. Ces portails, en plus d'offrir cet accès centralisé aux données de l'exploitation, mettent à disposition une batterie d'outils logiciels standard d'analyses (extraction des données, représentation graphique, etc.) et proposent un récapitulatif des droits et consentement de l'agriculteur sur ses données.
- Des sociétés dont la stratégie repose sur l'analyse des données massives et hétérogènes, mais également la construction de tableaux de bord personnalisés pour les agriculteurs. Elles répondent aux besoins des agriculteurs en quête d'une analyse fine et transparente des informations produites par l'ensemble des capteurs de leur exploitation et par leurs partenaires d'amont et d'aval.

La séparation des métiers a permis des stratégies plus cohérentes de la part des entreprises des filières. Certains fabricants de matériel sont venus chercher de la valeur sur ces deux métiers, tandis que d'autres ont préféré concentrer tout leur savoir-faire sur l'innovation autour de leur produit ou métier d'origine. Cette segmentation a permis également l'entrée sur le marché agricole de nouveaux acteurs, ce qui redynamise l'offre de service. Cette approche, renforcée par une volonté de faire émerger des standards (nature de données, codage, format, ...), a permis d'améliorer sensiblement l'exploitation des données.



*Depuis un an, Hervé et ses associés sont très satisfaits de leur dernier investissement : un robot de traite. La masse de données collectées est désormais considérable, et depuis quelques mois les 3 associés de l'exploitation ont un projet : mieux capitaliser sur ces données afin d'optimiser encore un peu le pilotage de l'exploitation, tant sur le plan technique qu'économique. Hervé est en relation depuis quelques semaines avec deux sociétés, FarmData, une société spécialisée dans le traitement des données de productions agricoles hétérogènes, et Agri-B.I., une filiale de sa coopérative, qui propose des services similaires avec une composante d'analyse économique plus prononcée (optimisation des flux entre les agriculteurs et leurs fournisseurs/clients d'amont et d'aval). La réunion de ce matin vise à auditer les deux entreprises. Le consultant d'ADW (AgroDataWarehouse), société qui assure la centralisation des données de leur exploitation depuis 4 ans au sein d'un même portail, sera également présent afin de valider la faisabilité des options techniques proposées par les deux sociétés candidates. Avec l'accroissement de la taille des structures, les agriculteurs sont de plus en plus nombreux à investir dans ces outils et analyses de « business intelligence ». Le projet s'annonce bien, Hervé et ses associés ont attendu de mûrir leur projet avant de se lancer, mais désormais, ils en sont certains, ces nouvelles possibilités d'optimisation valorisant les outils numériques vont leur permettre d'améliorer le fonctionnement de l'exploitation. Au final, cela sera au bénéfice de la performance économique et de la qualité du travail sur la ferme.*

Pour tous les acteurs de la R&D (privés et collectifs agricoles), cette organisation des flux de données issues des exploitations est un accélérateur d'innovation : les étapes fastidieuses de collecte, standardisation, harmonisation... sont maintenant assurées par les sociétés de « connexion de données » qui jouent le rôle d'assembleur et de fournisseur de données au service des producteurs. Les acteurs de la R&D peuvent donc se concentrer sur leur cœur de métier : produire de la référence et du conseil pour la ferme France, bénéficier des nouveaux flux de données pour explorer de nouvelles pistes de traitement et de valorisation, évaluer de manière objective les différentes innovations du secteur notamment les outils de prévision et d'aide à la décision, intégrer ces avancées dans des outils au service des agriculteurs. Par ailleurs, la gestion des consentements des producteurs est gérée très simplement au niveau de ces portails de données : l'agriculteur conserve donc la maîtrise et la connaissance de la destination des flux de données issus de son exploitation, et peut à tout moment décider de modifier la liste des bénéficiaires de ces données. Cette facilité et transparence accroît considérablement le volume et la représentativité des données mises à disposition de la R&D qui peut mieux répondre aux attentes des agriculteurs.

## Liste des personnes consultées

Ces personnes étaient toutes présentes au workshop du 8 octobre 2015.

Nom	Prénom	Institut
AUBRY	Alexia	IFIP
BALVAY	Béatrice	Institut de l'Elevage
BESNARD	Julie	ARVALIS - Institut du végétal
BIDANEL	Joël	IFIP
BRIEND	Guillaume	MEITO - Programme AgreTIC
BRUN	François	ACTA
CHANET	Jean-Pierre	Irstea
CIMINO	Mariane	ACTA
COLINET	Julie	ACTA
COLLOT	Laurent	CTIFL
DAMAVE-HENARD	Marie-Cécile	saf agr'iDées
DANGLEANT	Alexandre	ITSAP - Institut de l'abeille
DE COSNAC	Gérald	Terres Inovia
DIAZ	Alexandre	ISAGRI
GALLOT	Sylvain	ITAVI
GERVOIS	Sébastien	Terres Inovia
GOURDAIN	Emmanuelle	ARVALIS - Institut du végétal
HAEZEBROUCK	Théo-Paul	ACTA
HERMAN	Mathias	Orange
LAUGA	Bruno	ARVALIS - Institut du végétal
LE BRIS	Xavier	ARVALIS - Institut du végétal
LECOUVEY	Philippe	ACTA
MORCRETTE	Nathalie	INRA
MOUILLERON	Marc-Antoine	Orange
MUNIER-JOLAIN	Nicolas	INRA
NODE-LANGLAIS	Rémi	ATOS-WORLDBLINE
PAGES	Benoît	ARVALIS - Institut du végétal
PIEDNOIR	Benoist	Institut de l'Elevage
PINCHET	Xavier	Terres Inovia
REHBEIN	Erik	Institut de l'Elevage
SALVI	Frédéric	Terres Inovia
SINE	Mehdi	ARVALIS - Institut du végétal
TALES	Romain	Etalab
VAN RUYSKENSVELDE	Jean-Pierre	IFV (Institut Français de la Vigne et du Vin)
VISSAC	Philippe	ACTA
VOISIN	Thomas	SMAG
SAVARY	Alain	AXEMA

## Informations sur la publication

Cet ouvrage est une production du Réseau Numérique et Agriculture des Instituts Techniques Agricoles.

Les principaux contributeurs sont :



*François Brun,  
Julie Colinet,  
Mariane Cimino,  
Théo-Paul Haezebrouck*



*Mehdi Siné,  
Bruno Lauga,  
Julie Besnard*



*Sylvain Gallot*

Les participants du Workshop du 8 octobre 2015 ont contribué au document à travers leurs présentations d'exemples, leur participation aux différents ateliers organisés et la relecture du document.

<http://www.acta.asso.fr/numerique>  
[numerique@acta.asso.fr](mailto:numerique@acta.asso.fr)

Réalisation : Violaine Lejeune, François Brun, Francis Villain

ACTA – Les Instituts Techniques Agricoles  
149 rue de Bercy, 75595 Paris Cedex 12

 @ACTA\_asso  
[www.acta.asso.fr](http://www.acta.asso.fr)



Formaté typographiquement par IFIP-Institut du porc  
Imprimé en France en octobre 2016 par SEPEC

ISBN : 978-2-85794-298-6

# 10

## Recommandations pour favoriser l'accès et la valorisation des données

Le développement du numérique dans les exploitations agricoles conduit à une situation inédite en matière de collecte des données sur les exploitations. La création de données se déconcentre rapidement et les agriculteurs en deviennent des fournisseurs incontournables.

La problématique de l'accès aux données agricoles pour les activités de recherche et d'innovation devient centrale. Les Instituts Techniques Agricoles présentent leur analyse et proposent 10 recommandations autour de 3 axes : "Innover", "Fluidifier", "Rassurer".

#FLUIDIFIER

#INNOVER

#RASSURER



**acta**  
LES INSTITUTS  
TECHNIQUES  
AGRICOLES #

149, rue de Bercy  
75595 Paris Cedex 12  
tél. +33 (0)1 40 04 50 10  
 @ACTA\_asso  
[www.acta.asso.fr](http://www.acta.asso.fr)  
contact : [numerique@acta.asso.fr](mailto:numerique@acta.asso.fr)  
ISBN : 978-2-85794-298-6